

# Lineare Algebra I

Wintersemester 2011/2012

Universität Bayreuth

MICHAEL STOLL

## INHALTSVERZEICHNIS

1. Einige allgemeine Vorbemerkungen	2
2. Die Sprache der Mathematik: Logik und Mengenlehre	3
3. Algebraische Strukturen: Gruppen, Ringe, Körper	20
4. Der Körper der komplexen Zahlen	26
5. Vektorräume: Definition und Beispiele	30
6. Untervektorräume	35
7. Erzeugendensysteme	39
8. Lineare Unabhängigkeit, Basis und Dimension	44
9. Lineare Abbildungen	58
10. Matrizen	71
11. Der Normalformalgorithmus und Lineare Gleichungssysteme	76
12. Matrizen und lineare Abbildungen	88
13. Die Determinante	93
14. Eigenwerte und Eigenvektoren	102
15. Diagonalisierbarkeit	112
16. Bilinearformen	117
17. Volumina	125

## 1. EINIGE ALLGEMEINE VORBEMERKUNGEN

Die meisten von Ihnen kommen mehr oder weniger direkt von der Schule (die prominenteste Ausnahme sind die zukünftigen Realschullehrer/innen, die die Lineare Algebra I erst im dritten Semester hören). Das Erste, das Sie sich zu Beginn Ihres Mathematik-Studiums klar machen müssen, ist, dass das, was Sie in der Schule unter der Bezeichnung „Mathematik“ kennen gelernt haben, nicht wirklich Mathematik ist. Das bedeutet, dass Sie hier an der Universität im Grunde auf völlig andere Art und Weise gefordert sein werden als an der Schule. Das heißt jetzt nicht, dass Sie die ganze Schulmathematik vergessen können — manches kann als Beispielmateriale noch nützlich sein, und es schadet auch nicht, wenn man eine gewisse Fertigkeit im Rechnen hat, wie man sie an der Schule lernt.

Was folgt aus diesem in Deutschland leider traditionellen Bruch zwischen Schule und Universität?

- Die meisten von Ihnen werden sich erst einmal sehr schwer tun. Das ist völlig normal und kein Grund zur Beunruhigung.
- Wenn Sie in der Schule in Mathe sehr gut waren, heißt das nicht, dass Ihnen die Mathematik an der Universität auch leicht fällt. Umgekehrt kann es sein, dass Ihnen die Mathematik an der Schule langweilig war und Sie dann hier auf den Geschmack kommen.
- Sie sollten nicht erwarten, den Stoff sofort während der Vorlesung zu verstehen. Das Nacharbeiten der Vorlesung ist sehr wichtig, da man Mathematik nur verstehen kann, wenn man darüber nachdenkt. (Das Modulhandbuch sieht drei Stunden pro Woche dafür vor.) Ganz wichtig ist auch, dass Sie die Übungsaufgaben bearbeiten, denn richtig versteht man den Stoff erst, wenn man ihn anwendet. (Das Modulhandbuch sieht dafür fünf Stunden pro Woche vor.) Dabei hilft es, gemeinsam in kleinen Gruppen zu arbeiten, denn für das Verständnis ist es ungemein förderlich, wenn man versucht, jemand anderem etwas zu erklären.
- Für diejenigen von Ihnen, die Lehrer/innen werden wollen, heißt das umgekehrt auch, dass Sie den größten Teil von dem, was Sie hier lernen, in der Schule nicht direkt verwenden können. Es ist zu hoffen, dass sich das bald einmal ändert und Sie die Möglichkeit haben werden, die „richtige“ Mathematik Ihren Schülern nahezubringen. In jedem Fall sollte Sie die Ausbildung, die Sie an der Universität erhalten, in die Lage versetzen, Ihren Unterricht innerhalb der Mathematik einzuordnen und weiter gehende Fragen Ihrer Schüler/innen souverän zu beantworten.

Lassen Sie sich von den Schwierigkeiten am Anfang nicht zu sehr frustrieren! Bei den meisten von Ihnen wird in den ersten beiden Semestern der Groschen fallen. Falls Sie aber nach zwei Semestern immer noch das Gefühl haben, nichts zu verstehen, dann kann es auch sein, dass das Mathematikstudium doch nicht das Richtige für Sie ist.

Ich habe in dieses Skript an manchen Stellen Links zu Webseiten eingebaut, die so aussehen (dieser Link führt auf meine Homepage). Die meisten davon verweisen auf die Wikipedia, die für den Zweck einer ersten Orientierung meistens gut geeignet ist. (Als Hauptquelle für Zitate in einer wissenschaftlichen Arbeit wie z.B. einer Bachelor- oder Masterarbeit ist die Wikipedia aber nicht geeignet. Da müssen Sie Lehrbücher oder Fachartikel zitieren.)

## 2. DIE SPRACHE DER MATHEMATIK: LOGIK UND MENGENLEHRE

Worum geht es nun in der Mathematik?

Die Wikipedia schreibt:

Für *Mathematik* gibt es keine allgemein anerkannte Definition; heute wird sie üblicherweise als eine Wissenschaft beschrieben, die selbst geschaffene abstrakte Strukturen auf ihre Eigenschaften und Muster untersucht.

Es geht also unter anderem um *Abstraktion*. Man abstrahiert von den speziellen Eigenschaften, die man in verschiedenen Situationen vorliegen hat, und zieht das Gemeinsame heraus. Dann versucht man auf der Grundlage nur dieser wesentlichen Merkmale möglichst viele Aussagen abzuleiten, die dann auf *alle* Situationen zutreffen, die diese Merkmale aufweisen. Dies geschieht durch den zentralen Vorgang aller mathematischen Tätigkeit, nämlich durch das Führen eines *mathematischen Beweises*. Zugespitzt kann man sagen, dass ein *Mathematiker* der- oder diejenige ist, die oder der in der Lage ist, einen solchen Beweis zu führen:

- Das wichtigste „Lernziel“ in den Grundvorlesungen besteht darin, dass Sie lernen, wie man mathematische Beweise führt.

Sie sollen hier natürlich auch und nicht zuletzt Ergebnisse und Methoden der Linearen Algebra kennen lernen, aber ohne die mathematische Grundfertigkeit des Beweisens würde Ihnen das kaum etwas nützen.

Bevor wir damit beginnen können, müssen wir die Vokabeln und Grammatik der Sprache der Mathematik lernen. Mathematische Aussagen und Beweise werden in der Sprache der *Logik* formuliert; die Objekte, von denen die Rede ist, in der Sprache der *Mengenlehre*. Beide werden wir hier kurz einführen (oder wiederholen, je nachdem wie viel Sie davon schon aus der Schule kennen). Es handelt sich um das „Handwerkszeug“, mit dem Sie täglich zu tun haben werden, also passen Sie gut auf!

## 2.1. Aussagenlogik.

Die Aussagenlogik verknüpft mathematische *Aussagen* (die wahr oder falsch sein können) miteinander und untersucht, wie das Wahr- oder Falschsein einer zusammengesetzten Aussage von den beteiligten Aussagen abhängt.

Die logischen Verknüpfungen sind:

- |   |                       |
|---|-----------------------|
|   | <b>DEF</b>            |
|   | $\neg A$              |
| (1) Die <i>Negation</i> : wir schreiben „nicht $A$ “ oder „ $\neg A$ “ für die Verneinung der Aussage $A$ . $\neg A$ ist genau dann wahr, wenn $A$ falsch ist, und umgekehrt.                           | $A \wedge B$          |
|   | $A \vee B$            |
| (2) Die <i>Konjunktion</i> : wir schreiben „ $A$ und $B$ “ oder „ $A \wedge B$ “; diese Aussage ist genau dann wahr, wenn sowohl $A$ als auch $B$ wahr sind.  | $A \Rightarrow B$     |
|   | $A \Leftrightarrow B$ |
| (3) Die <i>Disjunktion</i> : wir schreiben „ $A$ oder $B$ “ oder „ $A \vee B$ “; diese Aussage ist genau dann wahr, wenn wenigstens eine der Aussagen $A$ und $B$ wahr ist.                             |                       |
| (4) Die <i>Implikation</i> : wir schreiben „aus $A$ folgt $B$ “, „ $A$ impliziert $B$ “ oder „ $A \Rightarrow B$ “; diese Aussage ist genau dann wahr, wenn $A$ falsch oder $B$ wahr ist (oder beides). |                       |

- (5) Die *Äquivalenz*: wir schreiben „ $A$  genau dann, wenn  $B$ “, „ $A$  und  $B$  sind äquivalent“ oder „ $A \Leftrightarrow B$ “; diese Aussage ist genau dann wahr, wenn entweder  $A$  und  $B$  beide wahr oder  $A$  und  $B$  beide falsch sind.

Alle hier aufgeführten Schreibweisen sind möglich und erlaubt; die Schreibweise „ $A \wedge B$ “ ist zum Beispiel nicht besser oder schlechter als „ $A$  und  $B$ “ (nur kürzer). Bei verschachtelten Verknüpfungen werden Klammern gesetzt, um die Bedeutung klar zu machen: Bei „ $A$  und  $B$  oder  $C$ “ ist sonst nicht klar, ob  $(A \wedge B) \vee C$  oder  $A \wedge (B \vee C)$  gemeint ist.

Die Definition der logischen Verknüpfungen lässt sich übersichtlich durch die entsprechenden *Wahrheitstafeln* zusammenfassen. Wir schreiben  $W$  für *wahr* und  $F$  für *falsch*. Dann lässt sich die Negation wie folgt definieren:

$A$	$\neg A$
$W$	$F$
$F$	$W$

Die übrigen Verknüpfungen sind gegeben durch:

$A$	$B$	$A \wedge B$	$A$	$B$	$A \vee B$	$A$	$B$	$A \Rightarrow B$	$A$	$B$	$A \Leftrightarrow B$
$W$	$W$	$W$	$W$	$W$	$W$	$W$	$W$	$W$	$W$	$W$	$W$
$W$	$F$	$F$	$W$	$F$	$W$	$W$	$F$	$F$	$W$	$F$	$F$
$F$	$W$	$F$	$F$	$W$	$W$	$F$	$W$	$W$	$F$	$W$	$F$
$F$	$F$	$F$	$F$	$F$	$F$	$F$	$F$	$W$	$F$	$F$	$W$

Die wichtigste (und gleichzeitig die am schwersten zu verstehende) dieser Verknüpfungen ist die Implikation. Sie ist wichtig, weil die große Mehrzahl aller mathematischen Sätze die Form einer Implikation haben: Wenn gewisse Voraussetzungen  $A$  gelten, dann folgt eine Aussage  $B$ . Sie ist ein wenig schwierig, weil mit ihr im täglichen Leben oft ungenau bis falsch umgegangen wird. Vor allem neigen viele Menschen dazu, zwischen „aus  $A$  folgt  $B$ “ und „aus  $B$  folgt  $A$ “ nicht sorgfältig zu unterscheiden. Diesen Unterschied zu begreifen, ist die erste wichtige Hürde für Sie als zukünftige Mathematiker. Machen Sie sich Folgendes klar:

- $A \Rightarrow B$  ist jedenfalls immer dann wahr, *wenn  $A$  falsch ist*.
- $A \Rightarrow B$  ist auch immer dann wahr, wenn  $B$  wahr ist.
- $A \Rightarrow B$  kann *nur dann* falsch sein, wenn  $A$  wahr, aber  $B$  falsch ist.

Wir verwenden manchmal die Schreibweise „ $\perp$ “ für das *Falsum*, also eine stets falsche Aussage oder einen Widerspruch. Analog gibt es die stets wahre Aussage „ $\top$ “. Dann können wir also schreiben

$$\perp \Rightarrow B \quad \text{und} \quad A \Rightarrow \top \quad \text{gelten stets.}$$

Für die Lateiner unter Ihnen: Die erste dieser Tatsachen ist auch unter dem schönen Namen *Ex falso quodlibet* bekannt.

Folgende Schlussweise ist *nicht* erlaubt:

Wir wollen  $A$  zeigen. Also nehmen wir einmal an, dass  $A$  stimmt.  
Dann müsste auch  $B$  gelten.  $B$  ist aber richtig, also muss auch  $A$  gelten.

Als Beispiel: Wir wollen  $0 = 1$  zeigen. Dazu formen wir um: Aus  $0 = 1$  folgt durch Verdoppeln  $0 = 2$ , dann durch Subtraktion von 1 auf beiden Seiten  $-1 = 1$ , schließlich durch Quadrieren  $1 = 1$ , was offensichtlich stimmt. Also gilt auch die ursprüngliche Gleichung  $0 = 1$ .

Hier ist alles korrekt bis auf das „Also“ im letzten Satz, denn der Schluss von  $A \Rightarrow B$  und  $B$  auf  $A$  ist nicht möglich.

Der Schluss von  $A \Rightarrow B$  und  $A$  auf  $B$  ist hingegen sehr wohl möglich und stellt eine der grundlegenden Schlussweisen in Beweisen dar. Häufig ist „ $A \Rightarrow B$ “ ein mathematischer Satz, der angewendet werden soll. Wir weisen nach, dass die Voraussetzung  $A$  gilt, und können dann auf  $B$  schließen. Die Korrektheit dieses Schlusses drückt sich darin aus, dass die Aussage

$$((A \Rightarrow B) \wedge A) \Rightarrow B$$

stets wahr ist. So eine Aussage heißt auch eine *Tautologie*. In den Tautologien

$$(A \wedge B) \Rightarrow A, \quad A \Rightarrow (A \vee B) \quad \text{und} \quad (A \Leftrightarrow B) \Leftrightarrow ((A \Rightarrow B) \wedge (B \Rightarrow A))$$

verbergen sich weitere Schlussregeln. Die letzte davon zeigt, dass man eine Äquivalenz  $A \Leftrightarrow B$  dadurch beweisen kann, dass man die beiden Implikationen  $A \Rightarrow B$  und  $B \Rightarrow A$  nachweist. Das wird uns häufig begegnen.

Wie zeigt man, dass eine Verknüpfung von Aussagen eine Tautologie ist? Das kann man mit Hilfe von Wahrheitstafeln tun, indem man alle möglichen Kombinationen von Wahrheitswerten der beteiligten Grundaussagen ausprobiert. Zum Beispiel:

$A$	$B$	$A \Rightarrow B$	$(A \Rightarrow B) \wedge A$	$((A \Rightarrow B) \wedge A) \Rightarrow B$
*	W	W	*	W
W	F	F	F	W
F	F	W	F	W

Der Stern \* steht dabei für einen nicht festgelegten Wahrheitswert; wir nutzen aus, dass die Implikation  $C \Rightarrow B$  immer wahr ist, wenn  $B$  wahr ist.

Weitere wichtige Schlussregeln kommen aus den Tautologien

$$\neg A \Leftrightarrow (A \Rightarrow \perp) \quad \text{und} \quad (A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A).$$

Die erste besagt, dass man die Negation von  $A$  dadurch beweisen kann, dass man die Annahme, dass  $A$  gilt, zum Widerspruch („ $\perp$ “) führt. Die zweite ist der klassische *Widerspruchsbeweis*: Um die Implikation  $A \Rightarrow B$  zu zeigen, nehme ich  $A$  an und will  $B$  zeigen. Für den Widerspruchsbeweis nehme ich nun an, dass  $B$  falsch ist (also dass  $\neg B$  gilt) und leite daraus den Widerspruch  $\neg A$  zu  $A$  ab. Das zeigt, dass  $\neg B$  unter der Annahme  $A$  nicht gelten kann, also muss  $B$  richtig sein. Die Implikation  $\neg B \Rightarrow \neg A$  wird auch die *Kontraposition* der zu ihr äquivalenten Implikation  $A \Rightarrow B$  genannt.

Hier sind ein paar weitere Tautologien, an denen Sie sich versuchen können. Sie zeigen, wie man eine Negation in andere Verknüpfungen „hineinziehen“ kann.

$$\neg(A \wedge B) \Leftrightarrow (\neg A \vee \neg B), \quad \neg(A \vee B) \Leftrightarrow (\neg A \wedge \neg B), \quad \neg(A \Rightarrow B) \Leftrightarrow (A \wedge \neg B).$$

Die ersten beiden davon sind als *de Morgansche Regeln* bekannt.

Als ein weiteres Beispiel möchte ich Ihnen vorführen, dass

$$((A \Rightarrow B) \wedge (B \Rightarrow C)) \Rightarrow (A \Rightarrow C)$$

eine Tautologie ist.

$A$	$B$	$C$	$A \Rightarrow B$	$B \Rightarrow C$	$(A \Rightarrow B) \wedge (B \Rightarrow C)$	$A \Rightarrow C$	$(\dots) \Rightarrow (A \Rightarrow C)$
F	*	*	W	*	*	W	W
*	*	W	*	W	*	W	W
W	W	F	W	F	F	F	W
W	F	F	F	W	F	F	W

Das entspricht einer Schlusskette: Wenn wir aus  $A$  folgern können, dass  $B$  gilt, und aus  $B$ , dass  $C$  gilt, dann ist es auch richtig, dass aus  $A$  die Richtigkeit von  $C$  folgt. Man kann also den Beweis von „ $A \Rightarrow C$ “ zerlegen in Beweise von „ $A \Rightarrow B$ “ und von „ $B \Rightarrow C$ “. Wenn man einen solchen Beweis aufschreibt, schreibt man dann auch einfach

$$A \Rightarrow B \Rightarrow C$$

oder

$$A \Rightarrow B_1 \Rightarrow B_2 \Rightarrow \dots \Rightarrow B_n \Rightarrow C,$$

wenn es über mehrere Zwischenschritte geht.

**Warnung.** Die Notation „ $A \Rightarrow B$ “ kann zweierlei bedeuten:

- Die *Aussage* „aus  $A$  folgt  $B$ “, und
- den *Beweisschritt* „wir schließen von  $A$  auf  $B$ “, der die *als wahr bekannte* Aussage  $A \Rightarrow B$  verwendet.

Eigentlich wäre es besser, dies auch in der Schreibweise zu unterscheiden, etwa indem man ein anderes Symbol (wie zum Beispiel  $\curvearrowright$ ) für Beweisschritte verwendet. Allerdings sind beide Verwendungen von „ $\Rightarrow$ “ ziemlich üblich, und so werden wir hier auch beide benutzen.

Gelegentlich finden Sie in diesem Skript kleiner gedruckten Text wie diesen hier. Dort wird Material behandelt, das über den eigentlichen Stoff der Vorlesung hinaus geht, aber vielleicht für den Einen oder die Andere von Ihnen interessant ist.

Hier geht es um die Frage, ob es schwierig ist, für eine gegebene aussagenlogische Formel zu entscheiden, ob sie eine Tautologie ist. Wir haben ja gesehen, dass man mit Hilfe einer Wahrheitstafel immer feststellen kann, ob eine Tautologie vorliegt oder nicht. Allerdings gibt es, wenn  $n$  verschiedene elementare Aussagen (wie  $A$  und  $B$  oben) beteiligt sind,  $2^n$  mögliche Kombinationen von Wahrheitswerten, die überprüft werden müssen. Diese Zahl wächst sehr schnell mit  $n$ :  $2^{100} = 1\,267\,650\,600\,228\,229\,401\,496\,703\,205\,376$  („exponentielles Wachstum“), so dass es praktisch unmöglich ist, alles durchzuprobieren. Auf der anderen Seite haben wir gesehen, dass man oft mehrere Möglichkeiten zusammenfassen kann, sodass man sich fragen kann, ob es auch eine einigermaßen effiziente (also mit vertretbarem Aufwand durchführbare) Methode gibt. Solche Fragen werden von der *Komplexitätstheorie* studiert, die im Bereich zwischen mathematischer Logik und theoretischer Informatik angesiedelt ist. Im vorliegenden Fall ist die Antwort „wahrscheinlich Nein“: Das eng verwandte Erfüllbarkeitsproblem ist *NP-vollständig* (eine Aussage ist genau dann nicht erfüllbar, wenn ihre Negation eine Tautologie ist), und für solche Probleme sind keine effizienten Lösungsverfahren (Algorithmen) *bekannt*. Die Frage danach, ob es tatsächlich keine *gibt*, ist der Inhalt des „ $P = NP?$ “-Problems, für dessen Lösung man eine Million Dollar bekommen würde.

## 2.2. Mengen.

Ich setze voraus, dass Sie in der Schule gelernt haben, mit *Mengen* umzugehen. Daher werde ich mich auf eine kurze Wiederholung bzw. Einführung von Schreibweisen und grundlegenden Operationen und Rechenregeln beschränken.

Endliche Mengen können durch Aufzählung ihrer Elemente angegeben werden:

$$\{1, 2, 4, 8\}, \quad \{\{1\}, \{1, 2\}\}.$$

Beachte: die Elemente einer Menge können selbst wieder Mengen sein. Sehr wichtig ist die *leere Menge*, die  $\{\}$  geschrieben werden kann. Es ist aber die Schreibweise  $\emptyset$  allgemein gebräuchlich; wir werden uns ebenfalls daran halten. Wir schreiben „ $x \in M$ “ für die Aussage „ $x$  ist Element der Menge  $M$ “ und „ $x \notin M$ “ für ihre

**DEF**  
 $\emptyset, x \in M$

Negation. Zum Beispiel ist  $x \in \emptyset$  stets falsch, da die leere Menge keine Elemente hat. Zwei Mengen sind *gleich*, wenn sie dieselben Elemente haben. Insbesondere kommt es nicht darauf an, wie oft man ein Element aufführt:

$$\{1, 1, 2, 2, 2, 3\} = \{1, 2, 3\}.$$

Man kann Mengen auch durch Angabe der Eigenschaften beschreiben, die ihre Elemente haben, wie zum Beispiel

$$\{n \mid n \text{ ist Primzahl}\}.$$

(Statt des senkrechten Strichs „|“ ist auch ein Doppelpunkt „:“ gebräuchlich.) Es gibt Symbole für gewisse häufig benötigte Mengen, wie

- die Menge  $\mathbb{N} = \{0, 1, 2, 3, \dots\}$  der *natürlichen Zahlen*,
- die Menge  $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$  der *ganzen Zahlen*,
- die Menge  $\mathbb{Q} = \{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\}$  der *rationalen Zahlen* und
- die Menge  $\mathbb{R}$  der *reellen Zahlen*.

**DEF**  
 $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$

Die Definition von  $\mathbb{N}$  ist in der Literatur nicht einheitlich; häufig wird auch  $\mathbb{N} = \{1, 2, 3, \dots\}$  (also ohne die Null) gesetzt. Hier gibt es kein Richtig oder Falsch; letzten Endes ist das eine Geschmacksfrage. Für mich sind die natürlichen Zahlen gerade die Mächtigkeiten (die *Mächtigkeit* einer Menge ist die Anzahl ihrer Elemente) von endlichen Mengen, und da die leere Menge endlich ist, sollte auch die Null eine natürliche Zahl sein. Für die echt positiven natürlichen Zahlen werden wir gelegentlich die Schreibweise  $\mathbb{N}_+$  verwenden.

Eine Menge  $T$  heißt *Teilmenge* der Menge  $M$ , geschrieben  $T \subset M$ , wenn jedes Element von  $T$  auch ein Element von  $M$  ist. Man beachte, dass der Fall  $T = M$  hier erlaubt ist. Um auszudrücken, dass  $T$  eine *echte* Teilmenge von  $M$  ist (also Teilmenge von  $M$ , aber nicht ganz  $M$ ), schreiben wir  $T \subsetneq M$ . Statt  $M \subset N$  kann man auch  $N \supset M$  schreiben. Die Teilmengenbeziehung heißt auch *Inklusion*.

**DEF**  
Teilmenge

**Warnung.** Die Schreibweise wird in der Literatur nicht einheitlich verwendet; oft findet man  $T \subset M$  für echte Teilmengen und  $T \subseteq M$  für beliebige Teilmengen. Machen Sie sich solche Unterschiede bewusst, wenn Sie Lehrbücher benutzen!

Einfache Beispiele von Teilmengen sind die leere Menge, die Teilmenge *jeder* Menge ist:  $\emptyset \subset M$ , und natürlich ist jede Menge Teilmenge von sich selbst:  $M \subset M$ . Für die oben eingeführten Zahlenmengen haben wir die Beziehungen  $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ .

Die Menge aller Teilmengen von  $M$  heißt die *Potenzmenge* von  $M$ ; wir schreiben

$$\mathcal{P}(M) = \{T \mid T \subset M\}$$

**DEF**  
Potenzmenge  
 $\mathcal{P}(M)$

dafür. Zum Beispiel gilt

$$\mathcal{P}(\emptyset) = \{\emptyset\}, \quad \mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\} \quad \text{und} \quad \mathcal{P}(\{1, 2\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}.$$

An dieser Stelle gleich noch ein wichtiger Hinweis:

- Man muss sorgfältig zwischen Mengen und ihren Elementen unterscheiden. Zum Beispiel haben „ $a \in M$ “ und „ $a \subset M$ “ völlig verschiedene Bedeutungen.
- Besonders schwer fällt die Unterscheidung zwischen dem Element  $a$  und der Einermenge  $\{a\}$ . Es ist sehr wichtig, sich diese Unterschiede gleich zu Beginn klar zu machen!

Auf der anderen Seite ist  $a \in M$  äquivalent zu  $\{a\} \subset M$  — wenn man also beides falsch macht, wird es wieder richtig.

Mengen können miteinander verknüpft werden. Es gibt die *Vereinigung*

$$M \cup N = \{x \mid x \in M \vee x \in N\}$$

und den *Durchschnitt*

$$M \cap N = \{x \mid x \in M \wedge x \in N\}.$$

Zwei Mengen, deren Durchschnitt leer ist ( $M \cap N = \emptyset$ ) heißen *disjunkt*. Schließlich kann man auch noch die *Mengendifferenz*

$$M \setminus N = \{x \mid x \in M \wedge x \notin N\}$$

bilden. Sie besteht aus den Elementen von  $M$ , die keine Elemente von  $N$  sind.

Für diese Verknüpfungen gelten gewisse Rechenregeln und es gibt Beziehungen zum Begriff der Teilmenge. Ich werde hier einige davon angeben und auch beweisen; andere können Sie sich selbst herleiten, was eine gute erste Übung darstellt, einfache Beweise zu führen. Einige solche Aufgaben finden Sie auf dem Übungsblatt.

- (1) Es gilt für alle Mengen  $M$  und  $N$ , dass  $M$  genau dann eine Teilmenge von  $N$  ist, wenn die Mengen  $M \cup N$  und  $N$  übereinstimmen:

$$M \subset N \iff M \cup N = N.$$

**DEF** $M \cup N$  $M \cap N$  $M \setminus N$ **SATZ**

Eigensch.

Mengen

*Beweis.* Zu zeigen ist eine Äquivalenz  $A \iff B$ . In den meisten Fällen ist es am besten, den Beweis in zwei Teile, nämlich den Beweis von  $A \Rightarrow B$  und den Beweis von  $B \Rightarrow A$ , zu zerlegen.

„ $\Rightarrow$ “: (D.h., wir beweisen die Richtung „von links nach rechts“, also die Aussage  $M \subset N \Rightarrow M \cup N = N$ .) Wir setzen voraus, dass  $M$  eine Teilmenge von  $N$  ist, und wir müssen zeigen, dass  $M \cup N = N$  ist. Dies ist wiederum eine Äquivalenz, nämlich die Aussage  $x \in M \cup N \iff x \in N$ . Wir zerlegen den Beweis wieder in zwei Schritte:

„ $M \cup N \subset N$ “: Sei  $x \in M \cup N$ . (Das ist die übliche Formulierung dafür, dass man annimmt, dass  $x \in M \cup N$  richtig ist.) Das bedeutet nach Definition  $x \in M$  oder  $x \in N$ . Im ersten Fall ( $x \in M$ ) folgt  $x \in N$ , da nach Voraussetzung  $M$  Teilmenge von  $N$  ist. Im zweiten Fall gilt  $x \in N$  bereits.

„ $N \subset M \cup N$ “: Das gilt immer, denn aus  $x \in N$  folgt  $x \in M \vee x \in N$ .

Damit ist die Gleichheit  $M \cup N = N$  gezeigt und der Beweis der einen Richtung beendet.

„ $\Leftarrow$ “: (Jetzt beweisen wir die Richtung „von rechts nach links“, also die Aussage  $M \cup N = N \Rightarrow M \subset N$ .) Es gelte  $M \cup N = N$ . Zu zeigen ist  $M \subset N$ , also die Implikation  $x \in M \Rightarrow x \in N$ . Sei also  $x \in M$ . Dann ist auch  $x \in M \cup N$ , aber  $M \cup N = N$ , also folgt  $x \in N$ . Damit ist gezeigt, dass  $M$  eine Teilmenge von  $N$  ist.  $\square$

- (2) Für je zwei Mengen  $X$  und  $Y$  gilt das „Absorptionsgesetz“

$$(X \cap Y) \cup Y = Y.$$

*Beweis.* Es ist eine Gleichheit von Mengen zu beweisen. Wir haben eben schon gesehen, dass das eine Äquivalenz  $a \in (X \cap Y) \cup Y \Leftrightarrow a \in Y$  ist, die man in der Regel am besten in zwei Implikationen aufspaltet. In diesem Fall verwenden wir häufig folgende Kurzschreibweise:

„ $\subset$ “: (Wir beweisen die Inklusion  $(X \cap Y) \cup Y \subset Y$ .) Sei  $a \in (X \cap Y) \cup Y$ . Das bedeutet  $a \in X \cap Y$  oder  $a \in Y$ . Der zweite Fall ( $a \in Y$ ) ist gerade, was wir zeigen wollen. Im ersten Fall gilt  $a \in X$  und  $a \in Y$ , also insbesondere (ein Wort, das Mathematiker gerne verwenden) wieder  $a \in Y$ .

„ $\supset$ “: (Wir beweisen die Inklusion  $(X \cap Y) \cup Y \supset Y$ , also  $Y \subset (X \cap Y) \cup Y$ .) Sei  $a \in Y$ . Dann gilt die schwächere Aussage „ $a \in X \cap Y$  oder  $a \in Y$ “, und das bedeutet gerade  $a \in (X \cap Y) \cup Y$ .  $\square$

Alternativ kann man auch argumentieren, dass Aussage (2) aus Aussage (1) folgt, wenn man dort  $M = X \cap Y$  und  $N = Y$  setzt, denn die Inklusion  $X \cap Y \subset Y$ , die dann auf der linken Seite der Äquivalenz steht, ist immer richtig.

- (3) Für alle Mengen  $A, B, C$  gilt das „Distributivgesetz“

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

*Beweis.* Wie eben teilen wir den Beweis der Gleichheit zweier Mengen auf in die Teile „ $\subset$ “ und „ $\supset$ “.

„ $\subset$ “: Sei  $x \in (A \cap B) \cup C$ . Das bedeutet  $x \in A \cap B$  oder  $x \in C$ . Im ersten Fall gilt  $x \in A$  und  $x \in B$ , daraus folgt  $x \in A \cup C$  und  $x \in B \cup C$  und damit  $x \in (A \cup C) \cap (B \cup C)$ . Im zweiten Fall (also  $x \in C$ ) gilt ebenfalls  $x \in A \cup C$  und  $x \in B \cup C$ , also folgt auch in diesem Fall  $x \in (A \cup C) \cap (B \cup C)$ . Damit ist die Inklusion  $(A \cap B) \cup C \subset (A \cup C) \cap (B \cup C)$  gezeigt.

„ $\supset$ “: Sei  $x \in (A \cup C) \cap (B \cup C)$ . Das bedeutet  $x \in A \cup C$  und  $x \in B \cup C$ . Die erste dieser Aussagen heißt  $x \in A$  oder  $x \in C$ . Wenn  $x \in C$  ist, dann ist auch  $x \in (A \cap B) \cup C$ . Wenn  $x \notin C$  ist, dann muss  $x \in A$  und  $x \in B$  sein, also ist  $x \in A \cap B$  und damit auch  $x \in (A \cap B) \cup C$ .  $\square$

Beachten Sie, dass wir in einigen der obigen Beweise eine *Fallunterscheidung* benutzt haben. Dahinter stecken die Tautologien

$$((A \Rightarrow C) \wedge (B \Rightarrow C)) \Rightarrow ((A \vee B) \Rightarrow C),$$

die man benutzt, um eine Implikation der Form  $(A \vee B) \Rightarrow C$  zu zeigen, und

$$((A \Rightarrow B) \wedge (\neg A \Rightarrow B)) \Rightarrow B,$$

die die klassische Fallunterscheidung darstellt: Wenn ich  $B$  sowohl unter der Annahme, dass  $A$  gilt, als auch unter der Annahme, dass  $A$  nicht gilt, zeigen kann, dann muss  $B$  richtig sein.

Da die Aussagen, die wir hier bewiesen haben, durch aussagenlogische Verknüpfungen aus endlich vielen „Elementaraussagen“ der Form  $x \in M$  zusammengesetzt sind, könnten wir sie auch durch Aufstellen einer Wahrheitstafel beweisen. Zweck der Übung sollte aber sein, zu einer gewissen Fingerfertigkeit im logischen Schließen zu kommen, denn später wird es meistens nicht mehr möglich sein, Beweise rein aussagenlogisch zu führen.

### 2.3. Prädikatenlogik.

Nun hat man es in der Mathematik nicht nur mit einfachen Aussagen zu tun, die man irgendwie verknüpft, sondern in aller Regel hängen die Aussagen noch von gewissen Parametern oder Variablen ab. Ein typisches Beispiel ist die Aussage „ $x \in M$ “, deren Wahrheitswert davon abhängt, wofür  $x$  und  $M$  stehen. (Man nennt solche parameterabhängigen Aussagen manchmal *Aussageformen*, weil sie erst dadurch zu einer Aussage mit festgelegtem Wahrheitswert werden, dass man den Parametern Werte zuweist. Auch der Begriff *Prädikat* ist gebräuchlich, was die Bezeichnung „Prädikatenlogik“ erklärt.) Um aus solchen von Variablen abhängigen Aussagen wiederum Aussagen zu machen, die nicht mehr von (einigen oder allen) Variablen abhängen, gibt es im Wesentlichen zwei Möglichkeiten. Sei dafür  $A(x)$  eine (möglicherweise) von der Variablen  $x$  abhängige Aussage.

**DEF**  
Quantoren  
 $\forall, \exists$

- Wir machen die Aussage „für alle  $x$  gilt  $A(x)$ “ oder kurz „ $\forall x : A(x)$ “.
- Wir machen die Aussage „es gibt ein  $x$ , sodass  $A(x)$  gilt“ oder kurz „ $\exists x : A(x)$ “.

Im Fachjargon spricht man von *Quantifizierung*, da man eine Aussage darüber macht, für wie viele  $x$  (alle oder wenigstens eines)  $A(x)$  stimmt. In diesem Zusammenhang heißen die Symbole  $\forall$  und  $\exists$  auch *Quantoren*, genauer *Allquantor* ( $\forall$ ) und *Existenzquantor* ( $\exists$ ).

In der Praxis kommen fast nur die Kombinationen

$$\forall x : x \in M \Rightarrow A(x) \quad \text{und} \quad \exists x : x \in M \wedge A(x)$$

vor, die man dann zu

$$\forall x \in M : A(x) \quad \text{„für alle } x \in M \text{ gilt } A(x)\text{“}$$

und

$$\exists x \in M : A(x) \quad \text{„es gibt ein } x \in M \text{ mit } A(x)\text{“}$$

abkürzt. An der ausführlicheren Form oben erkennt man, dass

$$\forall x \in \emptyset : A(x)$$

*immer wahr ist*, denn die Voraussetzung  $x \in \emptyset$  in der Implikation „ $x \in \emptyset \Rightarrow A(x)$ “ ist falsch. Entsprechend ist

$$\exists x \in \emptyset : A(x)$$

immer falsch, denn es gibt ja kein Element der leeren Menge, also erst recht keines mit zusätzlichen Eigenschaften.

Für den Umgang mit den Quantoren sind folgende Regeln wichtig:

$$\neg \forall x \in M : A(x) \quad \text{ist gleichbedeutend mit} \quad \exists x \in M : \neg A(x)$$

und

$$\neg \exists x \in M : A(x) \quad \text{ist gleichbedeutend mit} \quad \forall x \in M : \neg A(x).$$

Die erste zeigt, wie man eine „Allaussage“ widerlegt: *Man gibt ein Gegenbeispiel an*. Das macht auch verständlich, warum  $\forall x \in \emptyset : A(x)$  wahr sein muss: Es gibt kein Gegenbeispiel! Das klingt jetzt vielleicht wie esoterische Spielerei, das ist es aber keineswegs: *Es ist sehr wichtig, Grenzfälle zu verstehen*. Die leere Menge ist ein typischer Grenzfall in vielen Situationen, und nur wenn Sie diesen Grenzfall verstehen, haben Sie die Situation wirklich verstanden! (Für die Angst vor der leeren Menge gibt es auch einen lateinischen Ausdruck: *horror vacui*. Mathematiker

sollten nicht daran leiden.) Zum Beispiel gilt die auf den ersten Blick offensichtliche Implikation

$$(\forall x \in M : A(x)) \Rightarrow (\exists x \in M : A(x))$$

(„wenn alle Elemente von  $M$  die Eigenschaft  $A$  haben, dann hat wenigstens eines diese Eigenschaft“) *nur dann, wenn  $M$  nicht die leere Menge ist*. Das zu wissen kann einem manche Falltüren im Beweisgeschäft ersparen.

Sehr wichtig ist auch, dass es auf die *Reihenfolge* der Quantoren ankommt: Die Aussagen

$$\forall x \in M \exists y \in N : A(x, y) \quad \text{und} \quad \exists y \in N \forall x \in M : A(x, y)$$

haben unterschiedliche Bedeutung — in der ersten Aussage kann das  $y$ , dessen Existenz behauptet wird, von  $x$  abhängen, in der zweiten Aussage gibt es *ein festes  $y$* , das für alle  $x$  funktionieren muss. Hier ist ein einfaches Beispiel:

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

(„für jede natürliche Zahl  $x$  gibt es eine größere natürliche Zahl  $y$ “) ist sicher richtig, während

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

(„es gibt eine natürliche Zahl  $y$ , die größer ist als alle natürlichen Zahlen  $x$ “) falsch ist. Die Variante

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

(„es gibt eine natürliche Zahl  $x$ , die kleiner ist als alle natürlichen Zahlen  $y$ “) ist ebenfalls falsch, aber sozusagen nur knapp, denn mit „ $\geq$ “ statt „ $>$ “ wäre sie richtig (mit  $x = 0$ ).

Als Illustration dafür, wie man Beweise von quantifizierten Aussagen führt, zeige ich jetzt, dass es unendlich viele gerade natürliche Zahlen gibt. Dabei lernen Sie auch gleich eine mathematische „Redewendung“ kennen, mit der man ausdrücken kann, dass es unendlich viele natürliche Zahlen mit einer gewissen Eigenschaft gibt: Man sagt, dass es zu jeder gegebenen natürlichen Zahl eine größere gibt, die die Eigenschaft hat.

**Behauptung:**  $\forall m \in \mathbb{N} \exists n \in \mathbb{N} : n > m$  und  $n$  ist gerade.

*Beweis.* Sei  $m \in \mathbb{N}$  beliebig. Dann gilt für  $n_0 = 2m + 2 \in \mathbb{N}$ , dass  $n_0 > m$  ist (denn  $n_0 = m + (m + 2)$  und  $m + 2 > 0$ ), und  $n_0 = 2(m + 1)$  ist gerade. Also existiert ein  $n \in \mathbb{N}$  (nämlich zum Beispiel  $n_0$ ) mit  $n > m$  und  $n$  gerade:  $\exists n \in \mathbb{N} : n > m$  und  $n$  ist gerade. Da  $m \in \mathbb{N}$  beliebig war, gilt diese Aussage für alle  $m \in \mathbb{N}$ .  $\square$

Man beweist also eine Allaussage  $\forall x \in M : A(x)$ , indem man ein nicht näher spezifiziertes  $x \in M$  betrachtet („Sei  $x \in M$ “ — das Wort „beliebig“ steht oben nur zur Verdeutlichung und kann weggelassen werden) und für dieses  $x$  die Aussage  $A(x)$  zeigt. Eine Existenzaussage  $\exists x \in M : A(x)$  kann man zeigen, indem man für ein *bestimmtes*  $x_0 \in M$  die Aussage  $A(x_0)$  beweist.

Das nennt man dann auch einen *konstruktiven* Existenzbeweis, weil man ein geeignetes Element explizit angibt oder konstruiert. Alternativ kann man die äquivalente Aussage  $\neg \forall x \in M : \neg A(x)$  beweisen, indem man die Annahme, dass kein  $x \in M$  die Eigenschaft  $A$  hat, zum Widerspruch führt. Dabei muss kein Element von  $M$  angegeben werden, das tatsächlich die Eigenschaft  $A$  hat. Zum besseren Verständnis hier ein Beispiel: Sei  $N > 1$  eine gegebene natürliche Zahl. Sie wollen beweisen, dass  $N$  einen echten Teiler  $d > 1$  hat, also die Aussage

$$\exists d, m \in \mathbb{N} : d > 1 \wedge m > 1 \wedge N = d \cdot m.$$

Das kann man natürlich tun, indem man einen echten Teiler  $d$  findet. Man kann aber auch versuchen, die Negation der Aussage, nämlich „ $N$  ist Primzahl“ zum Widerspruch zu führen. Dazu kann man etwa den „kleinen Satz von Fermat“ verwenden, der aussagt, dass für jede Primzahl  $p$  und jede ganze Zahl  $a$  die Zahl  $a^p - a$  durch  $p$  teilbar ist. Wenn Sie also eine ganze Zahl  $a$  finden, sodass  $a^N - a$  nicht durch  $N$  teilbar ist, dann folgt daraus, dass  $N$  keine Primzahl ist, also muss  $N$  einen echten Teiler haben, ohne dass Sie einen angeben können. Dieser Unterschied ist durchaus auch praktisch relevant. Es gibt nämlich effiziente Algorithmen, die feststellen, ob  $N$  eine Primzahl ist oder nicht, aber es sind bisher keine effizienten Algorithmen bekannt, die eine zusammengesetzte Zahl faktorisieren können.

Weitere Beispiele für Beweise werden in großer Zahl im Lauf der Vorlesung folgen.

An den bisherigen Beispielen von Beweisen können Sie jedenfalls schon sehen, dass *die Struktur der zu beweisenden Aussage die Struktur des Beweises vorgibt*: Bis zu einem gewissen Grad gibt es auch für das Beweisen Rezepte, die man anwenden kann!

## 2.4. Geordnete Paare.

Häufig möchte man mit zwei (oder vielleicht auch mehr, siehe unten) Elementen von verschiedenen Mengen gemeinsam arbeiten, wobei es auf die Reihenfolge ankommt. (Wenn die Reihenfolge keine Rolle spielt, also bei *ungeordneten* Paaren, kann man Zweiermengen  $\{a, b\}$  verwenden.) Dazu führt man *geordnete Paare* ein: Sind  $a$  und  $b$  Elemente irgendwelcher Mengen, dann steht  $(a, b)$  für das daraus gebildete geordnete Paar. Die wesentliche Eigenschaft dieser geordneten Paare ist, dass zwei solche Paare genau dann gleich sind, wenn sie in beiden Komponenten übereinstimmen:

$$(a, b) = (x, y) \iff (a = x \text{ und } b = y).$$

Man kann geordnete Paare innerhalb der Mengenlehre definieren, indem man

$$(a, b) = \{\{a\}, \{a, b\}\}$$

setzt. (Man beachte den Sonderfall  $(a, a) = \{\{a\}\}$ .) Man muss dann zeigen, dass die so definierten Paare die obige Eigenschaft haben. Das sollten Sie als Aufforderung begreifen!

Sind  $M$  und  $N$  zwei Mengen, dann schreibt man

$$M \times N = \{(m, n) \mid m \in M, n \in N\}$$

(lies „ $M$  kreuz  $N$ “) für die Menge der geordneten Paare, deren erste Komponente aus  $M$  und deren zweite Komponente aus  $N$  kommt; die Menge  $M \times N$  heißt das *kartesische Produkt* der Mengen  $M$  und  $N$ . („Kartesisch“ leitet sich vom latinisierten Namen *Cartesius* des Mathematikers und Philosophen („ich denke, also bin ich“) René Descartes ab.)

Analog kann man (geordnete) Tripel  $(a, b, c)$ , Quadrupel  $(a, b, c, d)$ , Quintupel  $(a, b, c, d, e)$ , Sextupel  $(a, b, c, d, e, f)$ , Septupel  $(a, b, c, d, e, f, g)$  und ganz allgemein  $n$ -Tupel  $(a_1, a_2, \dots, a_n)$  einführen und kartesische Produkte mit mehr als zwei Faktoren definieren, zum Beispiel

$$A \times B \times C \times D = \{(a, b, c, d) \mid a \in A, b \in B, c \in C, d \in D\}.$$

Einen wichtigen Spezialfall erhalten wir, wenn alle beteiligten Mengen übereinstimmen. Dann schreibt man kurz  $M^2$  für  $M \times M$ ,  $M^3$  für  $M \times M \times M$  und allgemein

$$M^n = \{(m_1, m_2, \dots, m_n) \mid m_1, m_2, \dots, m_n \in M\}$$

**Def**  
geordnetes  
Paar  $(a, b)$

**DEF**  
 $M \times N$

**DEF**  
Tripel, ...  
 $n$ -Tupel

**DEF**  
 $M^n$

für die Menge der  $n$ -Tupel, deren Komponenten aus der Menge  $M$  kommen. Zum Beispiel ist  $\mathbb{R}^2$  die Menge der Paare reeller Zahlen. Wenn man die Komponenten als  $x$ - und  $y$ -Koordinate interpretiert, dann kann man  $\mathbb{R}^2$  als die Menge der Punkte der Ebene auffassen, und entsprechend  $\mathbb{R}^3$  als die Menge der Punkte des (dreidimensionalen) Raumes. Diese Mengen und ihre allgemeinere Form  $\mathbb{R}^n$  werden uns bald wieder als Standardbeispiele von „Vektorräumen“ begegnen.

Als Grenzfall haben wir zum einen  $M^1$  als Menge der „1-Tupel“; da sich Elemente  $m$  und 1-Tupel  $(m)$  eindeutig entsprechen, identifiziert man gerne  $M^1$  mit  $M$ . Zum anderen ist  $M^0 = \{()\}$  die Menge, deren einziges Element das „Nulltupel“  $()$  ist.

### 2.5. Abbildungen.

Der (vorläufig) letzte wichtige Begriff, den wir einführen müssen, ist der der *Abbildung* zwischen zwei Mengen. Seien  $M$  und  $N$  zwei Mengen. Dann ist eine *Abbildung*  $f$  von  $M$  nach  $N$  eine Vorschrift, die jedem  $x \in M$  ein eindeutig bestimmtes  $y \in N$  zuordnet; für dieses  $y$  schreiben wir dann  $f(x)$  („ $f$  von  $x$ “). Wir schreiben

**DEF**  
Abbildung

$$f : M \longrightarrow N$$

oder, wenn wir die Abbildungsvorschrift angeben wollen,

$$f : M \longrightarrow N, \quad x \longmapsto f(x),$$

wobei statt „ $f(x)$ “ meistens eine konkrete Formel oder Ähnliches steht. Beachten Sie die beiden unterschiedlichen Pfeile „ $\rightarrow$ “ und „ $\mapsto$ “! Der erste steht zwischen den *Mengen*  $M$  und  $N$ , der zweite zwischen den *Elementen*  $x \in M$  und  $f(x) \in N$ .  $f(x) \in N$  heißt dann das *Bild* von  $x \in M$  unter  $f$ . Gilt  $f(x) = y$  für ein  $y \in N$ , dann heißt  $x$  ein *Urbild* von  $y$  unter  $f$ . Man beachte: Es ist durchaus möglich, dass ein  $y \in N$  kein Urbild oder viele verschiedene Urbilder unter  $f$  hat.

Mit „Vorschrift“ ist hier nicht gemeint, dass das Bild von  $x$  unter  $f$  durch einen Rechenausdruck oder so etwas gegeben sein muss. Es kommt *nur* darauf an, dass *jedem*  $x \in M$  *genau ein*  $f(x) \in N$  zugeordnet ist. Man kann sich  $f$  als eine „Black Box“ vorstellen, die einem, wenn man ein  $x \in M$  hineinsteckt, ein Element  $f(x) \in N$  herausgibt (und zwar für dasselbe  $x$  immer dasselbe  $f(x)$ ):

$$M \ni x \longrightarrow \boxed{f} \longrightarrow f(x) \in N$$

(Die Pfeile hier dienen nur der Illustration und sind nicht Teil der oben eingeführten Notation  $f : M \rightarrow N, x \mapsto f(x)$ .)

$M$  heißt die *Definitionsmenge*, der *Definitionsbereich* oder die *Quelle* von  $f$ ,  $N$  dementsprechend der *Wertebereich* oder das *Ziel* von  $f$ . Wichtig ist dabei, dass zur Angabe einer Abbildung immer auch Quelle und Ziel gehören; die Abbildungsvorschrift alleine genügt nicht. Häufig (vor allem in der Analysis) verwendet man auch das Wort *Funktion* für *Abbildung* (was die häufig verwendete Bezeichnung „ $f$ “ für Abbildungen erklärt).

**DEF**  
Definitions-,  
Wertebereich

Zwei Abbildungen  $f$  und  $g$  sind genau dann gleich (und man schreibt  $f = g$ ), wenn ihre Definitions- und Wertebereiche übereinstimmen und für alle Elemente  $x$  des Definitionsbereichs gilt  $f(x) = g(x)$ : Abbildungen sind (bei gegebenem Definitions- und Wertebereich) durch ihre Werte festgelegt.

Beispiele von Abbildungen sind

$$n : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto 0$$

**BSP**  
Abbildungen

(die Nullfunktion; es gilt  $n(x) = 0 \in \mathbb{R}$  für alle  $x \in \mathbb{R}$ ),

$$p : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \longmapsto x^3 - 2x^2 + x - 5$$

(eine Polynomfunktion; es gilt zum Beispiel  $p(1) = p(0) = -5$ ),

$$s : \mathbb{R} \longrightarrow \{-1, 0, 1\}, \quad x \longmapsto \begin{cases} 1 & \text{falls } x > 0, \\ 0 & \text{falls } x = 0, \\ -1 & \text{falls } x < 0 \end{cases}$$

(die Vorzeichenfunktion). Für eine beliebige Menge  $M$  gibt es die „Einermengenabbildung“

$$e : M \longrightarrow \mathcal{P}(M), \quad x \longmapsto \{x\}.$$

Zum kartesischen Produkt  $M \times N$  gehören die *Projektionsabbildungen*

$$\text{pr}_1 : M \times N \longrightarrow M, \quad (a, b) \longmapsto a \quad \text{und} \quad \text{pr}_2 : M \times N \longrightarrow N, \quad (a, b) \longmapsto b.$$

Ist  $T$  eine Teilmenge von  $M$ , dann hat man die *Inklusionsabbildung*

$$i : T \longrightarrow M, \quad x \longmapsto x.$$

Für jede Menge  $X$  gibt es (als Spezialfall der Inklusionsabbildung) die *identische Abbildung* oder kurz *Identität*

$$\text{id}_X : X \longrightarrow X, \quad x \longmapsto x,$$

die jedes Element von  $X$  auf sich selbst abbildet. Als Grenzfälle haben wir für jede Menge  $X$  genau eine Abbildung  $\emptyset \rightarrow X$ ; eine Abbildung  $X \rightarrow \emptyset$  gibt es jedoch nur dann, wenn  $X$  selbst die leere Menge ist, denn wenn  $X$  ein Element  $x$  hat, könnte es auf kein Element abgebildet werden (denn die leere Menge hat keine Elemente). Manchmal schreiben wir  $\text{Abb}(X, Y)$  für die Menge aller Abbildungen mit Definitionsbereich  $X$  und Wertebereich  $Y$ .

Wenn Sie den Begriff „Vorschrift“, den wir oben verwendet haben, zu schwammig finden, dann erfahren Sie hier, wie man den Abbildungsbegriff auf eine solide Grundlage stellen kann. Man greift dazu auf die Mengenlehre zurück und identifiziert eine Abbildung  $f : M \rightarrow N$  mit ihrem *Graphen*

$$\Gamma(f) = \{(x, f(x)) \mid x \in M\} \subset M \times N.$$

(Das verallgemeinert die Funktionsgraphen von Funktionen  $\mathbb{R} \rightarrow \mathbb{R}$ , die Sie aus der Schule kennen.) Dann kann man sagen, dass eine Teilmenge  $F \subset M \times N$  genau dann einer Abbildung  $f : M \rightarrow N$  entspricht, wenn die Bedingungen

$$\forall x \in M \exists y \in N : (x, y) \in F$$

und

$$\forall x \in M \forall y_1, y_2 \in N : ((x, y_1) \in F \wedge (x, y_2) \in F) \Rightarrow y_1 = y_2$$

erfüllt sind. Die erste Bedingung drückt aus, dass *jedes*  $x \in M$  auf ein Element von  $N$  abgebildet werden muss, und die zweite Bedingung sagt, dass es *höchstens ein* solches Element von  $N$  gibt.

Es gibt gewisse wichtige Eigenschaften, die eine Abbildung haben kann oder nicht. Sei  $f : M \rightarrow N$  eine Abbildung.

- $f$  heißt *injektiv* oder eine *Injektion*, wenn  $f$  keine zwei verschiedenen Elemente von  $M$  auf dasselbe Element von  $N$  abbildet:

$$\forall x_1, x_2 \in M : f(x_1) = f(x_2) \Rightarrow x_1 = x_2$$

(So weist man auch nach, dass  $f$  injektiv ist: Man nimmt an, zwei Elemente hätten dasselbe Bild unter  $f$  und zeigt dann, dass diese beiden Elemente gleich sein müssen.)

**DEF**  
injektiv  
surjektiv  
bijektiv

- $f$  heißt *surjektiv* oder eine *Surjektion* (das „sur“ ist französisch für „auf“, daher ist die korrekte Aussprache „ßür“), wenn jedes Element von  $N$  als Bild unter  $f$  eines Elements von  $M$  auftritt:

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

- $f$  heißt *bijektiv* oder eine *Bijektion*, wenn  $f$  sowohl injektiv als auch surjektiv ist.

Man kann das auch so ausdrücken:

- $f$  ist injektiv genau dann, wenn jedes Element von  $N$  *höchstens* ein Urbild unter  $f$  hat.
- $f$  ist surjektiv genau dann, wenn jedes Element von  $N$  *mindestens* ein Urbild unter  $f$  hat.
- $f$  ist bijektiv genau dann, wenn jedes Element von  $N$  *genau* ein Urbild unter  $f$  hat.

Wenn  $f : M \rightarrow N$  bijektiv ist, dann kann man eine Abbildung  $f^{-1} : N \rightarrow M$  dadurch definieren, dass man für  $f^{-1}(y)$  das eindeutig bestimmte  $x \in M$  mit  $f(x) = y$  nimmt. Diese Abbildung  $f^{-1}$  heißt dann die *Umkehrabbildung* oder *inverse Abbildung* von  $f$ . Eine bijektive Abbildung  $f : X \rightarrow X$  heißt auch eine *Permutation* von  $X$ .

**DEF**  
Umkehrabb.

**Beispiele.** Wir schreiben  $\mathbb{R}_{\geq 0}$  für die Menge  $\{x \in \mathbb{R} \mid x \geq 0\}$  der nichtnegativen reellen Zahlen. Dann gilt:

**BSP**  
injektiv  
surjektiv

- $f_1 : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$ , ist weder injektiv noch surjektiv, denn es gilt zum Beispiel  $f_1(1) = f_1(-1) = 1$  und  $-1 \in \mathbb{R}$  hat kein Urbild.
- $f_2 : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}, x \mapsto x^2$ , ist injektiv, aber nicht surjektiv.
- $f_3 : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, x \mapsto x^2$ , ist surjektiv, aber nicht injektiv.
- $f_4 : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}, x \mapsto x^2$ , ist bijektiv.

Daran sieht man auch sehr schön, dass Definitions- und Wertebereich wesentlich für eine Abbildung sind. Weitere allgemeine Beispiele sind:

- Für jede Menge  $M$  ist die identische Abbildung  $\text{id}_M$  bijektiv.
- Für jede Menge  $M$  ist die „leere Abbildung“  $\emptyset \rightarrow M$  injektiv.
- Jede Abbildung  $\{a\} \rightarrow M$  ist injektiv.
- Eine Abbildung  $M \rightarrow \{a\}$  ist genau dann surjektiv, wenn  $M$  nicht leer ist.
- Die Einermengenabbildung  $e : M \rightarrow \mathcal{P}(M)$  ist injektiv, aber nicht surjektiv (Letzteres, weil zum Beispiel die leere Menge kein Urbild hat).

Abbildungen können verknüpft werden, indem man sie „hintereinanderschaltet“: Sind  $f : X \rightarrow Y$  und  $g : Y \rightarrow Z$  Abbildungen, sodass der Wertebereich von  $f$  mit dem Definitionsbereich von  $g$  übereinstimmt, dann kann man die zusammengesetzte Abbildung  $g \circ f : X \rightarrow Z$  bilden, die  $x \in X$  auf  $g(f(x)) \in Z$  abbildet:

**DEF**  
Komposition

$$x \longrightarrow \boxed{f} \longrightarrow f(x) \longrightarrow \boxed{g} \longrightarrow g(f(x))$$

Man muss sich merken, dass in  $g \circ f$  die Abbildung  $f$  *zuerst* ausgeführt wird, obwohl sie hinter  $g$  steht. Die Sprechweise „ $g$  nach  $f$ “ für  $g \circ f$  hilft dabei.

Diese Verknüpfung oder *Komposition* von Abbildungen hat einige wichtige Eigenschaften:

- (1) Sind  $f : W \rightarrow X$ ,  $g : X \rightarrow Y$  und  $h : Y \rightarrow Z$  Abbildungen, dann gilt  $(h \circ g) \circ f = h \circ (g \circ f)$ . Man lässt deswegen meistens die Klammern weg und schreibt  $h \circ g \circ f$ .

**SATZ**  
Eigensch.  
Abbildungen

$$w \longrightarrow \boxed{f} \longrightarrow f(w) \longrightarrow \boxed{g} \longrightarrow g(f(w)) \longrightarrow \boxed{h} \longrightarrow h(g(f(w)))$$

*Beweis.* Erst einmal ist klar, dass die Abbildungen den gemeinsamen Definitionsbereich  $W$  und den gemeinsamen Wertebereich  $Z$  haben. Die Aussage „ $(h \circ g) \circ f = h \circ (g \circ f)$ “ bedeutet dann

$$\forall w \in W : ((h \circ g) \circ f)(w) = (h \circ (g \circ f))(w).$$

Sei also  $w \in W$ . Dann ist

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

und ebenso

$$(h \circ (g \circ f))(w) = h((g \circ f)(w)) = h(g(f(w))),$$

also gilt die behauptete Gleichheit für  $w$ . Da  $w \in W$  beliebig war, gilt die Gleichheit für alle  $w \in W$ .  $\square$

- (2) Ist  $f : X \rightarrow Y$  eine Abbildung, dann gilt

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

*Beweis.* In beiden Fällen haben alle beteiligten Abbildungen denselben Definitionsbereich  $X$  und denselben Wertebereich  $Y$ . Für  $x \in X$  gilt

$$(f \circ \text{id}_X)(x) = f(\text{id}_X(x)) = f(x),$$

also ist  $f \circ \text{id}_X = f$ , und

$$(\text{id}_Y \circ f)(x) = \text{id}_Y(f(x)) = f(x),$$

also ist auch  $\text{id}_Y \circ f = f$ .  $\square$

- (3) Sind  $f : X \rightarrow Y$  und  $g : Y \rightarrow Z$  injektive Abbildungen, dann ist auch  $g \circ f : X \rightarrow Z$  injektiv.

*Beweis.* Übung.  $\square$

- (4) Sind  $f : X \rightarrow Y$  und  $g : Y \rightarrow Z$  surjektive Abbildungen, dann ist auch  $g \circ f : X \rightarrow Z$  surjektiv.

*Beweis.* Übung.  $\square$

- (5) Ist  $f : X \rightarrow Y$  bijektiv mit Umkehrabbildung  $f^{-1} : Y \rightarrow X$ , dann gilt

$$f^{-1} \circ f = \text{id}_X \quad \text{und} \quad f \circ f^{-1} = \text{id}_Y.$$

*Beweis.* Die Definitions- und Wertebereiche stimmen jeweils überein. Für  $x \in X$  gilt  $f^{-1}(f(x)) = x = \text{id}_X(x)$  nach Definition der Umkehrabbildung, also ist  $f^{-1} \circ f = \text{id}_X$ . Für  $y \in Y$  gilt  $f(f^{-1}(y)) = y = \text{id}_Y(y)$  ebenfalls nach Definition der Umkehrabbildung, also ist  $f \circ f^{-1} = \text{id}_Y$ .  $\square$

- (6) Sind  $f : X \rightarrow Y$  und  $g : Y \rightarrow Z$  zwei Abbildungen, dann gilt  
 $g \circ f$  injektiv  $\implies f$  injektiv      und       $g \circ f$  surjektiv  $\implies g$  surjektiv.

*Beweis.* Wir nehmen an, dass  $g \circ f$  injektiv ist; wir müssen zeigen, dass auch  $f$  injektiv ist. Seien dazu  $x_1, x_2 \in X$  mit  $f(x_1) = f(x_2)$ . Dann folgt  $(g \circ f)(x_1) = g(f(x_1)) = g(f(x_2)) = (g \circ f)(x_2)$ , und weil  $g \circ f$  injektiv ist, muss  $x_1 = x_2$  sein. Damit ist gezeigt, dass  $f$  injektiv ist.

Jetzt nehmen wir an, dass  $g \circ f$  surjektiv ist; wir müssen zeigen, dass auch  $g$  surjektiv ist. Sei dazu  $z \in Z$ . Da nach Voraussetzung  $g \circ f$  surjektiv ist, gibt es  $x \in X$  mit  $(g \circ f)(x) = z$ . Das heißt aber  $g(f(x)) = z$ , also gilt mit  $y = f(x) \in Y$  auch  $g(y) = z$ . Das zeigt, dass  $g$  surjektiv ist.  $\square$

- (7) Ist  $f : X \rightarrow Y$  eine Abbildung, dann ist  $f$  genau dann injektiv, wenn  $X$  leer ist oder es eine Abbildung  $g : Y \rightarrow X$  mit  $g \circ f = \text{id}_X$  gibt.

*Beweis.* Zu zeigen ist die Äquivalenz

$$f : X \rightarrow Y \text{ injektiv} \iff (X = \emptyset \vee \exists g \in \text{Abb}(Y, X) : g \circ f = \text{id}_X).$$

„ $\implies$ “: Wir nehmen an,  $f$  sei injektiv. Wenn  $X$  leer ist, dann gilt die rechte Seite. Wenn  $X$  nicht leer ist, dann sei  $x_0 \in X$  irgendein Element. Wir konstruieren eine passende Abbildung  $g : Y \rightarrow X$  wie folgt: Sei  $y \in Y$ . Wenn es ein  $x \in X$  gibt mit  $f(x) = y$ , dann setzen wir  $g(y) = x$ . Da es (weil  $f$  injektiv ist) dann genau ein solches  $x$  gibt, ist  $g(y)$  eindeutig bestimmt. Wenn es kein  $x \in X$  gibt mit  $f(x) = y$ , dann setzen wir  $g(y) = x_0$ . Jetzt müssen wir nachprüfen, dass  $g$  die geforderte Eigenschaft  $g \circ f = \text{id}_X$  hat. Definitions- und Wertebereich beider Seiten stimmen überein, und für  $x \in X$  gilt nach Definition von  $g$ , dass  $(g \circ f)(x) = g(f(x)) = x = \text{id}_X(x)$  ist. Damit ist die Gleichheit der Abbildungen gezeigt.

„ $\impliedby$ “: Wenn  $X = \emptyset$  ist, dann ist  $f$  injektiv. Wenn es  $g : Y \rightarrow X$  gibt mit  $g \circ f = \text{id}_X$ , dann ist  $f$  ebenfalls injektiv nach Teil (6), denn  $\text{id}_X$  ist injektiv.  $\square$

- (8) Ist  $f : X \rightarrow Y$  eine Abbildung, dann ist  $f$  genau dann surjektiv, wenn es eine Abbildung  $g : Y \rightarrow X$  gibt mit  $f \circ g = \text{id}_Y$ .

*Beweis.* „ $\implies$ “: Ist  $f$  surjektiv, dann können wir zu jedem  $y \in Y$  ein  $x_y \in X$  auswählen mit  $f(x_y) = y$  (denn es gibt ja immer mindestens ein Urbild). Wir setzen dann  $g(y) = x_y$  und es folgt  $f \circ g = \text{id}_Y$ .

„ $\impliedby$ “: Das folgt aus Teil (6), denn  $\text{id}_Y$  ist surjektiv.  $\square$

- (9) Ist  $f : X \rightarrow Y$  eine Abbildung, dann ist  $f$  genau dann bijektiv, wenn es eine Abbildung  $g : Y \rightarrow X$  gibt mit  $g \circ f = \text{id}_X$  und  $f \circ g = \text{id}_Y$ .

*Beweis.* „ $\implies$ “: Ist  $f$  bijektiv, dann hat  $g = f^{-1}$  die verlangte Eigenschaft.

„ $\impliedby$ “: Nach Teil (7) ist  $f$  injektiv und nach Teil (8) auch surjektiv, also bijektiv.  $\square$

Man kann also auf (mindestens) zwei verschiedene Arten beweisen, dass eine Abbildung  $f : X \rightarrow Y$  bijektiv ist:

- Man weist nach, dass  $f$  injektiv und surjektiv ist, oder
- man findet einen Kandidaten  $g$  für die Umkehrabbildung und rechnet nach, dass  $g \circ f = \text{id}_X$  und  $f \circ g = \text{id}_Y$  ist.

In vielen Fällen ist die zweite Methode einfacher durchzuführen.

Wenn man Abbildungen definieren möchte, die von zwei (oder mehr) Elementen möglicherweise verschiedener Mengen abhängen, dann kann man dies unter Zuhilfenahme von kartesischen Produkten tun: Möchte man einem Element von  $M_1$  und einem Element von  $M_2$  ein Element von  $N$  zuordnen, so entspricht das einer Abbildung  $M_1 \times M_2 \rightarrow N$ . Zum Beispiel kann man die Addition reeller Zahlen als eine Abbildung  $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $(x, y) \mapsto x + y$ , auffassen. Ist  $f : M_1 \times M_2 \rightarrow N$  eine Abbildung, dann schreibt man auch  $f(m_1, m_2)$  für  $f((m_1, m_2))$ .

Schließlich möchte ich noch eine weitere Interpretation und Schreibweise für Abbildungen einführen, die immer mal wieder vorkommt: Wenn  $a : I \rightarrow X$  eine Abbildung ist, dann schreibt man dafür auch  $(a_i)_{i \in I}$  und nennt das eine *Familie* mit der *Indexmenge*  $I$ . Dabei ist  $a_i = a(i)$  der Wert der Abbildung  $a$  an der Stelle  $i \in I$ . Sie kennen das von Folgen  $(a_n)_{n \in \mathbb{N}}$ . Die  $n$ -Tupel, die wir vor einer Weile eingeführt haben, kann man als den Spezialfall  $I = \{1, 2, \dots, n\}$  einer solchen Familie betrachten. In Analogie zur Schreibweise  $M^n$  für die Menge der  $n$ -Tupel mit Komponenten aus  $M$  schreibt man auch  $M^I$  für die Menge der Familien von Elementen von  $M$  mit Indexmenge  $I$ . Das ist nichts anderes als die Menge  $\text{Abb}(I, M)$  der Abbildungen von  $I$  nach  $M$ .

**DEF**  
Familie

An dieser Stelle bietet es sich an, etwas mehr zur Mengenlehre zu sagen. Was wir hier betreiben, ist „naive“ Mengenlehre; wir machen uns hier also keine Gedanken darüber, welche Konstruktionen mit Mengen tatsächlich möglich oder erlaubt sind. Das führt normalerweise auch nicht zu Problemen. Sie sollten aber wissen, dass die Mengenlehre durchaus nicht so harmlos ist, wie sie einem zunächst erscheinen mag. Wenn man bei der Bildung von Mengen zu viel erlaubt, kommt man in Schwierigkeiten, wie die berühmte *Russellsche Antinomie* zeigt. Denn dann könnte man die „Menge aller Mengen, die sich nicht selbst als Element enthalten“, also  $M = \{x \mid x \notin x\}$  konstruieren. Die Frage, ob  $M$  ein Element von  $M$  ist, führt auf einen unauflösbaren Widerspruch. (In der Unterhaltungsmathematik gibt es die Variante mit dem Dorfbarbier, der genau die Männer im Dorf rasiert, die sich nicht selbst rasieren. Rasiert sich nun der Barbier oder nicht?) Um diesen Widerspruch zu vermeiden, muss man genaue Regeln formulieren, wie man Mengen konstruieren darf. Das führt zur *axiomatischen Mengenlehre*.

Die meisten der Axiome sind recht „harmlos“; sie besagen etwa, dass die leere Menge existiert, dass man Einer- und Zweiermengen bilden kann, dass man immer Teilmengen bilden kann, und dass Vereinigungsmengen und Potenzmengen existieren. Es gibt aber ein Axiom, das *Auswahlaxiom*, das von einigen Mathematikern abgelehnt wurde. Es besagt, dass „es zu jeder Familie nichtleerer Mengen eine Auswahlfunktion gibt“. Genauer: Ist  $(X_i)_{i \in I}$  eine Familie von Mengen mit  $X_i \neq \emptyset$  für alle  $i \in I$ , dann gibt es eine *Auswahlfunktion*  $f : I \rightarrow X$ , wobei  $X = \{x \mid \exists i \in I : x \in X_i\}$  die Vereinigung aller Mengen  $X_i$  ist (die nach einem der harmlosen Axiome existiert), sodass für jedes  $i \in I$  das Bild  $f(i)$  ein Element von  $X_i$  ist. Die Auswahlfunktion wählt also aus jeder Menge  $X_i$  ein Element aus. Wir haben dieses Auswahlaxiom im Beweis von Teil (8) benutzt, als wir für jedes  $y \in Y$  ein Urbild  $x_y$  ausgewählt haben. Der Grund für die Ablehnung des Auswahlaxioms liegt darin, dass es nicht „konstruktiv“ ist: Es macht eine Existenzaussage („es gibt eine Auswahlfunktion“), sagt aber nicht, *wie* man eine Auswahlfunktion bekommt. Heutzutage vertreten die meisten Mathematiker den pragmatischen Standpunkt, dass

das Auswahlaxiom nützlich ist und es deswegen seine Berechtigung hat. Vor allem in der Analysis käme man ohne das Auswahlaxiom nicht weit. Es ist bekannt, dass die Hinzunahme des Auswahlaxioms nicht zu einem Widerspruch in der Mengenlehre führt (allerdings gilt das auch für seine Verneinung).

Zum Abschluss dieses Abschnitts über Grundlagen gibt es hier noch eine Tabelle mit griechischen Buchstaben. Als Mathematiker gehen einem schnell die Buchstaben aus, um die verschiedenen Objekte zu bezeichnen, mit denen man es zu tun hat. Darum wird gerne auf das griechische Alphabet zurückgegriffen.

klein	groß	Name
$\alpha$	A	Alpha
$\beta$	B	Beta
$\gamma$	$\Gamma$	Gamma
$\delta$	$\Delta$	Delta
$\varepsilon, \epsilon$	E	Epsilon
$\zeta$	Z	Zeta
$\eta$	H	Eta
$\theta, \vartheta$	$\Theta$	Theta

klein	groß	Name
$\iota$	I	Iota
$\kappa$	K	Kappa
$\lambda$	$\Lambda$	Lambda
$\mu$	M	My
$\nu$	N	Ny
$\xi$	$\Xi$	Xi
$o$	O	Omikron
$\pi$	$\Pi$	Pi

klein	groß	Name
$\rho, \varrho$	P	Rho
$\sigma$	$\Sigma$	Sigma
$\tau$	T	Tau
$\upsilon$	$\Upsilon$	Ypsilon
$\phi, \varphi$	$\Phi$	Phi
$\chi$	X	Chi
$\psi$	$\Psi$	Psi
$\omega$	$\Omega$	Omega

## 3. ALGEBRAISCHE STRUKTUREN: GRUPPEN, RINGE, KÖRPER

In diesem Abschnitt werden wir die wichtigsten algebraischen Strukturen einführen. Gruppen treten in vielen Zusammenhängen in der Mathematik auf, allerdings wird das hier in der Linearen Algebra noch nicht so deutlich werden. Für uns wichtig sind Körper (das sind Strukturen, in denen man die vier Grundrechenarten zusammen mit den üblichen Rechenregeln zur Verfügung hat), denn zu einem Vektorraum (das ist die Struktur, die in der Linearen Algebra hauptsächlich betrachtet wird) gehört immer ein Körper, aus dem die „Skalare“ kommen. Ringe sind gewissermaßen Körper ohne Division; sie sind als Zwischenschritt bei der Definition von Körpern praktisch und auch wichtig in der Algebra. Sie werden ausführlicher in der Vorlesung „Einführung in die Zahlentheorie und algebraische Strukturen“ untersucht.

Wir beginnen mit dem Minimum, das man für eine halbwegs interessante algebraische Struktur braucht.

**3.1. Definition.** Eine *Halbgruppe* ist ein Paar  $(H, *)$ , bestehend aus einer Menge  $H$  und einer Abbildung  $*$  :  $H \times H \rightarrow H$ ,  $(a, b) \mapsto a * b$ , die das *Assoziativgesetz* erfüllt:

**DEF**  
Halbgruppe

$$\forall a, b, c \in H : (a * b) * c = a * (b * c).$$

Die Halbgruppe heißt *kommutativ*, wenn zusätzlich das *Kommutativgesetz* gilt:

$$\forall a, b \in H : a * b = b * a.$$

Wenn die Verknüpfung  $*$  aus dem Kontext klar ist, spricht man der Einfachheit halber meist von „der Halbgruppe  $H$ “.  $\diamond$

Eine Bemerkung zur Notation: Verknüpfungen in algebraischen Strukturen wie  $*$  in obiger Definition werden gerne in „Infix-Notation“ geschrieben, also  $a * b$  statt  $*(a, b)$ .

Das Assoziativgesetz bewirkt, dass es nicht darauf ankommt, wie Ausdrücke, die drei oder mehr Elemente miteinander verknüpfen, geklammert sind. Zum Beispiel gilt für beliebige Elemente  $a, b, c, d, e$  von  $H$ :

$$\begin{aligned} a * ((b * c) * d) &= a * (b * (c * d)) = (a * b) * (c * d) \\ &= ((a * b) * c) * d = (a * (b * c)) * d \quad \text{und} \\ a * (b * (c * (d * e))) &= (a * b) * (c * (d * e)) = ((a * b) * (c * d)) * e = \dots \end{aligned}$$

Man kann deswegen einfach  $a * b * c * d$  bzw.  $a * b * c * d * e$  schreiben.

Hier ergibt sich die interessante kombinatorische Frage, *wie viele* verschiedene Klammierungen es für eine Verknüpfung von  $n$  Elementen gibt. Wir schreiben  $C_n$  für diese Zahl. Dann gilt offenbar  $C_1 = C_2 = 1$ ,  $C_3 = 2$  und  $C_4 = 5$ . Wenn man sich überlegt, dass man  $n$  Elemente dadurch verknüpfen kann, dass man eine Verknüpfung von  $k$  Elementen (mit  $1 \leq k < n$ ) mit einer Verknüpfung von  $n - k$  Elementen verknüpft, dann sieht man die folgende Rekursion für die Zahlen  $C_n$ :

$$C_n = \sum_{k=1}^{n-1} C_k C_{n-k} = C_1 C_{n-1} + C_2 C_{n-2} + \dots + C_{n-2} C_2 + C_{n-1} C_1 \quad \text{für alle } n \geq 2.$$

Damit kann man dann zum Beispiel  $C_5 = 1 \cdot 5 + 1 \cdot 2 + 2 \cdot 1 + 5 \cdot 1 = 14$ ,  $C_6 = 42$ ,  $C_7 = 132$  usw. berechnen. Es gibt auch eine Formel für  $C_n$ , nämlich

$$C_n = \frac{1}{n} \binom{2n-2}{n-1} = \frac{1}{2n-1} \binom{2n-1}{n-1} = \frac{(2n-2)!}{(n-1)!n!},$$

die aber direkt nicht so einfach zu beweisen ist (was Sie natürlich nicht von einem Versuch abhalten soll!). Die Zahlen  $C_n$  heißen *Catalan-Zahlen* (was die Bezeichnung erklärt; oft ist der Index verschoben und man fängt mit  $C_0 = C_1 = 1$  an) und treten in der Kombinatorik in vielen verschiedenen Zusammenhängen auf.

Wenn die Halbgruppe kommutativ ist, dann kommt es auch nicht auf die Reihenfolge an:

$$a * b * c = b * a * c = b * c * a = c * b * a = c * a * b = a * c * b.$$

**3.2. Beispiele.** Das Trivialbeispiel einer Halbgruppe ist  $(\emptyset, *)$ , wobei  $* : \emptyset \times \emptyset \rightarrow \emptyset$  die leere Abbildung ist (beachte:  $\emptyset \times \emptyset = \emptyset$ ).

**BSP**  
Halbgruppen

Beispiele von kommutativen Halbgruppen sind  $(\mathbb{N}_+, +)$ ,  $(\mathbb{N}, +)$ ,  $(\mathbb{Z}, +)$ ,  $(\mathbb{N}_+, \cdot)$ ,  $(\mathbb{N}, \cdot)$ ,  $(\mathbb{Z}, \cdot)$ . Die Halbgruppe  $(\text{Abb}(X, X), \circ)$  für eine beliebige Menge  $X$ , mit der Komposition von Abbildungen als Verknüpfung, ist im Allgemeinen nicht kommutativ. (Diese Halbgruppe ist genau dann kommutativ, wenn  $X$  höchstens ein Element hat — Übung!)



Mit Halbgruppen kann man allerdings noch nicht allzu viel anfangen. Deshalb fordern wir zusätzliche Eigenschaften.

**3.3. Definition.** Ein *Monoid* ist ein Tripel  $(M, *, e)$ , bestehend aus einer Menge  $M$ , einer Abbildung  $* : M \times M \rightarrow M$  und einem Element  $e \in M$ , sodass  $(M, *)$  eine Halbgruppe mit *neutralem Element*  $e$  ist:

**DEF**  
Monoid

$$\forall a \in M : e * a = a = a * e.$$

Das Monoid heißt *kommutativ*, wenn die Halbgruppe  $(M, *)$  kommutativ ist.  $\diamond$

Wenn es ein neutrales Element gibt, dann ist es eindeutig bestimmt, wie das folgende Lemma zeigt. (Ein *Lemma* ist eine Hilfsaussage oder ein weniger wichtiger mathematischer Satz.)

**3.4. Lemma.** Sei  $(H, *)$  eine Halbgruppe. Ist  $e$  ein links- und  $e'$  ein rechtsneutrales Element in dieser Halbgruppe, also

**LEMMA**  
Eindeutigkeit  
des neutrales  
Elements

$$\forall a \in H : e * a = a \text{ und } a * e' = a,$$

dann gilt  $e = e'$ .

*Beweis.* Da  $e$  linksneutral ist, gilt  $e * e' = e'$ . Da  $e'$  rechtsneutral ist, gilt  $e * e' = e$ . Es folgt  $e = e'$ .  $\square$

Aus diesem Grund lässt man meistens die Angabe des neutralen Elements weg und spricht vom „Monoid  $(M, *)$ “ oder auch nur vom „Monoid  $M$ “, wenn die Verknüpfung aus dem Kontext klar ist.

Es ist allerdings möglich, dass es in einer Halbgruppe zum Beispiel mehrere linksneutrale Elemente (und dann natürlich kein rechtsneutrales Element) gibt. Wenn etwa  $M$  beliebig ist und man als Verknüpfung  $\text{pr}_2$  wählt (also  $a * b = b$ ), dann hat man eine Halbgruppe, in der *alle* Elemente linksneutral sind.

**3.5. Beispiele.** Da die Definition von „Monoid“ ein neutrales Element fordert, kann die leere Menge kein Monoid sein. Das triviale Monoid ist dann  $(\{e\}, *, e)$ , wobei  $*$  die einzige Abbildung  $\{e\} \times \{e\} \rightarrow \{e\}$  ist (es ist also  $e * e = e$ ).

**BSP**  
Monoide

Bis auf  $(\mathbb{N}_+, +)$ , wo es kein neutrales Element gibt, lassen sich alle Beispiele von Halbgruppen aus 3.2 als Monoide  $(\mathbb{N}, +, 0)$ ,  $(\mathbb{Z}, +, 0)$ ,  $(\mathbb{N}_+, \cdot, 1)$ ,  $(\mathbb{N}, \cdot, 1)$ ,  $(\mathbb{Z}, \cdot, 1)$  und  $(\text{Abb}(X, X), \circ, \text{id}_X)$  betrachten. ♣

Noch schöner ist es, wenn sich die Verknüpfung mit einem Element durch die Verknüpfung mit einem (in der Regel) anderen Element wieder rückgängig machen lässt. Das führt auf den Begriff der Gruppe.

**3.6. Definition.** Eine *Gruppe* ist ein Quadrupel  $(G, *, e, i)$ , bestehend aus einer Menge  $G$ , einer Abbildung  $* : G \times G \rightarrow G$ , einem Element  $e \in G$  und einer Abbildung  $i : G \rightarrow G$ , sodass  $(G, *, e)$  ein Monoid ist und für jedes  $g \in G$  das Element  $i(g) \in G$  ein *Inverses* von  $g$  ist:

**DEF**  
Gruppe

$$\forall g \in G : i(g) * g = e = g * i(g).$$

Die Gruppe heißt *kommutativ* oder *abelsch*, wenn das Monoid  $(G, *, e)$  kommutativ ist. ◇

Die Bezeichnung „abelsch“ ehrt den norwegischen Mathematiker Niels Henrik Abel, nach dem auch der *Abelpreis* benannt ist, ein dem Nobelpreis vergleichbarer Preis für Mathematik, der seit 2003 jährlich verliehen wird.

Auch Inverse sind eindeutig bestimmt:

**3.7. Lemma.** Sei  $(M, *, e)$  ein Monoid und sei  $a \in M$ . Ist  $b \in M$  ein *Linksinverse* und  $c \in M$  ein *Rechtsinverse* von  $a$ , also

**LEMMA**  
Eindeutigkeit  
des Inversen

$$b * a = e = a * c,$$

dann gilt  $b = c$ .

*Beweis.* Wir haben

$$b = b * e = b * (a * c) = (b * a) * c = e * c = c. \quad \square$$

Analog zu Monoiden spricht man deshalb auch einfach von „der Gruppe  $(G, *)$ “ oder auch von „der Gruppe  $G$ “, wenn die Verknüpfung aus dem Kontext klar ist.

Gruppen schreibt man gerne „multiplikativ“, dann ist die Verknüpfung  $a \cdot b$  oder kurz  $ab$ , das neutrale Element heißt 1 und das Inverse von  $a$  wird  $a^{-1}$  geschrieben. Kommutative Gruppen schreibt man auch häufig „additiv“, dann ist die Verknüpfung  $a + b$ , das neutrale Element heißt 0 und das Inverse von  $a$  wird als das Negative von  $a$  geschrieben:  $-a$ . Dann schreibt man auch kurz  $a - b$  für  $a + (-b)$ .

**3.8. Beispiele.** Das triviale Monoid lässt sich auch als Gruppe betrachten, denn das einzige Element  $e$  ist sein eigenes Inverses.

**BSP**  
Gruppe

Von den übrigen Beispielen von Monoiden in 3.5 kann nur  $(\mathbb{Z}, +, 0, -)$  auch als Gruppe betrachtet werden (und im letzten Beispiel  $\text{Abb}(X, X)$ , wenn  $X$  höchstens ein Element hat; dann hat man eine triviale Gruppe). Ein weiteres Beispiel einer kommutativen Gruppe ist  $(\mathbb{R}_{>0}, \cdot, 1, x \mapsto 1/x)$ , wobei  $\mathbb{R}_{>0}$  die Menge der positiven reellen Zahlen ist.

Wenn man sich bei den Abbildungen  $X \rightarrow X$  auf die bijektiven Abbildungen beschränkt, dann erhält man eine Gruppe  $(S(X), \circ, \text{id}_X, f \mapsto f^{-1})$ , die auch die *symmetrische Gruppe* von  $X$  heißt. Dabei ist

$$S(X) = \{f : X \rightarrow X \mid f \text{ bijektiv}\}.$$

Diese Gruppe ist genau dann kommutativ, wenn  $X$  höchstens zwei Elemente enthält (Übung).

Gruppen treten häufig in der Mathematik als „Symmetriegruppen“ von irgendwelchen Objekten auf. Zum Beispiel bilden die Drehungen und Spiegelungen der Ebene, die ein regelmäßiges  $n$ -Eck auf sich abbilden, eine Gruppe, oder die Drehungen des dreidimensionalen Raumes, die ein reguläres Tetraeder, einen Würfel (oder ein reguläres Oktaeder) oder ein reguläres Dodekaeder (oder Ikosaeder) in sich abbilden, bilden jeweils eine Gruppe, die Tetraeder-, Oktaeder- und Ikosaedergruppe. In einem recht allgemeinen Sinn ist die symmetrische Gruppe  $S(X)$  die Symmetriegruppe der Menge  $X$  ohne weitere Struktur. In der Algebra treten Symmetriegruppen als „Automorphismengruppen“ auf. Zum Beispiel bildet für eine Gruppe  $(G, *)$  die Menge

$$\text{Aut}(G) = \{f : G \rightarrow G \mid f \text{ bijektiv und } \forall g, g' \in G : f(g * g') = f(g) * f(g')\}$$

mit der Komposition von Abbildungen eine Gruppe, die *Automorphismengruppe* von  $G$ . Sie besteht aus den bijektiven Abbildungen  $G \rightarrow G$ , die mit der Struktur von  $G$  als Gruppe verträglich sind. ♣

Damit eine Halbgruppe sogar eine Gruppe ist, genügt es, die Existenz eines linksneutralen Elements  $e$  und für jedes Element  $x$  die Existenz eines Linksinversen  $i(x)$  (also mit  $i(x) * x = e$ ) zu fordern. Dann folgt zunächst, dass  $e$  auch rechtsneutral ist, denn es gilt  $x * e = e * x * e = i(i(x)) * i(x) * x * e = i(i(x)) * e * e = i(i(x)) * e = i(i(x)) * i(x) * x = e * x = x$ .

Daraus ergibt sich auch  $i(i(x)) = x$ . Damit kann man dann zeigen, dass  $i(x)$  auch Rechtsinverses von  $x$  ist:

$$x * i(x) = i(i(x)) * i(x) = e.$$

Ganz analog funktioniert das natürlich auch, wenn man „links“ jeweils durch „rechts“ ersetzt. Auf der anderen Seite gibt es aber Halbgruppen mit linksneutralen und rechtsinversen Elementen, die keine Gruppen sind. Finden Sie ein Beispiel!

Eine wichtige Eigenschaft von Gruppen ist, dass sich gewisse Gleichungen stets eindeutig lösen lassen. Zuerst beweisen wir aber eine Kürzungsregel.

**3.9. Lemma.** Sei  $(G, *, e, i)$  eine Gruppe und seien  $a, b, c \in G$ . Dann gilt

$$a * c = b * c \iff a = b \iff c * a = c * b.$$

**LEMMA**  
Kürzungsregel  
in Gruppen

*Beweis.* Wir beweisen die erste Äquivalenz; der Beweis der zweiten ist analog.

„ $\Leftarrow$ “ ist klar. Für „ $\Rightarrow$ “ haben wir

$$\begin{aligned} a * c = b * c &\implies (a * c) * i(c) = (b * c) * i(c) \implies a * (c * i(c)) = b * (c * i(c)) \\ &\implies a * e = b * e \implies a = b. \end{aligned} \quad \square$$

**3.10. Lemma.** Sei  $(G, *, e, i)$  eine Gruppe und seien  $a, b \in G$ . Dann haben die Gleichungen

$$a * x = b \quad \text{und} \quad x * a = b$$

jeweils eine eindeutige Lösung  $x \in G$ , nämlich  $x = i(a) * b$  bzw.  $x = b * i(a)$ .

**LEMMA**  
Gleichungen  
in Gruppen

*Beweis.* Wir führen den Beweis exemplarisch für die erste Gleichung:

$$a * x = b \iff i(a) * a * x = i(a) * b \iff e * x = i(a) * b \iff x = i(a) * b.$$

Für die erste Äquivalenz haben wir Lemma 3.9 benutzt.  $\square$

Als Nächstes betrachten wir Strukturen mit zwei Verknüpfungen.

**3.11. Definition.** Ein *Ring* ist ein Sextupel  $(R, +, 0, -, \cdot, 1)$ , bestehend aus einer Menge  $R$ , Abbildungen  $+, \cdot : R \times R \rightarrow R$ , Elementen  $0, 1 \in R$  und einer Abbildung  $- : R \rightarrow R$ , sodass  $(R, +, 0, -)$  eine kommutative Gruppe und  $(R, \cdot, 1)$  ein Monoid ist und die *Distributivgesetze*

**DEF**  
Ring

$$\forall a, b, c \in R : a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{und} \quad (a + b) \cdot c = a \cdot c + b \cdot c$$

gelten. Der Ring heißt *kommutativ*, wenn das Monoid  $(R, \cdot, 1)$  kommutativ ist.  $\diamond$

Da die neutralen und inversen Elemente eindeutig bestimmt sind, spricht man oft nur vom „Ring  $(R, +, \cdot)$ “ oder sogar vom „Ring  $R$ “, wenn die Verknüpfungen aus dem Kontext klar sind. Ist der Ring kommutativ, dann genügt es, eines der beiden Distributivgesetze zu fordern. Für das Produkt  $a \cdot b$  zweier Elemente schreibt man auch kurz  $ab$ .

In einem Ring kann man also addieren, subtrahieren und multiplizieren, und die üblichen Rechenregeln gelten, wie zum Beispiel  $0 \cdot a = a \cdot 0 = 0$ ,  $-(a + b) = -a - b$ ,  $(-a) \cdot (-b) = a \cdot b$ . Was aber im Allgemeinen *nicht* gelten muss, ist die Implikation  $a \cdot b = 0 \Rightarrow a = 0 \vee b = 0$ .

**3.12. Beispiele.** Das Trivialbeispiel für einen Ring ist der sogenannte *Nullring*  $(\{0\}, +, 0, -, \cdot, 0)$ , in dem  $0 = 1$  und  $0 + 0 = -0 = 0 \cdot 0 = 0$  gelten. Jeder Ring  $R$ , in dem  $0_R = 1_R$  gilt, ist so ein Nullring, denn für alle  $r \in R$  gilt dann  $r = 1_R \cdot r = 0_R \cdot r = 0_R$ .

**BSP**  
Ringe

Das Standardbeispiel für einen (kommutativen) Ring ist der Ring  $\mathbb{Z}$  der ganzen Zahlen mit der üblichen Addition und Multiplikation als Verknüpfungen. Ein etwas anders geartetes Beispiel ist  $(\mathcal{P}(X), \Delta, \emptyset, \text{id}_{\mathcal{P}(X)}, \cap, X)$  für eine beliebige Menge  $X$ ; dabei ist  $T_1 \Delta T_2 = (T_1 \setminus T_2) \cup (T_2 \setminus T_1)$  die „symmetrische Differenz“ der Mengen  $T_1$  und  $T_2$  (Übung).

Falls Sie aus der Schule Matrizen kennen und wissen, wie man sie addiert und multipliziert, dann können Sie nachprüfen, dass die Menge der  $2 \times 2$ -Matrizen mit Einträgen aus  $\mathbb{R}$  zusammen mit der Addition und Multiplikation von Matrizen einen nicht-kommutativen Ring bildet.  $\clubsuit$

Schließlich kommen wir zu den Körpern.

\*

**3.13. Definition.** Ein *Körper* ist ein Septupel  $(K, +, 0, -, \cdot, 1, i)$ , bestehend aus einer Menge  $K$ , Abbildungen  $+, \cdot : K \times K \rightarrow K$ , Elementen  $0, 1 \in K$  und Abbildungen  $-, i : K \rightarrow K$ , sodass  $(K, +, 0, -, \cdot, 1)$  ein kommutativer Ring und  $(K \setminus \{0\}, \cdot, 1, i)$  eine (kommutative) Gruppe ist. Für  $i(a)$  schreibt man  $a^{-1}$ .  $\diamond$

**DEF**  
Körper

Wie üblich spricht man meistens einfach von dem „Körper  $(K, +, \cdot)$ “ oder von dem „Körper  $K$ “. Aus der Definition folgt, dass 0 und 1 in einem Körper verschieden sein müssen, denn 1 soll das neutrale Element der Gruppe  $K \setminus \{0\}$  sein. Für diese Gruppe  $(K \setminus \{0\}, \cdot)$  schreibt man auch  $K^\times$  und nennt sie die *multiplikative Gruppe* von  $K$ . (Häufig findet man auch die Schreibweise  $K^*$  dafür.)

Man kann natürlich auch ohne Rückgriff auf Ringe und Gruppen definieren, was ein Körper ist. Dann hat man für alle  $a, b, c \in K$  die folgenden Axiome:

$$\begin{array}{ll} (a+b)+c = a+(b+c), & a+b = b+a \\ a+0 = a, & a+(-a) = 0 \\ (a \cdot b) \cdot c = a \cdot (b \cdot c) & a \cdot b = b \cdot a \\ a \cdot 1 = a, & a \neq 0 \Rightarrow a \cdot a^{-1} = 1 \\ 0 \neq 1, & a \cdot (b+c) = a \cdot b + a \cdot c \end{array}$$

Für  $a, b \in K$ ,  $b \neq 0$ , kann man die Division definieren durch  $a/b = a \cdot b^{-1}$ . Dann hat man die vier Grundrechenarten zur Verfügung und die üblichen Rechenregeln dafür gelten, denn man kann sie aus den Körperaxiomen ableiten. Zum Beispiel gilt in einem Körper stets, dass aus  $a \cdot b = 0$  folgt, dass  $a = 0$  oder  $b = 0$  ist. (Denn ist  $a \neq 0$ , dann folgt  $0 = a^{-1} \cdot 0 = a^{-1} \cdot a \cdot b = 1 \cdot b = b$ .)

**3.14. Beispiele.** Das kleinste Beispiel für einen Körper hat nur die beiden Elemente 0 und 1, die in der Definition gefordert werden. Für die Addition und Multiplikation folgt  $0+0=0$ ,  $0+1=1+0=1$ ,  $0 \cdot 0=0 \cdot 1=1 \cdot 0=0$  und  $1 \cdot 1=1$  direkt aus der Definition; für die verbleibende Summe  $1+1$  bleibt nur der Wert 0, da die Gleichung  $a+1=0$  lösbar sein muss. Man kann (einfach, aber länglich) nachprüfen, dass dieser Körper, der mit  $\mathbb{F}_2$  bezeichnet wird, die Axiome erfüllt.

**BSP**  
Körper

Es gibt noch weitere endliche Körper: Zu jeder Potenz  $p^e$  einer Primzahl  $p$  (mit  $e \geq 1$ ) gibt es im Wesentlichen genau einen Körper mit  $p^e$  Elementen, und es gibt keine anderen endlichen Körper. Das wird in der „Einführung in die Algebra“ genauer besprochen.

Standardbeispiele für Körper sind der Körper  $\mathbb{Q}$  der rationalen Zahlen und der Körper  $\mathbb{R}$  der reellen Zahlen, jeweils mit der bekannten Addition und Multiplikation. Im nächsten Abschnitt werden wir einen weiteren Körper konstruieren, den Körper  $\mathbb{C}$  der komplexen Zahlen.  $\clubsuit$

## 4. DER KÖRPER DER KOMPLEXEN ZAHLEN

Der Körper  $\mathbb{R}$  der reellen Zahlen hat, wie Sie in der Analysis lernen, viele schöne Eigenschaften. Eine Eigenschaft allerdings fehlt ihm: Es sind nicht alle Gleichungen der Form

$$x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 = 0$$

(mit  $n \geq 1$  und  $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$ ) in  $\mathbb{R}$  lösbar.

Für ungerades  $n$  folgt aus dem Zwischenwertsatz, dass es stets eine Lösung geben muss; das lernen Sie bald in der Analysis.

Die einfachste Gleichung dieser Art ohne Lösung ist  $x^2 + 1 = 0$ : Die linke Seite ist stets  $\geq 1$ , kann also niemals null werden. Wir werden jetzt einen  $\mathbb{R}$  umfassenden Körper konstruieren, in dem diese Gleichung eine Lösung hat.

Um zu sehen, wie man dabei vorgehen kann, stellen wir uns einfach einmal vor, dass wir schon so einen Körper  $\mathbb{C}$  hätten. Dann haben wir eine Lösung  $i$  obiger Gleichung, also ein Element  $i \in \mathbb{C}$  mit  $i^2 = -1$ . Wir haben natürlich auch die reellen Zahlen in  $\mathbb{C}$ . Mit  $a, b \in \mathbb{R}$  können wir dann das Element  $a + bi \in \mathbb{C}$  erzeugen. Muss es noch weitere Elemente geben? Dazu müssen wir überprüfen, ob die vier Grundrechenarten aus der Menge der Elemente der Form  $a + bi$  herausführen. Seien  $a, b, a', b' \in \mathbb{R}$ . Dann gilt, wenn  $\mathbb{C}$  ein Körper ist,

$$\begin{aligned} (a + bi) + (a' + b'i) &= (a + a') + (b + b')i && \text{und} \\ (a + bi) \cdot (a' + b'i) &= aa' + ab'i + ba'i + bb'i^2 = (aa' - bb') + (ab' + ba')i. \end{aligned}$$

Dabei haben wir  $i^2 = -1$  benutzt. Offensichtlich ist das additive Inverse (also das Negative) von  $a + bi$  gerade  $(-a) + (-b)i$ . Wie sieht es mit dem multiplikativen Inversen aus (also dem Kehrwert)? Dazu überlegen wir uns erst, dass genau dann  $a + bi = 0$  ist, wenn  $a = b = 0$  gilt. Eine Richtung („ $\Leftarrow$ “) ist klar. Umgekehrt sei  $a + bi = 0$ . Dann folgt

$$0 = (a - bi) \cdot 0 = (a - bi) \cdot (a + bi) = a^2 + b^2.$$

Da  $a$  und  $b$  reelle Zahlen sind, ist das nur möglich, wenn  $a = b = 0$  gilt. Seien also  $a$  und  $b$  nicht beide null. Dann sollte gelten (das ist der alte Trick, wie man „Quadratwurzeln aus dem Nenner entfernt“; man beachte, dass  $i = \sqrt{-1}$ ):

$$\frac{1}{a + bi} = \frac{a - bi}{(a - bi)(a + bi)} = \frac{a - bi}{a^2 + b^2} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2}i.$$

Offenbar brauchen wir also keine zusätzlichen Elemente.

Um das Ganze formal auf eine solide Grundlage zu stellen, ersetzen wir einen Ausdruck der Form  $a + bi$  durch das Paar  $(a, b) \in \mathbb{R} \times \mathbb{R}$ . Wir schreiben  $\mathbb{C}$  für  $\mathbb{R} \times \mathbb{R}$  und definieren die folgenden Abbildungen:

$$\begin{aligned} +_{\mathbb{C}} : \mathbb{C} \times \mathbb{C} &\longrightarrow \mathbb{C}, && ((a, b), (a', b')) \longmapsto (a + a', b + b') \\ \cdot_{\mathbb{C}} : \mathbb{C} \times \mathbb{C} &\longrightarrow \mathbb{C}, && ((a, b), (a', b')) \longmapsto (aa' - bb', ab' + ba') \\ -_{\mathbb{C}} : \mathbb{C} &\longrightarrow \mathbb{C}, && (a, b) \longmapsto (-a, -b) \\ i_{\mathbb{C}} : \mathbb{C} \setminus \{(0, 0)\} &\longrightarrow \mathbb{C}, && (a, b) \longmapsto \left( \frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right) \end{aligned}$$

Außerdem schreiben wir  $0_{\mathbb{C}}$  und  $1_{\mathbb{C}}$  für  $(0, 0)$  und  $(1, 0)$ .

**4.1. Satz.** Die Menge  $\mathbb{C} = \mathbb{R} \times \mathbb{R}$  zusammen mit den oben definierten Abbildungen und Elementen bildet einen Körper.

**SATZ**  
Körper  $\mathbb{C}$

*Beweis.* Es sind die verschiedenen Axiome nachzuprüfen. Für die additive Gruppe  $(\mathbb{C}, +_{\mathbb{C}}, 0_{\mathbb{C}}, -_{\mathbb{C}})$  ist das sehr leicht; darum lassen wir das hier weg (es sei Ihnen aber als Übung empfohlen). Wir prüfen Assoziativität und Kommutativität der Multiplikation. Dabei benutzen wir, dass  $\mathbb{R}$  ein Körper ist, dass also dort die bekannten Rechenregeln gelten.

$$\begin{aligned} ((a, b) \cdot_{\mathbb{C}} (a', b')) \cdot_{\mathbb{C}} (a'', b'') &= (aa' - bb', ab' + a'b) \cdot_{\mathbb{C}} (a'', b'') \\ &= ((aa' - bb')a'' - (ab' + a'b)b'', (aa' - bb')b'' + (ab' + a'b)a'') \\ &= (aa'a'' - ab'b'' - ba'b'' - bb'a'', aa'b'' + ab'a'' + ba'a'' - bb'b'') \end{aligned}$$

und dasselbe Resultat erhalten wir aus  $(a, b) \cdot_{\mathbb{C}} ((a', b') \cdot_{\mathbb{C}} (a'', b''))$ . Ebenso gilt

$$(a, b) \cdot_{\mathbb{C}} (a', b') = (aa' - bb', ab' + ba') = (a'a - b'b, ba' + ab') = (a', b') \cdot_{\mathbb{C}} (a, b).$$

Dass  $1_{\mathbb{C}} = (1, 0)$  neutrales Element der Multiplikation ist, folgt aus

$$(1, 0) \cdot_{\mathbb{C}} (a, b) = (1 \cdot a - 0 \cdot b, 1 \cdot b + 0 \cdot a) = (a, b).$$

Wir rechnen nach, dass  $i_{\mathbb{C}}((a, b))$  das multiplikative Inverse von  $(a, b) \neq (0, 0)$  ist:

$$\begin{aligned} (a, b) \cdot_{\mathbb{C}} i_{\mathbb{C}}((a, b)) &= (a, b) \cdot_{\mathbb{C}} \left( \frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right) \\ &= \left( \frac{a^2}{a^2 + b^2} - \frac{-b^2}{a^2 + b^2}, \frac{-ab}{a^2 + b^2} + \frac{ba}{a^2 + b^2} \right) \\ &= (1, 0) = 1_{\mathbb{C}}. \end{aligned}$$

$0_{\mathbb{C}} \neq 1_{\mathbb{C}}$  ist klar. Es bleibt das Distributivgesetz nachzuprüfen:

$$\begin{aligned} (a, b) \cdot_{\mathbb{C}} ((a', b') +_{\mathbb{C}} (a'', b'')) &= (a, b) \cdot_{\mathbb{C}} (a' + a'', b' + b'') \\ &= (a(a' + a'') - b(b' + b''), a(b' + b'') + b(a' + a'')) \\ &= (aa' + aa'' - bb' - bb'', ab' + ab'' + ba' + ba'') \\ &= (aa' - bb' + aa'' - bb'', ab' + ba' + ab'' + ba'') \\ &= (aa' - bb', ab' + ba') +_{\mathbb{C}} (aa'' - bb'', ab'' + ba'') \\ &= (a, b) \cdot_{\mathbb{C}} (a', b') +_{\mathbb{C}} (a, b) \cdot_{\mathbb{C}} (a'', b''). \quad \square \end{aligned}$$

Ist  $a$  eine reelle Zahl, dann haben wir das Element  $a_{\mathbb{C}} = (a, 0) \in \mathbb{C}$ . Für  $a, b \in \mathbb{R}$  gilt

$$a = b \iff a_{\mathbb{C}} = b_{\mathbb{C}}, \quad (a + b)_{\mathbb{C}} = a_{\mathbb{C}} +_{\mathbb{C}} b_{\mathbb{C}} \quad \text{und} \quad (ab)_{\mathbb{C}} = a_{\mathbb{C}} \cdot b_{\mathbb{C}}.$$

Mit den Elementen  $a_{\mathbb{C}}$  rechnet man also genauso wie mit den zugehörigen reellen Zahlen  $a$ . Deswegen macht man keinen Unterschied zwischen  $a$  und  $a_{\mathbb{C}}$  und betrachtet  $\mathbb{R}$  als eine Teilmenge von  $\mathbb{C}$ . Wir schreiben also einfach  $a$  für das Element  $a_{\mathbb{C}} = (a, 0)$  von  $\mathbb{C}$ . Außerdem schreiben wir ab jetzt der Einfachheit halber meistens  $+$ ,  $\cdot$  und so weiter statt  $+_{\mathbb{C}}$ ,  $\cdot_{\mathbb{C}}$  etc.

**4.2. Definition.** Der in Satz 4.1 eingeführte Körper  $\mathbb{C}$  heißt der *Körper der komplexen Zahlen*. Wir schreiben  $i$  für das Element  $(0, 1) \in \mathbb{C}$ . Dann gilt  $i^2 = -1$ , und jedes Element  $z = (a, b) \in \mathbb{C}$  kann geschrieben werden als  $z = a + bi$  (oder  $a + ib$ ) mit  $a, b \in \mathbb{R}$ . Dann heißt  $a$  der *Realteil*  $\operatorname{Re} z$  und  $b$  der *Imaginärteil*  $\operatorname{Im} z$  von  $z$ . Gilt  $\operatorname{Re} z = 0$ , dann heißt  $z$  *rein imaginär*.  $\diamond$

**DEF**  
Körper der  
komplexen  
Zahlen

Die letzten beiden Behauptungen sollten wir nachprüfen:

$$i^2 = (0, 1) \cdot_{\mathbb{C}} (0, 1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0) = (-1)_{\mathbb{C}} = -1$$

und

$$a + bi = (a, 0) +_{\mathbb{C}} (b, 0) \cdot_{\mathbb{C}} (0, 1) = (a, 0) +_{\mathbb{C}} (0, b) = (a, b).$$

Wir können jetzt immerhin zeigen, dass man quadratische Gleichungen in  $\mathbb{C}$  stets lösen kann.

**4.3. Satz.** Seien  $a, b, c \in \mathbb{C}$  mit  $a \neq 0$ . Dann hat die Gleichung

$$az^2 + bz + c = 0$$

mindestens eine Lösung  $z \in \mathbb{C}$ .

**SATZ**  
quadratische  
Gleichungen  
in  $\mathbb{C}$

*Beweis.* Die Gleichung ist äquivalent zu  $(2az + b)^2 = b^2 - 4ac$ . Es genügt also zu zeigen, dass jede komplexe Zahl eine Quadratwurzel in  $\mathbb{C}$  hat. Sei also  $u + vi \in \mathbb{C}$  (mit  $u, v \in \mathbb{R}$ ). Wir wollen  $z = x + yi \in \mathbb{C}$  finden mit  $z^2 = u + vi$ . Das bedeutet

$$x^2 - y^2 + 2xyi = u + vi, \quad \text{also} \quad x^2 - y^2 = u \quad \text{und} \quad 2xy = v.$$

Aus

$$u^2 + v^2 = (x^2 - y^2)^2 + (2xy)^2 = x^4 - 2x^2y^2 + y^4 + 4x^2y^2 = x^4 + 2x^2y^2 + y^4 = (x^2 + y^2)^2$$

erhalten wir  $x^2 + y^2 = \sqrt{u^2 + v^2}$  (die Quadratwurzel existiert in  $\mathbb{R}_{\geq 0}$ , da  $u^2 + v^2 \geq 0$  ist). Damit sollte gelten:

$$x^2 = \frac{(x^2 + y^2) + (x^2 - y^2)}{2} = \frac{\sqrt{u^2 + v^2} + u}{2} \quad \text{und}$$

$$y^2 = \frac{(x^2 + y^2) - (x^2 - y^2)}{2} = \frac{\sqrt{u^2 + v^2} - u}{2}.$$

Weil  $\sqrt{u^2 + v^2} \geq |u|$  ist, sind beide Ausdrücke  $\geq 0$ . Wir können also  $x$  und  $y$  wie folgt definieren (das Vorzeichen  $s(v)$  ist nötig, damit  $2xy$  den richtigen Wert bekommt):

$$x = \sqrt{\frac{\sqrt{u^2 + v^2} + u}{2}} \quad \text{und} \quad y = s(v) \sqrt{\frac{\sqrt{u^2 + v^2} - u}{2}}.$$

mit  $s(v) = 1$ , wenn  $v \geq 0$ , und  $s(v) = -1$ , wenn  $v < 0$ . Dann haben wir

$$x^2 - y^2 = \frac{\sqrt{u^2 + v^2} + u}{2} - \frac{\sqrt{u^2 + v^2} - u}{2} = u$$

und

$$2xy = 2s(v) \sqrt{\frac{\sqrt{u^2 + v^2} + u}{2}} \cdot \frac{\sqrt{u^2 + v^2} - u}{2}$$

$$= 2s(v) \sqrt{\frac{u^2 + v^2 - u^2}{4}} = s(v) \sqrt{v^2} = s(v)|v| = v,$$

also ist  $z = x + yi$  die gesuchte Quadratwurzel von  $u + vi$ .  $\square$

Es gilt sogar noch viel mehr.

**4.4. Satz.** *Jede Gleichung*

$$z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0 = 0$$

mit  $n \geq 1$  und  $a_0, a_1, \dots, a_{n-1} \in \mathbb{C}$  hat mindestens eine Lösung  $z \in \mathbb{C}$ .

Beweisen können wir diesen Satz hier nicht. Es gibt verschiedene Beweise; der wohl einfachste verwendet den *Satz von Liouville* aus der Funktionentheorie. Sie werden ihn in der „Einführung in die Funktionentheorie“ kennenlernen.

Ein Körper  $K$ , sodass jede Gleichung

$$x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 = 0$$

mit  $n \geq 1$  und  $a_0, a_1, \dots, a_{n-1} \in K$  eine Lösung  $x \in K$  hat, heißt *algebraisch abgeschlossen*. Der „Fundamentalsatz der Algebra“ lässt sich also auch so formulieren:

*Der Körper der komplexen Zahlen ist algebraisch abgeschlossen.*

Demgegenüber ist der Körper der reellen Zahlen *nicht* algebraisch abgeschlossen, wie wir gesehen haben. In dieser Hinsicht ist  $\mathbb{C}$  also „besser“ als  $\mathbb{R}$ . Auf der anderen Seite ist  $\mathbb{C}$  kein angeordneter Körper mehr; man verliert also auch etwas beim Übergang von  $\mathbb{R}$  zu  $\mathbb{C}$ . (In einem angeordneten Körper  $K$  gilt  $x^2 \geq 0$  für  $x \in K$ . Damit müsste in  $\mathbb{C}$  gelten, dass  $-1 = i^2 \geq 0$  ist, aber  $-1$  ist in einem angeordneten Körper immer negativ, und wir haben einen Widerspruch.)

Man kann sich die komplexen Zahlen ganz gut veranschaulichen, wenn man sich daran erinnert, dass  $\mathbb{C} = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$  der Menge der Punkte der Ebene entspricht. Wenn man die Ebene so interpretiert, spricht man auch von der *komplexen (Zahlen-)Ebene*. Die Addition entspricht dann dem, was Sie aus der Physik als „Kräfteparallelogramm“ kennen.

Auch die Multiplikation lässt sich geometrisch interpretieren. Wir betrachten dazu  $z = a + bi \in \mathbb{C}$ . Dann ist  $a^2 + b^2 \geq 0$ ; man setzt  $|z| = \sqrt{a^2 + b^2}$  und nennt das den *Absolutbetrag* von  $z$ . Das entspricht dem Abstand des Punktes  $z$  in der komplexen Ebene vom Ursprung  $0 \in \mathbb{C}$ . Für  $z \in \mathbb{R}$  (also  $b = 0$ ) bekommt man den bekannten Absolutbetrag auf  $\mathbb{R}$ . Ist  $z \neq 0$ , dann hat  $w = z/|z|$  den Absolutbetrag 1. Wenn wir  $w = u + vi$  schreiben, dann gilt  $u^2 + v^2 = 1$ , also liegt der Punkt  $(u, v)$  auf dem Einheitskreis. Es gibt dann  $\alpha \in \mathbb{R}$  mit  $u = \cos \alpha$ ,  $v = \sin \alpha$ . Dieser Winkel  $\alpha$  heißt auch das *Argument* von  $w$  und von  $z$ . Da die Beziehung

$$(\cos \alpha + i \sin \alpha)(\cos \beta + i \sin \beta) = \cos(\alpha + \beta) + i \sin(\alpha + \beta)$$

gilt (Übung), addieren sich die Winkel bei Multiplikation. Man kann das dann so formulieren: Multiplikation mit  $z \neq 0$  bewirkt eine *Drehstreckung* der komplexen Ebene mit dem Drehwinkel  $\alpha$  und dem Streckfaktor  $|z|$ .

Da mit  $i$  auch  $-i$  eine Lösung von  $x^2 + 1 = 0$  ist, könnte man überall  $i$  durch  $-i$  ersetzen und alles würde genauso funktionieren. Das führt auf den Begriff der komplexen Konjugation. Für  $z = a + bi \in \mathbb{C}$  setzen wir  $\bar{z} = a - bi$ ; die Abbildung  $\mathbb{C} \rightarrow \mathbb{C}$ ,  $z \mapsto \bar{z}$ , heißt die *komplexe Konjugation*. Es gilt  $\overline{w + z} = \bar{w} + \bar{z}$  und  $\overline{wz} = \bar{w} \cdot \bar{z}$  (leichte Übung); außerdem  $z\bar{z} = a^2 + b^2 = |z|^2$  (das haben wir schon benutzt). Daraus bekommt man die Formel  $z^{-1} = \bar{z}/|z|^2$  für den Kehrwert einer komplexen Zahl  $z \neq 0$ ; das ist derselbe Ausdruck, den wir bereits hergeleitet haben, in einer etwas abgekürzten Form. Außerdem hat die komplexe Konjugation noch die folgenden Eigenschaften:

$$z \in \mathbb{R} \iff z = \bar{z}, \quad \operatorname{Re} z = \frac{z + \bar{z}}{2}, \quad \operatorname{Im} z = \frac{z - \bar{z}}{2i}.$$

**SATZ**  
Fundamental-  
satz der  
Algebra

## 5. VEKTORRÄUME: DEFINITION UND BEISPIELE

In diesem Abschnitt beginnen wir mit dem Studium der Linearen Algebra. Was ist „Lineare Algebra“? Die Lineare Algebra befasst sich mit „linearen Strukturen“, genauer mit *Vektorräumen* und *linearen Abbildungen* zwischen ihnen. Diese Begriffe sind zunächst einmal sehr abstrakt, aber darin liegt gerade die Stärke der Linearen Algebra: Vektorräume und lineare Abbildungen treten sehr häufig in der Mathematik in den unterschiedlichsten Zusammenhängen auf. Gerade weil man von den jeweils konkreten individuellen Umständen abstrahiert und sich auf die wesentlichen gemeinsamen Eigenschaften beschränkt, lassen sich die Ergebnisse der Linearen Algebra in all diesen unterschiedlichen Situationen anwenden. Es war, historisch gesehen, ein langwieriger Prozess, zu dieser Abstraktion zu gelangen, aber am Endpunkt dieser Entwicklung steht eine sehr leistungsfähige, allgemein anwendbare und erfolgreiche Theorie. Das hat dazu geführt, dass *lineare* Probleme als *einfach* gelten, während *nichtlineare* Probleme sehr häufig besonders *schwierig* sind. In Ausschreibungen für Mathematik-Professuren findet man häufiger das Wort „nichtlinear“ (etwa im Kontext von „nichtlinearen partiellen Differentialgleichungen“), aber so gut wie niemals das Wort „linear“. Zwei Beispiele mit physikalischem Hintergrund: Die *Wärmeleitungsgleichung*, die die zeitliche Entwicklung der Temperaturverteilung in einem Körper beschreibt, ist eine *lineare* partielle Differentialgleichung. Die zugehörige Lösungstheorie wurde bereits von Jean-Baptiste-Joseph Fourier entwickelt („Théorie analytique de la chaleur“, 1822). Im Gegensatz dazu sind die *Navier-Stokes-Gleichungen*, die die Bewegung von Flüssigkeiten beschreiben, *nichtlineare* partielle Differentialgleichungen, und die Frage, ob sie für vernünftige Anfangsbedingungen im dreidimensionalen Raum immer eindeutig lösbar sind, ist eines der sieben Millenniumprobleme der Clay Foundation; für die Lösung bekommt man eine Million US-Dollar.

Was bedeutet nun „linear“? Dazu als Beispiel drei lineare Gleichungen (oder Gleichungssysteme):

- (1) Wir suchen  $w, x, y, z \in \mathbb{R}$  mit

$$w + x + y + z = 0 \quad \text{und} \quad x + 2y + 3z = 0.$$

Wahrscheinlich haben Sie in der Schule gelernt, wie man solche Gleichungssysteme löst (und in jedem Fall werden wir das auch in dieser Vorlesung besprechen). Als Lösungen erhält man

$$(w, x, y, z) = (a, -2a + b, a - 2b, b) \quad \text{mit } a, b \in \mathbb{R}.$$

- (2) Wir suchen Folgen  $(a_n)_{n \in \mathbb{N}}$  reeller Zahlen, für die gilt

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

Die Folge  $(0, 1, 1, 2, 3, 5, 8, \dots)$  der Fibonacci-Zahlen ist eine Lösung, aber es gibt noch mehr. Alle Lösungen lassen sich darstellen in der Form

$$a_n = a \left( \frac{1 + \sqrt{5}}{2} \right)^n + b \left( \frac{1 - \sqrt{5}}{2} \right)^n \quad \text{mit } a, b \in \mathbb{R}.$$

- (3) Wir suchen (zweimal differenzierbare) Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$ , für die gilt

$$f''(x) + f(x) = 0 \quad \text{für alle } x \in \mathbb{R}.$$

Hier sind die Lösungen gegeben durch

$$f(x) = a \cos x + b \sin x \quad \text{mit } a, b \in \mathbb{R}.$$

Obwohl die betrachteten Objekte ganz unterschiedlich sind (Quadrupel von reellen Zahlen, Folgen reeller Zahlen, zweimal differenzierbare reelle Funktionen), ist die Struktur der Lösungsmenge in allen drei Fällen sehr ähnlich. Dass dies so sein muss, ist ein allgemeines Resultat über lineare Gleichungen. Etwas konkreter äußert sich die Linearität darin, dass die *Summe* zweier Lösungen wieder eine Lösung ist, und dass *Vielfache* einer Lösung wieder Lösungen sind. Diese beiden Operationen, also Addition und Vervielfachung, d.h. Multiplikation mit einem „Skalar“ (in den Beispielen ist das jeweils eine reelle Zahl), ergeben die lineare Struktur, die in der folgenden Definition formalisiert ist.

\*

**5.1. Definition.** Sei  $K$  ein Körper. Ein  $K$ -Vektorraum oder *Vektorraum über  $K$*  oder *linearer Raum über  $K$*  ist ein Quintupel  $(V, +, \mathbf{0}, -, \cdot)$ , bestehend aus einer Menge  $V$ , einer Abbildung  $+$  :  $V \times V \rightarrow V$  (genannt Addition), einem Element  $\mathbf{0} \in V$ , einer Abbildung  $-$  :  $V \rightarrow V$  und einer Abbildung  $\cdot$  :  $K \times V \rightarrow V$  (Skalarmultiplikation), sodass  $(V, +, \mathbf{0}, -)$  eine kommutative Gruppe ist und die folgenden weiteren Bedingungen („Axiome“) erfüllt sind:

**DEF**  
Vektorraum

- (1)  $\forall v \in V : 1 \cdot v = v$  (hier ist  $1 \in K$  das Einselement des Körpers  $K$ ).
- (2) (Assoziativität der Skalarmultiplikation)  
 $\forall \lambda, \mu \in K \forall v \in V : \lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v.$
- (3) (Distributivgesetze)  
 $\forall \lambda, \mu \in K \forall v \in V : (\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$       und  
 $\forall \lambda \in K \forall v, w \in V : \lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w.$

Statt  $\lambda \cdot v$  schreibt man oft kurz  $\lambda v$ . Die Elemente eines Vektorraums werden auch *Vektoren* genannt.  $\mathbf{0} \in V$  heißt der *Nullvektor* des Vektorraums  $V$ .  $\diamond$

Machen Sie sich klar, dass „+“ in diesen Axiomen zwei verschiedene Bedeutungen hat: Es kann die Addition im Körper  $K$  gemeint sein oder die Addition im Vektorraum  $V$ !

Der Vollständigkeit halber und zur Erinnerung sind hier noch einmal die vier Axiome für eine kommutative Gruppe  $(V, +, \mathbf{0}, -)$  angegeben:

- (1) (Assoziativität der Addition)  
 $\forall v_1, v_2, v_3 \in V : (v_1 + v_2) + v_3 = v_1 + (v_2 + v_3).$
- (2) (Kommutativität der Addition)  
 $\forall v, w \in V : v + w = w + v.$
- (3) (Nullelement)  
 $\forall v \in V : v + \mathbf{0} = v.$
- (4) (Negative Elemente)  
 $\forall v \in V : v + (-v) = \mathbf{0}.$

Wir kürzen  $v + (-w)$  zu  $v - w$  ab.

Wie üblich kann man sich auf die Angabe von Addition und Skalarmultiplikation beschränken, da das Nullelement und die Negation eindeutig bestimmt sind. Wenn die Verknüpfungen aus dem Kontext klar sind, spricht man einfach nur vom „ $K$ -Vektorraum  $V$ “; wenn auch der Körper  $K$  aus dem Kontext klar ist, vom „Vektorraum  $V$ “.

Ein  $\mathbb{R}$ -Vektorraum heißt auch *reeller Vektorraum*, ein  $\mathbb{C}$ -Vektorraum *komplexer Vektorraum*.

Wir kommen zu einigen einfachen Eigenschaften.

5.2. **Lemma.** Sei  $(V, +, \mathbf{0}, -, \cdot)$  ein  $K$ -Vektorraum. Dann gilt:

- (1)  $\forall v \in V : 0 \cdot v = \mathbf{0}$ .
- (2)  $\forall \lambda \in K : \lambda \cdot \mathbf{0} = \mathbf{0}$ .
- (3)  $\forall v \in V : (-1) \cdot v = -v$ .
- (4)  $\forall \lambda \in K \forall v \in V : \lambda \cdot v = \mathbf{0} \iff \lambda = 0 \text{ oder } v = \mathbf{0}$ .

**LEMMA**  
Rechenregeln  
Vektorraum

*Beweis.*

(1) Wir haben

$$\mathbf{0} = 0 \cdot v - 0 \cdot v = (0 + 0) \cdot v - 0 \cdot v = 0 \cdot v + 0 \cdot v - 0 \cdot v = 0 \cdot v.$$

(2) Das geht analog unter Verwendung des anderen Distributivgesetzes:

$$\mathbf{0} = \lambda \cdot \mathbf{0} - \lambda \cdot \mathbf{0} = \lambda \cdot (\mathbf{0} + \mathbf{0}) - \lambda \cdot \mathbf{0} = \lambda \cdot \mathbf{0} + \lambda \cdot \mathbf{0} - \lambda \cdot \mathbf{0} = \lambda \cdot \mathbf{0}.$$

(3) Es gilt

$$v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 + (-1)) \cdot v = 0 \cdot v = \mathbf{0},$$

also muss  $(-1) \cdot v$  das eindeutig bestimmte Negative  $-v$  von  $v$  sein.

(4) Sei  $\lambda \in K$  und  $v \in V$ . Die Implikation „ $\Leftarrow$ “ wurde bereits in den ersten beiden Teilen des Lemmas bewiesen. Es gelte also  $\lambda \cdot v = \mathbf{0}$ . Ist  $\lambda = 0$ , dann gilt die rechte Seite. Anderenfalls gibt es  $\lambda^{-1} \in K$  und es folgt (mit Teil (2) und der Assoziativität der Skalarmultiplikation)

$$\mathbf{0} = \lambda^{-1} \cdot \mathbf{0} = \lambda^{-1} \cdot (\lambda \cdot v) = (\lambda^{-1} \lambda) \cdot v = 1 \cdot v = v. \quad \square$$

Es ist Zeit für Beispiele.

5.3. **Beispiele.** Sei  $K$  ein Körper.

Der kleinste  $K$ -Vektorraum besteht nur aus dem Nullvektor:  $V = \{\mathbf{0}\}$  und es gilt  $\mathbf{0} + \mathbf{0} = \mathbf{0}$  und  $\lambda \cdot \mathbf{0} = \mathbf{0}$  für alle  $\lambda \in K$ . Dieser Vektorraum heißt der *Null-Vektorraum*. Er ist als Vektorraum nicht besonders interessant, spielt aber in der Linearen Algebra eine ähnliche Rolle wie die leere Menge in der Mengenlehre.

Das nächste Beispiel ist der Körper  $K$  selbst mit seiner Addition und Multiplikation. Die Vektorraum-Axiome entsprechen einem Teil der Körper-Axiome.

Sehr wichtig ist die folgende Klasse von Beispielen, denn es sind die Standardbeispiele für  $K$ -Vektorräume. Als Menge nimmt man  $K^n$ , die Menge der  $n$ -Tupel von Elementen von  $K$ , und die Verknüpfungen definiert man „komponentenweise“:

$$\begin{aligned} (x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) &= (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \quad \text{und} \\ \lambda \cdot (x_1, x_2, \dots, x_n) &= (\lambda x_1, \lambda x_2, \dots, \lambda x_n). \end{aligned}$$

Dann kann man die Axiome leicht nachprüfen. Wir führen das hier exemplarisch für eines der Distributivgesetze durch:

$$\begin{aligned} \lambda \cdot ((x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n)) &= \lambda \cdot (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \\ &= (\lambda(x_1 + y_1), \lambda(x_2 + y_2), \dots, \lambda(x_n + y_n)) \\ &= (\lambda x_1 + \lambda y_1, \lambda x_2 + \lambda y_2, \dots, \lambda x_n + \lambda y_n) \\ &= (\lambda x_1, \lambda x_2, \dots, \lambda x_n) + (\lambda y_1, \lambda y_2, \dots, \lambda y_n) \\ &= \lambda \cdot (x_1, x_2, \dots, x_n) + \lambda \cdot (y_1, y_2, \dots, y_n). \end{aligned}$$

**BSP**  
Vektorräume

Man sieht, dass das direkt aus dem Distributivgesetz  $\lambda(x + y) = \lambda x + \lambda y$  von  $K$  folgt. Für die übrigen Axiome geht das ganz analog. In diesem Beispiel sind die beiden vorigen Beispiele als Grenzfälle enthalten: Für  $n = 0$  hat die Menge  $K^0$  nur ein Element (das Nulltupel, das keine Komponenten hat) und ist somit ein Nullvektorraum. Für  $n = 1$  kann man  $K^1$  mit  $K$  identifizieren und bekommt  $K$  als Vektorraum über  $K$ . Für  $K = \mathbb{R}$  und  $K = \mathbb{C}$  haben wir den reellen Vektorraum  $\mathbb{R}^n$  und den komplexen Vektorraum  $\mathbb{C}^n$  für jedes  $n \in \mathbb{N}$ .

Man kann das vorige Beispiel noch verallgemeinern:  $K^n$  kann als der Spezialfall  $I = \{1, 2, \dots, n\}$  der Menge  $K^I$  der Familien von Elementen von  $K$  mit Indexmenge  $I$  aufgefasst werden. (Zur Erinnerung: Familien  $(x_i)_{i \in I}$  mit  $x_i \in K$  sind nur eine andere Schreibweise für Abbildungen  $I \rightarrow K$ .) Man macht  $K^I$  zu einem  $K$ -Vektorraum, indem man Addition und Skalarmultiplikation „punktweise“ definiert:

$$(x_i)_{i \in I} + (y_i)_{i \in I} = (x_i + y_i)_{i \in I} \quad \text{und} \\ \lambda \cdot (x_i)_{i \in I} = (\lambda x_i)_{i \in I}.$$

Wenn man statt mit Familien mit Abbildungen  $I \rightarrow K$  arbeitet, dann sieht das so aus:

$$f + g : I \longrightarrow K, i \longmapsto f(i) + g(i), \quad \text{d.h.} \quad (f + g)(i) = f(i) + g(i) \quad \text{und} \\ \lambda \cdot f : I \longrightarrow K, i \longmapsto \lambda f(i), \quad \text{d.h.} \quad (\lambda \cdot f)(i) = \lambda f(i).$$

Das Nachprüfen der Axiome funktioniert im Wesentlichen genauso wie für die  $n$ -Tupel. Als Beispiel hier das andere Distributivgesetz (in der Abbildungs-Schreibweise): Seien  $\lambda, \mu \in K$  und  $f : I \rightarrow K$  eine Abbildung. Dann gilt für  $i \in I$ :

$$((\lambda + \mu) \cdot f)(i) = (\lambda + \mu)f(i) = \lambda f(i) + \mu f(i) \\ = (\lambda \cdot f)(i) + (\mu \cdot f)(i) = (\lambda \cdot f + \mu \cdot f)(i),$$

also folgt  $(\lambda + \mu) \cdot f = \lambda \cdot f + \mu \cdot f$ . Zum Beispiel können wir den reellen Vektorraum  $\mathbb{R}^{\mathbb{R}} = \text{Abb}(\mathbb{R}, \mathbb{R})$  aller reellen Funktionen betrachten oder den Vektorraum  $\mathbb{R}^{\mathbb{N}}$  aller Folgen reeller Zahlen.

Ein auf den ersten Blick ganz anders gearteter Vektorraum ist der folgende: Sei  $X$  eine Menge. Dann definieren wir eine Addition auf der Potenzmenge  $\mathcal{P}(X)$  durch

$$A + B = A \triangle B = (A \setminus B) \cup (B \setminus A)$$

(symmetrische Differenz, siehe Beispiel 3.12) und eine Skalarmultiplikation mit Elementen des Körpers  $\mathbb{F}_2 = \{0, 1\}$  in der einzig möglichen Form, nämlich durch  $0 \cdot A = \mathbf{0} = \emptyset$  und  $1 \cdot A = A$ . Dann erhält man einen  $\mathbb{F}_2$ -Vektorraum. Man kann die Axiome wieder nachrechnen, aber man tut sich etwas leichter, wenn man sich klar macht, dass die Potenzmenge  $\mathcal{P}(X)$  und die Menge  $\mathbb{F}_2^X$  der Abbildungen  $X \rightarrow \mathbb{F}_2$  einander bijektiv entsprechen durch

$$\mathcal{P}(X) \longrightarrow \mathbb{F}_2^X, \quad A \longmapsto \left( x \mapsto \begin{cases} 0 & \text{falls } x \notin A, \\ 1 & \text{falls } x \in A, \end{cases} \right) \\ \mathbb{F}_2^X \longrightarrow \mathcal{P}(X), \quad f \longmapsto \{x \in X \mid f(x) = 1\}.$$

Dann entsprechen sich auch Addition und Skalarmultiplikation auf beiden Seiten, also folgt die Gültigkeit der Axiome für  $\mathcal{P}(X)$  aus ihrer Gültigkeit für  $\mathbb{F}_2^X$ . ♣

Weitere Beispiele von Vektorräumen erhalten wir als *Untervektorräume* von anderen Vektorräumen; das werden wir im nächsten Abschnitt genauer betrachten.

In den Beispielen für lineare Gleichungen vom Beginn dieses Abschnitts sind Lösungen in gewissen reellen Vektorräumen gesucht: Im ersten Beispiel in  $\mathbb{R}^4$ , im zweiten Beispiel in  $\mathbb{R}^N$  und im dritten Beispiel in einem Untervektorraum von  $\mathbb{R}^R$ .

## 6. UNTERVEKTORRÄUME

Häufig möchte man, wenn man einen Vektorraum  $V$  gegeben hat, nicht mit dem ganzen Vektorraum arbeiten, sondern mit einer Teilmenge. Damit stellt sich die Frage, wann so eine Teilmenge (wenn man die Addition und Skalarmultiplikation darauf einschränkt) selbst wieder ein Vektorraum ist. Damit diese Frage sinnvoll ist, müssen die Addition und Skalarmultiplikation auf der Teilmenge *wohldefiniert* sein, das heißt, dass Summen und Vielfache von Elementen der Teilmenge wieder in der Teilmenge liegen müssen. Außerdem brauchen wir natürlich das Nullelement. Das führt auf folgende Definition:

\*

**6.1. Definition.** Sei  $K$  ein Körper,  $V$  ein  $K$ -Vektorraum und  $U \subset V$  eine Teilmenge von  $V$ . Dann heißt  $U$  ein *Untervektorraum* oder *linearer Unterraum* von  $V$ , wenn  $U$  die folgenden Bedingungen erfüllt:

**DEF**  
Unter-  
vektorraum

$$(1) \mathbf{0} \in U,$$

$$(2) \forall u_1, u_2 \in U : u_1 + u_2 \in U$$

(„ $U$  ist abgeschlossen unter der Addition“),

$$(3) \forall \lambda \in K \forall u \in U : \lambda \cdot u \in U$$

(„ $U$  ist abgeschlossen unter der Skalarmultiplikation“).

◇

Wir zeigen gleich, dass diese Definition sinnvoll ist.

**6.2. Lemma.** Sei  $K$  ein Körper,  $V$  ein  $K$ -Vektorraum und  $U \subset V$  ein Untervektorraum. Dann gilt für alle  $u \in U$ , dass auch  $-u$  ein Element von  $U$  ist.

**LEMMA**  
Unter-VR ist  
Vektorraum

Wir schreiben  $+_U$  für die auf  $U$  eingeschränkte Addition  $U \times U \rightarrow U$ ,  $(u_1, u_2) \mapsto u_1 + u_2$ ,  $-_U$  für die auf  $U$  eingeschränkte Negationsabbildung  $U \rightarrow U$ ,  $u \mapsto -u$ , und  $\cdot_U$  für die auf  $U$  eingeschränkte Skalarmultiplikation  $K \times U \rightarrow U$ ,  $(\lambda, u) \mapsto \lambda \cdot u$ . Dann ist  $(U, +_U, \mathbf{0}, -_U, \cdot_U)$  ein  $K$ -Vektorraum.

*Beweis.* Die erste Behauptung ist  $\forall u \in U : -u \in U$ . Das folgt aber aus der Definition, denn  $-u = (-1) \cdot u$ , vgl. Lemma 5.2. Deshalb und nach der Definition können wir  $+_U$ ,  $-_U$  und  $\cdot_U$  wie angegeben definieren (denn die Bilder liegen jeweils in  $U$ ). Es bleiben die Vektorraum-Axiome für  $U$  nachzuprüfen. Diese haben aber alle die Form von „Allaussagen“, es wird also verlangt, dass eine Aussage für alle Elemente  $u_1, u_2, \dots$  von  $U$  gilt. Da  $V$  ein Vektorraum ist, gelten diese Aussagen aber sogar für alle Elemente von  $V$ , also erst recht für alle Elemente von  $U$ . □

In der Literatur finden Sie meistens eine Definition von „Vektorraum“ (und analog für Gruppen, Ringe, Körper, ...), die von dem Tripel  $(V, +, \cdot)$  ausgeht und dann die *Existenz* eines Nullelements und von Inversen bezüglich der Addition fordert. Im Gegensatz dazu haben wir hier das Nullelement und die Negationsabbildung mit in die „Daten“ des Vektorraums aufgenommen. Der Vorteil ist, dass die Axiome dann alle zu Allaussagen werden, die man leichter nachprüfen kann, wie im obigen Beweis. Auf der anderen Seite muss man sich aber vorher überlegen, was das Nullelement ist und wie die Negationsabbildung aussieht. Im gerade bewiesenen Lemma geschieht dies dadurch, dass wir zeigen, dass  $U$  auch unter der Negation abgeschlossen ist, sodass wir die Negationsabbildung  $-_U$  definieren können. Wenn man die andere Formulierung der Axiome benutzt, dann muss man diesen Beweisschritt ausführen, wenn man die Existenz des zu  $u$  negativen Elements zeigt. Im Endeffekt muss man also das Gleiche tun, nur die Reihenfolge ist etwas anders.

Die Schreibweise  $+_U$  usw. für die auf  $U$  eingeschränkten Abbildungen diene nur der Verdeutlichung für die Formulierung des Lemmas. Wir schreiben normalerweise einfach  $+$  usw. für die Addition usw. auf  $U$ .

**6.3. Beispiele.** Jeder Vektorraum  $V$  hat die Untervektorräume  $U = \{\mathbf{0}\} \subset V$  (ein Null-Vektorraum) und  $U = V$ .

**BSP**  
♣ triviale  
Unter-VR

**6.4. Beispiel.** Sei  $a \in \mathbb{R}$ . Wir betrachten den reellen Vektorraum  $V = \mathbb{R}^2$  und setzen  $U_a = \{(x, y) \in \mathbb{R}^2 \mid x + y = a\}$ . Für welche  $a$  ist  $U_a$  ein Untervektorraum von  $\mathbb{R}^2$ ?

**BSP**  
Unter-VR  
von  $\mathbb{R}^2$

Dazu müssen wir die Bedingungen in der Definition nachprüfen. Die erste davon sagt, dass der Nullvektor  $\mathbf{0} = (0, 0)$  ein Element von  $U_a$  sein muss. Das bedeutet  $0 + 0 = a$ , also ist das nur für  $a = 0$  möglich. Wir prüfen die beiden anderen Bedingungen:

- $U_0$  ist abgeschlossen unter der Addition, denn für Elemente  $u_1 = (x_1, y_1)$  und  $u_2 = (x_2, y_2)$  von  $U_0$  gilt  $u_1 + u_2 = (x_1 + x_2, y_1 + y_2)$  und

$$(x_1 + x_2) + (y_1 + y_2) = (x_1 + y_1) + (x_2 + y_2) = 0 + 0 = 0,$$

also ist  $u_1 + u_2 \in U_0$ .

- $U_0$  ist abgeschlossen unter der Skalarmultiplikation, denn für ein Element  $u = (x, y) \in U_0$  und  $\lambda \in \mathbb{R}$  ist  $\lambda \cdot u = (\lambda x, \lambda y)$  und es gilt

$$\lambda x + \lambda y = \lambda(x + y) = \lambda \cdot 0 = 0,$$

also ist  $\lambda \cdot u \in U_0$ . ♣

Weitere interessante Beispiele sind „Folgenräume“ und „Funktionsräume“, die als Untervektorräume des Vektorraums  $\mathbb{R}^{\mathbb{N}}$  der Folgen reeller Zahlen oder des Vektorraums  $\text{Abb}(\mathbb{R}, \mathbb{R})$  der reellen Funktionen auftreten.

**6.5. Beispiele.** Sei  $V = \mathbb{R}^{\mathbb{N}}$  der reelle Vektorraum, dessen Elemente alle Folgen reeller Zahlen sind.

**BSP**  
Folgenräume

- (1) Sei  $U_b = \{(a_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid (a_n)_{n \in \mathbb{N}} \text{ ist beschränkt}\}$ . Dann ist  $U_b$  ein Untervektorraum von  $\mathbb{R}^{\mathbb{N}}$ .

*Beweis.* Wir prüfen die Bedingungen nach. Die konstante Nullfolge (mit  $a_n = 0$  für alle  $n \in \mathbb{N}$ ) ist beschränkt, also gilt  $\mathbf{0} \in U_b$ . Seien  $(a_n)_{n \in \mathbb{N}}$  und  $(b_n)_{n \in \mathbb{N}}$  zwei beschränkte Folgen. Dann gibt es  $A, B \in \mathbb{R}$  mit  $|a_n| \leq A$  und  $|b_n| \leq B$  für alle  $n \in \mathbb{N}$ . Es folgt  $|a_n + b_n| \leq A + B$ , also ist auch die Summenfolge  $(a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}} = (a_n + b_n)_{n \in \mathbb{N}}$  beschränkt. Ist zusätzlich  $\lambda \in \mathbb{R}$ , dann gilt  $|\lambda a_n| \leq |\lambda|A$ , also ist auch die Folge  $\lambda \cdot (a_n)_{n \in \mathbb{N}} = (\lambda a_n)_{n \in \mathbb{N}}$  beschränkt. □

- (2) Sei  $U_n = \{(a_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid (a_n)_{n \in \mathbb{N}} \text{ ist eine Nullfolge}\}$ . Dann ist  $U_n$  ein Untervektorraum von  $\mathbb{R}^{\mathbb{N}}$  (oder auch von  $U_b$ ).

*Beweis.* Übung. □

- (3) Sei  $U_k = \{(a_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid (a_n)_{n \in \mathbb{N}} \text{ konvergiert}\}$ . Dann ist  $U_k$  ein Untervektorraum von  $\mathbb{R}^{\mathbb{N}}$  (oder auch von  $U_b$ ).

*Beweis.* Die konstante Nullfolge konvergiert (gegen 0), also ist sie in  $U_k$ . In der Analysis lernen Sie, dass die Summe zweier konvergenter Folgen wieder konvergiert und dass jedes Vielfache einer konvergenten Folge konvergiert. Damit sind die drei Bedingungen erfüllt.  $\square$

Für diese drei Untervektorräume gilt  $U_n \subset U_k \subset U_b$  (denn jede Nullfolge konvergiert gegen 0 und jede konvergente Folge ist beschränkt, vgl. Analysis).  $\clubsuit$

**6.6. Beispiele.** Sei  $V = \text{Abb}(\mathbb{R}, \mathbb{R})$  der reelle Vektorraum, dessen Elemente alle Funktionen  $\mathbb{R} \rightarrow \mathbb{R}$  sind.

**BSP**  
Funktionen-  
räume

- (1) Sei  $\mathcal{C}(\mathbb{R}) = \{f \in \text{Abb}(\mathbb{R}, \mathbb{R}) \mid f \text{ ist stetig}\}$ . Dann ist  $\mathcal{C}(\mathbb{R})$  ein Untervektorraum von  $V$ .

*Beweis.* Die Nullfunktion  $x \mapsto 0$  ist stetig. In der Analysis werden Sie lernen, dass Summen und Vielfache stetiger Funktionen wieder stetig sind.  $\square$

- (2) Sei  $n \in \mathbb{N}$  und  $\mathcal{C}^n(\mathbb{R}) = \{f \in \text{Abb}(\mathbb{R}, \mathbb{R}) \mid f \text{ ist } n\text{-mal differenzierbar und } f^{(n)} \text{ ist stetig}\}$  der Raum der  $n$ -mal stetig differenzierbaren Funktionen. Aus Ergebnissen der Analysis wird folgen, dass  $\mathcal{C}^n(\mathbb{R})$  ein Untervektorraum von  $V$  ist.

- (3) Sei  $a > 0$  und  $\mathcal{P}(a) = \{f \in \text{Abb}(\mathbb{R}, \mathbb{R}) \mid \forall x \in \mathbb{R} : f(x+a) = f(x)\}$  die Menge der periodischen Funktionen mit Periode  $a$  (zum Beispiel sind  $\sin$  und  $\cos$  Elemente von  $\mathcal{P}(2\pi)$ ). Dann ist  $\mathcal{P}(a)$  ein Untervektorraum von  $V$ .

*Beweis.* Die Nullfunktion ist periodisch, also ein Element von  $\mathcal{P}(a)$ . Seien  $f, g \in \mathcal{P}(a)$  und  $\lambda \in \mathbb{R}$ . Wir zeigen  $f+g, \lambda f \in \mathcal{P}(a)$ : Für alle  $x \in \mathbb{R}$  gilt

$$(f+g)(x+a) = f(x+a) + g(x+a) \stackrel{f, g \in \mathcal{P}(a)}{=} f(x) + g(x) = (f+g)(x) \quad \text{und}$$

$$(\lambda f)(x+a) = \lambda f(x+a) \stackrel{f \in \mathcal{P}(a)}{=} \lambda f(x) = (\lambda f)(x).$$

Damit sind alle drei Bedingungen erfüllt.  $\square$



Auch in der *Codierungstheorie* spielt der Begriff des Untervektorraums eine sehr wichtige Rolle.

**6.7. Beispiel.** Sei  $F$  ein endlicher Körper (zum Beispiel  $F = \mathbb{F}_2$ ) und  $n \in \mathbb{N}$ . Dann heißt ein Untervektorraum von  $F^n$  ein *linearer Code* der Länge  $n$  über  $F$ . Ein Beispiel ist der *Hamming-Code* der Länge 7 über  $\mathbb{F}_2$ , der gegeben ist durch

**BSP**  
Lineare  
Codes

$$H = \{(x_1, x_2, x_3, x_4, x_1+x_2+x_4, x_1+x_3+x_4, x_2+x_3+x_4) \in \mathbb{F}_2^7 \mid x_1, x_2, x_3, x_4 \in \mathbb{F}_2\}.$$

In der Codierungstheorie interessiert man sich dann für die „Größe“ (genauer: die *Dimension*, die wir bald einführen werden) des Codes und dafür, wie viele Fehler er korrigieren kann. Dafür ist wichtig, dass je zwei verschiedene Codewörter (also Elemente des Codes) sich an möglichst vielen Stellen unterscheiden. Wegen der linearen Struktur kann man Differenzen bilden und daher annehmen, dass eines der Codewörter null ist. Dann ist die Frage, an mindestens wie vielen Stellen ein von  $\mathbf{0}$  verschiedenes Codewort eine von 0 verschiedene Komponente hat. Für den Hamming-Code  $H$  ist diese „Minimaldistanz“ 3, was bedeutet, dass er „einen Fehler korrigieren“ kann. (Wenn ein Codewort an einer Stelle verändert wird, kann

man es rekonstruieren, da sich jedes andere Codewort von dem veränderten Wort an mindestens zwei Stellen unterscheidet.) ♣

## 7. ERZEUGENDENSYSTEME

Wir erinnern uns an die Beispiele von Funktionenräumen im letzten Abschnitt. Dort hatten wir gesehen, dass der Raum  $\mathcal{C}(\mathbb{R})$  der stetigen reellen Funktionen und der Raum  $\mathcal{P}(a)$  der  $a$ -periodischen reellen Funktionen beides Untervektorräume von  $\text{Abb}(\mathbb{R}, \mathbb{R})$  sind. Wie sieht es mit stetigen periodischen Funktionen aus? Muss  $\mathcal{C}(\mathbb{R}) \cap \mathcal{P}(a)$  auch ein Untervektorraum sein?

Bevor wir ein entsprechendes Resultat beweisen, führen wir eine Schreibweise für Vereinigungen und Durchschnitte von vielen Mengen ein.

**7.1. Definition.** Ist  $(A_i)_{i \in I}$  eine Familie von Mengen, dann schreiben wir

$$\bigcup_{i \in I} A_i = \{x \mid \exists i \in I : x \in A_i\}$$

für die Vereinigung aller Mengen  $A_i$ . (Ist  $I$  die leere Menge, dann ist diese Vereinigung ebenfalls leer.) Ist  $I \neq \emptyset$ , dann schreiben wir analog

$$\bigcap_{i \in I} A_i = \{x \mid \forall i \in I : x \in A_i\}$$

für den Durchschnitt aller Mengen  $A_i$ .

Ist  $\mathcal{M}$  eine Menge, deren Elemente selbst Mengen sind, dann schreiben wir

$$\bigcup \mathcal{M} = \bigcup_{A \in \mathcal{M}} A = \{x \mid \exists A \in \mathcal{M} : x \in A\}$$

für die Vereinigung all dieser Mengen und, falls  $\mathcal{M}$  nicht leer ist,

$$\bigcap \mathcal{M} = \bigcap_{A \in \mathcal{M}} A = \{x \mid \forall A \in \mathcal{M} : x \in A\}$$

für ihren Durchschnitt. ◇

Im Fall  $I = \emptyset$  wäre die Bedingung  $\forall i \in I : x \in A_i$  für alle  $x$  erfüllt und man bekäme die Menge, die alles enthält. Diese Menge kann es aber nicht geben, denn sie würde die Menge enthalten, die zur Russellschen Antinomie führt, siehe die Bemerkungen zur Mengenlehre am Ende von Abschnitt 2.

Damit können wir jetzt die folgende wichtige Aussage formulieren:

**7.2. Lemma.** Sei  $V$  ein  $K$ -Vektorraum und sei  $(U_i)_{i \in I}$  eine Familie von Untervektorräumen von  $V$  mit  $I \neq \emptyset$ . Dann ist

$$U = \bigcap_{i \in I} U_i$$

ebenfalls ein Untervektorraum von  $V$ .

Für  $I = \{1, 2, 3, \dots, n\}$  haben wir den Spezialfall

$$\begin{aligned} U_1, U_2, \dots, U_n \subset V \text{ Untervektorräume} \\ \implies U_1 \cap U_2 \cap \dots \cap U_n \subset V \text{ Untervektorraum.} \end{aligned}$$

*Beweis.* Wir müssen die Bedingungen aus Definition 6.1 für  $U$  nachprüfen.

- (1) Da jede Teilmenge  $U_i$  ein Untervektorraum von  $V$  ist, gilt  $\forall i \in I : \mathbf{0} \in U_i$ . Das bedeutet gerade  $\mathbf{0} \in U$ .

**DEF**

$$\begin{aligned} \bigcup_{i \in I} A_i \\ \bigcap_{i \in I} A_i \\ \bigcup \mathcal{M} \\ \bigcap \mathcal{M} \end{aligned}$$

**LEMMA**

Durchschnitt  
von Unter-VR

- (2) Seien  $u_1, u_2 \in U$ . Nach Definition von  $U$  bedeutet das  $\forall i \in I : u_1, u_2 \in U_i$ . Da alle  $U_i$  Untervektorräume von  $V$  sind, folgt  $\forall i \in I : u_1 + u_2 \in U_i$ , also  $u_1 + u_2 \in U$ .
- (3) Sei  $\lambda \in K$  und  $u \in U$ . Dann gilt  $\forall i \in I : u \in U_i$ . Da alle  $U_i$  Untervektorräume von  $V$  sind, folgt  $\forall i \in I : \lambda u \in U_i$ , also  $\lambda u \in U$ .  $\square$

**7.3. Beispiel.** Der Raum

$$\mathcal{C}(\mathbb{R}) \cap \mathcal{P}(a) = \{f \in \text{Abb}(\mathbb{R}, \mathbb{R}) \mid f \text{ ist stetig und } a\text{-periodisch}\}$$

ist ein Untervektorraum von  $\text{Abb}(\mathbb{R}, \mathbb{R})$ .

**BSP**



Wie sieht es mit Vereinigungen von Untervektorräumen aus? Im Allgemeinen erhält man daraus *keinen* Untervektorraum. Die Vereinigung von zwei Untervektorräumen  $U_1$  und  $U_2$  zum Beispiel ist nur dann wieder ein Untervektorraum, wenn einer der beiden im anderen enthalten ist (Übung). Man hat aber immerhin das folgende Resultat.

**7.4. Lemma.** Sei  $V$  ein  $K$ -Vektorraum und sei  $(U_n)_{n \in \mathbb{N}}$  eine aufsteigende Folge von Untervektorräumen von  $V$  (d.h.  $U_n \subset U_{n+1}$  für alle  $n \in \mathbb{N}$ ). Dann ist

$$U = \bigcup_{n \in \mathbb{N}} U_n$$

ebenfalls ein Untervektorraum von  $V$ .

**LEMMA**  
aufsteigende  
Vereinigung  
von Unter-VR

*Beweis.* Wir prüfen die Bedingungen für  $U$ .

- (1)  $\mathbf{0} \in U_0$ , also ist auch  $\mathbf{0} \in U$ .
- (2) Seien  $u_1, u_2 \in U$ . Dann gibt es  $n_1, n_2 \in \mathbb{N}$  mit  $u_1 \in U_{n_1}$  und  $u_2 \in U_{n_2}$ . Sei  $n$  die größere der beiden Zahlen  $n_1$  und  $n_2$ . Da wir eine aufsteigende Folge von Untervektorräumen haben, gilt dann  $U_{n_1} \subset U_n$  und  $U_{n_2} \subset U_n$  und damit  $u_1, u_2 \in U_n$ . Da  $U_n$  ein Untervektorraum ist, folgt  $u_1 + u_2 \in U_n \subset U$ .
- (3) Sei  $\lambda \in K$  und  $u \in U$ . Dann gibt es  $n \in \mathbb{N}$ , sodass  $u \in U_n$  ist. Da  $U_n$  ein Untervektorraum ist, folgt  $\lambda u \in U_n \subset U$ .  $\square$

Lemma 7.2 erlaubt es uns nun, den kleinsten Untervektorraum zu konstruieren, der eine gegebene Teilmenge eines Vektorraums  $V$  enthält.

**\* 7.5. Definition.** Sei  $V$  ein  $K$ -Vektorraum und  $A \subset V$  eine beliebige Teilmenge von  $V$ . Dann heißt der Untervektorraum

$$\langle A \rangle = \langle A \rangle_K = \bigcap \{U \subset V \mid U \text{ Untervektorraum von } V \text{ und } A \subset U\}$$

(also der Durchschnitt aller  $A$  enthaltenden Untervektorräume von  $V$ ) der *von  $A$  erzeugte* oder *aufgespannte* Untervektorraum von  $V$ , die *( $K$ -)lineare Hülle von  $A$*  oder der *( $K$ -)Spann von  $A$* . Ist  $A = \{v_1, v_2, \dots, v_n\}$  endlich, dann schreiben wir auch

$$\langle v_1, v_2, \dots, v_n \rangle \quad \text{oder} \quad \langle v_1, v_2, \dots, v_n \rangle_K$$

an Stelle von  $\langle A \rangle$  oder  $\langle A \rangle_K$ .



**DEF**  
Lineare  
Hülle

Lemma 7.2 garantiert uns, dass  $\langle A \rangle$  tatsächlich ein Untervektorraum von  $V$  ist, denn  $\langle A \rangle$  ist definitionsgemäß der Durchschnitt einer nichtleeren Menge von Untervektorräumen (nichtleer, weil  $V$  selbst immer ein  $A$  enthaltender Untervektorraum von  $V$  ist).

Wir benutzen die Schreibweise  $\langle A \rangle_K$ , um zu verdeutlichen, welcher Körper zugrunde gelegt wird. Zum Beispiel gilt mit  $V = \mathbb{C}$  als  $\mathbb{R}$ -Vektorraum, dass  $\langle 1 \rangle = \langle 1 \rangle_{\mathbb{R}} = \mathbb{R}$  ist. Wird  $\mathbb{C}$  aber als  $\mathbb{C}$ -Vektorraum betrachtet, dann haben wir  $\langle 1 \rangle = \langle 1 \rangle_{\mathbb{C}} = \mathbb{C}$ .

**7.6. Beispiel.** In Definition 7.5 können wir für  $A$  die leere Menge wählen. Was ist der von  $A$  erzeugte Untervektorraum?

**BSP**  
 $\langle \emptyset \rangle$

Da *jeder* Untervektorraum von  $V$  die leere Menge enthält, müssen wir den Durchschnitt über *alle* Untervektorräume von  $V$  bilden. Da jeder Untervektorraum den Nullvektor enthält und  $\{\mathbf{0}\}$  ein Untervektorraum ist, folgt  $\langle \emptyset \rangle = \{\mathbf{0}\}$ . ♣

\* **7.7. Definition.** Sei  $V$  ein  $K$ -Vektorraum und  $E \subset V$  eine Teilmenge von  $V$ . Dann heißt  $E$  ein  $(K)$ -Erzeugendensystem von  $V$ , wenn  $V = \langle E \rangle$  gilt. ◇

**DEF**  
Erzeugendensystem

Zum Beispiel ist die leere Menge ein Erzeugendensystem des Null-Vektorraums.

Definition 7.5 ist sehr elegant, aber nicht besonders praktisch, weil sie uns nicht sagt, „wie die lineare Hülle von  $A$  aussieht“, also was ihre Elemente sind. In gewisser Weise ist es eine Definition „von oben“ — wir betrachten alle Untervektorräume, die mindestens so groß sind wie gewünscht, und wählen dann den kleinsten (im Sinne der Inklusion von Mengen) aus. (Das ist übrigens völlig analog zur Definition des Supremums einer Menge reeller Zahlen in der Analysis als kleinste obere Schranke.) Was wir aber gerne hätten, ist eine Definition „von unten“, die die Elemente von  $\langle A \rangle$  aus den Elementen von  $A$  konstruiert.

Dafür betrachten wir als Beispiel eine zweielementige Menge  $A = \{v_1, v_2\} \subset V$ . Welche Elemente muss  $\langle A \rangle$  mindestens enthalten? Nun, wir wissen, dass  $v_1$  und  $v_2$  Elemente von  $\langle A \rangle$  sind, außerdem ist  $\langle A \rangle$  ein Untervektorraum, also unter Addition und Skalarmultiplikation abgeschlossen. Es müssen also insbesondere Summen von Vielfachen von  $v_1$  und  $v_2$  in  $\langle A \rangle$  enthalten sein:

$$\{\lambda_1 v_1 + \lambda_2 v_2 \mid \lambda_1, \lambda_2 \in K\} \subset \langle A \rangle.$$

Auf der anderen Seite überlegt man sich leicht, dass diese Menge selbst schon ein Untervektorraum von  $V$  ist. Da dieser Untervektorraum  $A$  enthält und gleichzeitig in allen  $A$  enthaltenden Untervektorräumen enthalten ist, muss er gleich  $\langle A \rangle$  sein. Diese Beobachtung lässt sich verallgemeinern.

**7.8. Satz.** Sei  $V$  ein  $K$ -Vektorraum.

**SATZ**  
Beschreibung  
von  $\langle A \rangle$

(1) Sind  $v_1, v_2, \dots, v_n \in V$ , dann gilt

$$\langle v_1, v_2, \dots, v_n \rangle = \{\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n \mid \lambda_1, \lambda_2, \dots, \lambda_n \in K\}.$$

(2) Ist  $A \subset V$  beliebig, dann gilt

$$\langle A \rangle = \{\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n \mid n \in \mathbb{N}, v_1, v_2, \dots, v_n \in A, \lambda_1, \lambda_2, \dots, \lambda_n \in K\}.$$

Für  $n = 0$  setzen wir dabei  $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = \mathbf{0}$  („leere Summe“).

*Beweis.*

(1) Sei  $U$  die Menge auf der rechten Seite der Gleichung. Da  $v_1, v_2, \dots, v_n \in \langle A \rangle$  und  $\langle A \rangle$  unter Skalarmultiplikation und Addition abgeschlossen ist, muss jedes Element der Form  $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n$  ebenfalls in  $\langle A \rangle$  liegen. Es gilt also  $U \subset \langle A \rangle$ .

Auf der anderen Seite gilt  $A \subset U$  (wähle  $\lambda_j = 1$  und  $\lambda_i = 0$  für alle  $i \in \{1, 2, \dots, n\} \setminus \{j\}$ , um zu sehen, dass  $v_j \in U$  ist) und  $U$  ist ein Untervektorraum von  $V$ :

- $\mathbf{0} \in U$  (setze  $\lambda_i = 0$  für alle  $i$ ).

- $U$  ist abgeschlossen unter der Addition, denn

$$\begin{aligned} (\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n) + (\mu_1 v_1 + \mu_2 v_2 + \dots + \mu_n v_n) \\ = (\lambda_1 + \mu_1) v_1 + (\lambda_2 + \mu_2) v_2 + \dots + (\lambda_n + \mu_n) v_n. \end{aligned}$$

- $U$  ist abgeschlossen unter der Skalarmultiplikation, denn

$$\lambda(\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n) = (\lambda \lambda_1) v_1 + (\lambda \lambda_2) v_2 + \dots + (\lambda \lambda_n) v_n.$$

Da  $U$  ein  $A$  enthaltender Untervektorraum von  $V$  ist, folgt  $\langle A \rangle \subset U$ ; insgesamt erhalten wir die behauptete Gleichheit.

- (2) Sei wieder  $U$  die Menge auf der rechten Seite der Gleichung. Wie in Teil (1) ist klar, dass  $U \subset \langle A \rangle$  ist. Es gilt wieder, dass  $U$  ein  $A$  enthaltender Untervektorraum ist. Die einzige Schwierigkeit tritt beim Nachweis der Abgeschlossenheit unter der Addition auf, denn in den beiden zu addierenden Elementen können verschiedene Elemente von  $A$  auftreten. Da aber nicht vorausgesetzt ist, dass die auftretenden Elemente paarweise verschieden<sup>1</sup> sein müssen, können wir die beiden Summen einfach „formal“ addieren:

$$\begin{aligned} (\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n) + (\mu_1 w_1 + \mu_2 w_2 + \dots + \mu_m w_m) \\ = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n + \lambda_{n+1} v_{n+1} + \dots + \lambda_{n+m} v_{n+m}, \end{aligned}$$

wenn wir  $\lambda_{n+j} = \mu_j$  und  $v_{n+j} = w_j$  setzen für  $j \in \{1, 2, \dots, m\}$ . □

Es ist eine gute Übung, sich zu überlegen, an welcher Stelle in diesem Beweis welche der Vektorraum-Axiome verwendet werden.

Weil die Ausdrücke der Form  $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n$  so wichtig sind, haben sie einen eigenen Namen.

**7.9. Definition.** Sei  $K$  ein Körper und  $V$  ein  $K$ -Vektorraum.

**DEF**  
Linear-  
kombination

- (1) Sind  $v_1, v_2, \dots, v_n \in V$  und  $\lambda_1, \lambda_2, \dots, \lambda_n \in K$ , dann heißt

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n$$

eine  $(K)$ -Linearkombination von  $v_1, v_2, \dots, v_n$ .

- (2) Ist  $A \subset V$  eine beliebige Teilmenge von  $V$ , dann heißt jede  $K$ -Linearkombination von Elementen  $v_1, v_2, \dots, v_n \in A$  eine  $(K)$ -Linearkombination von Elementen von  $A$ .

Dabei heißt  $\lambda_j$  der *Koeffizient* von  $v_j$  in der Linearkombination. ◇

Satz 7.8 kann dann so formuliert werden:

*Die lineare Hülle von  $A \subset V$  besteht genau aus allen Linearkombinationen von Elementen von  $A$ .*

Eine Teilmenge  $E \subset V$  ist genau dann ein Erzeugendensystem von  $V$ , wenn jedes Element von  $V$  eine Linearkombination von Elementen von  $E$  ist.

<sup>1</sup>„ $v_1, v_2, \dots, v_n$  sind paarweise verschieden“ bedeutet „ $\forall i, j \in \{1, 2, \dots, n\} : i \neq j \Rightarrow v_i \neq v_j$ “.

**Wichtig:** In einer Linearkombination kommen immer nur **endlich viele** Elemente vor! In der Linearen Algebra gibt es (im Gegensatz zur Analysis) keine unendlichen Summen!

7.10. **Definition.** Analog zur in der Analysis eingeführten Summenschreibweise schreiben wir

$$\sum_{i \in I} a_i \quad \text{bzw.} \quad \sum_{i=1}^n a_i$$

für die Summe der Glieder der Familie  $(a_i)_{i \in I}$  bzw. für die Summe der Komponenten des  $n$ -Tupels  $(a_1, a_2, \dots, a_n)$ . Dabei sind die  $a_i$  aus einer kommutativen Gruppe (bei uns fast immer Elemente eines Vektorraums) und die Menge  $I$  ist endlich. Ist  $I$  leer (bzw.  $n = 0$ ), dann ist der Wert dieser „leeren Summe“ das Nullelement der Gruppe. Eine Linearkombination kann dann in der Form

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

geschrieben werden. ◇

7.11. **Beispiel.** Sei  $K$  ein Körper und  $n \in \mathbb{N}$ . Im Standard-Vektorraum  $K^n$  haben wir die Elemente

$$\mathbf{e}_1 = (1, 0, 0, \dots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \dots, 0), \quad \dots, \quad \mathbf{e}_n = (0, 0, 0, \dots, 0, 1).$$

Dabei sind alle Komponenten von  $\mathbf{e}_j$  null mit Ausnahme der  $j$ -ten, die den Wert 1 hat. Die Menge  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  ist ein Erzeugendensystem von  $K^n$ , denn jedes Element von  $K^n$  ist eine Linearkombination dieser Elemente:

$$(x_1, x_2, \dots, x_n) = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_n \mathbf{e}_n. \quad \clubsuit$$

7.12. **Beispiel.** Ein Vektorraum hat im Allgemeinen viele Erzeugendensysteme. Zum Beispiel sind

$$\{\mathbf{e}_1, \mathbf{e}_2\}, \quad \{(1, 1), (1, -1)\}, \quad \{(1, 2), (2, 3), (3, 4)\}, \quad \mathbb{Z} \times \mathbb{Z} \quad \text{und} \quad \mathbb{R} \times \mathbb{R}$$

alles Erzeugendensysteme von  $V = \mathbb{R}^2$ . ♣

7.13. **Beispiel.** Im Vektorraum  $V = \text{Abb}(\mathbb{R}, \mathbb{R})$  betrachten wir die *Potenzfunktionen*

$$f_0 : x \mapsto 1, \quad f_1 : x \mapsto x, \quad f_2 : x \mapsto x^2, \quad \dots, \quad f_n : x \mapsto x^n, \quad \dots$$

Wie sieht der von  $\{f_0, f_1, f_2, \dots\} = \{f_n \mid n \in \mathbb{N}\}$  erzeugte Untervektorraum  $P$  von  $V$  aus?

Seine Elemente sind gerade die Linearkombinationen von endlich vielen der Potenzfunktionen. Indem wir eventuell Potenzfunktionen mit Koeffizient 0 hinzufügen (was am Wert der Linearkombination nichts ändert) und gleichartige Terme zusammenfassen, können wir annehmen, dass die Linearkombination die Form

$$f = a_0 f_0 + a_1 f_1 + \dots + a_n f_n$$

hat mit  $n \in \mathbb{N}$  und  $a_0, a_1, \dots, a_n \in \mathbb{R}$ . Dann gilt

$$f(x) = a_0 f_0(x) + a_1 f_1(x) + \dots + a_n f_n(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n.$$

Die Elemente von  $P$  sind also gerade die *Polynomfunktionen*. ♣

**DEF**

$$\sum_{i \in I} a_i$$

**BSP**  
Erzeugendensystem  
von  $K^n$

**BSP**  
Viele Erzeugendensysteme

**BSP**  
Vektorraum der Polynomfunktionen

## 8. LINEARE UNABHÄNGIGKEIT, BASIS UND DIMENSION

Wir haben gesehen, dass ein  $K$ -Vektorraum  $V$  sehr viele Erzeugendensysteme haben kann; eines davon ist zum Beispiel die Menge  $V$  selbst. Das erscheint aber ein wenig verschwenderisch, sodass sich die Frage stellt, ob es auch minimale Erzeugendensysteme gibt und wie sie gegebenenfalls charakterisiert werden können. Dazu überlegen wir Folgendes: Sei  $E$  ein Erzeugendensystem von  $V$ , das nicht minimal ist in dem Sinn, dass es ein Element  $v_0 \in E$  gibt, sodass  $E_0 = E \setminus \{v_0\}$  auch schon ein Erzeugendensystem von  $V$  ist. Dann können wir  $v_0$  als Linearkombination von Elementen von  $E_0$  schreiben:

$$v_0 = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n$$

mit  $v_1, v_2, \dots, v_n \in E_0$  und  $\lambda_1, \lambda_2, \dots, \lambda_n \in K$ . Dabei können wir annehmen, dass  $v_1, v_2, \dots, v_n$  paarweise verschieden sind (sonst fassen wir die Terme entsprechend zusammen). Wenn wir  $\lambda_0 = -1$  setzen, dann können wir das auch in symmetrischer Form schreiben als

$$\lambda_0 v_0 + \lambda_1 v_1 + \dots + \lambda_n v_n = \mathbf{0}.$$

Es gibt also eine *nichttriviale* Linearkombination (das ist eine, in der nicht alle Koeffizienten null sind; hier ist  $\lambda_0 = -1 \neq 0$ ) von Elementen von  $E$ , die den Nullvektor ergibt.

Umgekehrt gilt: Gibt es eine solche nichttriviale Linearkombination von Elementen von  $E$ , deren Wert der Nullvektor ist, etwa

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = \mathbf{0}$$

mit  $v_1, v_2, \dots, v_n \in E$  paarweise verschieden, dann ist  $\lambda_j \neq 0$  für wenigstens ein  $j \in \{1, 2, \dots, n\}$ . Wir können dann (falls nötig) die Nummerierung so ändern, dass  $\lambda_n \neq 0$  ist. Dann ist die Gleichung äquivalent zu

$$v_n = -(\lambda_n^{-1} \lambda_1) v_1 - (\lambda_n^{-1} \lambda_2) v_2 - \dots - (\lambda_n^{-1} \lambda_{n-1}) v_{n-1}.$$

Wir können also ein Element  $v_0$  von  $E$  (nämlich  $v_n$ ) als Linearkombination von Elementen von  $E \setminus \{v_0\}$  schreiben. Daraus folgt, dass  $E_0 = E \setminus \{v_0\}$  immer noch ein Erzeugendensystem von  $V$  ist. Das sieht man so: Wir nehmen an, dass es eine Darstellung

$$v_0 = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n$$

gibt mit  $v_1, v_2, \dots, v_n \in E_0$  und  $\lambda_1, \lambda_2, \dots, \lambda_n \in K$ . Jetzt müssen wir zeigen, dass jedes Element  $v \in V$  als Linearkombination von Elementen von  $E_0$  geschrieben werden kann. Wir können (da ja  $E$  ein Erzeugendensystem ist)  $v$  jedenfalls als Linearkombination von Elementen von  $E$  schreiben:

$$v = \mu_1 w_1 + \mu_2 w_2 + \dots + \mu_m w_m$$

mit  $\mu_1, \mu_2, \dots, \mu_m \in K$  und  $w_1, w_2, \dots, w_m \in E$ ; durch Zusammenfassen gleichartiger Terme können wir annehmen, dass  $w_1, w_2, \dots, w_m$  paarweise verschieden sind. Kommt  $v_0$  nicht unter diesen Elementen vor, dann haben wir bereits eine Linearkombination von Elementen von  $E_0$ . Wenn  $v_0$  vorkommt, dann können wir (möglicherweise nach Änderung der Nummerierung) annehmen, dass  $v_0 = w_m$  ist. Dann haben wir

$$\begin{aligned} v &= \mu_1 w_1 + \mu_2 w_2 + \dots + \mu_{m-1} w_{m-1} + \mu_m w_m \\ &= \mu_1 w_1 + \mu_2 w_2 + \dots + \mu_{m-1} w_{m-1} + \mu_m v_0 \\ &= \mu_1 w_1 + \mu_2 w_2 + \dots + \mu_{m-1} w_{m-1} + \mu_m (\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n) \\ &= \mu_1 w_1 + \mu_2 w_2 + \dots + \mu_{m-1} w_{m-1} + (\mu_m \lambda_1) v_1 + (\mu_m \lambda_2) v_2 + \dots + (\mu_m \lambda_n) v_n; \end{aligned}$$

dies ist eine Linearkombination von Elementen von  $E_0$  (denn wir haben  $v_0$  durch Elemente von  $E_0$  ersetzt).

$E$  ist also genau dann ein minimales Erzeugendensystem, wenn der Nullvektor *nicht* als nichttriviale Linearkombination von (paarweise verschiedenen) Elementen von  $E$  geschrieben werden kann. Diese Eigenschaft ist sehr wichtig und hat einen eigenen Namen.

\* 8.1. **Definition.** Sei  $V$  ein  $K$ -Vektorraum.

**DEF**  
Linear  
unabhängig

(1) Sei  $n \in \mathbb{N}$ . Die Vektoren  $v_1, v_2, \dots, v_n \in V$  heißen *(K-)linear unabhängig*, wenn gilt:

$$\forall \lambda_1, \lambda_2, \dots, \lambda_n \in K : \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = \mathbf{0} \Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

Anderenfalls heißen die Vektoren *(K-)linear abhängig*.

(2) Sei  $I$  eine Menge. Eine Familie  $(v_i)_{i \in I}$  von Elementen von  $V$  heißt *(K-)linear unabhängig*, wenn für jede endliche Teilmenge  $\{i_1, i_2, \dots, i_n\} \subset I$  (mit  $i_1, i_2, \dots, i_n$  paarweise verschieden) die Vektoren  $v_{i_1}, v_{i_2}, \dots, v_{i_n}$  linear unabhängig sind. Anderenfalls heißt  $(v_i)_{i \in I}$  *(K-)linear abhängig*.

(3) Eine Teilmenge  $A \subset V$  heißt *(K-)linear unabhängig*, wenn die Familie  $(v)_{v \in A}$  linear unabhängig ist, sonst *(K-)linear abhängig*.  $\diamond$

Eine Familie oder Menge von Vektoren ist also genau dann linear abhängig, wenn man den Nullvektor als nichttriviale Linearkombination von Vektoren aus der Familie oder der Menge schreiben kann.

Der Unterschied zwischen Familien und Mengen ist, dass die Elemente in einer Familie gewissermaßen durch die Indexmenge nummeriert sind und sich wiederholen können, während die Elemente einer Menge keine weitere Ordnung haben und nicht mehrfach vorkommen. Eine Menge  $A$  von Vektoren ist genau dann linear unabhängig, wenn jede *endliche* Teilmenge von  $A$  linear unabhängig ist.

Wie wir oben gesehen haben, ist ein Erzeugendensystem genau dann minimal, wenn es linear unabhängig ist. Aus unseren Überlegungen hat sich auch Folgendes ergeben:

*$v_1, v_2, \dots, v_n \in V$  sind genau dann linear abhängig, wenn sich einer der Vektoren als Linearkombination der übrigen schreiben lässt.*

**Wichtig:** Die Definition der Linearen Unabhängigkeit ist zentral für die Lineare Algebra. Es ist äußerst wichtig, dass Sie sie verstehen!

8.2. **Beispiel.** Wir betrachten den Grenzfall: Ist die leere Menge linear unabhängig oder linear abhängig?

**BSP**  
 $\emptyset$  ist l.u.

Die einzige Linearkombination der leeren Menge ist die leere Summe mit dem Wert  $\mathbf{0}$ . Ist diese Linearkombination trivial oder nicht? Da „trivial“ bedeutet, dass alle Koeffizienten null sind, muss die leere Linearkombination trivial sein, denn da es keine Koeffizienten gibt, ist jede Allaussage über die Koeffizienten wahr. Die leere Menge ist also linear unabhängig.

Das passt auch mit der obigen Beobachtung zusammen, dass ein Erzeugendensystem genau dann minimal ist, wenn es linear unabhängig ist, denn die leere Menge kann man ja nicht verkleinern.  $\clubsuit$

8.3. **Beispiel.** Wann ist ein einzelner Vektor  $v$  linear unabhängig?

**BSP**  
Wann ist  
 $v$  l.u.?

Die Linear„kombinationen“ haben die Form  $\lambda v$  mit  $\lambda$  aus dem jeweiligen Körper. Aus  $\lambda v = \mathbf{0}$  folgt  $\lambda = 0$  oder  $v = \mathbf{0}$  (vgl. Lemma 5.2). Das zeigt, dass  $v$  linear unabhängig ist, wenn  $v$  nicht der Nullvektor ist. Auf der anderen Seite ist  $1 \cdot \mathbf{0} = \mathbf{0}$  eine nichttriviale Linearkombination, die den Nullvektor ergibt, also ist  $\mathbf{0}$  linear abhängig. ♣

8.4. **Beispiel.** Nach unseren Überlegungen vom Anfang dieses Abschnitts sind zwei Vektoren  $v_1, v_2 \in V$  genau dann linear abhängig, wenn einer der beiden ein Vielfaches des anderen ist:  $v_2 = \lambda v_1$  oder  $v_1 = \lambda v_2$  für ein  $\lambda \in K$ . (Ist  $v_1 = \mathbf{0}$ ,  $v_2 \neq \mathbf{0}$ , dann ist  $v_1$  ein Vielfaches von  $v_2$ , aber nicht umgekehrt.) ♣

**BSP**  
L.U. von  
zwei Vektoren

8.5. **Beispiel.** Hier ist ein sehr konkretes (und typisches) Beispiel. Sind die Vektoren  $v_1 = (1, 1, 1, 1)$ ,  $v_2 = (1, 2, 3, 4)$  und  $v_3 = (1, 3, 5, 7)$  in  $V = \mathbb{R}^4$  linear unabhängig oder nicht?

**BSP**

Wir müssen die Bedingung überprüfen. Seien also  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$  mit

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = \mathbf{0} = (0, 0, 0, 0).$$

Die Frage ist, ob daraus zwingend  $\lambda_1 = \lambda_2 = \lambda_3 = 0$  folgt. Ausgeschrieben lautet die Gleichung

$$(\lambda_1 + \lambda_2 + \lambda_3, \lambda_1 + 2\lambda_2 + 3\lambda_3, \lambda_1 + 3\lambda_2 + 5\lambda_3, \lambda_1 + 4\lambda_2 + 7\lambda_3) = (0, 0, 0, 0);$$

das ist äquivalent zu den vier Gleichungen

$$\begin{aligned}\lambda_1 + \lambda_2 + \lambda_3 &= 0 \\ \lambda_1 + 2\lambda_2 + 3\lambda_3 &= 0 \\ \lambda_1 + 3\lambda_2 + 5\lambda_3 &= 0 \\ \lambda_1 + 4\lambda_2 + 7\lambda_3 &= 0\end{aligned}$$

Dieses Gleichungssystem hat  $(\lambda_1, \lambda_2, \lambda_3) = (1, -2, 1)$  als eine nichttriviale Lösung. Das bedeutet, dass die Vektoren linear abhängig sind. ♣

8.6. **Beispiel.** Das Erzeugendensystem  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  von  $K^n$  ist linear unabhängig, denn

**BSP**

$$\lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \dots + \lambda_n \mathbf{e}_n = (\lambda_1, \lambda_2, \dots, \lambda_n)$$

ist genau dann der Nullvektor, wenn alle Koeffizienten null sind. ♣

8.7. **Beispiel.** Die Funktionen  $x \mapsto 1, \sin, \cos, \sin^2, \cos^2$  aus dem Raum  $\mathcal{C}(\mathbb{R})$  der stetigen reellen Funktionen sind linear abhängig, denn es gilt

**BSP**

$$\forall x \in \mathbb{R} : \sin^2(x) + \cos^2(x) - 1 = 0,$$

also haben wir eine nichttriviale Linearkombination, die die Nullfunktion darstellt:

$$(-1) \cdot (x \mapsto 1) + 0 \cdot \sin + 0 \cdot \cos + 1 \cdot \sin^2 + 1 \cdot \cos^2 = \mathbf{0}. \quad \clubsuit$$

**8.8. Beispiel.** Die Potenzfunktionen  $f_n : x \mapsto x^n$  für  $n \in \mathbb{N}$  sind dagegen linear unabhängig. Das bedeutet **BSP**

$$\forall n \in \mathbb{N} \forall a_0, a_1, \dots, a_n \in \mathbb{R} : \\ (\forall x \in \mathbb{R} : a_0 + a_1x + \dots + a_nx^n = 0) \Rightarrow a_0 = a_1 = \dots = a_n = 0.$$

Das kann man durch vollständige Induktion beweisen. Für  $n = 0$  reduziert sich die Behauptung auf die triviale Aussage  $a_0 = 0 \Rightarrow a_0 = 0$ . Sei also jetzt  $n > 0$ . Einsetzen von  $x = 0$  liefert  $a_0 = 0$ , also haben wir

$$\forall x \in \mathbb{R} : x(a_1 + a_2x + \dots + a_nx^{n-1}) = 0,$$

was bedeutet

$$\forall x \in \mathbb{R} \setminus \{0\} : a_1 + a_2x + \dots + a_nx^{n-1} = 0.$$

Weil Polynomfunktionen stetig sind, gilt dies dann auch für  $x = 0$ . Aus der Induktionsvoraussetzung folgt dann  $a_1 = a_2 = \dots = a_n = 0$  wie gewünscht.

Man kann diese Aussage auch beweisen, indem man die (aus der Schule bekannte?) Tatsache verwendet, dass ein Polynom vom Grad  $n$  (also eine Polynomfunktion wie oben mit  $a_n \neq 0$ ) höchstens  $n$  Nullstellen hat. Das bedeutet, dass es nicht die Nullfunktion sein kann (denn die hat unendlich viele Nullstellen). Die einzige Möglichkeit, die Nullfunktion zu bekommen, ist dann, dass man alle Koeffizienten null setzt. ♣

Wir schreiben noch eine einfache, aber nützliche Beobachtung auf, die unsere Überlegungen vom Beginn dieses Abschnitts formalisiert.

**8.9. Lemma.** Sei  $V$  ein Vektorraum und  $A \subset V$  linear unabhängig. Dann gilt für alle  $v \in V$ :

$$v \in \langle A \rangle \iff v \in A \text{ oder } A \cup \{v\} \text{ linear abhängig.}$$

**LEMMA**  
Erzeugnis  
einer l.u.  
Menge

*Beweis.* „ $\Rightarrow$ “:  $v \in \langle A \rangle$  bedeutet, dass  $v = \lambda_1v_1 + \lambda_2v_2 + \dots + \lambda_nv_n$  eine Linearkombination von paarweise verschiedenen Elementen von  $A$  ist. Wenn  $v \notin A$ , dann ist

$$(-1)v + \lambda_1v_1 + \lambda_2v_2 + \dots + \lambda_nv_n = \mathbf{0}$$

und diese Linearkombination von Elementen von  $A \cup \{v\}$  ist nichttrivial. (Beachte, dass  $v$  von den  $v_1, v_2, \dots, v_n$  verschieden ist.) Also ist  $A \cup \{v\}$  linear abhängig.

„ $\Leftarrow$ “: Aus  $v \in A$  folgt  $v \in \langle A \rangle$ . Sei jetzt  $v \notin A$ . Wenn  $A \cup \{v\}$  linear abhängig ist, dann gibt es eine nichttriviale Linearkombination

$$\lambda v + \lambda_1v_1 + \lambda_2v_2 + \dots + \lambda_nv_n = \mathbf{0}$$

mit  $v_1, v_2, \dots, v_n \in A$  paarweise verschieden. Dann kann  $\lambda$  nicht null sein, denn sonst hätten wir eine nichttriviale Linearkombination von Elementen von  $A$ , die den Nullvektor darstellt, im Widerspruch zur linearen Unabhängigkeit von  $A$ . Dann können wir die Gleichung aber nach  $v$  auflösen:

$$v = -\lambda^{-1}\lambda_1v_1 - \lambda^{-1}\lambda_2v_2 - \dots - \lambda^{-1}\lambda_nv_n,$$

was  $v \in \langle A \rangle$  zeigt. □

Minimale, also linear unabhängige, Erzeugendensysteme spielen eine fundamentale Rolle in der Linearen Algebra.

\* **8.10. Definition.** Sei  $V$  ein  $K$ -Vektorraum. Eine Familie  $(v_i)_{i \in I}$  von Elementen von  $V$  heißt  $(K)$ -Basis(familie) von  $V$ , wenn sie linear unabhängig ist und die Menge  $\{v_i \mid i \in I\}$  ein Erzeugendensystem von  $V$  ist. Eine Teilmenge  $B \subset V$  heißt  $(K)$ -Basis(menge) von  $V$ , wenn sie ein linear unabhängiges Erzeugendensystem von  $V$  ist.  $\diamond$

**DEF**  
Basis

Manchmal ist es praktischer, mit Familien (also „nummerierten Mengen“) zu arbeiten, und manchmal ist es praktischer, mit Mengen zu arbeiten, darum haben wir den Begriff der Basis in beiden Versionen definiert. Der Unterschied ist gering, denn in einer linear unabhängigen Familie kann kein Element mehrfach auftreten.

### 8.11. Beispiele.

**BSP**  
Basen

- Ist  $V$  ein Vektorraum und  $A \subset V$  linear unabhängig, dann ist  $A$  eine Basis von  $\langle A \rangle$  (denn  $A$  ist ein linear unabhängiges Erzeugendensystem von  $\langle A \rangle$ ).
- Die leere Menge ist Basis des Null-Vektorraums  $\{0\}$ .
- Das Tupel  $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$  ist eine  $K$ -Basis von  $K^n$ , die sogenannte *Standardbasis* von  $K^n$ .
- Die Folge  $(f_n)_{n \in \mathbb{N}}$  der Potenzfunktionen ist eine Basis des Vektorraums  $P$  der Polynomfunktionen.  $\clubsuit$

Wir hatten zu Beginn dieses Abschnitts gesehen, dass ein Erzeugendensystem genau dann minimal ist, wenn es linear unabhängig (also eine Basis) ist. Wir formulieren das hier noch einmal und ergänzen es um eine ähnliche Aussage über linear unabhängige Mengen.

\* **8.12. Lemma.** Sei  $V$  ein Vektorraum und  $B \subset V$  eine Teilmenge. Dann sind die folgenden Aussagen äquivalent:

**LEMMA**  
Charakterisierung  
von Basen

- (1)  $B$  ist eine Basis von  $V$ .
- (2)  $B$  ist ein minimales Erzeugendensystem von  $V$ .
- (3)  $B$  ist eine maximale linear unabhängige Teilmenge von  $V$ .

„Maximal“ heißt dabei, dass für jedes  $v \in V \setminus B$  die (echt) größere Menge  $B \cup \{v\}$  nicht mehr linear unabhängig ist.

*Beweis.* Nach Definition 8.10 ist eine Basis ein linear unabhängiges Erzeugendensystem. Die Äquivalenz von (1) und (2) ist schon gezeigt. Um die Äquivalenz von (1) und (3) zu zeigen, müssen wir nachprüfen, dass eine linear unabhängige Teilmenge genau dann maximal ist, wenn sie ein Erzeugendensystem ist.

Für die eine Richtung nehmen wir an, dass keine echt größere Teilmenge von  $V$  linear unabhängig ist. Wir zeigen, dass  $B$  ein Erzeugendensystem ist. Dazu sei  $v \in V$ . Ist  $v \in B$ , dann ist  $v \in \langle B \rangle$ . Ist  $v \notin B$ , dann ist  $B \cup \{v\}$  linear abhängig, nach Lemma 8.9 also  $v \in \langle B \rangle$ . Da  $v \in V$  beliebig war, folgt  $\langle B \rangle = V$ , also ist  $B$  ein Erzeugendensystem von  $V$ .

Für die Gegenrichtung nehmen wir an, dass  $B$  linear unabhängig und ein Erzeugendensystem ist. Wir zeigen, dass jede echt größere Menge linear abhängig sein muss. Sei dazu  $v \in V \setminus B$ . Da  $\langle B \rangle = V$ , folgt mit Lemma 8.9, dass  $B \cup \{v\}$  linear abhängig ist.  $\square$

Wir können die Eigenschaften, ein Erzeugendensystem, linear unabhängig oder eine Basis zu sein, auch durch die Anzahl der Linearkombinationen ausdrücken, die ein gegebenes Element von  $V$  darstellen. Wir formulieren das hier für endlich viele Vektoren.

**8.13. Lemma.** Sei  $V$  ein  $K$ -Vektorraum und seien  $v_1, v_2, \dots, v_n \in V$ . Wir definieren die zugehörige „Linearkombinationenabbildung“

$$\phi_{v_1, v_2, \dots, v_n} : K^n \longrightarrow V, \quad (\lambda_1, \lambda_2, \dots, \lambda_n) \longmapsto \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n.$$

Dann gilt:

- (1)  $\{v_1, v_2, \dots, v_n\}$  ist ein **Erzeugendensystem** von  $V$  genau dann, wenn jeder Vektor  $v \in V$  auf **mindestens** eine Weise als Linearkombination von  $v_1, v_2, \dots, v_n$  geschrieben werden kann, also genau dann, wenn  $\phi_{v_1, v_2, \dots, v_n}$  **surjektiv** ist.
- (2)  $v_1, v_2, \dots, v_n$  sind **linear unabhängig** genau dann, wenn jeder Vektor  $v \in V$  auf **höchstens** eine Weise als Linearkombination von  $v_1, v_2, \dots, v_n$  geschrieben werden kann, also genau dann, wenn  $\phi_{v_1, v_2, \dots, v_n}$  **injektiv** ist.
- (3)  $(v_1, v_2, \dots, v_n)$  ist eine **Basis** von  $V$  genau dann, wenn jeder Vektor  $v \in V$  auf **genau** eine Weise als Linearkombination von  $v_1, v_2, \dots, v_n$  geschrieben werden kann, also genau dann, wenn  $\phi_{v_1, v_2, \dots, v_n}$  **bijektiv** ist.

**LEMMA**  
EZS/LU/Basis  
über Anzahl  
Lin.komb.

*Beweis.* Teil (1) folgt direkt aus Definition 7.7.

Wir beweisen Teil (2). „ $\Rightarrow$ “: Wir nehmen an, dass  $v_1, v_2, \dots, v_n$  linear unabhängig sind. Sei  $v \in V$ . Wenn wir zwei Linearkombinationen haben, also

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = \mu_1 v_1 + \mu_2 v_2 + \dots + \mu_n v_n$$

mit  $\lambda_1, \lambda_2, \dots, \lambda_n, \mu_1, \mu_2, \dots, \mu_n \in K$ , dann bilden wir die Differenz:

$$(\lambda_1 - \mu_1)v_1 + (\lambda_2 - \mu_2)v_2 + \dots + (\lambda_n - \mu_n)v_n = \mathbf{0}.$$

Weil  $v_1, v_2, \dots, v_n$  linear unabhängig sind, muss das die triviale Linearkombination sein, also folgt  $\lambda_1 = \mu_1, \lambda_2 = \mu_2, \dots, \lambda_n = \mu_n$ .

„ $\Leftarrow$ “: Wir nehmen an, dass jedes  $v \in V$  höchstens auf eine Weise als Linearkombination von  $v_1, v_2, \dots, v_n$  darstellbar ist. Das gilt dann auch für  $v = \mathbf{0}$ . Da die triviale Linearkombination  $\mathbf{0}$  darstellt, muss es die einzige sein. Damit ist gezeigt, dass  $v_1, v_2, \dots, v_n$  linear unabhängig sind.

Teil (3) folgt dann aus (1) und (2). □

Um das vorstehende Lemma auch für beliebige Familien  $(v_i)_{i \in I}$  von Vektoren formulieren zu können, definieren wir

$$K^{(I)} = \{(\lambda_i)_{i \in I} \in K^I \mid \{i \in I \mid \lambda_i \neq 0\} \text{ ist endlich}\}.$$

Das ist also die Menge derjenigen Familien von Elementen von  $K$  mit Indexmenge  $I$ , die nur endlich viele von null verschiedene Komponenten haben. Dann können wir wieder die Linearkombinationenabbildung definieren als

$$\phi_{(v_i)_{i \in I}} : K^{(I)} \longrightarrow V, \quad (\lambda_i)_{i \in I} \longmapsto \sum_{i \in I} \lambda_i v_i,$$

wobei wir die Definition der Summe etwas ausgedehnt haben: Wir setzen

$$\sum_{i \in I} a_i = \sum_{i \in J} a_i \quad \text{mit } J = \{i \in I \mid a_i \neq 0\}$$

und verlangen, dass die rechte Summe endlich ist (was in unserem Fall genau durch die Definition von  $K^{(I)}$  gesichert wird). Dann gilt wieder:

- (1)  $(v_i)_{i \in I}$  Erzeugendensystem  $\iff \phi_{(v_i)_{i \in I}}$  surjektiv.
- (2)  $(v_i)_{i \in I}$  linear unabhängig  $\iff \phi_{(v_i)_{i \in I}}$  injektiv.
- (3)  $(v_i)_{i \in I}$  Basis  $\iff \phi_{(v_i)_{i \in I}}$  bijektiv.

$K^{(I)}$  ist übrigens genau der  $K$ -Untervektorraum von  $K^I$ , der durch die Familien  $\mathbf{e}_i = (\delta_{ij})_{j \in I}$  für  $i \in I$  erzeugt wird. Dabei ist  $\delta_{ij} = 1$  für  $i = j$  und  $\delta_{ij} = 0$  für  $i \neq j$  (das sogenannte *Kronecker-Delta*); die Familie  $\mathbf{e}_i$  hat also als  $i$ -te Komponente eine Eins, alle anderen Komponenten sind null.

Eine Basis  $(v_1, v_2, \dots, v_n)$  von  $V$  verhilft uns also zu einer bijektiven Abbildung  $K^n \rightarrow V$ . Damit können wir die Elemente von  $V$  durch ihr Koeffiziententupel  $(\lambda_1, \lambda_2, \dots, \lambda_n) \in K^n$  beschreiben (und Addition und Skalarmultiplikation von  $V$  verhalten sich genauso wie die von  $K^n$ ). Das ist natürlich eine schöne Sache. Es stellt sich dann die Frage, ob jeder Vektorraum eine Basis hat. Wir werden das hier für endlich erzeugte Vektorräume positiv beantworten. Dafür beweisen wir sogar eine stärkere Aussage, die viele nützliche Anwendungen haben wird.

\*

**8.14. Satz.** *Sei  $V$  ein Vektorraum und seien  $v_1, v_2, \dots, v_n$  und  $w_1, w_2, \dots, w_m$  Elemente von  $V$ , sodass die Vektoren  $v_1, v_2, \dots, v_n$  linear unabhängig sind und die Menge  $\{v_1, v_2, \dots, v_n, w_1, w_2, \dots, w_m\}$  ein Erzeugendensystem von  $V$  ist. Dann kann man  $(v_1, v_2, \dots, v_n)$  durch Hinzunahme geeigneter Vektoren  $w_j$  zu einer Basis von  $V$  ergänzen.*

**SATZ**  
Basis-  
ergänzungs-  
satz

*Genauer bedeutet das: Es gibt  $k \in \mathbb{N}$  und Indizes  $j_1, j_2, \dots, j_k \in \{1, 2, \dots, m\}$ , sodass*

$$(v_1, v_2, \dots, v_n, w_{j_1}, w_{j_2}, \dots, w_{j_k})$$

*eine Basis von  $V$  ist.*

Die natürlichen Zahlen  $n$  und  $m$  dürfen und  $k$  kann auch null sein. Wenn  $m = 0$  ist, dann ist  $(v_1, v_2, \dots, v_n)$  schon eine Basis, und es ist nichts zu tun (dann ist auch  $k = 0$ ). Das werden wir im Beweis als Induktionsanfang benutzen.

Wenn  $n = 0$  ist, dann sagt der Satz, dass jedes endliche Erzeugendensystem eine Basis enthält. Das ist plausibel, denn man kann ja immer Elemente entfernen, solange das Erzeugendensystem nicht minimal ist. Irgendwann (nach spätestens  $m$ -maligem Entfernen eines Elements) muss man bei einem minimalen Erzeugendensystem ankommen; das ist dann eine Basis.

Wenn sich  $k = 0$  ergibt, dann bedeutet das, dass  $(v_1, v_2, \dots, v_n)$  bereits eine Basis ist.

*Beweis.* Der Beweis benutzt vollständige Induktion nach  $m$ . Er basiert auf der anschaulichen Idee, dass man nacheinander Vektoren  $w_j$  zu den  $v_1, v_2, \dots, v_n$  hinzunimmt, solange das entstehende Tupel linear unabhängig ist. Ist das nicht mehr möglich, dann muss man eine Basis haben.

Man beachte, dass im Beweis die Zahl  $n$  nicht fixiert ist.

Der Induktionsanfang, also der Fall  $m = 0$ , ist klar, denn dann ist  $(v_1, v_2, \dots, v_n)$  bereits ein linear unabhängiges Erzeugendensystem, also eine Basis. Die Behauptung gilt also mit  $k = 0$ .

Für den Induktionsschritt nehmen wir an, dass die Aussage für ein gegebenes  $m$  stimmt, und beweisen sie für  $m + 1$ . Seien also  $v_1, v_2, \dots, v_n \in V$  linear unabhängig

und  $w_1, w_2, \dots, w_m, w_{m+1} \in V$ , sodass  $\{v_1, v_2, \dots, v_n, w_1, w_2, \dots, w_m, w_{m+1}\}$  ein Erzeugendensystem von  $V$  ist. Wir unterscheiden zwei Fälle:

- (1)  $w_{m+1} \in \langle v_1, v_2, \dots, v_n \rangle$ .  
Dann ist  $\{v_1, v_2, \dots, v_n, w_1, w_2, \dots, w_m\}$  auch schon ein Erzeugendensystem; die Behauptung folgt direkt aus der Induktionsannahme.
- (2)  $w_{m+1} \notin \langle v_1, v_2, \dots, v_n \rangle$ .  
Wir schreiben  $v_{n+1}$  für  $w_{m+1}$ . Dann sind  $v_1, v_2, \dots, v_n, v_{n+1}$  linear unabhängig (wir benutzen hier wieder Lemma 8.9) und die Menge

$$\{v_1, v_2, \dots, v_n, v_{n+1}, w_1, w_2, \dots, w_m\}$$

ist ein Erzeugendensystem (dasselbe wie vorher). Nach der Induktionsannahme gibt es  $j'_1, \dots, j'_{k'} \in \{1, 2, \dots, m\}$ , sodass

$$(v_1, v_2, \dots, v_n, v_{n+1}, w_{j'_1}, w_{j'_2}, \dots, w_{j'_{k'}}) = (v_1, \dots, v_n, w_{m+1}, w_{j'_1}, \dots, w_{j'_{k'}})$$

eine Basis von  $V$  ist. Wir setzen

$$k = k' + 1, \quad j_1 = m + 1 \quad \text{und} \quad j_2 = j'_1, \quad j_3 = j'_2, \quad \dots, \quad j_k = j'_{k'}$$

und erhalten die Behauptung. □

**8.15. Folgerung.** *Jeder Vektorraum, der ein endliches Erzeugendensystem besitzt, hat eine Basis.*

**FOLG**  
Existenz  
einer Basis

*Beweis.* Das folgt aus Satz 8.14, wenn man  $n = 0$  nimmt. Genauer erhalten wir die Aussage, dass man eine Basis finden kann, die aus Elementen eines gegebenen endlichen Erzeugendensystems besteht. □

Was passiert, wenn es kein endliches Erzeugendensystem gibt? Dann gibt es auch noch einen Basisergänzungssatz, den wir hier für Mengen formulieren:

**Satz.** *Sei  $V$  ein Vektorraum und seien  $A$  und  $E$  Teilmengen von  $V$ , sodass  $A$  linear unabhängig und  $A \cup E$  ein Erzeugendensystem von  $V$  ist. Dann gibt es eine Teilmenge  $B \subset E$ , sodass  $A \cup B$  eine Basismenge von  $V$  ist.*

**SATZ**  
Basis-  
ergänzungs-  
satz

Den Beweis kann man jetzt natürlich nicht mehr durch vollständige Induktion führen. Man braucht ein anderes Werkzeug dafür, zum Beispiel das sogenannte *Zornsche Lemma*. Es besagt Folgendes.

**Satz.** *Sei  $X$  eine Menge und  $\mathcal{M} \subset \mathcal{P}(X)$  eine Menge von Teilmengen von  $X$ . Eine **Kette** in  $\mathcal{M}$  ist eine Teilmenge  $\mathcal{K} \subset \mathcal{M}$ , sodass je zwei Elemente von  $\mathcal{K}$  vergleichbar sind, das heißt*

$$\forall T_1, T_2 \in \mathcal{K} : T_1 \subset T_2 \quad \text{oder} \quad T_2 \subset T_1.$$

**SATZ**  
Zornsches  
Lemma

*Wenn jede solche Kette  $\mathcal{K}$  eine **obere Schranke** in  $\mathcal{M}$  hat, wenn es also zu  $\mathcal{K}$  ein Element  $S \in \mathcal{M}$  gibt, so dass*

$$\forall T \in \mathcal{K} : T \subset S,$$

*dann hat  $\mathcal{M}$  **maximale** Elemente. Es gibt dann also (mindestens) ein  $M \in \mathcal{M}$ , sodass gilt*

$$\forall T \in \mathcal{M} : M \subset T \Rightarrow M = T$$

*(es gibt also keine echt größere Menge in  $\mathcal{M}$ ).*

Man kann zeigen, dass das Zornsche Lemma (wenn man die „harmlosen“ Axiome der Mengenlehre als gegeben annimmt) zum Auswahlaxiom (siehe die Diskussion im Kleingedruckten auf Seite 18) äquivalent ist.

Der Beweis des Basisergänzungssatzes geht dann so:  $E$  ist die Menge  $X$  im Zornschen Lemma und  $\mathcal{M} = \{B \subset E \mid A \cup B \text{ linear unabhängig}\}$ . Wir müssen die Voraussetzung des Zornschen Lemmas nachprüfen. Sei dazu  $\mathcal{K} \subset \mathcal{M}$  eine Kette. Wir setzen  $S = \bigcup \mathcal{K}$  (das ist also die Vereinigung all der Teilmengen von  $E$ , die Elemente der Kette  $\mathcal{K}$  sind). Es ist dann klar, dass  $T \subset S$  für alle  $T \in \mathcal{K}$  gilt. Wir müssen noch zeigen, dass  $S \in \mathcal{M}$  ist, dass also  $A \cup S$  linear unabhängig ist. Angenommen, das wäre falsch, dann gäbe es eine nichttriviale Linearkombination von Elementen von  $A \cup S$ , die den Nullvektor darstellt. In dieser Linearkombination kommen nur endlich viele Elemente  $v_1, v_2, \dots, v_n$  von  $S$  vor. Da  $S = \bigcup \mathcal{K}$ , gibt es für jedes  $v_j$  ein  $T_j \in \mathcal{K}$  mit  $v_j \in T_j$ . Nach eventueller Umnummerierung können wir annehmen, dass  $K_1 \subset K_2 \subset \dots \subset K_n$  ist (hier wird verwendet, dass  $\mathcal{K}$  eine Kette ist). Dann sind aber  $v_1, v_2, \dots, v_n \in K_n$ , und es würde folgen, dass  $A \cup K_n$  linear abhängig ist. Weil  $K_n \in \mathcal{M}$  ist, ist das ein Widerspruch, also muss  $A \cup S$  linear unabhängig sein. (Für dieses Argument ist die Endlichkeit von Linearkombinationen entscheidend!) Damit ist  $S$  eine obere Schranke von  $\mathcal{K}$  in  $\mathcal{M}$  und die Voraussetzung im Zornschen Lemma ist erfüllt. Es folgt, dass  $\mathcal{M}$  ein maximales Element  $B$  hat. Da  $B \in \mathcal{M}$  ist, ist  $A \cup B$  linear unabhängig. Wäre  $A \cup B$  kein Erzeugendensystem, dann gäbe es  $v \in E$  mit  $v \notin \langle A \cup B \rangle$ . Dann wäre aber  $A \cup (B \cup \{v\})$  ebenfalls linear unabhängig. Das würde  $B \cup \{v\} \in \mathcal{M}$  bedeuten, aber das kann nicht sein, da  $B$  maximal ist ( $v$  kann kein Element von  $B$  sein, sonst wäre  $v \in \langle A \cup B \rangle$ ). Also ist  $A \cup B$  auch ein Erzeugendensystem und somit eine Basis.

Wir erhalten daraus sofort (mit  $A = \emptyset$  und  $E = V$ ):

**Folgerung.** *Jeder Vektorraum hat eine Basis.*

**FOLG**  
Existenz  
von Basen

Aus dem Auswahlaxiom folgt also zum Beispiel, dass  $\mathbb{R}$  als  $\mathbb{Q}$ -Vektorraum (als den man  $\mathbb{R}$  mit seiner Addition und der auf  $\mathbb{Q} \times \mathbb{R}$  eingeschränkten Multiplikation betrachten kann) eine Basis hat. Gesehen hat so eine Basis aber noch niemand. Wie schon früher erwähnt ist das Auswahlaxiom (und damit auch das Zornsche Lemma) inhärent inkonstruktiv, sodass unser Beweis oben (im Gegensatz zum endlichen Fall) keinerlei Hinweis darauf gibt, wie die gesuchte Teilmenge  $B$  zu finden wäre.

Eine weitere wichtige Folgerung besagt, dass man (in einem endlich erzeugten Vektorraum) beliebige linear unabhängige Vektoren stets zu einer Basis ergänzen kann.

**8.16. Folgerung.** *Sei  $V$  ein Vektorraum mit endlichem Erzeugendensystem und seien  $v_1, v_2, \dots, v_n \in V$  linear unabhängig. Dann gibt es  $k \in \mathbb{N}$  und Vektoren  $v_{n+1}, v_{n+2}, \dots, v_{n+k}$ , sodass  $(v_1, v_2, \dots, v_{n+k})$  eine Basis von  $V$  ist.*

**FOLG**  
Erweiterung  
zu Basis

*Beweis.* Sei  $\{w_1, w_2, \dots, w_m\}$  ein endliches Erzeugendensystem von  $V$ . Dann sind für  $v_1, \dots, v_n$  und  $w_1, \dots, w_m$  die Voraussetzungen von Satz 8.14 erfüllt. Die Aussage des Satzes liefert dann die Behauptung, wenn man  $v_{n+1} = w_{j_1}, \dots, v_{n+k} = w_{j_k}$  setzt.  $\square$

**8.17. Beispiel.** Wir finden eine Basis des Untervektorraums

$$U = \{(x, y, z) \in \mathbb{R}^3 \mid z = x + y\} \subset \mathbb{R}^3.$$

**BSP**  
Basis

Dazu finden wir möglichst viele linear unabhängige Vektoren und prüfen dann, ob wir ein Erzeugendensystem haben. Zum Beispiel sind  $(1, 0, 1)$  und  $(0, 1, 1)$  linear unabhängige Elemente von  $U$ , denn

$$\lambda(1, 0, 1) + \mu(0, 1, 1) = \mathbf{0} \iff (\lambda, \mu, \lambda + \mu) = (0, 0, 0) \iff \lambda = \mu = 0.$$

Diese beiden Vektoren bilden auch ein Erzeugendensystem, denn für  $(x, y, z) \in U$  gilt  $z = x + y$ , also

$$(x, y, z) = (x, y, x + y) = x(1, 0, 1) + y(0, 1, 1) \in \langle (1, 0, 1), (0, 1, 1) \rangle.$$

Damit ist  $((1, 0, 1), (0, 1, 1))$  eine Basis von  $U$ . ♣

Eine weitere wichtige Konsequenz des Basisergänzungssatzes ist der *Basisaustauschsatz*.

\* **8.18. Satz.** Sei  $V$  ein Vektorraum und seien  $(v_1, v_2, \dots, v_n)$  und  $(w_1, w_2, \dots, w_m)$  zwei Basen von  $V$ . Für jedes  $i \in \{1, 2, \dots, n\}$  gibt es ein  $j \in \{1, 2, \dots, m\}$ , sodass  $(v_1, \dots, v_{i-1}, w_j, v_{i+1}, \dots, v_n)$  ebenfalls eine Basis von  $V$  ist.

**SATZ**  
Basis-  
austausch-  
satz

Man tauscht also das Basiselement  $v_i$  der ersten Basis durch ein Element der zweiten Basis aus.

*Beweis.* Wir können ohne Einschränkung  $i = n$  annehmen (sonst ändere man die Nummerierung entsprechend). Wir wenden den Basisergänzungssatz 8.14 an mit  $v_1, v_2, \dots, v_{n-1}$  und  $w_1, w_2, \dots, w_m$ . Die Voraussetzungen sind erfüllt, da Teilmengen von linear unabhängigen Mengen immer linear unabhängig sind und die  $w_j$  schon alleine ein Erzeugendensystem bilden. Es gibt also  $k \in \mathbb{N}$  und Indizes  $j_1, \dots, j_k \in \{1, 2, \dots, m\}$ , sodass  $(v_1, \dots, v_{n-1}, w_{j_1}, w_{j_2}, \dots, w_{j_k})$  eine Basis von  $V$  ist. Die Behauptung bedeutet gerade  $k = 1$ ; wir setzen dann  $j = j_1$ . Es ist klar, dass  $k > 0$  sein muss, denn  $\{v_1, v_2, \dots, v_{n-1}\}$  ist kein Erzeugendensystem mehr ( $\{v_1, v_2, \dots, v_{n-1}, v_n\}$  ist ein minimales Erzeugendensystem, aus dem wir ein Element entfernt haben). Wir zeigen, dass  $\{v_1, v_2, \dots, v_{n-1}, w_{j_1}\}$  ein Erzeugendensystem ist; daraus folgt  $k = 1$ . Wir haben  $w_{j_1} \in \langle v_1, v_2, \dots, v_n \rangle = V$ . Nach Lemma 8.9 bedeutet das, dass  $v_1, v_2, \dots, v_{n-1}, v_n, w_{j_1}$  linear abhängig sind. Da  $v_1, v_2, \dots, v_{n-1}, w_{j_1}$  als Teil der Basis  $(v_1, \dots, v_{n-1}, w_{j_1}, w_{j_2}, \dots, w_{j_k})$  linear unabhängig sind, folgt dann wieder mit Lemma 8.9, dass  $v_n \in \langle v_1, v_2, \dots, v_{n-1}, w_{j_1} \rangle$  ist. Da natürlich auch  $v_1, v_2, \dots, v_{n-1}$  in diesem Untervektorraum enthalten sind, enthält er ein Erzeugendensystem von  $V$ ; es folgt  $\langle v_1, v_2, \dots, v_{n-1}, w_{j_1} \rangle = V$  wie behauptet. □

**8.19. Folgerung.** Sei  $V$  ein Vektorraum und seien  $(v_1, \dots, v_n)$  und  $(w_1, \dots, w_m)$  zwei (endliche) Basen von  $V$ . Dann ist  $n = m$ .

**FOLG**  
Größe  
von Basen

Je zwei Basen haben also gleich viele Elemente.

*Beweis.* Wir nehmen  $n > m$  an und leiten einen Widerspruch her (der Fall  $n < m$  geht genauso). Durch  $n$ -malige Anwendung von Satz 8.18 erhalten wir Indizes  $j_1, j_2, \dots, j_n \in \{1, 2, \dots, m\}$ , sodass  $(w_{j_1}, w_{j_2}, \dots, w_{j_n})$  eine Basis von  $V$  ist. Da  $m$  kleiner als  $n$  ist, müssen sich in diesem Tupel Vektoren wiederholen. Dann sind  $w_{j_1}, w_{j_2}, \dots, w_{j_n}$  aber nicht linear unabhängig. Dies ist der gewünschte Widerspruch. □

Wir führen eine Schreibweise für die Anzahl der Elemente einer Menge ein.

**8.20. Definition.** Sei  $M$  eine Menge. Wir schreiben  $\#M$  für die Anzahl der Elemente von  $M$ . Wenn  $M$  unendlich ist, dann setzen wir  $\#M = \infty$ . ◇

**DEF**  
 $\#M$

Eine andere häufig anzutreffende Schreibweise ist  $|M|$ . Ich bevorzuge  $\#M$ , weil es dabei keine Verwechslungsgefahr gibt.

Wir können jetzt die Dimension eines Vektorraums einführen.

\* 8.21. **Definition.** Sei  $V$  ein Vektorraum. Wenn  $V$  eine endliche Basis  $(v_1, \dots, v_n)$  hat, dann sagen wir, dass  $V$  *Dimension  $n$*  hat oder  *$n$ -dimensional* ist und schreiben  $\dim V = n$ . Hat  $V$  keine endliche Basis, dann sagen wir, dass  $V$  *unendlich-dimensional* ist und schreiben  $\dim V = \infty$ . Hat  $V$  Dimension  $n$  für ein  $n \in \mathbb{N}$ , dann heißt  $V$  *endlich-dimensional* und wir schreiben  $\dim V < \infty$ .

**DEF**  
Dimension

Wenn wir betonen wollen, dass es um die Dimension von  $V$  als  $K$ -Vektorraum geht, dann schreiben wir genauer  $\dim_K V$ .  $\diamond$

Zum Beispiel ist  $\dim_{\mathbb{C}} \mathbb{C} = 1$  ( $\mathbb{C}$ -Basis  $(1)$ ), aber  $\dim_{\mathbb{R}} \mathbb{C} = 2$  ( $\mathbb{R}$ -Basis  $(1, i)$ ).

Folgerung 8.19 sagt uns, dass diese Definition sinnvoll ist, weil alle endlichen Basen von  $V$  (wenn es sie gibt) dieselbe Anzahl von Elementen haben.

### 8.22. Beispiele.

**BSP**  
Dimension

- Die leere Menge ist Basis des Null-Vektorraums, also ist  $\dim\{\mathbf{0}\} = 0$ . Ist umgekehrt  $V$  ein Vektorraum mit  $\dim V = 0$ , dann hat  $V$  eine Basis aus null Vektoren, also ist  $V = \{\mathbf{0}\}$ .
- Für  $n \in \mathbb{N}$  gilt  $\dim K^n = n$ , denn  $K^n$  hat die  $n$ -elementige Standardbasis  $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ .
- Für den Vektorraum der Polynomfunktionen gilt  $\dim P = \infty$ , denn er hat eine unendliche Basis und kann deswegen nicht endlich-dimensional sein (siehe Folgerung 8.24 unten).  $\clubsuit$

Die Dimension eines Vektorraums ist eine wichtige Größe, wie die folgenden Aussagen zeigen.

8.23. **Satz.** Seien  $m, n \in \mathbb{N}$ , sei  $V$  ein  $n$ -dimensionaler Vektorraum und seien  $v_1, v_2, \dots, v_m \in V$ .

**SATZ**  
Eigensch.  
Dimension

- (1) Wenn  $v_1, v_2, \dots, v_m$  linear unabhängig sind, dann ist  $m \leq n$ . Im Fall  $m = n$  ist  $(v_1, v_2, \dots, v_m)$  eine Basis von  $V$ .
- (2) Wenn  $\{v_1, v_2, \dots, v_m\}$  ein Erzeugendensystem von  $V$  ist, dann ist  $m \geq n$ . Im Fall  $m = n$  ist  $(v_1, v_2, \dots, v_m)$  eine Basis von  $V$ .

Weil dieser Satz so wichtig ist, gebe ich eine weitere Formulierung.

Man kann den ersten Teil der beiden Aussagen auch so ausdrücken:

- (1) In einem  **$n$ -dimensionalen** Vektorraum sind **mehr als  $n$**  Vektoren **immer linear abhängig**.
- (2) Die **lineare Hülle** von  **$m$**  Vektoren hat **Dimension höchstens  $m$** :

$$\dim\langle v_1, v_2, \dots, v_m \rangle \leq m.$$

Die erste dieser Aussagen ist eine starke *Existenzaussage*. Sie besagt nämlich Folgendes: Sind  $v_1, v_2, \dots, v_m \in V$  mit  $m > \dim V$ , dann gibt es eine nichttriviale *Linearkombination*

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n = \mathbf{0}.$$

Der zweite Teil der beiden Aussagen im Satz oben bedeutet:

- (1) In einem  $n$ -dimensionalen Vektorraum sind  $n$  linear unabhängige Vektoren immer schon eine Basis.

- (2) *In einem  $n$ -dimensionalen Vektorraum ist ein Erzeugendensystem mit  $n$  Elementen immer schon eine Basis.*

*Beweis.*

- (1) Nach Folgerung 8.16 können wir  $(v_1, v_2, \dots, v_m)$  durch Hinzunehmen von geeigneten Vektoren von  $V$  zu einer Basis von  $V$  ergänzen. Diese Basis hat  $n$  Elemente, also gilt  $m \leq n$ . Wenn  $m = n$  ist, dann werden keine Elemente hinzugefügt, also liegt bereits eine Basis vor.
- (2) Nach dem Basisergänzungssatz 8.14 (mit  $n = 0$  in der dortigen Notation) gibt es eine Basis, die durch Weglassen von geeigneten Vektoren  $v_j$  aus  $(v_1, v_2, \dots, v_m)$  entsteht. Diese Basis hat Länge  $n$ , also gilt  $m \geq n$ . Wenn  $m = n$  ist, dann kann nichts weggelassen werden, also liegt bereits eine Basis vor.  $\square$

Linear unabhängige Mengen geben also untere Schranken und Erzeugendensysteme geben obere Schranken für die Dimension. Es ist daher plausibel, dass wir unendlich-dimensionale Vektorräume wie folgt charakterisieren können.

**8.24. Folgerung.** *Sei  $V$  ein Vektorraum. Die folgenden Aussagen sind äquivalent:*

**FOLG**  
 $\dim = \infty$

- (1) *Es gibt in  $V$  eine unendliche Menge linear unabhängiger Vektoren.*
- (2)  $\dim V = \infty$ .

*Beweis.* „(1)  $\Rightarrow$  (2)“: Es gebe eine unendliche Menge linear unabhängiger Vektoren in  $V$ . Wäre  $\dim V < \infty$ , also  $\dim V = n$  mit einem  $n \in \mathbb{N}$ , dann könnten nach Satz 8.23 nicht mehr als  $n$  Vektoren linear unabhängig sein. Wir könnten aus der nach Voraussetzung existierenden unendlichen linear unabhängigen Teilmenge von  $V$  aber (z. B.)  $n + 1$  Vektoren nehmen, die dann ebenfalls linear unabhängig wären. Wir erhalten also einen Widerspruch, d. h., die Annahme, dass  $V$  endlich-dimensional ist, muss falsch sein. Also ist  $V$  unendlich-dimensional.

„(2)  $\Rightarrow$  (1)“: Sei  $V$  unendlich-dimensional. Das bedeutet, dass  $V$  keine endliche Basis hat; damit kann eine endliche linear unabhängige Teilmenge von  $V$  kein Erzeugendensystem sein. Ist also  $T \subset V$  endlich und linear unabhängig, dann gibt es  $v_T \in V \setminus \langle T \rangle$ . Es folgt (nach Lemma 8.9), dass  $T \cup \{v_T\}$  linear unabhängig ist. Auf diese Weise kann man sukzessive eine unendliche linear unabhängige Teilmenge konstruieren.

Wir konstruieren dafür zunächst rekursiv eine aufsteigende Folge  $(A_n)_{n \in \mathbb{N}}$  von linear unabhängigen Teilmengen  $A_n$  von  $V$  mit  $\#A_n = n$ :

$$A_0 = \emptyset \quad \text{und} \quad A_{n+1} = A_n \cup \{v_{A_n}\}.$$

Dann ist  $A = \bigcup_{n \in \mathbb{N}} A_n$  die gesuchte unendliche Menge von linear unabhängigen Vektoren in  $V$ . Denn sei

$$\lambda_1 w_1 + \lambda_2 w_2 + \dots + \lambda_m w_m = \mathbf{0}$$

eine Linearkombination von (paarweise verschiedenen) Elementen von  $A$ , die den Nullvektor darstellt. Dann gibt es zu jedem  $j \in \{1, 2, \dots, m\}$  einen Index  $n_j$ , sodass  $w_j \in A_{n_j}$  ist. Weil die Folge der  $A_n$  aufsteigend ist (also  $A_0 \subset A_1 \subset A_2 \subset \dots$ ), gilt dann  $w_1, w_2, \dots, w_m \in A_n$  mit  $n = \max\{n_1, n_2, \dots, n_m\}$ . Weil  $A_n$  linear unabhängig ist, muss die Linearkombination trivial sein. Also ist  $A$  linear unabhängig.  $\square$

Wir können also sagen:

- Die Dimension von  $V$  ist die **maximale** Anzahl **linear unabhängiger** Vektoren in  $V$ .
- Die Dimension von  $V$  ist die **minimale** Anzahl von **Erzeugern** von  $V$ .

Hier ist eine Anwendung der Aussage, dass  $n+1$  Vektoren in einem  $n$ -dimensionalen Vektorraum linear abhängig sein müssen.

8.25. **Definition.** Wir sagen, eine Polynomfunktion  $f \in P$  habe  $\text{Grad} \leq n$  (und wir schreiben  $\deg(f) \leq n$ ), wenn sie in der Form

$$f(x) = a_0 + a_1x + \dots + a_nx^n$$

(mit  $a_0, a_1, \dots, a_n \in \mathbb{R}$ ) geschrieben werden kann.  $f$  hat  $\text{Grad} n$  ( $\deg(f) = n$ ), wenn  $a_n \neq 0$  ist, wenn also  $f$  nicht  $\text{Grad} \leq n-1$  hat.  $f$  hat  $\text{Grad} < n$  ( $\deg(f) < n$ ), wenn  $f$   $\text{Grad} \leq n-1$  hat.  $\diamond$

**DEF**  
Grad einer  
Polynomfkt.

Sie wissen aus der Schule, dass eine Polynomfunktion vom Grad  $n$  höchstens  $n$  reelle Nullstellen haben kann. Das kann man auch so ausdrücken:

8.26. **Lemma.** *Ist  $f$  eine Polynomfunktion mit  $\deg(f) < n$ , die mindestens  $n$  reelle Nullstellen hat, dann ist  $f$  die Nullfunktion.*

**LEMMA**  
Polynom = 0

8.27. **Beispiel.** *Seien  $x_1, \dots, x_n \in \mathbb{R}$  paarweise verschieden und  $y_1, \dots, y_n \in \mathbb{R}$ . Dann gibt es eine Polynomfunktion  $f$  mit  $\deg(f) < n$ , sodass  $f(x_j) = y_j$  ist für alle  $j \in \{1, 2, \dots, n\}$ .*

**BSP**  
Interpolation

*Beweis.* Wir betrachten die folgenden  $n+1$  Vektoren in  $\mathbb{R}^n$ :

$$\begin{aligned} v_0 &= (1, 1, 1, \dots, 1) \\ v_1 &= (x_1, x_2, x_3, \dots, x_n) \\ v_2 &= (x_1^2, x_2^2, x_3^2, \dots, x_n^2) \\ &\vdots \\ v_{n-1} &= (x_1^{n-1}, x_2^{n-1}, x_3^{n-1}, \dots, x_n^{n-1}) \\ v_n &= (y_1, y_2, y_3, \dots, y_n) \end{aligned}$$

Dann wissen wir, dass  $v_0, v_1, \dots, v_n$  linear abhängig sein müssen, denn es ist  $\dim \mathbb{R}^n = n < n+1$ . Es gibt also  $\lambda_0, \lambda_1, \dots, \lambda_n \in \mathbb{R}$ , nicht alle null, mit

$$\lambda_0 + \lambda_1x_j + \lambda_2x_j^2 + \dots + \lambda_{n-1}x_j^{n-1} + \lambda_ny_j = 0$$

für alle  $j \in \{1, 2, \dots, n\}$ . Ich behaupte, dass  $\lambda_n$  nicht null sein kann. Denn sonst hätte die Polynomfunktion

$$x \mapsto \lambda_0 + \lambda_1x + \dots + \lambda_{n-1}x^{n-1}$$

vom Grad  $< n$  mindestens die  $n$  Nullstellen  $x_1, x_2, \dots, x_n$ , müsste also nach Lemma 8.26 die Nullfunktion sein, was  $\lambda_0 = \lambda_1 = \dots = \lambda_{n-1} = 0$  bedeuten würde. Dann wäre die obige Linearkombination aber trivial, ein Widerspruch. Also ist  $\lambda_n \neq 0$ . Wir setzen

$$a_0 = -\frac{\lambda_0}{\lambda_n}, \quad a_1 = -\frac{\lambda_1}{\lambda_n}, \quad \dots, \quad a_{n-1} = -\frac{\lambda_{n-1}}{\lambda_n}$$

und

$$f(x) = a_0 + a_1x + \dots + a_{n-1}x^{n-1};$$

dann gilt

$$f(x_j) = a_0 + a_1x_j + \dots + a_{n-1}x_j^{n-1} = y_j$$

wie gewünscht. □

Als Nebenprodukt unserer Überlegungen ergab sich, dass die Vektoren  $v_j = (x_1^j, \dots, x_n^j)$  für  $j \in \{0, 1, \dots, n-1\}$  linear unabhängig sind. ♣

Die Dimension ist ein Maß für die „Größe“ eines Vektorraums. Das wird deutlich, wenn man die Dimension eines Untervektorraums betrachtet.

**8.28. Satz.** *Sei  $V$  ein Vektorraum und  $U \subset V$  ein Untervektorraum. Dann gilt  $\dim U \leq \dim V$ . Ist  $V$  endlich-dimensional und gilt  $\dim U = \dim V$ , dann ist  $U = V$ .*

**SATZ**  
Dimension  
von Unter-VR

Dabei gelte  $n \leq \infty$  für alle  $n \in \mathbb{N}$  und  $\infty \leq \infty$ .

*Beweis.* Im Fall  $\dim V = \infty$  ist die Aussage trivialerweise richtig. Sei jetzt also  $\dim V = n \in \mathbb{N}$ . Wäre  $\dim U = \infty$ , dann gäbe es nach Folgerung 8.24 unendlich viele linear unabhängige Elemente in  $U$  und damit auch in  $V$ , ein Widerspruch. Also ist  $U$  endlich-dimensional mit  $\dim U = m \in \mathbb{N}$ . Eine Basis von  $U$  besteht aus  $m$  linear unabhängigen Vektoren von  $V$ . Nach Satz 8.23 folgt  $\dim U = m \leq n = \dim V$ . Gilt  $m = n$ , dann ist die Basis von  $U$  bereits eine Basis von  $V$  und es folgt  $U = V$ . □

**8.29. Beispiel.** Ein unendlich-dimensionaler Vektorraum kann durchaus echte Untervektorräume haben, die ihrerseits unendlich-dimensional sind. Zum Beispiel können wir im Vektorraum  $P$  der Polynomfunktionen den Untervektorraum  $P_g$  der geraden Polynomfunktionen betrachten:

$$P_g = \{f \in P \mid \forall x \in \mathbb{R} : f(-x) = f(x)\}.$$

(Prüfen Sie nach, dass  $P_g$  tatsächlich ein Untervektorraum von  $P$  ist!) Da die Funktion  $x \mapsto x$ , die ein Element von  $P$  ist, nicht in  $P_g$  liegt, gilt  $P_g \neq P$ . Auf der anderen Seite sind die geraden Potenzfunktionen  $x \mapsto x^{2n}$  für  $n \in \mathbb{N}$  alle linear unabhängig, also ist  $\dim P_g = \infty$ . ♣

**BSP**  
 $\dim U =$   
 $\dim V = \infty$   
 $U \subsetneq V$

## 9. LINEARE ABBILDUNGEN

Sei  $V$  ein  $K$ -Vektorraum und seien  $v_1, v_2, \dots, v_n \in V$ . Sei weiter  $\phi$  die zugehörige Linearkombinationenabbildung

$$\phi : K^n \longrightarrow V, \quad (x_1, x_2, \dots, x_n) \longmapsto x_1 v_1 + x_2 v_2 + \dots + x_n v_n.$$

Dann gilt für  $\mathbf{x} = (x_1, x_2, \dots, x_n) \in K^n$ ,  $\mathbf{y} = (y_1, y_2, \dots, y_n) \in K^n$  und  $\lambda \in K$ :

$$\begin{aligned} \phi(\mathbf{x} + \mathbf{y}) &= \phi((x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n)) \\ &= \phi(x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \\ &= (x_1 + y_1)v_1 + (x_2 + y_2)v_2 + \dots + (x_n + y_n)v_n \\ &= (x_1 v_1 + x_2 v_2 + \dots + x_n v_n) + (y_1 v_1 + y_2 v_2 + \dots + y_n v_n) \\ &= \phi(x_1, x_2, \dots, x_n) + \phi(y_1, y_2, \dots, y_n) \\ &= \phi(\mathbf{x}) + \phi(\mathbf{y}) \end{aligned}$$

und

$$\begin{aligned} \phi(\lambda \mathbf{x}) &= \phi(\lambda(x_1, x_2, \dots, x_n)) \\ &= \phi(\lambda x_1, \lambda x_2, \dots, \lambda x_n) \\ &= (\lambda x_1)v_1 + (\lambda x_2)v_2 + \dots + (\lambda x_n)v_n \\ &= \lambda(x_1 v_1 + x_2 v_2 + \dots + x_n v_n) \\ &= \lambda \phi(x_1, x_2, \dots, x_n) \\ &= \lambda \phi(\mathbf{x}). \end{aligned}$$

(Man beachte, dass Addition und Skalarmultiplikation hier einmal in  $K^n$  und einmal in  $V$  stattfinden.)

Die Abbildung  $\phi$  ist also mit Addition und Skalarmultiplikation verträglich: Das Bild einer Summe ist die Summe der Bilder und das Bild eines skalaren Vielfachen ist das entsprechende Vielfache des Bildes. Solche mit der linearen Struktur verträgliche Abbildungen heißen *lineare Abbildungen*.

\*

**9.1. Definition.** Sei  $K$  ein Körper und seien  $V_1$  und  $V_2$  zwei  $K$ -Vektorräume. Eine Abbildung  $\phi : V_1 \rightarrow V_2$  heißt *(K-)linear* oder ein *Homomorphismus (von K-Vektorräumen)*, wenn sie die folgenden beiden Bedingungen erfüllt:

- (1)  $\forall v, w \in V_1 : \phi(v + w) = \phi(v) + \phi(w)$ .
- (2)  $\forall \lambda \in K \forall v \in V_1 : \phi(\lambda v) = \lambda \phi(v)$ .

**DEF**  
Lineare  
Abbildung  
Homo-  
morphismus

Eine lineare Abbildung heißt ein *Monomorphismus*, wenn sie injektiv ist, ein *Epi-morphismus*, wenn sie surjektiv ist, und ein *Isomorphismus*, wenn sie bijektiv ist. Eine lineare Abbildung  $\phi : V \rightarrow V$  heißt ein *Endomorphismus* von  $V$ ;  $\phi$  heißt ein *Automorphismus* von  $V$ , wenn  $\phi$  außerdem bijektiv ist. Zwei Vektorräume  $V_1$  und  $V_2$  heißen (zueinander) *isomorph*,  $V_1 \cong V_2$ , wenn es einen Isomorphismus  $\phi : V_1 \rightarrow V_2$  gibt.  $\diamond$

Mono-, Epi-,  
Iso-, Endo-,  
Automorph.  
isomorph

**9.2. Beispiele.**

- Für beliebige  $K$ -Vektorräume  $V_1$  und  $V_2$  ist die *Nullabbildung*  $V_1 \rightarrow V_2$ ,  $v \mapsto \mathbf{0}$ , eine lineare Abbildung.
- Für jeden  $K$ -Vektorraum  $V$  ist die identische Abbildung  $\text{id}_V : V \rightarrow V$  ein Automorphismus von  $V$ .

**BSP**  
lineare  
Abbildungen

- Ist  $V$  ein Vektorraum und  $U \subset V$  ein Untervektorraum, dann ist die Inklusionsabbildung  $U \rightarrow V$  linear. ♣

**9.3. Beispiel.** Seien  $V, v_1, v_2, \dots, v_n$  und  $\phi$  wie zum Beginn dieses Abschnitts. Dann ist  $\phi$  ein Homomorphismus. Nach Lemma 8.13 gilt außerdem:

- $v_1, v_2, \dots, v_n$  linear unabhängig  $\iff \phi$  ist ein Monomorphismus.
- $\langle v_1, v_2, \dots, v_n \rangle = V \iff \phi$  ist ein Epimorphismus.
- $(v_1, v_2, \dots, v_n)$  Basis von  $V \iff \phi$  ist ein Isomorphismus.

**BSP**  
Linearkomb.-  
Abbildung  
ist linear

Aus dem letzten Punkt ergibt sich die Aussage

$$\dim V < \infty \implies V \cong K^{\dim V}.$$

♣

Wir überzeugen uns noch davon, dass eine lineare Abbildung wirklich mit der gesamten Struktur verträglich ist und dass sich lineare Abbildungen bezüglich Komposition und Inversion gut verhalten.

**9.4. Lemma.**  $V_1, V_2$  und  $V_3$  seien  $K$ -Vektorräume.

**LEMMA**  
Lin. Abb.:  
Eigensch.

- (1) Sei  $\phi : V_1 \rightarrow V_2$  eine lineare Abbildung. Dann gilt
 
$$\phi(\mathbf{0}) = \mathbf{0} \quad \text{und} \quad \forall v \in V_1 : \phi(-v) = -\phi(v).$$
- (2) Seien  $\phi_1 : V_1 \rightarrow V_2$  und  $\phi_2 : V_2 \rightarrow V_3$  lineare Abbildungen. Dann ist auch  $\phi_2 \circ \phi_1 : V_1 \rightarrow V_3$  linear.
- (3) Sei  $\phi : V_1 \rightarrow V_2$  ein Isomorphismus. Dann ist die Umkehrabbildung  $\phi^{-1} : V_2 \rightarrow V_1$  ebenfalls ein Isomorphismus.

Teil (3) zeigt, dass es in der Definition von „isomorph“ nicht darauf ankommt, ob man einen Isomorphismus  $V_1 \rightarrow V_2$  oder einen Isomorphismus  $V_2 \rightarrow V_1$  fordert.

*Beweis.*

- (1) Es gilt  $\phi(\mathbf{0}) = \phi(\mathbf{0} + \mathbf{0}) = \phi(\mathbf{0}) + \phi(\mathbf{0})$ . Durch Addition von  $-\phi(\mathbf{0})$  folgt  $\phi(\mathbf{0}) = \mathbf{0}$ .  
Außerdem gilt für  $v \in V_1$ :  $\phi(-v) = \phi((-1)v) = (-1)\phi(v) = -\phi(v)$ .

- (2) Wir müssen die beiden Eigenschaften aus Definition 9.1 für  $\phi_2 \circ \phi_1$  nachweisen. Seien dazu  $v, w \in V_1$  und  $\lambda \in K$ . Dann gilt

$$\begin{aligned} (\phi_2 \circ \phi_1)(v + w) &= \phi_2(\phi_1(v + w)) = \phi_2(\phi_1(v) + \phi_1(w)) \\ &= \phi_2(\phi_1(v)) + \phi_2(\phi_1(w)) = (\phi_2 \circ \phi_1)(v) + (\phi_2 \circ \phi_1)(w) \end{aligned}$$

und

$$(\phi_2 \circ \phi_1)(\lambda v) = \phi_2(\phi_1(\lambda v)) = \phi_2(\lambda \phi_1(v)) = \lambda \phi_2(\phi_1(v)) = \lambda (\phi_2 \circ \phi_1)(v).$$

- (3) Wir weisen die Eigenschaften aus Definition 9.1 für  $\phi^{-1}$  nach. Seien dazu  $v, w \in V_2$  und  $\lambda \in K$ . Wir setzen  $v' = \phi^{-1}(v)$  und  $w' = \phi^{-1}(w)$ , sodass  $v = \phi(v')$  und  $w = \phi(w')$ . Dann gilt

$$\phi^{-1}(v + w) = \phi^{-1}(\phi(v') + \phi(w')) = \phi^{-1}(\phi(v' + w')) = v' + w' = \phi^{-1}(v) + \phi^{-1}(w)$$

und

$$\phi^{-1}(\lambda v) = \phi^{-1}(\lambda \phi(v')) = \phi^{-1}(\phi(\lambda v')) = \lambda v' = \lambda \phi^{-1}(v). \quad \square$$

Bevor wir weitere Eigenschaften untersuchen, führen wir noch eine Schreibweise ein.

**9.5. Definition.** Sei  $f : X \rightarrow Y$  eine Abbildung zwischen beliebigen Mengen  $X$  und  $Y$ . Ist  $T$  eine Teilmenge von  $X$ , dann schreiben wir

$$f(T) = \{f(x) \mid x \in T\} \subset Y$$

für die Menge der Bilder der Elemente von  $T$  und nennen  $f(T)$  das *Bild von  $T$  unter  $f$* . Im Spezialfall  $T = X$  schreiben wir auch  $\text{im}(f)$  für  $f(X)$ ;  $\text{im}(f)$  heißt das *Bild* oder die *Bildmenge* von  $f$ . Ist  $U$  eine Teilmenge von  $Y$ , dann schreiben wir

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

für die Menge der Urbilder der Elemente von  $U$  und nennen  $f^{-1}(U)$  das *Urbild von  $U$  unter  $f$* .  $\diamond$

Auf dem zweiten Übungsblatt haben wir  $f_{\mathcal{P}}(T)$  für  $f(T)$  und  $f_{\mathcal{P}}^{-1}(U)$  für  $f^{-1}(U)$  geschrieben, um den Unterschied zwischen  $f : X \rightarrow Y$  und  $f_{\mathcal{P}} : \mathcal{P}(X) \rightarrow \mathcal{P}(Y)$  deutlich zu machen. Es ist (leider) üblich, diesen Unterschied in der Notation zu verwischen; deshalb muss man umso genauer aufpassen, was gemeint ist. Häufig wird auch  $f^{-1}(y) = \{x \in X \mid f(x) = y\}$  für die Menge der Urbilder eines *Elements*  $y \in Y$  geschrieben. Wir werden die „Datentypen“ (Elemente bzw. Teilmengen) hier aber sorgfältig auseinanderhalten und immer  $f^{-1}(\{y\})$  für diese Menge schreiben.

Wenn  $f$  bijektiv ist, dann hat  $f^{-1}(U)$  zwei mögliche Bedeutungen: einerseits ausgehend von  $f$  wie oben definiert und andererseits ausgehend von der Umkehrfunktion  $f^{-1}$ . Zum Glück stimmen beide Versionen überein.

Noch eine **Warnung**: Die hier eingeführte Schreibweise kann einen dazu verführen zu denken, dass  $f^{-1}(f(T)) = T$  und  $f(f^{-1}(U)) = U$  sein muss. Das ist aber im Allgemeinen **falsch**! Es gilt immer  $f^{-1}(f(T)) \supset T$  und  $f(f^{-1}(U)) \subset U$ ; die Inklusionen können jedoch echt sein.

Da der Nullvektor eine ausgezeichnete Rolle in einem Vektorraum spielt, ist die Menge seiner Urbilder unter einer linearen Abbildung ein wichtiges Datum.

\* **9.6. Definition.** Sei  $\phi : V_1 \rightarrow V_2$  eine lineare Abbildung. Der *Kern* von  $\phi$  ist die Menge der Urbilder von  $\mathbf{0} \in V_2$ : **DEF**  
Kern

$$\ker(\phi) = \phi^{-1}(\{\mathbf{0}\}) = \{v \in V_1 \mid \phi(v) = \mathbf{0}\} \subset V_1. \quad \diamond$$

Nach Lemma 9.4 gilt stets  $\mathbf{0} \in \ker(\phi)$ .

**9.7. Beispiel.** Sei  $V \subset \mathbb{R}^{\mathbb{N}}$  der Vektorraum der konvergenten Folgen. Dann ist **BSP**  
Limes ist  
linear

$$\lim : V \longrightarrow \mathbb{R}, \quad (a_n)_{n \in \mathbb{N}} \longmapsto \lim_{n \rightarrow \infty} a_n$$

eine  $\mathbb{R}$ -lineare Abbildung. Das folgt aus den Rechenregeln für Grenzwerte.

Der Kern  $\ker(\lim)$  ist gerade die Menge der Nullfolgen, denn das sind definitionsgemäß die Folgen mit Limes null.  $\clubsuit$

Eine wichtige Eigenschaft des Kerns ist die folgende:

9.8. **Lemma.** Sei  $\phi : V_1 \rightarrow V_2$  eine lineare Abbildung. Dann gilt:

$$\phi \text{ ist injektiv} \iff \ker(\phi) = \{\mathbf{0}\}.$$

**LEMMA**  
injektiv  
 $\iff$   
 $\ker = \{\mathbf{0}\}$

Man sagt in diesem Fall auch, der Kern sei *trivial*.

*Beweis.* „ $\Rightarrow$ “: Sei  $\phi$  injektiv und  $v \in \ker(\phi)$ . Dann ist  $\phi(v) = \mathbf{0} = \phi(\mathbf{0})$ , also  $v = \mathbf{0}$ .  
„ $\Leftarrow$ “: Es gelte  $\ker(\phi) = \{\mathbf{0}\}$ . Seien weiter  $v, w \in V_1$  mit  $\phi(v) = \phi(w)$ . Dann folgt  $\mathbf{0} = \phi(v) - \phi(w) = \phi(v - w)$ , also ist  $v - w \in \ker(\phi) = \{\mathbf{0}\}$  und damit  $v - w = \mathbf{0}$ ; das bedeutet  $v = w$ .  $\square$

Wie zu erwarten, vertragen sich lineare Abbildungen sehr gut mit Untervektorräumen.

9.9. **Satz.** Sei  $\phi : V_1 \rightarrow V_2$  eine  $K$ -lineare Abbildung.

**SATZ**  
lin. Abb.  
und UVR

- (1) Ist  $U_1 \subset V_1$  ein Untervektorraum, dann ist  $\phi(U_1) \subset V_2$  wieder ein Untervektorraum. Insbesondere ist  $\text{im}(\phi) = \phi(V_1) \subset V_2$  ein Untervektorraum von  $V_2$ . Außerdem ist die auf  $U_1$  eingeschränkte Abbildung  $\phi|_{U_1} : U_1 \rightarrow V_2$  ebenfalls linear.
- (2) Ist  $U_2 \subset V_2$  ein Untervektorraum, dann ist  $\phi^{-1}(U_2) \subset V_1$  wieder ein Untervektorraum. Insbesondere ist  $\ker(\phi) = \phi^{-1}(\{\mathbf{0}\}) \subset V_1$  ein Untervektorraum von  $V_1$ .
- (3) Die Abbildungen  $U_1 \mapsto \phi(U_1)$  und  $U_2 \mapsto \phi^{-1}(U_2)$  sind zueinander inverse Bijektionen zwischen der Menge der  $\ker(\phi)$  enthaltenden Untervektorräume von  $V_1$  und der Menge der in  $\text{im}(\phi)$  enthaltenen Untervektorräume von  $V_2$ .

Die Aussage, dass der Kern einer linearen Abbildung ein Untervektorraum ist, ist oft nützlich, wenn man zeigen möchte, dass eine Teilmenge eines Vektorraums ein Untervektorraum ist. Oft kann man nämlich Untervektorräume in natürlicher Weise als Kerne schreiben.

Die dritte Aussage kann man etwas genauer so formulieren: Seien

$$M_1 = \{U_1 \mid U_1 \subset V_1 \text{ Untervektorraum mit } \ker(\phi) \subset U_1\} \quad \text{und}$$

$$M_2 = \{U_2 \mid U_2 \subset V_2 \text{ Untervektorraum mit } U_2 \subset \text{im}(\phi)\}$$

und  $F$  und  $G$  folgende Abbildungen:

$$F : M_1 \longrightarrow M_2, \quad U_1 \longmapsto \phi(U_1), \quad G : M_2 \longrightarrow M_1, \quad U_2 \longmapsto \phi^{-1}(U_2).$$

Dann sind  $F$  und  $G$  bijektiv und Umkehrabbildungen voneinander. Dazu äquivalent ist, dass  $G \circ F = \text{id}_{M_1}$  und  $F \circ G = \text{id}_{M_2}$  ist.

*Beweis.*

- (1) Wir müssen die Bedingungen für einen Untervektorraum für  $\phi(U_1)$  nachprüfen:
  - $\mathbf{0} = \phi(\mathbf{0}) \in \phi(U_1)$ , da  $\mathbf{0} \in U_1$ .
  - Seien  $v, w \in \phi(U_1)$ . Dann gibt es  $v', w' \in U_1$  mit  $\phi(v') = v$  und  $\phi(w') = w$ . Es folgt  $v + w = \phi(v') + \phi(w') = \phi(v' + w') \in \phi(U_1)$ , denn  $v' + w' \in U_1$ .
  - Sei  $v \in \phi(U_1)$  und  $\lambda \in K$ . Dann gibt es  $v' \in U_1$  mit  $\phi(v') = v$ . Es folgt  $\lambda v = \lambda \phi(v') = \phi(\lambda v') \in \phi(U_1)$ , denn  $\lambda v' \in U_1$ .

Da  $V_1$  ein Untervektorraum von  $V_1$  ist, folgt, dass  $\text{im}(\phi)$  ein Untervektorraum von  $V_2$  ist.

Dass  $\phi|_{U_1}$  linear ist, folgt aus Definition 9.1, da die geforderten Eigenschaften die Form „für alle...“ haben.

(2) Wir prüfen die Bedingungen für  $\phi^{-1}(U_2)$ :

- $\mathbf{0} \in \phi^{-1}(U_2)$ , da  $\phi(\mathbf{0}) = \mathbf{0} \in U_2$ .
- Seien  $v, w \in \phi^{-1}(U_2)$ . Dann sind  $\phi(v), \phi(w) \in U_2$ . Es folgt  $\phi(v+w) = \phi(v) + \phi(w) \in U_2$  und damit  $v+w \in \phi^{-1}(U_2)$ .
- Sei  $v \in \phi^{-1}(U_2)$  und  $\lambda \in K$ . Dann ist  $\phi(v) \in U_2$ , also auch  $\phi(\lambda v) = \lambda\phi(v) \in U_2$ , und damit  $\lambda v \in \phi^{-1}(U_2)$ .

Da  $\{\mathbf{0}\}$  ein Untervektorraum von  $V_2$  ist, folgt, dass  $\ker(\phi)$  ein Untervektorraum von  $V_1$  ist.

(3) Wir überlegen zunächst, dass die Abbildungen wohldefiniert sind: Für  $U_1 \subset V_1$  gilt  $\phi(U_1) \subset \phi(V_1) = \text{im}(\phi)$  und für  $U_2 \subset V_2$  gilt  $\phi^{-1}(U_2) \supset \phi^{-1}(\{\mathbf{0}\}) = \ker(\phi)$ . Nach den Teilen (1) und (2) werden Untervektorräume auf Untervektorräume abgebildet. Damit haben wir tatsächlich Abbildungen zwischen den beiden angegebenen Mengen.

Wir zeigen jetzt, dass die Abbildungen zueinander invers sind. Daraus folgt dann auch, dass sie bijektiv sind. Sei also  $U_1 \subset V_1$  ein Untervektorraum mit  $\ker(\phi) \subset U_1$ . Dann gilt

$$\begin{aligned} v \in \phi^{-1}(\phi(U_1)) &\iff \phi(v) \in \phi(U_1) \\ &\iff \exists v' \in U_1 : \phi(v) = \phi(v') \\ &\iff \exists v' \in U_1 : \phi(v - v') = \mathbf{0} \\ &\iff \exists v' \in U_1 : v - v' \in \ker(\phi) \\ &\iff v \in U_1. \end{aligned}$$

(Die letzte Äquivalenz sieht man so: „ $\Leftarrow$ “: wähle  $v' = v$ . „ $\Rightarrow$ “: Sei  $v'' = v - v' \in \ker(\phi) \subset U_1$ , dann ist  $v = v' + v'' \in U_1$ .) Das zeigt  $\phi^{-1}(\phi(U_1)) = U_1$ . Sei jetzt  $U_2 \subset \text{im}(\phi)$  ein Untervektorraum von  $V_2$ . Dann gilt

$$\begin{aligned} v \in \phi(\phi^{-1}(U_2)) &\iff \exists v' \in \phi^{-1}(U_2) : \phi(v') = v \\ &\iff \exists v' \in V_1 : \phi(v') \in U_2 \text{ und } \phi(v') = v \\ &\iff v \in U_2 \quad \text{und} \quad v \in \text{im}(\phi) \\ &\iff v \in U_2 \cap \text{im}(\phi) \\ &\iff v \in U_2. \end{aligned}$$

Das zeigt  $\phi(\phi^{-1}(U_2)) = U_2$ . □

**9.10. Beispiel.** Sei  $K$  ein Körper,  $X$  eine Menge und  $V$  ein Untervektorraum von  $K^X = \text{Abb}(X, K)$  (zum Beispiel können wir  $X = K = \mathbb{R}$  setzen und für  $V$  den Vektorraum der stetigen reellen Funktionen nehmen). Sei weiter  $x \in X$ . Dann ist die *Auswertungsabbildung*

**BSP**  
Auswertungs-  
abbildung

$$\text{ev}_x : V \longrightarrow K, \quad f \longmapsto f(x)$$

linear. Das ergibt sich direkt aus der Definition der Addition und skalaren Multiplikation von Funktionen:

$$\begin{aligned} \text{ev}_x(f + g) &= (f + g)(x) = f(x) + g(x) = \text{ev}_x(f) + \text{ev}_x(g) \quad \text{und} \\ \text{ev}_x(\lambda f) &= (\lambda f)(x) = \lambda f(x) = \lambda \text{ev}_x(f). \end{aligned}$$

(Man kann sagen, dass die Addition und Skalarmultiplikation in  $K^X$  gerade so definiert sind, *damit* die Auswertungsabbildungen linear werden!)

Sei  $T$  eine Teilmenge von  $X$ . Dann ist

$$\{f \in V \mid \forall x \in T : f(x) = 0\} = \bigcap_{x \in T} \ker(\text{ev}_x)$$

ein Untervektorraum von  $V$ .

Im Spezialfall  $X = \{1, 2, \dots, n\}$  haben wir  $K^X = K^n$ ; dann heißen die Abbildungen  $\text{ev}_j$  (für  $j \in \{1, 2, \dots, n\}$ ) *Projektionen* und werden  $\text{pr}_j$  geschrieben:

$$\text{pr}_j : K^n \longrightarrow K, \quad (a_1, a_2, \dots, a_n) \longmapsto a_j$$

Sie sind also ebenfalls linear. ♣

Wir zeigen jetzt, dass eine lineare Abbildung dadurch festgelegt ist, was sie auf einer Basis macht.

**\* 9.11. Satz.** *Sei  $V$  ein  $K$ -Vektorraum mit Basis  $(b_1, b_2, \dots, b_n)$  und sei  $W$  ein weiterer  $K$ -Vektorraum. Seien weiter  $w_1, w_2, \dots, w_n \in W$ . Dann gibt es genau eine  $K$ -lineare Abbildung  $\phi : V \rightarrow W$  mit  $\phi(b_j) = w_j$  für alle  $j \in \{1, 2, \dots, n\}$ .*

**SATZ**  
Basen und  
lin. Abb.

*Beweis.* Wir beweisen zuerst die Eindeutigkeit. Seien also  $\phi_1, \phi_2 : V \rightarrow W$  lineare Abbildungen mit  $\phi_1(b_j) = w_j = \phi_2(b_j)$  für alle  $j \in \{1, 2, \dots, n\}$ . Sei  $v \in V$  beliebig. Dann ist  $v$  eine Linearkombination der Basisvektoren:

$$v = \lambda_1 b_1 + \lambda_2 b_2 + \dots + \lambda_n b_n.$$

Es folgt

$$\begin{aligned} \phi_1(v) &= \phi_1(\lambda_1 b_1 + \lambda_2 b_2 + \dots + \lambda_n b_n) \\ &= \lambda_1 \phi_1(b_1) + \lambda_2 \phi_1(b_2) + \dots + \lambda_n \phi_1(b_n) \\ &= \lambda_1 w_1 + \lambda_2 w_2 + \dots + \lambda_n w_n \\ &= \lambda_1 \phi_2(b_1) + \lambda_2 \phi_2(b_2) + \dots + \lambda_n \phi_2(b_n) \\ &= \phi_2(\lambda_1 b_1 + \lambda_2 b_2 + \dots + \lambda_n b_n) \\ &= \phi_2(v), \end{aligned}$$

also ist  $\phi_1 = \phi_2$ .

Dieser Eindeutigkeitsbeweis zeigt uns, wie wir die Existenz beweisen können: Wenn es eine lineare Abbildung  $\phi : V \rightarrow W$  gibt mit  $\phi(b_j) = w_j$  für alle  $j \in \{1, 2, \dots, n\}$ , dann muss für  $v \in V$  wie oben gelten

$$\phi(v) = \lambda_1 w_1 + \lambda_2 w_2 + \dots + \lambda_n w_n.$$

Wir müssen prüfen,

- (1) dass  $\phi$  wohldefiniert ist, dass also  $\phi(v)$  nicht davon abhängt, wie  $v$  als Linearkombination der  $b_j$  geschrieben wurde, und
- (2) dass die so definierte Abbildung  $\phi$  linear ist.

Die Wohldefiniertheit folgt daraus, dass  $v$  nur auf genau eine Weise als Linearkombination der Basisvektoren geschrieben werden kann (vgl. Lemma 8.13). Die Linearität rechnet man nach. Etwas eleganter ist es, wenn man bemerkt, dass  $\phi = \phi_{w_1, w_2, \dots, w_n} \circ \phi_{b_1, b_2, \dots, b_n}^{-1}$  ist (mit den zu  $b_1, b_2, \dots, b_n$  und zu  $w_1, w_2, \dots, w_n$  gehörigen Linearkombinationenabbildungen  $K^n \rightarrow V$  bzw.  $K^n \rightarrow W$  — man beachte, dass  $\phi_{b_1, b_2, \dots, b_n}$  hier ein Isomorphismus ist). Die Linearität von  $\phi$  folgt dann daraus, dass die Linearkombinationenabbildungen linear sind (Beispiel 9.3) und aus Lemma 9.4.  $\square$

Das analoge Resultat gilt auch für (nicht unbedingt endliche) Basismengen:

*Sind  $V$  und  $W$   $K$ -Vektorräume und ist  $B \subset V$  eine Basis, dann gibt es zu jeder Abbildung  $f : B \rightarrow W$  genau eine lineare Abbildung  $\phi : V \rightarrow W$  mit  $\phi(b) = f(b)$  für alle  $b \in B$  (oder kurz:  $\phi|_B = f$ ).*

Der Beweis geht im Wesentlichen genauso unter Verwendung der allgemeinen Linearkombinationenabbildungen  $K^{(B)} \rightarrow V$  bzw.  $K^{(B)} \rightarrow W$ .

Da eine lineare Abbildung also durch das Bild einer Basis eindeutig bestimmt ist, sollten sich auch Eigenschaften wie injektiv oder surjektiv zu sein durch das Bild der Basis ausdrücken lassen.

\* **9.12. Satz.** *Seien  $V$  und  $W$   $K$ -Vektorräume und  $\phi : V \rightarrow W$  linear. Sei weiter  $(b_1, b_2, \dots, b_n)$  eine Basis von  $V$ .*

**SATZ**  
inj./surj.  
lin. Abb.

- (1)  $\phi$  ist genau dann injektiv, wenn  $\phi(b_1), \phi(b_2), \dots, \phi(b_n) \in W$  linear unabhängig sind.
- (2)  $\phi$  ist genau dann surjektiv, wenn  $\phi(b_1), \phi(b_2), \dots, \phi(b_n)$  den Vektorraum  $W$  erzeugen.
- (3)  $\phi$  ist genau dann ein Isomorphismus, wenn  $(\phi(b_1), \phi(b_2), \dots, \phi(b_n))$  eine Basis von  $W$  ist.

*Beweis.* Seien  $w_1 = \phi(b_1)$ ,  $w_2 = \phi(b_2)$ ,  $\dots$ ,  $w_n = \phi(b_n)$ . Wie im Beweis von Satz 9.11 ist dann  $\phi = \phi_{w_1, w_2, \dots, w_n} \circ \phi_{b_1, b_2, \dots, b_n}^{-1}$ . Da  $\phi_{b_1, b_2, \dots, b_n}$  bijektiv ist, ist  $\phi$  injektiv bzw. surjektiv genau dann, wenn  $\phi_{w_1, w_2, \dots, w_n}$  die entsprechende Eigenschaft hat (beachte dafür  $\phi \circ \phi_{b_1, b_2, \dots, b_n} = \phi_{w_1, w_2, \dots, w_n}$ ). Die Behauptungen folgen dann sofort aus den Aussagen von Beispiel 9.3.  $\square$

Daraus können wir gleich zwei wichtige Folgerungen ziehen.

**9.13. Folgerung.** *Sind  $V$  und  $W$  zwei  $K$ -Vektorräume derselben endlichen Dimension  $n$ , dann sind  $V$  und  $W$  isomorph.*

**FOLG**  
endl.-dim.  
VR gleicher  
Dimension  
sind  
isomorph

*Beweis.* Sei  $(b_1, b_2, \dots, b_n)$  eine Basis von  $V$  und  $(b'_1, b'_2, \dots, b'_n)$  eine Basis von  $W$ . Dann gibt es nach Satz 9.11 eine lineare Abbildung  $\phi : V \rightarrow W$  mit  $\phi(b_j) = b'_j$  für alle  $j \in \{1, 2, \dots, n\}$ . Nach Satz 9.12 ist  $\phi$  ein Isomorphismus.  $\square$

**9.14. Folgerung.** Seien  $V$  und  $W$  zwei  $K$ -Vektorräume derselben endlichen Dimension  $n$  und sei  $\phi : V \rightarrow W$  eine lineare Abbildung. Dann sind die folgenden Aussagen äquivalent:

**FOLG**  
lin. Abb.  
bei gleicher  
Dimension

- (1)  $\phi$  ist ein Isomorphismus.
- (2)  $\phi$  ist injektiv.
- (3)  $\phi$  ist surjektiv.

*Beweis.* Es ist klar, dass aus (1) die beiden Aussagen (2) und (3) folgen. Sei  $(b_1, \dots, b_n)$  eine Basis von  $V$ . Nach Satz 9.12 ist  $\phi$  genau dann injektiv, wenn  $\phi(b_1), \dots, \phi(b_n)$  linear unabhängig sind.  $n$  linear unabhängige Vektoren bilden aber eine Basis (wegen  $\dim W = n$ , siehe Satz 8.23); wiederum nach Satz 9.12 ist  $\phi$  dann ein Isomorphismus. Analog ist  $\phi$  genau dann surjektiv, wenn  $\phi(b_1), \dots, \phi(b_n)$  den Vektorraum  $W$  erzeugen. Ein Erzeugendensystem aus  $n$  Elementen ist aber wieder eine Basis, also ist  $\phi$  dann ein Isomorphismus.  $\square$

Diese Folgerung besagt, dass sich endlich-dimensionale Vektorräume bezüglich linearer Abbildungen so verhalten wie endliche Mengen bezüglich beliebiger Abbildungen. Es gilt nämlich (wie leicht einzusehen ist):

Seien  $X$  und  $Y$  zwei endliche Mengen mit  $\#X = \#Y = n$  und sei  $f : X \rightarrow Y$  eine Abbildung. Dann sind die folgenden Aussagen äquivalent:

- (1)  $f$  ist bijektiv.
- (2)  $f$  ist injektiv.
- (3)  $f$  ist surjektiv.

**9.15. Beispiel.** Der Vektorraum  $P_{<n}$  der Polynomfunktionen vom Grad  $< n$  wird von den  $n$  linear unabhängigen Potenzfunktionen  $x \mapsto x^j$  für  $j \in \{0, 1, \dots, n-1\}$  erzeugt, hat also Dimension  $n$ .

**BSP**  
Interpolation

Seien  $x_1, x_2, \dots, x_n \in \mathbb{R}$  paarweise verschieden. Wir definieren für  $j \in \{1, 2, \dots, n\}$  die Polynomfunktion  $p_j \in P_{<n}$  durch

$$p_j(x) = \prod_{i \in \{1, \dots, n\} \setminus \{j\}} \frac{x - x_i}{x_j - x_i} = \frac{x - x_1}{x_j - x_1} \cdots \frac{x - x_{j-1}}{x_j - x_{j-1}} \cdot \frac{x - x_{j+1}}{x_j - x_{j+1}} \cdots \frac{x - x_n}{x_j - x_n}.$$

Wir definieren außerdem

$$\phi : P_{<n} \longrightarrow \mathbb{R}^n, \quad f \longmapsto (f(x_1), f(x_2), \dots, f(x_n))$$

( $\phi$  ist aus Auswertungsabbildungen zusammengesetzt und daher linear) und eine lineare Abbildung  $\psi : \mathbb{R}^n \rightarrow P_{<n}$  durch Festlegung der Bilder der Standardbasis:

$$\psi(\mathbf{e}_j) = p_j \quad \text{für alle } j \in \{1, 2, \dots, n\}.$$

( $\psi$  ist gerade die Linearkombinationenabbildung  $\phi_{p_1, \dots, p_n}$ .) Dann gilt  $\phi \circ \psi = \text{id}_{\mathbb{R}^n}$ :

$$\phi(\psi(\mathbf{e}_j)) = \phi(p_j) = (p_j(x_1), \dots, p_j(x_n)) = \mathbf{e}_j,$$

denn  $p_j(x_i) = 0$  für  $i \neq j$  und  $p_j(x_j) = 1$ .  $\phi \circ \psi$  und die identische Abbildung stimmen auf einer Basis überein, also sind sie gleich. Dann muss  $\psi$  injektiv sein und  $\phi$  surjektiv. Nach Folgerung 9.14 sind wegen  $\dim \mathbb{R}^n = n = \dim P_{<n}$  beide Abbildungen (zueinander inverse) Isomorphismen. Das bedeutet zum Beispiel:

Seien  $x_1, x_2, \dots, x_n \in \mathbb{R}$  paarweise verschieden und  $y_1, y_2, \dots, y_n \in \mathbb{R}$ . Dann gibt es **genau eine** Polynomfunktion  $f \in P_{<n}$  mit  $f(x_j) = y_j$  für alle  $j \in \{1, 2, \dots, n\}$ , nämlich  $f = y_1 p_1 + y_2 p_2 + \dots + y_n p_n$ .

Das sieht man so: Die Bedingung an  $f$  bedeutet  $f \in P_{<n}$  und  $\phi(f) = (y_1, y_2, \dots, y_n)$ . Letzteres ist aber äquivalent zu

$$f = \psi(y_1, y_2, \dots, y_n) = y_1 p_1 + y_2 p_2 + \dots + y_n p_n.$$

Diese Formel für das Interpolationspolynom heißt Lagrangesche Interpolationsformel. ♣

9.16. **Beispiele.** Wir bleiben bei den Polynomfunktionen und geben weitere Beispiele für lineare Abbildungen.

**BSP**  
lin. Abb. auf  
Polynomfkt.

- (1) Wir haben schon in Beispiel 9.10 gesehen, dass für  $a \in \mathbb{R}$  die Auswertungsabbildung

$$\text{ev}_a : P \longrightarrow \mathbb{R}, \quad f \longmapsto f(a)$$

linear ist.

- (2) Die *Differentiation* von Polynomfunktionen ist linear:

$$D : P \longrightarrow P, \quad f \longmapsto f'.$$

Für  $f(x) = a_0 + a_1 x + \dots + a_n x^n$  gilt dabei  $f'(x) = a_1 + 2a_2 x + \dots + n a_n x^{n-1}$ . Man könnte  $D$  also definieren als diejenige lineare Abbildung, die für  $n > 0$  die Potenzfunktion  $f_n : x \mapsto x^n$  auf  $n f_{n-1}$  abbildet und  $f_0$  auf die Nullfunktion.

$D$  ist surjektiv (also ein Epimorphismus) und der Kern von  $D$  besteht genau aus den konstanten Funktionen. (An diesem Beispiel kann man sehen, dass die Aussage von Satz 9.12 für unendlich-dimensionale Vektorräume nicht gelten muss.)

- (3) Die Berechnung des *bestimmten Integrals* von  $a$  bis  $b$  ist linear:

$$I_{a,b} : P \longrightarrow \mathbb{R}, \quad f \longmapsto \int_a^b f(x) dx.$$

Für die Potenzfunktionen  $f_n$  gilt  $I_{a,b}(f_n) = \frac{b^{n+1} - a^{n+1}}{n+1}$ .

- (4) Auch die *unbestimmte Integration* mit Anfangspunkt  $a \in \mathbb{R}$  ist linear:

$$I_a : P \longrightarrow P, \quad f \longmapsto \left( x \mapsto \int_a^x f(t) dt \right).$$

Das ist die lineare Abbildung mit  $I_a(f_n) = \left( x \mapsto \frac{1}{n+1} (x^{n+1} - a^{n+1}) \right)$ . Die Abbildung  $I_a$  ist injektiv, aber nicht surjektiv: ihr Bild ist gerade der Kern von  $\text{ev}_a$  (die Integralfunktionen verschwinden alle an der Stelle  $a$ ).

- (5) Die *Translation* (also Verschiebung) um  $a \in \mathbb{R}$  ist linear:

$$T_a : P \longrightarrow P, \quad f \longmapsto (x \mapsto f(x - a)).$$

$T_a$  ist ein Automorphismus von  $P$ , der inverse Automorphismus ist  $T_{-a}$ .

Zwischen diesen Abbildungen bestehen eine Reihe von Relationen, wie zum Beispiel

$$\begin{aligned} \text{ev}_b \circ I_a &= I_{a,b}, & D \circ I_a &= \text{id}_P, & (I_a \circ D)(f) &= f - \text{ev}_a(f) f_0, \\ T_a \circ D &= D \circ T_a, & T_a \circ T_b &= T_{a+b}, & I_a \circ T_b &= T_b \circ I_{a-b}, \\ I_{a,b} \circ T_c &= I_{a-c,b-c}, & \text{ev}_a \circ T_b &= \text{ev}_{a-b}. \end{aligned}$$

Man kann sie leicht auf den Potenzfunktionen nachprüfen (das genügt, weil die Potenzfunktionen eine Basis von  $P$  sind).

In der Analysis werden Sie lernen, dass Differentiation und Integration ganz allgemein lineare Abbildungen sind. ♣

Kern und Bild einer linearen Abbildung sind wichtige Daten. Für die Dimension des Bildes gibt es sogar einen eigenen Namen.

\* 9.17. **Definition.** Ist  $\phi : V \rightarrow W$  eine lineare Abbildung, dann heißt

$$\text{rk}(\phi) = \dim \text{im}(\phi)$$

der *Rang* von  $\phi$ . ◇

**DEF**  
Rang einer  
linearen Abb.

Zwischen dem Rang und der Dimension des Kerns besteht ein einfacher Zusammenhang.

\* 9.18. **Satz.** Sei  $\phi : V \rightarrow W$  eine lineare Abbildung. Dann gilt

$$\dim \ker(\phi) + \text{rk}(\phi) = \dim V.$$

**SATZ**  
 $\dim(\ker)$   
und Rang

Dabei sei  $n + \infty = \infty + n = \infty + \infty = \infty$  für  $n \in \mathbb{N}$ .

*Beweis.* Ist  $\dim \ker(\phi) = \infty$ , dann muss auch  $\dim V = \infty$  sein, denn  $\ker(\phi)$  ist ein Untervektorraum von  $V$  (siehe Satz 8.28). Also ist die Behauptung in diesem Fall richtig. Ist  $\text{rk}(\phi) = \infty$ , dann können wir unendlich viele linear unabhängige Vektoren  $w_j \in \text{im}(\phi)$  finden ( $j \in \mathbb{N}$ ). Sei  $v_j \in V$  ein Urbild von  $w_j$ ; dann sind auch die  $v_j$  linear unabhängig. Denn sei  $\lambda_0 v_0 + \lambda_1 v_1 + \dots + \lambda_n v_n = \mathbf{0}$ , dann folgt durch Anwenden von  $\phi$  auch

$$\begin{aligned} \lambda_0 w_0 + \lambda_1 w_1 + \dots + \lambda_n w_n &= \lambda_0 \phi(v_0) + \lambda_1 \phi(v_1) + \dots + \lambda_n \phi(v_n) \\ &= \phi(\lambda_0 v_0 + \lambda_1 v_1 + \dots + \lambda_n v_n) = \mathbf{0}. \end{aligned}$$

Weil  $w_0, w_1, \dots, w_n$  linear unabhängig sind, müssen alle Koeffizienten  $\lambda_j$  null sein, was zeigt, dass  $v_0, v_1, \dots, v_n$  linear unabhängig sind. Es gibt also unendlich viele linear unabhängige Vektoren in  $V$ ; damit ist  $\dim V = \infty$  und die Behauptung des Satzes stimmt. Wir können also jetzt annehmen, dass  $\dim \ker(\phi)$  und  $\text{rk}(\phi)$  beide endlich sind.

Seien  $k = \dim \ker(\phi)$ ,  $r = \text{rk}(\phi)$  und  $n = \dim V$ . Wir können eine Basis  $(b_1, \dots, b_k)$  von  $\ker(\phi)$  wählen und sie zu einer Basis  $(b_1, \dots, b_k, b_{k+1}, \dots, b_n)$  von  $V$  ergänzen (Basisergänzungssatz 8.14). Ich behaupte jetzt, dass  $(\phi(b_{k+1}), \dots, \phi(b_n))$  eine Basis von  $\text{im}(\phi)$  ist. Daraus folgt  $n - k = r$ , also die Behauptung des Satzes.

- $\langle \phi(b_{k+1}), \dots, \phi(b_n) \rangle = \text{im}(\phi)$ :  
„ $\subset$ “ ist klar. „ $\supset$ “: Sei  $w \in \text{im}(\phi)$ . Dann gibt es  $v \in V$  mit  $\phi(v) = w$ . Wir schreiben  $v$  als Linearkombination unserer Basisvektoren:

$$v = \lambda_1 b_1 + \dots + \lambda_k b_k + \lambda_{k+1} b_{k+1} + \dots + \lambda_n b_n.$$

Dann ist

$$\begin{aligned} w &= \phi(v) \\ &= \phi(\lambda_1 b_1 + \dots + \lambda_k b_k + \lambda_{k+1} b_{k+1} + \dots + \lambda_n b_n) \\ &= \lambda_1 \phi(b_1) + \dots + \lambda_k \phi(b_k) + \lambda_{k+1} \phi(b_{k+1}) + \dots + \lambda_n \phi(b_n) \\ &= \lambda_{k+1} \phi(b_{k+1}) + \dots + \lambda_n \phi(b_n) \\ &\in \langle \phi(b_{k+1}), \dots, \phi(b_n) \rangle. \end{aligned}$$

(Denn  $\phi(b_j) = 0$  für  $j \in \{1, 2, \dots, k\}$  wegen  $b_j \in \ker(\phi)$ .)

- $\phi(b_{k+1}), \dots, \phi(b_n)$  sind linear unabhängig:  
Seien  $\lambda_{k+1}, \dots, \lambda_n$  Skalare mit

$$\lambda_{k+1}\phi(b_{k+1}) + \dots + \lambda_n\phi(b_n) = \mathbf{0}.$$

Dann ist

$$\phi(\lambda_{k+1}b_{k+1} + \dots + \lambda_nb_n) = \mathbf{0},$$

also ist  $\lambda_{k+1}b_{k+1} + \dots + \lambda_nb_n \in \ker(\phi)$  und damit eine Linearkombination von  $b_1, \dots, b_k$ . Wir können das schreiben als

$$\lambda_1b_1 + \dots + \lambda_kb_k + \lambda_{k+1}b_{k+1} + \dots + \lambda_nb_n = \mathbf{0}.$$

Weil  $b_1, \dots, b_n$  linear unabhängig sind, folgt daraus  $\lambda_1 = \dots = \lambda_n = 0$ , also war auch unsere ursprüngliche Linearkombination trivial.  $\square$

**9.19. Beispiel.** Die auf  $P_{<n}$  eingeschränkte Differentiation  $D_n : P_{<n} \rightarrow P_{<n}$  hat (für  $n \geq 1$ ) als Kern die konstanten Funktionen, also ist  $\dim \ker(D_n) = 1$ . Da  $\dim P_{<n} = n$ , folgt  $\text{rk}(D_n) = n - 1$ .  $\clubsuit$

**BSP**

In vielen Fällen ist es einfacher, den Kern und seine Dimension direkt zu bestimmen als den Rang. Mit Satz 9.18 kann man daraus dann den Rang berechnen.

Die Konstruktion des Vektorraums  $K^X = \text{Abb}(X, K)$  lässt sich verallgemeinern.

**9.20. Definition.** Sei  $K$  ein Körper und  $V$  ein  $K$ -Vektorraum. Sei weiter  $X$  eine Menge. Dann können wir auf  $V^X = \text{Abb}(X, V)$  eine Struktur als  $K$ -Vektorraum definieren durch

**DEF**  
Vektorraum  
 $V^X$ 

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

Der Beweis ist analog zu dem für  $K^X$ .

Für  $X = \{1, 2, \dots, n\}$  identifizieren wir  $V^X$  mit  $V^n$ .  $\diamond$

Wir können also insbesondere zwei lineare Abbildungen  $V \rightarrow W$  addieren oder eine solche Abbildung mit einem Skalar multiplizieren.

**9.21. Satz.** Seien  $V$  und  $W$  zwei  $K$ -Vektorräume. Die Menge der linearen Abbildungen  $V \rightarrow W$  bildet einen  $K$ -Untervektorraum von  $\text{Abb}(V, W)$ .

**SATZ**  
Vektorraum  
der lin. Abb.

**9.22. Definition.** Dieser Vektorraum wird mit  $\text{Hom}(V, W)$  (oder  $\text{Hom}_K(V, W)$ ) bezeichnet. Im Fall  $V = W$  schreiben wir auch  $\text{End}(V) = \text{Hom}(V, V)$  (oder auch  $\text{End}_K(V)$ ) für den Vektorraum der Endomorphismen von  $V$ .  $\diamond$

**DEF**  
 $\text{Hom}(V, W)$   
 $\text{End}(V)$ 

*Beweis.* Wir müssen die Bedingungen für einen Untervektorraum nachprüfen.

- Die Nullabbildung ist linear, also ist  $\mathbf{0} \in \text{Hom}(V, W)$ .
- Seien  $\phi, \psi : V \rightarrow W$  linear. Wir müssen zeigen, dass  $\phi + \psi$  ebenfalls linear ist. Seien dazu  $v, v' \in V$ ,  $\lambda \in K$ . Dann haben wir

$$\begin{aligned} (\phi + \psi)(v + v') &= \phi(v + v') + \psi(v + v') = \phi(v) + \phi(v') + \psi(v) + \psi(v') \\ &= \phi(v) + \psi(v) + \phi(v') + \psi(v') = (\phi + \psi)(v) + (\phi + \psi)(v') \end{aligned}$$

und

$$\begin{aligned} (\phi + \psi)(\lambda v) &= \phi(\lambda v) + \psi(\lambda v) = \lambda\phi(v) + \lambda\psi(v) \\ &= \lambda(\phi(v) + \psi(v)) = \lambda(\phi + \psi)(v). \end{aligned}$$

- Sei  $\phi : V \rightarrow W$  linear und  $\lambda \in K$ , Wir müssen zeigen, dass  $\lambda\phi$  ebenfalls linear ist. Seien dazu  $v, v' \in V$ ,  $\mu \in K$ . Dann haben wir

$$\begin{aligned}(\lambda\phi)(v + v') &= \lambda\phi(v + v') = \lambda(\phi(v) + \phi(v')) \\ &= \lambda\phi(v) + \lambda\phi(v') = (\lambda\phi)(v) + (\lambda\phi)(v')\end{aligned}$$

und

$$\begin{aligned}(\lambda\phi)(\mu v) &= \lambda\phi(\mu v) = \lambda \cdot \mu\phi(v) \\ &= \mu \cdot \lambda\phi(v) = \mu(\lambda\phi)(v).\end{aligned}\quad \square$$

**9.23. Satz.** Seien  $V$  und  $W$  zwei  $K$ -Vektorräume mit  $\dim V = n < \infty$ . Sei weiter  $(b_1, b_2, \dots, b_n)$  eine Basis von  $V$ . Dann ist

$$\Phi : \text{Hom}(V, W) \longrightarrow W^n, \quad \phi \longmapsto (\phi(b_1), \phi(b_2), \dots, \phi(b_n))$$

ein Isomorphismus. Insbesondere ist im Fall von  $\dim W = m < \infty$

$$\dim \text{Hom}(V, W) = \dim W^n = n \dim W = mn = (\dim V)(\dim W).$$

*Beweis.* Es ist klar, dass  $\Phi$  linear ist (denn  $\Phi$  setzt sich aus Auswertungsabbildungen zusammen). Nach Satz 9.11 gibt es zu jeder Wahl der Bilder von  $b_1, \dots, b_n$  in  $W$  genau eine lineare Abbildung; das bedeutet, dass  $\Phi$  bijektiv ist. Isomorphe Vektorräume haben dieselbe Dimension; der Beweis von  $\dim W^n = n \dim W$  ist eine Übungsaufgabe.  $\square$

**9.24. Folgerung.** Ist  $(b_1, b_2, \dots, b_n)$  eine Basis von  $V$  und  $(b'_1, b'_2, \dots, b'_m)$  eine Basis von  $W$ , dann bilden die linearen Abbildungen  $\phi_{ij} : V \rightarrow W$  für  $i \in \{1, 2, \dots, m\}$  und  $j \in \{1, 2, \dots, n\}$  eine Basis von  $\text{Hom}(V, W)$ , wobei  $\phi_{ij}(b_k) = \mathbf{0}$  für  $k \neq j$  und  $\phi_{ij}(b_j) = b'_i$ .

**SATZ**  
 $\text{Hom}(V, W)$   
 $\cong W^{\dim V}$

**FOLG**  
Basis von  
 $\text{Hom}(V, W)$

*Beweis.* Nach Satz 9.11 existieren eindeutig bestimmte  $\phi_{ij}$  wie angegeben. Wir zeigen, dass die  $\phi_{ij} \in \text{Hom}(V, W)$  linear unabhängig sind. Seien dazu  $\lambda_{ij}$  Skalare mit

$$\sum_{i=1}^m \sum_{j=1}^n \lambda_{ij} \phi_{ij} = \mathbf{0}.$$

Sei  $k \in \{1, 2, \dots, n\}$ . Einsetzen von  $b_k$  liefert dann

$$\mathbf{0} = \left( \sum_{i=1}^m \sum_{j=1}^n \lambda_{ij} \phi_{ij} \right) (b_k) = \sum_{i=1}^m \sum_{j=1}^n \lambda_{ij} \phi_{ij}(b_k) = \sum_{i=1}^m \lambda_{ik} b'_i$$

Da die  $b'_i$  linear unabhängig sind, folgt  $\lambda_{ik} = 0$  für alle  $i$ . Da  $k$  beliebig war, sind also alle  $\lambda_{ij} = 0$ , was zu zeigen war. Nach Satz 9.23 ist  $\dim \text{Hom}(V, W) = nm$  gleich der Anzahl der linear unabhängigen Elemente  $\phi_{ij} \in \text{Hom}(V, W)$ , nach Satz 8.23 sind die  $\phi_{ij}$  dann bereits eine Basis von  $\text{Hom}(V, W)$ .  $\square$

Im Fall  $V = K^n$ ,  $W = K^m$  mit den Standardbasen kann man das, was  $\phi_{ij}$  bewirkt, so beschreiben: Man nimmt die  $j$ -te Komponente von  $(x_1, x_2, \dots, x_n) \in K^n$  und steckt sie in die  $i$ -te Komponente des Resultats in  $K^m$ ; die übrigen Komponenten sind null.

9.25. **Beispiel.** Als einfaches Beispiel betrachten wir  $V = \mathbb{R}^3$ ,  $W = \mathbb{R}^2$ , jeweils mit der Standardbasis  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  bzw.  $(\mathbf{e}'_1, \mathbf{e}'_2)$ . Die Basis von  $\text{Hom}(\mathbb{R}^3, \mathbb{R}^2)$  aus Folgerung 9.24 sieht in diesem Fall so aus:

**BSP**  
Basis von  
 $\text{Hom}(\mathbb{R}^3, \mathbb{R}^2)$

$$\phi_{11} : (x, y, z) \mapsto (x, 0)$$

$$\phi_{12} : (x, y, z) \mapsto (y, 0)$$

$$\phi_{13} : (x, y, z) \mapsto (z, 0)$$

$$\phi_{21} : (x, y, z) \mapsto (0, x)$$

$$\phi_{22} : (x, y, z) \mapsto (0, y)$$

$$\phi_{23} : (x, y, z) \mapsto (0, z)$$

Jede lineare Abbildung  $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  lässt sich als Linearkombination dieser sechs Abbildungen schreiben; es gibt also  $a, b, c, d, e, f \in \mathbb{R}$ , sodass

$$\phi = a\phi_{11} + b\phi_{12} + c\phi_{13} + d\phi_{21} + e\phi_{22} + f\phi_{23},$$

also

$$\phi(x, y, z) = (ax + by + cz, dx + ey + fz).$$



Die Endomorphismen eines Vektorraums  $V$  bilden sogar einen Ring, den *Endomorphismenring* von  $V$ :

9.26. **Satz.** Sei  $V$  ein  $K$ -Vektorraum. Dann ist  $\text{End}(V)$  ein Ring mit der Addition des  $K$ -Vektorraums  $\text{End}(V) = \text{Hom}(V, V)$  und der Komposition von Abbildungen als Multiplikation; das Einselement ist die identische Abbildung  $\text{id}_V$ .

**SATZ**  
 $\text{End}(V)$  ist  
ein Ring

*Beweis.* Die Vektorraum-Axiome, die in  $\text{End}(V)$  gelten, liefern uns die Ring-Axiome für die Addition. Es bleibt noch zu zeigen, dass die Multiplikation assoziativ ist mit Einselement  $\text{id}_V$  und dass die beiden Ring-Distributivgesetze gelten. Seien also  $f, g, h \in \text{End}(V)$ . Die Assoziativität  $(f \circ g) \circ h = f \circ (g \circ h)$  gilt für Abbildungen ganz allgemein, ebenso wie  $\text{id}_V \circ f = f = f \circ \text{id}_V$ . Zum Nachweis der Distributivgesetze rechnen wir für  $v \in V$ :

$$\begin{aligned} ((f + g) \circ h)(v) &= (f + g)(h(v)) = f(h(v)) + g(h(v)) \\ &= (f \circ h)(v) + (g \circ h)(v) = (f \circ h + g \circ h)(v), \end{aligned}$$

also ist  $(f + g) \circ h = f \circ h + g \circ h$ , und

$$\begin{aligned} (f \circ (g + h))(v) &= f((g + h)(v)) = f(g(v) + h(v)) \\ &= f(g(v)) + f(h(v)) = (f \circ g)(v) + (f \circ h)(v) \\ &= (f \circ g + f \circ h)(v), \end{aligned}$$

also ist  $f \circ (g + h) = f \circ g + f \circ h$  (dabei haben wir verwendet, dass  $f$  linear ist).  $\square$

Der Endomorphismenring ist nicht kommutativ, wenn  $\dim V \geq 2$  ist (Übung!). Für  $\dim V = 1$  ist  $\text{End}(V) = K$ , da alle Endomorphismen durch Multiplikation mit Skalaren gegeben sind; für  $\dim V = 0$  ist  $\text{End}(V)$  der Nullring.

Die Automorphismen von  $V$  bilden eine Gruppe, die *Automorphismengruppe*  $\text{Aut}(V)$  von  $V$  (das ist auch die Gruppe der invertierbaren Elemente des Rings  $\text{End}(V)$ ).

## 10. MATRIZEN

Die Ergebnisse des letzten Abschnitts zeigen uns, dass wir lineare Abbildungen zwischen zwei endlich-dimensionalen  $K$ -Vektorräumen  $V$  und  $W$  der Dimensionen  $n$  und  $m$  durch  $mn$  Koeffizienten aus  $K$  beschreiben können. Dazu müssen wir Basen von  $V$  und  $W$  wählen; daraus bekommen wir eine Basis von  $\text{Hom}(V, W)$  wie in Folgerung 9.24 und die gesuchten Koeffizienten sind dann die Koeffizienten in der Darstellung der gegebenen linearen Abbildung als Linearkombination bezüglich dieser Basis. Für diese Koeffizienten führt man eine spezielle Form der Darstellung ein.

**10.1. Definition.** Sei  $K$  ein Körper und seien  $m, n \in \mathbb{N}$ . Eine  $m \times n$ -Matrix mit Einträgen aus  $K$  (oder kurz über  $K$ ) ist ein rechteckiges Schema aus  $mn$  Elementen von  $K$ , das wie folgt notiert wird:

**DEF**  
Matrix

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

Zur Abkürzung schreiben wir auch  $(a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$  (oder auch  $(a_{ij})_{i,j}$ , falls die Zahlen  $m$  und  $n$  aus dem Kontext klar sind) für diese Matrix. Im Fall  $m = n$  heißt die Matrix *quadratisch*. Für  $i \in \{1, 2, \dots, m\}$  heißt das  $n$ -Tupel  $(a_{i1}, a_{i2}, \dots, a_{in})$  die  $i$ -te Zeile der Matrix, für  $j \in \{1, 2, \dots, n\}$  heißt das  $m$ -Tupel  $(a_{1j}, a_{2j}, \dots, a_{mj})$  die  $j$ -te Spalte der Matrix.

Wir schreiben  $\text{Mat}(m \times n, K)$  für die Menge aller  $m \times n$ -Matrizen mit Einträgen aus  $K$ ; im Fall  $m = n$  auch kürzer  $\text{Mat}(n, K)$  für  $\text{Mat}(n \times n, K)$ .  $\diamond$

Im Grunde ist eine  $m \times n$ -Matrix über  $K$  nichts anderes als eine Familie von Elementen von  $K$  mit der Indexmenge  $\{1, 2, \dots, m\} \times \{1, 2, \dots, n\}$ , also

$$\text{Mat}(m \times n, K) = K^{\{1,2,\dots,m\} \times \{1,2,\dots,n\}}.$$

Da wir auf beliebigen Mengen der Form  $K^I$  eine Struktur als  $K$ -Vektorraum definiert haben, folgt sofort:

**10.2. Lemma.** Sei  $K$  ein Körper. Die Menge  $\text{Mat}(m \times n, K)$  mit komponentenweise definierter Addition und Skalarmultiplikation ist ein  $K$ -Vektorraum der Dimension  $mn$ .

**LEMMA**  
Vektorraum  
der  $m \times n$ -  
Matrizen

Ist  $m = 0$  oder  $n = 0$  (oder beides), dann ist  $\text{Mat}(m \times n, K)$  ein Null-Vektorraum; sein einziges Element ist eine leere Matrix (mit null Zeilen und  $n$  Spalten oder mit  $m$  Zeilen und null Spalten).

Matrizen (mit der gleichen Anzahl an Zeilen und Spalten) werden also wie folgt addiert und mit Skalaren multipliziert:

$$\begin{aligned} & \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix} \end{aligned}$$

und

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

Wie zu Beginn dieses Abschnitts beschrieben, können wir linearen Abbildungen Matrizen zuordnen. Wir betrachten zunächst  $V = K^n$  und  $W = K^m$  mit den Standardbasen  $B = (\mathbf{e}_1, \dots, \mathbf{e}_n)$  von  $V$  und  $B' = (\mathbf{e}'_1, \dots, \mathbf{e}'_m)$  von  $W$  (wir schreiben hier  $\mathbf{e}'_i$  für den  $i$ -ten Standard-Basisvektor in  $K^m$  zur Unterscheidung von den Basisvektoren  $\mathbf{e}_j$  in  $K^n$ ). Wir haben dann die Basis  $(\phi_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$  von  $\text{Hom}(K^n, K^m)$  wie in Folgerung 9.24 mit  $\phi_{ij}(\mathbf{e}_k) = \mathbf{0}$  für  $k \neq j$  und  $\phi_{ij}(\mathbf{e}_j) = \mathbf{e}'_i$ . Ist  $\phi : K^n \rightarrow K^m$  eine lineare Abbildung, dann schreiben wir  $\phi$  als Linearkombination

$$\phi = \sum_{i=1}^m \sum_{j=1}^n a_{ij} \phi_{ij} \quad \text{mit } a_{ij} \in K.$$

Die zugehörige Matrix ist dann  $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ .

**10.3. Beispiel.** Wie wir gesehen haben, hat eine lineare Abbildung  $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  die Form  $\phi(x, y, z) = (ax + by + cz, dx + ey + fz)$  mit geeigneten  $a, b, c, d, e, f \in \mathbb{R}$ . Dann ist  $\phi = a\phi_{11} + b\phi_{12} + c\phi_{13} + d\phi_{21} + e\phi_{22} + f\phi_{23}$  (vergleiche Beispiel 9.25), also ist die zugehörige Matrix

**BSP**  
Matrix für  
 $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$

$$A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.$$



Die  $j$ -te Spalte der zu  $\phi : K^n \rightarrow K^m$  gehörigen  $m \times n$ -Matrix enthält gerade die Koeffizienten des Bildes des  $j$ -ten Standard-Basisvektors  $\mathbf{e}_j$ , denn

$$\phi(\mathbf{e}_j) = \sum_{i=1}^m \sum_{k=1}^n a_{ik} \phi_{ik}(\mathbf{e}_j) = \sum_{i=1}^m a_{ij} \mathbf{e}'_i = (a_{1j}, a_{2j}, \dots, a_{mj})$$

ähnlich wie im Beweis von Folgerung 9.24.

**10.4. Lemma.** Die oben beschriebene Zuordnung definiert einen Isomorphismus  $\text{Hom}(K^n, K^m) \rightarrow \text{Mat}(m \times n, K)$ . Wenn man  $\text{Mat}(m \times n, K)$  mit  $K^{\{1,2,\dots,m\} \times \{1,2,\dots,n\}}$  identifiziert, dann ist dieser Isomorphismus invers zu der Linearkombinationenabbildung  $K^{\{1,2,\dots,m\} \times \{1,2,\dots,n\}} \rightarrow \text{Hom}(K^n, K^m)$ , die zur Basis  $(\phi_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$  von  $\text{Hom}(K^n, K^m)$  gehört.

**LEMMA**  
 $\text{Mat}(m \times n, K)$   
 $\cong$   
 $\text{Hom}(K^n, K^m)$

*Beweis.* Die erwähnte Linearkombinationenabbildung

$$\Phi : \text{Mat}(m \times n, K) = K^{\{1,2,\dots,m\} \times \{1,2,\dots,n\}} \rightarrow \text{Hom}(K^n, K^m)$$

bildet eine Matrix  $(a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$  auf die Linearkombination  $\sum_{i,j} a_{ij} \phi_{ij}$  ab; sie ist ein Isomorphismus, da sie eine Basis auf eine Basis abbildet. Die Abbildung, die einer linearen Abbildung  $\phi : K^n \rightarrow K^m$  ihre Matrix zuordnet, ist offenbar die Umkehrabbildung von  $\Phi$ , insbesondere also ebenfalls ein Isomorphismus.  $\square$

Wie stellt sich die Anwendung der linearen Abbildung  $\phi : K^n \rightarrow K^m$  dar, wenn wir die zugehörige Matrix  $A = (a_{ij})_{i,j}$  verwenden? Es gilt

$$\phi(x_1, x_2, \dots, x_n) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} \phi_{ij}(x_1, x_2, \dots, x_n) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j \mathbf{e}'_i,$$

also ist die  $i$ -te Komponente von  $\phi(x_1, x_2, \dots, x_n)$  gegeben durch

$$\sum_{j=1}^n a_{ij}x_j = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n.$$

Man schreibt das dann gerne als Multiplikation der Matrix  $A$  mit dem  $(x_1, \dots, x_n)$  entsprechenden *Spaltenvektor*: Man identifiziert also  $K^n$  mit  $\text{Mat}(n \times 1, K)$  und  $K^m$  mit  $\text{Mat}(m \times 1, K)$ . Dann haben wir für das Resultat der Anwendung von  $\phi$ :

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}.$$

Das Ergebnis ist wieder ein Spaltenvektor, diesmal der Länge  $m$ . Seine  $i$ -te Komponente ergibt sich aus der  $i$ -ten Zeile der Matrix und dem Spaltenvektor zu  $(x_1, \dots, x_n)$  als das *Skalarprodukt*

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n.$$

(Das Skalarprodukt heißt so, weil sein Wert ein Skalar ist: „Vektor mal Vektor = Skalar“. Man beachte den Unterschied zur Skalarmultiplikation „Skalar mal Vektor = Vektor“!)

**10.5. Beispiele.**  $2 \times 3$ -Matrizen mit Einträgen in  $\mathbb{R}$  entsprechen linearen Abbildungen  $\mathbb{R}^3 \rightarrow \mathbb{R}^2$ . In diesem Fall sieht obige Formel so aus:

**BSP**

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ax + by + cz \\ dx + ey + fz \end{pmatrix}$$

$3 \times 2$ -Matrizen über  $\mathbb{R}$  entsprechen linearen Abbildungen  $\mathbb{R}^2 \rightarrow \mathbb{R}^3$ . Dann haben wir:

$$\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \\ ex + fy \end{pmatrix} \quad \clubsuit$$

Die Komposition von linearen Abbildungen entspricht der Multiplikation von Matrizen.

**10.6. Definition.** Sei  $K$  ein Körper und seien  $l, m, n \in \mathbb{N}$ . Für Matrizen  $A \in \text{Mat}(l \times m, K)$  und  $B \in \text{Mat}(m \times n, K)$  ist das *Produkt*  $AB = A \cdot B \in \text{Mat}(l \times n, K)$  definiert als die zu  $f \circ g$  gehörende Matrix, wobei  $f : K^m \rightarrow K^l$  und  $g : K^n \rightarrow K^m$  die den Matrizen  $A$  und  $B$  entsprechenden linearen Abbildungen sind.  $\diamond$

**DEF**  
Matrix-  
multiplikation

So wie man Abbildungen nur dann miteinander verknüpfen kann, wenn der Wertebereich der einen Abbildung mit dem Definitionsbereich der anderen übereinstimmt, kann man Matrizen nur dann miteinander multiplizieren, wenn sie in der Größe „zueinander passen“, wenn also die Spaltenanzahl des linken Faktors gleich der Zeilenanzahl des rechten Faktors ist.

Wie sieht diese Matrixmultiplikation konkret aus? Seien  $A = (a_{ij})_{1 \leq i \leq l, 1 \leq j \leq m}$ ,  $B = (b_{jk})_{1 \leq j \leq m, 1 \leq k \leq n}$  und  $C = (c_{ik})_{1 \leq i \leq l, 1 \leq k \leq n} = AB$ . Dann sollte  $c_{ik}$  die  $i$ -te Komponente von  $f(g(\mathbf{e}_k))$  sein. Es ist

$$f(g(\mathbf{e}_k)) = f(b_{1k}\mathbf{e}'_1 + b_{2k}\mathbf{e}'_2 + \dots + b_{mk}\mathbf{e}'_m) = b_{1k}f(\mathbf{e}'_1) + b_{2k}f(\mathbf{e}'_2) + \dots + b_{mk}f(\mathbf{e}'_m)$$

und die  $i$ -te Komponente von  $f(\mathbf{e}'_j)$  ist  $a_{ij}$ . Also ist

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{im}b_{mk} = \sum_{j=1}^m a_{ij}b_{jk}$$

das **Skalarprodukt der  $i$ -ten Zeile von  $A$  mit der  $k$ -ten Spalte von  $B$ .**

Die oben eingeführte Multiplikation „Matrix mal Spaltenvektor“ ist dann also ein Spezialfall dieser allgemeinen Matrixmultiplikation.

**10.7. Beispiel.** Wir berechnen das Produkt zweier Matrizen über  $\mathbb{R}$ :

**BSP**

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 3 + 3 \cdot 5 & 1 \cdot 2 + 2 \cdot 4 + 3 \cdot 6 \\ 4 \cdot 1 + 5 \cdot 3 + 6 \cdot 5 & 4 \cdot 2 + 5 \cdot 4 + 6 \cdot 6 \end{pmatrix} = \begin{pmatrix} 22 & 28 \\ 49 & 64 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 4 & 1 \cdot 2 + 2 \cdot 5 & 1 \cdot 3 + 2 \cdot 6 \\ 3 \cdot 1 + 4 \cdot 4 & 3 \cdot 2 + 4 \cdot 5 & 3 \cdot 3 + 4 \cdot 6 \\ 5 \cdot 1 + 6 \cdot 4 & 5 \cdot 2 + 6 \cdot 5 & 5 \cdot 3 + 6 \cdot 6 \end{pmatrix} = \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \\ 29 & 40 & 51 \end{pmatrix}$$



Zur identischen Abbildung gehört eine spezielle Matrix.

**10.8. Definition.** Sei  $K$  ein Körper und  $n \in \mathbb{N}$ . Die Matrix  $I_n \in \text{Mat}(n, K)$ , die der identischen Abbildung  $\text{id}_{K^n}$  entspricht, heißt die *Einheitsmatrix (der Größe  $n$  über  $K$ )*.

**DEF**  
Einheits-  
matrix

In der  $j$ -ten Spalte von  $I_n$  muss der  $j$ -te Standard-Basisvektor stehen, also sieht  $I_n$  so aus:

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Man schreibt das auch  $I_n = (\delta_{ij})_{1 \leq i, j \leq n}$  mit dem *Kronecker-Delta*

$$\delta_{ij} = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases}$$

**10.9. Lemma.** Sei  $K$  ein Körper. Die Matrixmultiplikation ist assoziativ und hat die Einheitsmatrix als neutrales Element; sie erfüllt die Distributivgesetze bezüglich der Matrixaddition:

**LEMMA**  
Eigensch.  
Matrixmult.

- (1) Für alle  $A \in \text{Mat}(k \times l, K)$ ,  $B \in \text{Mat}(l \times m, K)$ ,  $C \in \text{Mat}(m \times n, K)$  gilt  $(AB)C = A(BC)$ .
- (2) Für alle  $A \in \text{Mat}(m \times n, K)$  gilt  $I_m A = A = A I_n$ .
- (3) Für alle  $A \in \text{Mat}(l \times m, K)$  und  $B, C \in \text{Mat}(m \times n, K)$  gilt  $A(B + C) = AB + AC$ .
- (4) Für alle  $A, B \in \text{Mat}(l \times m, K)$  und  $C \in \text{Mat}(m \times n, K)$  gilt  $(A + B)C = AC + BC$ .

Insbesondere ist  $\text{Mat}(n, K)$  mit der Matrixaddition und Matrixmultiplikation als Verknüpfungen ein Ring.

*Beweis.* Das ist eine unmittelbare Übersetzung der entsprechenden Aussagen für lineare Abbildungen, vergleiche den Beweis von Satz 9.26 (die Beweise etwa für die Distributivgesetze gehen auch in der etwas allgemeineren Situation, die hier vorliegt).  $\square$

**10.10. Definition.** Der Ring  $\text{Mat}(n, K)$  heißt der *Matrizenring* (der Größe  $n$  über  $K$ ). Eine Matrix  $A \in \text{Mat}(n, K)$  heißt *invertierbar*, wenn es eine Matrix  $B \in \text{Mat}(n, K)$  gibt mit  $AB = I_n$ . Dann gilt auch  $BA = I_n$ ; wir schreiben  $A^{-1}$  für  $B$  und nennen  $B$  die *Inverse* von  $A$ .  $\diamond$

**DEF**  
Matrizen-  
ring  
invertierbare  
Matrix

Für die zu  $A$  und  $B$  gehörenden linearen Abbildungen  $f, g : K^n \rightarrow K^n$  bedeutet  $AB = I_n$ , dass  $f \circ g = \text{id}_{K^n}$  ist. Dann ist  $f$  surjektiv, also ein Isomorphismus (siehe Folgerung 9.14) und  $g = f^{-1}$ , also ist auch  $g \circ f = \text{id}_{K^n}$ , d.h.,  $BA = I_n$ . Die Matrix  $B = A^{-1}$  ist also die zu  $f^{-1}$  gehörende Matrix.

**10.11. Beispiel.** Die Matrix  $A = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \in \text{Mat}(2, K)$  (mit  $t \in K$  beliebig) ist invertierbar, denn **BSP**

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

$\clubsuit$

Im nächsten Abschnitt werden wir lernen, wie wir Basen von Kern und Bild einer linearen Abbildung  $f : K^n \rightarrow K^m$  anhand der zugehörigen Matrix berechnen können. Wir werden auch sehen, wie man feststellt, ob eine Matrix invertierbar ist, und wie man gegebenenfalls ihre Inverse findet.

11. DER NORMALFORMALGORITHMUS UND LINEARE GLEICHUNGSSYSTEME

Wie können wir den Rang einer durch eine Matrix  $A \in \text{Mat}(m \times n, K)$  gegebenen linearen Abbildung  $f : K^n \rightarrow K^m$  bestimmen und eine Basis ihres Kerns finden? Dazu überlegen wir uns, wie man die Matrix verändern kann, ohne dass sich der Kern ändert. Dann können wir versuchen, die Matrix in eine Form zu bringen, aus der sich zum Beispiel eine Basis des Kerns leicht ablesen lässt. Eine solche Form ist die *Zeilenstufenform*:

11.1. **Definition.** Seien  $K$  ein Körper,  $m, n \in \mathbb{N}$  und  $A = (a_{ij}) \in \text{Mat}(m \times n, K)$ . Die Matrix  $A$  ist in *Zeilenstufenform*, wenn sie folgende Form hat:

**DEF**  
Zeilen-  
stufen-  
form

$$A = \begin{pmatrix} 0 \cdots 0 & \mathbf{1} & * \cdots * & * & * \cdots * & * & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & \mathbf{1} & * \cdots * & * & * \cdots * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & \mathbf{1} & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 \end{pmatrix}$$

Formal bedeutet das, dass es  $0 \leq r \leq m$  und Indizes  $1 \leq j_1 < j_2 < \cdots < j_r \leq n$  gibt, sodass  $a_{ij} = 0$ , wenn  $i > r$  oder  $i \leq r$  und  $j < j_i$ , und  $a_{ij_i} = 1$  für alle  $i \in \{1, 2, \dots, r\}$ .

$A$  ist in *reduzierter Zeilenstufenform*, wenn zusätzlich  $a_{ijk} = 0$  ist für alle  $1 \leq i < k$  und alle  $k \in \{1, 2, \dots, r\}$ :

$$A = \begin{pmatrix} 0 \cdots 0 & \mathbf{1} & * \cdots * & 0 & * \cdots * & 0 & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & \mathbf{1} & * \cdots * & 0 & * \cdots * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & \mathbf{1} & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & 0 \cdots 0 \end{pmatrix}$$

Die „führenden Einsen“ der ersten  $r$  Zeilen (in den Spalten  $j_1, j_2, \dots, j_r$ ) sind durch Fettdruck hervorgehoben, die zugehörigen Spalten in der Bildschirmversion auch farblich abgesetzt. ◇

Zur Vereinfachung führen wir folgende Sprechweise ein:

11.2. **Definition.** Sei  $A \in \text{Mat}(m \times n, K)$ . Dann ist der *Rang* von  $A$ ,  $\text{rk}(A)$ , der Rang der zugehörigen linearen Abbildung  $f : K^n \rightarrow K^m$  und der *Kern* von  $A$ ,  $\ker(A)$ , der Kern von  $f$ . ◇

**DEF**  
Rang, Kern  
einer  
Matrix

11.3. **Lemma.** Sei  $A \in \text{Mat}(m \times n, K)$  in reduzierter Zeilenstufenform mit  $r$  und  $j_1, j_2, \dots, j_r$  wie in Definition 11.1. Dann ist  $\text{rk}(A) = r$  und wir erhalten eine Basis von  $\ker(A)$  wie folgt: Sei  $J = \{1, 2, \dots, n\} \setminus \{j_1, j_2, \dots, j_r\}$  die Menge der Indizes von Spalten ohne „führende Eins“ und sei für  $j \in J$  der Vektor  $b_j \in K^n$  definiert als

**LEMMA**  
Rang und  
Kern einer  
Matrix  
in ZSF

$$b_j = \mathbf{e}_j - \sum_{i=1}^r a_{ij} \mathbf{e}_{j_i}.$$

Dann ist  $(b_j)_{j \in J}$  eine Basis von  $\ker(A)$ .

Etwas anschaulicher bekommen wir die Basis des Kerns so: Die Indizes in  $J$ , die den Spalten ohne führende Eins einer Zeile entsprechen, sind Positionen, für die wir die Komponenten frei wählen können. Wir setzen eine (die Position  $j \in J$ ) davon auf 1, die anderen auf 0 und lösen die aus  $Ab_j = \mathbf{0}$  entstehenden Gleichungen nach den übrigen Komponenten auf.

*Beweis.* Das Bild der zu  $A$  gehörenden linearen Abbildung  $f$  wird von den Spalten der Matrix erzeugt (denn das sind die Bilder der Standardbasis von  $K^n$ ). In den Spalten mit den Nummern  $j_1, j_2, \dots, j_r$  stehen die Standard-Basisvektoren  $\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_r$  von  $K^m$  und alle übrigen Spalten sind Linearkombinationen dieser Vektoren, also ist das Bild  $\langle \mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_r \rangle$  und hat Dimension  $r$ . Das zeigt  $\text{rk}(A) = r$ .

Aus der Dimensionsformel in Satz 9.18 folgt, dass der Kern von  $A$  Dimension  $n - r = \#J$  hat. Es genügt also zu zeigen, dass die  $b_j$  im Kern liegen und linear unabhängig sind. Wir schreiben  $A_j$  für die  $j$ -te Spalte von  $A$ . Dann ist

$$f(b_j) = A_j - \sum_{i=1}^r a_{ij} A_{j_i} = \sum_{i=1}^r a_{ij} \mathbf{e}'_i - \sum_{i=1}^r a_{ij} \mathbf{e}'_i = \mathbf{0},$$

also ist  $b_j$  im Kern. (Wir haben verwendet, dass  $A_{j_i} = \mathbf{e}'_i$  ist und dass  $a_{ij} = 0$  ist für  $i > r$ .) Um zu zeigen, dass die  $b_j$  linear unabhängig sind, betrachten wir eine Linearkombination:

$$\mathbf{0} = \sum_{j \in J} \lambda_j b_j = \sum_{j \in J} \lambda_j \mathbf{e}_j - \sum_{i=1}^r \left( \sum_{j \in J} a_{ij} \lambda_j \right) \mathbf{e}_{j_i}$$

Da  $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$  eine Basis von  $K^n$  ist, folgt  $\lambda_j = 0$  für alle  $j \in J$ . □

11.4. **Beispiel.** Sei  $K = \mathbb{R}$  und  $A$  die folgende Matrix über  $\mathbb{R}$ :

**BSP**

$$A = \begin{pmatrix} 0 & \mathbf{1} & 2 & 0 & 0 & -2 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 1 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Dann ist  $A$  in reduzierter Zeilenstufenform mit  $r = 3$  (das ist die Anzahl der Zeilen, die keine Null-Zeilen sind) und  $j_1 = 2, j_2 = 4, j_3 = 5$ . Der Rang ist also 3,  $J = \{1, 3, 6\}$  und eine Basis des Kerns ist gegeben durch

$$b_1 = (\mathbf{1}, 0, \mathbf{0}, 0, 0, \mathbf{0}), \quad b_3 = (\mathbf{0}, -2, \mathbf{1}, 0, 0, \mathbf{0}), \quad b_6 = (\mathbf{0}, 2, \mathbf{0}, -1, -5, \mathbf{1}).$$

Die frei wählbaren Komponenten (Positionen 1, 3, 6) sind durch Fettdruck hervorgehoben. Die restlichen Komponenten von  $b_j$  ergeben sich aus den Negativen der ersten  $r$  Einträge der  $j$ -ten Spalte von  $A$ . ♣

Wie bekommen wir nun eine Matrix in diese Zeilenstufenform, ohne ihren Kern zu ändern? Dazu gehen wir schrittweise vor und führen kleine Veränderungen durch, von denen man leicht einsehen kann, dass sie diese Eigenschaft haben.

11.5. **Definition.** Seien  $K$  ein Körper,  $m, n \in \mathbb{N}$  und  $A \in \text{Mat}(m \times n, K)$ .

**DEF**  
elementare  
Zeilen-  
umformungen

- (1) Eine *elementare Zeilenumformung vom Typ I* an der Matrix  $A$  besteht darin, die  $i$ -te Zeile von  $A$  mit  $\lambda$  zu multiplizieren. Dabei ist  $i \in \{1, 2, \dots, m\}$  und  $\lambda \in K^\times$ . Wir schreiben  $\mathbf{I}_i(\lambda)$  für diese Umformung.

- (2) Eine *elementare Zeilenumformung vom Typ II* an der Matrix  $A$  besteht darin, das  $\lambda$ -fache der  $j$ -ten Zeile von  $A$  zur  $i$ -ten Zeile zu addieren. Dabei sind  $i, j \in \{1, 2, \dots, m\}$  mit  $i \neq j$  und  $\lambda \in K$ . Wir schreiben  $\mathbf{II}_{i,j}(\lambda)$  für diese Umformung.
- (3) Eine *elementare Zeilenumformung vom Typ III* an der Matrix  $A$  besteht darin, in  $A$  zwei Zeilen miteinander zu vertauschen. Dabei sind  $i, j \in \{1, 2, \dots, m\}$  mit  $i \neq j$ . Wir schreiben  $\mathbf{III}_{i,j}$  für diese Umformung.

Eine *Zeilenumformung* an der Matrix  $A$  ist eine Abfolge von sukzessiven elementaren Zeilenumformungen, beginnend mit der Matrix  $A$ .  $\diamond$

Eine elementare Zeilenumformung vom Typ III kann durch eine Abfolge geeigneter Umformungen der Typen I und II erreicht werden (Übung). Diese Art der Umformung ist also eigentlich nicht nötig, stellt aber häufig eine praktische Abkürzung dar.

**11.6. Lemma.** *Seien  $K$  ein Körper,  $m, n \in \mathbb{N}$  und  $A \in \text{Mat}(m \times n, K)$ . Sei weiter  $A'$  eine Matrix, die aus  $A$  durch eine elementare Zeilenumformung hervorgeht. Dann ist  $\ker(A') = \ker(A)$  und daher auch  $\text{rk}(A') = \text{rk}(A)$ .*

**LEMMA**  
Zeilenumf.  
erhalten  
Kern

*Beweis.* Ein Vektor  $v = (x_1, x_2, \dots, x_n) \in K^n$  ist genau dann im Kern von  $A = (a_{ij})$ , wenn für alle  $i \in \{1, 2, \dots, m\}$  gilt  $\sum_{j=1}^n a_{ij}x_j = 0$ . Eine elementare Zeilenumformung vom Typ I ersetzt eine dieser Gleichungen durch ihr  $\lambda$ -faches mit  $\lambda \neq 0$ , was ihre Gültigkeit nicht ändert. Bei einer elementaren Zeilenumformung vom Typ II wird zu einer der Gleichungen das  $\lambda$ -fache einer anderen Gleichung addiert, die neuen Gleichungen sind also gültig, wenn die alten es sind. Da man die Umformung rückgängig machen kann (durch Subtraktion des  $\lambda$ -fachen der  $j$ -ten Zeile von der  $i$ -ten), gelten die neuen Gleichungen genau dann, wenn die alten gelten. (Umformungen vom Typ III brauchen nicht extra betrachtet zu werden; da sie aber nur die Reihenfolge der Gleichungen ändern, ist klar, dass der Kern dabei erhalten bleibt.) Das zeigt, dass  $v$  genau dann im Kern von  $A$  ist, wenn  $v$  im Kern von  $A'$  ist. Die Gleichheit der Ränge folgt aus der Dimensionsformel von Satz 9.18.  $\square$

Wir zeigen jetzt, dass man jede Matrix durch Zeilenumformungen in Zeilenstufenform überführen kann.

\* **11.7. Satz.** *Seien  $K$  ein Körper,  $m, n \in \mathbb{N}$  und  $A \in \text{Mat}(m \times n, K)$ . Dann lässt sich  $A$  durch sukzessive elementare Zeilenumformungen in eine Matrix  $A'$  in reduzierter Zeilenstufenform überführen.*

**SATZ**  
Normal-  
form von  
Matrizen

*Beweis.* Wir zeigen zuerst, dass sich  $A$  in (nicht notwendig reduzierte) Zeilenstufenform bringen lässt. Der Beweis dafür geht durch Induktion nach der Zeilenanzahl  $m$ . Im Fall  $m = 0$  ist die Matrix (trivialerweise) bereits in Zeilenstufenform. Sei also  $m > 0$  und die Behauptung für alle Matrizen mit weniger als  $m$  Zeilen schon gezeigt. Ist  $A$  die Nullmatrix, dann ist  $A$  in Zeilenstufenform und es ist nichts zu zeigen. Wir können also annehmen, dass  $A$  einen von null verschiedenen Eintrag hat. Sei  $j_1$  der kleinste Index einer Spalte mit einem solchen Eintrag. Ist  $a_{1j_1} = 0$ , dann können wir durch eine Typ-III-Umformung erreichen, dass  $a_{1j_1} \neq 0$

ist. Eine Umformung vom Typ I mit  $\lambda = a_{1j_1}^{-1}$ , angewandt auf die erste Zeile, ergibt  $a_{1j_1} = 1$ . Die Matrix hat jetzt die Form

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & a_{2j_1} & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots & \\ 0 & \cdots & 0 & a_{mj_1} & * & \cdots & * \end{pmatrix}$$

Durch die Umformungen  $\mathbf{II}_{2,1}(-a_{2j_1})$ ,  $\mathbf{II}_{3,1}(-a_{3j_1})$ ,  $\dots$ ,  $\mathbf{II}_{m,1}(-a_{mj_1})$  können wir die  $j_1$ -te Spalte unterhalb der ersten Zeile „ausräumen“, sodass wir nun die Form

$$\left( \begin{array}{cccc|ccc} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & & \\ 0 & \cdots & 0 & 0 & & & \end{array} \right) \begin{array}{c} \\ \\ A' \\ \end{array}$$

haben mit einer  $(m-1) \times (n-j_1)$ -Matrix  $A'$ . Zeilenumformungen an  $A'$  können auch als Zeilenumformungen an dieser Matrix ausgeführt werden, ohne dass sich am linken Teil der Matrix etwas ändert. Nach Induktionsannahme kann nun  $A'$  durch Zeilenumformungen in Zeilenstufenform gebracht werden. Damit hat die gesamte Matrix ebenfalls Zeilenstufenform.

Wir führen jetzt noch für  $k = 1, 2, \dots, r$  und  $i = 1, 2, \dots, k-1$  die Umformungen  $\mathbf{II}_{i,k}(-a_{ij_k})$  aus (mit dem jeweils aktuellen Wert des Eintrags  $a_{ij_k}$ ) und räumen auf diese Weise auch noch den Teil der Spalten oberhalb der führenden Einsen aus. Wir erhalten so die reduzierte Zeilenstufenform.  $\square$

Dieser Beweis liefert uns sogar einen Algorithmus. Wir werden die Umformungen an einer Beispielmatrix durchführen.

**11.8. Beispiel.** Wir bestimmen die reduzierte Zeilenstufenform der folgenden Matrix über  $\mathbb{R}$ :

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$$

Die erste Spalte ist keine Null-Spalte, also ist  $j_1 = 1$ . Der oberste Eintrag in der ersten Spalte ist bereits 1, also sind keine Umformungen vom Typ III oder I nötig. Wir räumen den Rest der Spalte aus, indem wir das Fünffache der ersten Zeile von der zweiten und das Neunfache der ersten Zeile von der dritten Zeile abziehen. Dann bekommen wir die neue Matrix

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & -8 & -16 & -24 \end{pmatrix}$$

Wir machen mit der rechten unteren  $2 \times 3$ -Matrix weiter. Ihre erste Spalte  $(-4, -8)$  ist keine Null-Spalte, also ist  $j_2 = 2$ . Wir multiplizieren die zweite Zeile der gesamten Matrix mit  $-1/4$  und bekommen

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & -8 & -16 & -24 \end{pmatrix}.$$

Dann addieren wir das Achtfache der zweiten Zeile zur dritten:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

**BSP**  
Umformung  
in ZSF

Damit haben wir die Zeilenstufenform erreicht (mit  $r = 2$ ). Für die reduzierte Zeilenstufenform müssen wir noch das Doppelte der zweiten Zeile von der ersten abziehen; das liefert schließlich

$$A' = \begin{pmatrix} \mathbf{1} & 0 & -1 & -2 \\ 0 & \mathbf{1} & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Wir können jetzt eine Basis des Kerns von  $A$  (der gleich dem Kern von  $A'$  ist) ablesen, nämlich

$$b_3 = (1, -2, \mathbf{1}, \mathbf{0}) \quad \text{und} \quad b_4 = (2, -3, \mathbf{0}, \mathbf{1}). \quad \clubsuit$$

Elementare Zeilenumformungen lassen sich durch Multiplikation mit gewissen invertierbaren Matrizen von links beschreiben. Diese Matrizen sind die sogenannten *Elementarmatrizen*  $E_i(\lambda)$  mit  $\lambda \in K^\times$  und  $i \in \{1, 2, \dots, m\}$  und  $E_{ij}(\lambda)$  mit  $\lambda \in K$  und  $i, j \in \{1, 2, \dots, m\}$ ,  $i \neq j$ . Um sie zu definieren, führen wir  $M_{kl} = (\delta_{ik}\delta_{jl})_{1 \leq i, j \leq m}$  ein; in dieser Matrix sind alle Einträge null bis auf den Eintrag in der  $k$ -ten Zeile und  $l$ -ten Spalte, der den Wert 1 hat. (Die Matrizen  $M_{kl}$  entsprechen der Basis  $(\phi_{kl})_{1 \leq k, l \leq m}$  von  $\text{Hom}(K^m, K^m)$  wie in Folgerung 9.24.) Dann ist

$$E_i(\lambda) = I_m + (\lambda - 1)M_{ii} \quad \text{und} \quad E_{ij}(\lambda) = I_m + \lambda M_{ij}.$$

$E_i(\lambda)$  unterscheidet sich von der Einheitsmatrix  $I_m$  dadurch, dass an der  $i$ -ten Position auf der Diagonalen statt 1 der Eintrag  $\lambda$  steht. In  $E_{ij}(\lambda)$  steht außerhalb der Diagonalen an der Position  $(i, j)$  der Eintrag  $\lambda$ . Wegen

$$E_i(\lambda)E_i(\lambda^{-1}) = I_m \quad \text{und} \quad E_{ij}(\lambda)E_{ij}(-\lambda) = I_m$$

sind diese Elementarmatrizen invertierbar. Was bewirkt die Multiplikation von links mit so einer Elementarmatrix? Dazu überlegen wir, dass

$$M_{kl}A = \left( \sum_{h=1}^m \delta_{ik}\delta_{hl}a_{hj} \right)_{1 \leq i \leq m, 1 \leq j \leq n} = (\delta_{ik}a_{lj})_{1 \leq i \leq m, 1 \leq j \leq n};$$

in dieser Matrix sind alle Zeilen null bis auf die  $k$ -te Zeile, in welcher sich die  $l$ -te Zeile von  $A$  befindet. Multiplikation von links mit  $M_{kl}$  setzt also die  $l$ -te Zeile von  $A$  in die  $k$ -te Zeile und löscht alle anderen Zeilen.

Damit ergibt sich, dass die Zeilen von  $E_i(\lambda)A$  mit den entsprechenden Zeilen von  $A$  übereinstimmen bis auf die  $i$ -te Zeile, die mit  $\lambda$  multipliziert wird. Der Effekt ist also die elementare Zeilenumformung  $\mathbf{I}_i(\lambda)$  vom Typ I. Ebenso stimmen die Zeilen von  $E_{ij}(\lambda)A$  mit denen von  $A$  überein mit Ausnahme der  $i$ -ten Zeile, zu der das  $\lambda$ -fache der  $j$ -ten Zeile addiert wird. Der Effekt ist also die elementare Zeilenumformung  $\mathbf{II}_{i,j}(\lambda)$  vom Typ II.

Wir veranschaulichen das für die  $2 \times 3$ -Matrix  $A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$ :

$$\begin{aligned} E_1(\lambda)A &= \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b & \lambda c \\ d & e & f \end{pmatrix} \\ E_2(\lambda)A &= \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} a & b & c \\ \lambda d & \lambda e & \lambda f \end{pmatrix} \\ E_{12}(\lambda)A &= \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} a + \lambda d & b + \lambda e & c + \lambda f \\ d & e & f \end{pmatrix} \\ E_{21}(\lambda)A &= \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} a & b & c \\ d + \lambda a & e + \lambda b & f + \lambda c \end{pmatrix} \end{aligned}$$

Der Inhalt von Satz 11.7 ist also, dass es zu jeder Matrix  $A \in \text{Mat}(m \times n, K)$  eine invertierbare Matrix  $P \in \text{Mat}(m, K)$  gibt, sodass  $PA$  reduzierte Zeilenstufenform hat, wobei  $P$  ein Produkt von Elementarmatrizen ist.

Sei jetzt  $A \in \text{Mat}(m, K)$  invertierbar. Wendet man diese Aussage an auf  $A^{-1}$  und beachtet, dass die reduzierte Zeilenstufenform einer invertierbaren  $m \times m$ -Matrix gerade die Einheitsmatrix  $I_m$  ist (siehe Lemma 11.19 unten), dann erhält man ein Produkt  $P$  von Elementarmatrizen mit  $PA^{-1} = I_m$ . Es folgt  $A = P$ . Wir haben bewiesen:

**Satz.** *Jede invertierbare Matrix ist ein Produkt von Elementarmatrizen.*

Daraus folgt:

*Zwei Matrizen  $A, B \in \text{Mat}(m \times n, K)$  lassen sich durch Zeilenumformungen ineinander überführen genau dann, wenn es eine invertierbare Matrix  $P \in \text{Mat}(m, K)$  gibt mit  $B = PA$ .*

Statt Zeilenumformungen kann man ganz analog *Spaltenumformungen* betrachten. Sie werden durch Multiplikation mit Elementarmatrizen von *rechts* bewirkt. Man hat dann die folgende analoge Aussage:

*Zwei Matrizen  $A, B \in \text{Mat}(m \times n, K)$  lassen sich durch Spaltenumformungen ineinander überführen genau dann, wenn es eine invertierbare Matrix  $Q \in \text{Mat}(n, K)$  gibt mit  $B = AQ$ .*

Wir kommen zu linearen Gleichungen und Gleichungssystemen.

**11.9. Definition.** Seien  $V$  und  $W$  zwei  $K$ -Vektorräume und  $f : V \rightarrow W$  eine lineare Abbildung. Ist  $b \in W$  ein gegebener Vektor, dann heißt die Gleichung  $f(x) = b$ , deren Lösungen  $x \in V$  gesucht sind, eine *lineare Gleichung*. Die Gleichung heißt *homogen*, wenn  $b = \mathbf{0}$  ist, sonst *inhomogen*.

Ist  $V = K^n$  und  $W = K^m$ , dann kann die Gleichung unter Benutzung der zu  $f$  gehörenden Matrix  $A \in \text{Mat}(m \times n, K)$  auch geschrieben werden als  $A\mathbf{x} = \mathbf{b}$  mit Spaltenvektoren  $\mathbf{x} \in K^n$  und  $\mathbf{b} \in K^m$ . In diesem Fall spricht man auch von einem *linearen Gleichungssystem* (mit  $m$  Gleichungen in  $n$  Unbestimmten).  $\diamond$

Wir können schon recht genau sagen, welche Struktur die Lösungsmenge einer linearen Gleichung hat.

\*

**11.10. Satz.** *Seien  $V$  und  $W$  zwei  $K$ -Vektorräume und  $f : V \rightarrow W$  eine lineare Abbildung.*

- (1) *Die Lösungsmenge der homogenen linearen Gleichung  $f(x) = \mathbf{0}$  ist ein Untervektorraum von  $V$ , nämlich der Kern von  $f$ .*
- (2) *Sei  $\mathbf{0} \neq b \in W$ . Ist  $b \notin \text{im}(f)$ , dann hat die inhomogene lineare Gleichung  $f(x) = b$  keine Lösung. Anderenfalls sei  $x_0 \in V$  mit  $f(x_0) = b$ . Dann ist die Lösungsmenge gegeben durch  $x_0 + \ker(f) = \{x_0 + v \mid v \in \ker(f)\}$ .*

*Beweis.* Die erste Aussage folgt direkt aus der Definition des Kerns und der Tatsache, dass  $\ker(f)$  ein Untervektorraum von  $V$  ist. In der zweiten Aussage ist klar, dass es genau dann Lösungen gibt, wenn  $b \in \text{im}(f)$  ist (das ist die Definition von  $\text{im}(f)$ ). Es bleibt die letzte Behauptung zu zeigen. Sei dazu  $x \in V$ . Dann gilt

$$\begin{aligned} f(x) = b &\iff f(x) = f(x_0) \iff f(x - x_0) = \mathbf{0} \\ &\iff x - x_0 \in \ker(f) \iff x \in x_0 + \ker(f). \quad \square \end{aligned}$$

Das allgemeine Rezept für die Lösung einer linearen Gleichung  $f(x) = b$  lautet also:

- (1) Prüfe, ob  $b \in \text{im}(f)$ . Falls nein, dann gibt es keine Lösung.

**SATZ**  
Elementar-  
matrizen  
erzeugen  
invertierbare  
Matrizen

**DEF**  
Lineare  
Gleichung

**SATZ**  
Lösungs-  
menge  
einer  
linearen  
Gleichung

- (2) Bestimme eine „spezielle Lösung“  $x_0 \in V$ .  
 (3) Bestimme  $\ker(f)$ .  
 (4) Die Lösungsmenge ist  $x_0 + \ker(f)$ . Ist  $\ker(f)$  endlich-dimensional mit Basis  $(x_1, x_2, \dots, x_n)$ , dann ist die „allgemeine Lösung“

$$x = x_0 + \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n$$

mit  $\lambda_1, \lambda_2, \dots, \lambda_n \in K$ .

Die ersten beiden Schritte wird man im Regelfall zusammen ausführen, denn wenn man feststellt, dass  $b \in \operatorname{im}(f)$  ist, dann wird man meistens auch ein Urbild gefunden haben.

Im homogenen Fall (also  $b = \mathbf{0}$ ) gilt stets  $b \in \operatorname{im}(f)$  und wir können  $x_0 = \mathbf{0}$  nehmen; die Lösungsmenge ist dann  $\ker(f)$ .

**11.11. Beispiel.** Wir betrachten die folgende inhomogene lineare Differentialgleichung erster Ordnung:

$$y'(x) + y(x) = x.$$

Dabei sei  $y \in \mathcal{C}^1(\mathbb{R})$ . Hier ist  $K = \mathbb{R}$ ,  $V = \mathcal{C}^1(\mathbb{R})$ ,  $W = \mathcal{C}(\mathbb{R})$  und  $f : y \mapsto y' + y$ . Die Gleichung ist  $f(y) = \operatorname{id}_{\mathbb{R}}$ . Wir suchen nach einer speziellen Lösung. Mit etwas Probieren finden wir  $y_0(x) = x - 1$ . (In der Vorlesung über *Gewöhnliche Differentialgleichungen* werden Sie lernen, wie man solche Lösungen systematisch findet.) Jetzt müssen wir den Kern von  $f$  bestimmen, also die Menge aller Funktionen  $y$  mit  $y' + y = \mathbf{0}$ . Ich behaupte, dass  $\ker(f) = \langle x \mapsto e^{-x} \rangle$  ist, die Funktionen  $y$  mit  $y' + y = \mathbf{0}$  haben also die Form  $y(x) = Ce^{-x}$  mit  $C \in \mathbb{R}$ . Zum Beweis betrachten wir  $z(x) = e^x y(x)$ ; dann gilt

$$z'(x) = e^x y(x) + e^x y'(x) = e^x (y(x) + y'(x)) = \mathbf{0},$$

also ist  $z(x) = C$  konstant und damit  $y(x) = Ce^{-x}$ . Umgekehrt sind diese Funktionen auch Lösungen von  $y' + y = \mathbf{0}$ . Die allgemeine Lösung ist also

$$y(x) = x - 1 + Ce^{-x}, \quad C \in \mathbb{R}. \quad \clubsuit$$

Wie sieht das obige Rezept konkret aus, wenn wir ein lineares Gleichungssystem lösen wollen? Sei  $A\mathbf{x} = \mathbf{b}$  ein lineares Gleichungssystem mit  $A \in \operatorname{Mat}(m \times n, K)$ . Im homogenen Fall  $\mathbf{b} = \mathbf{0}$  müssen wir eine Basis von  $\ker(A)$  bestimmen. Dazu bringen wir  $A$  in reduzierte Zeilenstufenform und lesen eine Basis des Kerns ab wie in Lemma 11.3. Im inhomogenen Fall sei  $A' = (A \mid \mathbf{b})$  die *erweiterte Matrix* des Systems; wir erhalten sie, indem wir an die Matrix  $A$  den Spaltenvektor  $\mathbf{b}$  als  $(n+1)$ -te Spalte anfügen. Wir schreiben  $\operatorname{im}(A)$  für das Bild der zu  $A$  gehörenden linearen Abbildung; das ist der *Spaltenraum* von  $A$ , also der von den Spalten von  $A$  erzeugte Untervektorraum von  $K^m$ .

**11.12. Satz.** Sei  $K$  ein Körper, seien  $m, n \in \mathbb{N}$ , sei  $A \in \operatorname{Mat}(m \times n, K)$  und  $\mathbf{b} \in K^m$  ein Spaltenvektor. Sei weiter  $A' = (A \mid \mathbf{b})$ . Dann gilt

$$\mathbf{b} \in \operatorname{im}(A) \iff \operatorname{rk}(A') = \operatorname{rk}(A).$$

Dies kann geprüft werden, indem  $A'$  in reduzierte Zeilenstufenform  $\tilde{A}'$  gebracht wird.  $\operatorname{rk}(A') = \operatorname{rk}(A)$  ist dann dazu äquivalent, dass die letzte Spalte von  $\tilde{A}'$  keine

**BSP**  
inhomogene  
lineare  
Diff.gleichung

**SATZ**  
inhom.  
LGS

führende Eins einer Zeile enthält (das bedeutet  $j_r \leq n$  in der Notation von Definition 11.1). In diesem Fall kann eine spezielle Lösung von  $A\mathbf{x} = \mathbf{b}$  aus  $\tilde{A}'$  wie folgt abgelesen werden: Die letzte Spalte von  $\tilde{A}'$  sei  $(\tilde{b}_1, \dots, \tilde{b}_r, 0, \dots, 0)$ . Dann ist

$$\mathbf{x}_0 = \sum_{i=1}^r \tilde{b}_i \mathbf{e}_{j_i}$$

eine Lösung des Gleichungssystems.

*Beweis.* Seien  $A_1, \dots, A_n$  die Spalten von  $A$ . Es gilt

$$\begin{aligned} \mathbf{b} \in \text{im}(A) = \langle A_1, \dots, A_n \rangle &\iff \langle A_1, \dots, A_n, \mathbf{b} \rangle = \langle A_1, \dots, A_n \rangle \\ &\iff \text{im}(A') = \text{im}(A). \end{aligned}$$

Die letzte Aussage impliziert  $\text{rk}(A') = \text{rk}(A)$ . Es gilt immer  $\text{im}(A) \subset \text{im}(A')$ , also folgt aus  $\text{rk}(A') = \text{rk}(A)$  auch die Gleichheit von  $\text{im}(A')$  und  $\text{im}(A)$ . Damit ist die erste Behauptung gezeigt.

Sei nun  $\tilde{A}'$  die reduzierte Zeilenstufenform von  $A'$ . Dann bilden die ersten  $n$  Spalten von  $\tilde{A}'$  die reduzierte Zeilenstufenform  $\tilde{A}$  von  $A$ . Der Rang von  $A'$  ist genau dann größer als der Rang von  $A$ , wenn  $\tilde{A}'$  mehr Nichtnull-Zeilen hat als  $\tilde{A}$ . Das bedeutet aber gerade, dass die letzte Spalte von  $\tilde{A}'$  eine führende Eins enthalten muss. Das zeigt die zweite Aussage. Für die letzte Aussage beachten wir, dass die Zeilenumformungen, die im Zuge der Herstellung der reduzierten Zeilenstufenform durchgeführt werden, die ursprünglichen Gleichungen durch äquivalente Gleichungen ersetzen. Mit  $\tilde{A}' = (\tilde{A} \mid \tilde{\mathbf{b}})$  hat also das lineare Gleichungssystem  $\tilde{A}\mathbf{x} = \tilde{\mathbf{b}}$  dieselben Lösungen wie das ursprüngliche Gleichungssystem. Da die  $j_i$ -te Spalte von  $\tilde{A}$  gerade der Standard-Basisvektor  $\mathbf{e}'_i$  ist, ergibt sich

$$\tilde{A}\mathbf{x}_0 = \sum_{i=1}^r \tilde{b}_i \tilde{A}\mathbf{e}_{j_i} = \sum_{i=1}^r \tilde{b}_i \mathbf{e}'_i = \tilde{\mathbf{b}}. \quad \square$$

**11.13. Beispiel.** Wir lösen das folgende lineare Gleichungssystem (mit  $K = \mathbb{Q}$  oder  $\mathbb{R}$ ):

**BSP  
LGS**

$$\begin{array}{cccccc} x_1 & & & - & x_3 & - & 2x_4 & = & 3 \\ -x_1 & + & x_2 & + & x_3 & & & = & -2 \\ & & x_2 & & & - & x_4 & = & 0 \end{array}$$

oder, in Matrixschreibweise,

$$\begin{pmatrix} 1 & 0 & -1 & -2 \\ -1 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix}.$$

Die erweiterte Matrix ist

$$A' = \begin{pmatrix} 1 & 0 & -1 & -2 & 3 \\ -1 & 1 & 1 & 0 & -2 \\ 0 & 1 & 0 & -1 & 0 \end{pmatrix}.$$

Ihre reduzierte Zeilenstufenform ergibt sich als

$$\tilde{A}' = \begin{pmatrix} \mathbf{1} & 0 & -1 & 0 & 1 \\ 0 & \mathbf{1} & 0 & 0 & -1 \\ 0 & 0 & 0 & \mathbf{1} & -1 \end{pmatrix}.$$

Das zugehörige lineare Gleichungssystem ist:

$$\begin{array}{rclcl} x_1 & & - x_3 & = & 1 \\ & x_2 & & = & -1 \\ & & & x_4 = & -1 \end{array}$$

Als spezielle Lösung erhalten wir daraus  $\mathbf{x}_0 = (\mathbf{1}, -\mathbf{1}, 0, -\mathbf{1})$ . Außerdem lesen wir ab:  $\text{rk}(A) = 3$ ,  $\dim \ker(A) = 4 - 3 = 1$ , und eine Basis von  $\ker(A)$  ist gegeben durch  $\mathbf{x}_1 = (1, 0, \mathbf{1}, 0)$ . Die allgemeine Lösung des Gleichungssystems ist also

$$\mathbf{x} = \mathbf{x}_0 + \lambda_1 \mathbf{x}_1 = (1 + \lambda_1, -1, \lambda_1, -1)$$

mit  $\lambda_1 \in K$ . ♣

Hier ist das Rezept noch einmal ganz konkret:

- (1) Die erweiterte Matrix  $A'$  aufstellen.
- (2)  $A'$  in reduzierte Zeilenstufenform  $\tilde{A}'$  bringen. Sei  $r = \text{rk}(\tilde{A}')$  und seien  $1 \leq j_1 < j_2 < \dots < j_r \leq n + 1$  die Positionen der führenden Einsen der ersten  $r$  Zeilen von  $\tilde{A}'$ .
- (3)  $j_r = n + 1 \Rightarrow$  keine Lösung. Anderenfalls:
- (4) Sei  $J = \{1, 2, \dots, n\} \setminus \{j_1, j_2, \dots, j_r\}$  die Menge der „freien“ Positionen. Setze  $x_j = \lambda_j \in K$  beliebig für  $j \in J$  und löse das der Matrix  $\tilde{A}'$  entsprechende Gleichungssystem nach  $x_{j_i}$ ,  $i \in \{1, 2, \dots, r\}$  auf. Das ergibt die allgemeine Lösung.

Im Beispiel oben ist  $r = 3$ ,  $j_1 = 1$ ,  $j_2 = 2$ ,  $j_3 = 4 < 5 = n + 1$ ,  $J = \{3\}$ . Wir setzen also  $x_3 = \lambda$  und lösen das System nach  $x_1, x_2, x_4$  auf.

Diese Methode für die Lösung linearer Gleichungssysteme (und ihre Varianten) heißt *gaußsches Eliminationsverfahren* oder kürzer *Gauß-Elimination*. Eine Variante besteht darin, statt der reduzierten Zeilenstufenform nur die Zeilenstufenform herzustellen und dann das System schrittweise „von unten her“ durch Einsetzen zu lösen. Diese Version ist etwas effizienter im Hinblick auf die Zahl der nötigen Rechenoperationen, dafür aber auch etwas umständlicher durchzuführen.

Wir haben den Rang einer Matrix als den Rang der zugehörigen linearen Abbildung definiert, also als die Dimension ihres Spaltenraums. Man sollte also eigentlich genauer vom „Spaltenrang“ sprechen, denn man könnte genauso gut die Dimension des Zeilenraums (das ist der von den Zeilen der Matrix erzeugte Untervektorraum von  $K^n$ ), also den „Zeilenrang“ betrachten. Zum Glück macht das keinen Unterschied, wie wir jetzt zeigen werden.

**11.14. Lemma.** *Sei  $K$  ein Körper, seien  $m, n \in \mathbb{N}$  und  $A, B \in \text{Mat}(m \times n, K)$ . Lässt sich  $B$  aus  $A$  durch Zeilenumformungen erhalten, dann haben  $A$  und  $B$  denselben Zeilenraum.*

**LEMMA**  
Zeilenumf.  
erhalten  
Zeilenraum

*Beweis.* Bei einer elementaren Zeilenumformung werden Zeilen der Matrix durch Linearkombinationen von Zeilen ersetzt. Daraus folgt, dass der Zeilenraum der neuen Matrix im Zeilenraum der alten Matrix enthalten ist. Da sich elementare Zeilenumformungen rückgängig machen lassen, gilt auch die umgekehrte Inklusion. □

**11.15. Satz.** Sei  $K$  ein Körper, seien  $m, n \in \mathbb{N}$  und  $A \in \text{Mat}(m \times n, K)$ . Dann ist die Dimension des Zeilenraums von  $A$  gleich der Dimension des Spaltenraums von  $A$ .

**SATZ**  
Zeilenrang =  
Spaltenrang

*Beweis.* Nach Lemma 11.14 und Satz 11.7 können wir annehmen, dass  $A$  reduzierte Zeilenstufenform hat. Sei  $r = \text{rk}(A)$  die Dimension des Spaltenraums von  $A$ . Dann hat  $A$  genau  $r$  Zeilen, die keine Null-Zeilen sind, und diese Zeilen sind linear unabhängig, denn das gilt bereits, wenn man nur die Spalten  $j_1, j_2, \dots, j_r$  (Notation wie in Definition 11.1) betrachtet — die Matrix  $(a_{i,j_k})_{1 \leq i, k \leq r}$  ist die Einheitsmatrix  $I_r$ .  $\square$

Der Normalformalgorithmus aus Satz 11.7 berechnet demnach auch die Dimension und eine Basis des Zeilenraums der gegebenen Matrix. Wenn man also die Dimension und eine Basis des von Vektoren  $v_1, \dots, v_m \in K^n$  erzeugten Untervektorraums bestimmen möchte, dann schreibt man diese Vektoren als Zeilen in eine Matrix und bestimmt ihre (reduzierte) Zeilenstufenform. Die von null verschiedenen Zeilen der resultierenden Matrix bilden dann eine Basis.

Man kann den Satz kurz und elegant formulieren, wenn man folgende Definition verwendet.

\* **11.16. Definition.** Sei  $K$  ein Körper, seien  $m, n \in \mathbb{N}$  und sei  $A = (a_{ij}) \in \text{Mat}(m \times n, K)$ . Die *Transponierte* von  $A$  oder die zu  $A$  *transponierte Matrix* ist  $A^\top = (a_{ji})_{1 \leq i \leq n, 1 \leq j \leq m} \in \text{Mat}(n \times m, K)$ .

**DEF**  
Transponierte  
Matrix  $\diamond$

Da es leicht zu Verwirrung führt: Die Schreibweise

$$A^\top = (a_{ji})_{1 \leq i \leq n, 1 \leq j \leq m}$$

bedeutet Folgendes: Der erste Index unten hinter der Klammer (hier  $i$ ) ist der Zeilenindex und der zweite (hier  $j$ ) ist der Spaltenindex. Die Matrix  $A^\top$  hat also  $n$  Zeilen und  $m$  Spalten. Der Eintrag in Zeile  $i$  und Spalte  $j$  ist  $a_{ji}$  und damit derselbe Eintrag wie in *Spalte*  $i$  und *Zeile*  $j$  der Matrix  $A$ . Gleichbedeutend könnte man auch

$$A^\top = (a_{ij})_{1 \leq j \leq n, 1 \leq i \leq m}$$

schreiben. In diesem Fall wäre  $j$  der Zeilen- und  $i$  der Spaltenindex.

Die Matrix wird also „an der Hauptdiagonale gespiegelt“.

**11.17. Beispiel.**

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^\top = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

**BSP**  
Transponierte  
Matrix  $\clubsuit$

Der Zeilenraum von  $A$  ist der Spaltenraum von  $A^\top$  und umgekehrt. Satz 11.15 sagt also

$$\text{rk}(A^\top) = \text{rk}(A).$$

Hier sind die wichtigsten Rechenregeln für transponierte Matrizen:

11.18. **Lemma.** Sei  $K$  ein Körper und seien  $l, m, n \in \mathbb{N}$ .

- (1) Die Abbildung  $\text{Mat}(m \times n, K) \rightarrow \text{Mat}(n \times m, K)$ ,  $A \mapsto A^\top$  ist ein Isomorphismus (es gilt also  $(A + B)^\top = A^\top + B^\top$  und  $(\lambda A)^\top = \lambda A^\top$  für  $A, B \in \text{Mat}(m \times n, K)$ ,  $\lambda \in K$ ; die Bijektivität ist klar).
- (2) Für  $A \in \text{Mat}(l \times m, K)$  und  $B \in \text{Mat}(m \times n, K)$  gilt  $(AB)^\top = B^\top A^\top$ .
- (3) Für  $A \in \text{Mat}(m \times n, K)$  gilt  $(A^\top)^\top = A$ .

**LEMMA**  
Rechenregeln  
für  $A^\top$

*Beweis.* Übung. □

Wir haben von „der“ reduzierten Zeilenstufenform einer Matrix gesprochen. Tatsächlich ist das Ergebnis eines Verfahrens, das eine Matrix in reduzierte Zeilenstufenform überführt, eindeutig bestimmt, wie der folgende Satz zeigt. Es ist also letztlich ganz egal, welche Zeilenumformungen man in welcher Reihenfolge macht, um zur reduzierten Zeilenstufenform zu gelangen.

**Satz.** Sind  $A, B \in \text{Mat}(m \times n, K)$  zwei Matrizen in reduzierter Zeilenstufenform mit demselben Zeilenraum, dann gilt  $A = B$ .

**SATZ**  
Eindeutigkeit  
der Zeilen-  
stufenform

*Beweis.* Sei  $U \subset K^n$  der Zeilenraum von  $A$  und  $B$ . Für  $0 \leq k \leq n$  sei

$$V_k = \{(x_1, \dots, x_n) \in K^n \mid x_1 = x_2 = \dots = x_k = 0\} \subset K^n$$

und  $d_k = \dim(U \cap V_k)$ . Dann ist  $d_0 = r = \dim U$ ,  $d_n = 0$  und  $d_k - 1 \leq d_{k+1} \leq d_k$ . Es gibt also genau  $r$  „Sprungstellen“  $j_i$  mit  $d_{j_i-1} = r + 1 - i$  und  $d_{j_i} = r - i$  für  $i \in \{1, 2, \dots, r\}$ . Aus der Definition der Zeilenstufenform ergibt sich, dass  $j_i$  genau die Position der führenden Eins in der  $i$ -ten Zeile von  $A$  und von  $B$  ist. Die lineare Abbildung

$$\phi : U \longrightarrow K^r, \quad (x_1, x_2, \dots, x_n) \longmapsto (x_{j_1}, x_{j_2}, \dots, x_{j_r})$$

ist dann ein Isomorphismus, und die ersten  $r$  Zeilen von  $A$  und  $B$  müssen die Urbilder  $\phi^{-1}(\mathbf{e}_1), \phi^{-1}(\mathbf{e}_2), \dots, \phi^{-1}(\mathbf{e}_r)$  der Standard-Basisvektoren von  $K^r$  sein. Insbesondere sind  $A$  und  $B$  gleich. □

Abschließend wollen wir noch überlegen, wie man die Inverse einer Matrix berechnen kann. Dazu beachten wir, dass ein lineares Gleichungssystem  $A\mathbf{x} = \mathbf{b}$  für jedes  $\mathbf{b}$  eine eindeutige Lösung hat, wenn  $A$  invertierbar ist; diese Lösung ist  $\mathbf{x} = A^{-1}\mathbf{b}$ . (Die Umkehrung gilt ebenfalls — gibt es für jedes  $\mathbf{b}$  eine eindeutige Lösung, dann ist  $A$  invertierbar — Übung.) Wenn wir für  $\mathbf{b}$  den Standard-Basisvektor  $\mathbf{e}_j$  einsetzen, dann bekommen wir als Lösung gerade die  $j$ -te Spalte von  $A^{-1}$ . Wir können also  $A^{-1}$  finden, indem wir die linearen Gleichungssysteme  $A\mathbf{x} = \mathbf{e}_j$  für  $j \in \{1, 2, \dots, n\}$  alle lösen. Dies geht im Wesentlichen in einem Rutsch, wie im nächsten Satz beschrieben wird. Zuerst aber noch ein Lemma, das auch für sich interessant ist.

11.19. **Lemma.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $A \in \text{Mat}(n, K)$  eine quadratische Matrix. Dann ist  $A$  genau dann invertierbar, wenn ihre reduzierte Zeilenstufenform die Einheitsmatrix  $I_n$  ist.

**LEMMA**  
ZSF einer  
invertierbaren  
Matrix

*Beweis.*  $A$  ist genau dann invertierbar, wenn die zugehörige lineare Abbildung  $f : K^n \rightarrow K^n$  ein Isomorphismus ist. Da Definitions- und Wertebereich dieselbe Dimension haben, ist das dazu äquivalent, dass  $f$  surjektiv ist, also Rang  $n$  hat. Das bedeutet, dass es in der reduzierten Zeilenstufenform von  $A$  keine Null-Zeile gibt, also ist  $r = n$  und  $j_1 = 1, j_2 = 2, \dots, j_n = n$ . In der  $j$ -ten Spalte steht also der  $j$ -te Standard-Basisvektor, und die Matrix ist die Einheitsmatrix. □

\* 11.20. **Satz.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $A \in \text{Mat}(n, K)$  eine quadratische Matrix. Sei weiter  $A' = (A \mid I_n) \in \text{Mat}(n \times 2n, K)$  und  $\tilde{A}'$  ihre reduzierte Zeilenstufenform.  $A$  ist genau dann invertierbar, wenn  $\tilde{A}'$  die Form  $(I_n \mid B)$  hat; in diesem Fall ist  $B = A^{-1}$ .

**SATZ**  
Berechnung  
von  $A^{-1}$

*Beweis.* Sei  $\tilde{A}' = (\tilde{A} \mid B)$ , dann ist  $\tilde{A}$  die reduzierte Zeilenstufenform von  $A$ . Nach Lemma 11.19 ist  $A$  invertierbar genau dann, wenn  $\tilde{A} = I_n$  ist. Die Matrix  $A'$  repräsentiert das Gleichungssystem  $A(\mathbf{x}_1 \mid \mathbf{x}_2 \mid \cdots \mid \mathbf{x}_n) = (\mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_n)$  oder kurz  $AX = I_n$  mit  $X = (\mathbf{x}_1 \mid \mathbf{x}_2 \mid \cdots \mid \mathbf{x}_n) \in \text{Mat}(n, K)$ . Die Zeilenumformungen, die zur reduzierten Zeilenstufenform führen, ergeben das dazu äquivalente Gleichungssystem  $I_n X = B$ , also ist  $X = B$  die Lösung von  $AX = I_n$ ; damit ist  $B = A^{-1}$ .  $\square$

11.21. **Beispiel.** Sei  $K = \mathbb{Q}$  und

$$A = \begin{pmatrix} -1 & -2 & 2 & 2 \\ 2 & 5 & -4 & -4 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}.$$

**BSP**  
Berechnung  
von  $A^{-1}$

Wir überführen

$$A' = \begin{pmatrix} -1 & -2 & 2 & 2 & 1 & 0 & 0 & 0 \\ 2 & 5 & -4 & -4 & 0 & 1 & 0 & 0 \\ -1 & -1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

in reduzierte Zeilenstufenform:

$$\begin{aligned} A' &\longrightarrow \begin{pmatrix} \mathbf{1} & 2 & -2 & -2 & -1 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 3 & 1 & 0 & 0 & \mathbf{1} \end{pmatrix} \\ &\longrightarrow \begin{pmatrix} \mathbf{1} & 0 & -2 & -2 & -5 & -2 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & -3 & -1 & 1 & 0 \\ 0 & 0 & 2 & 3 & 3 & 1 & 0 & \mathbf{1} \end{pmatrix} \\ &\longrightarrow \begin{pmatrix} \mathbf{1} & 0 & 0 & 0 & 1 & 0 & -2 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 1 & 3 & 1 & -1 & 0 \\ 0 & 0 & 0 & \mathbf{1} & -3 & -1 & 2 & \mathbf{1} \end{pmatrix} \\ &\longrightarrow \begin{pmatrix} \mathbf{1} & 0 & 0 & 0 & 1 & 0 & -2 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 & 6 & 2 & -3 & -1 \\ 0 & 0 & 0 & \mathbf{1} & -3 & -1 & 2 & \mathbf{1} \end{pmatrix} \end{aligned}$$

Es folgt

$$A^{-1} = \begin{pmatrix} 1 & 0 & -2 & 0 \\ 2 & 1 & 0 & 0 \\ 6 & 2 & -3 & -1 \\ -3 & -1 & 2 & 1 \end{pmatrix}.$$

♣

12. MATRIZEN UND LINEARE ABBILDUNGEN

Wir haben bisher Matrizen als zu linearen Abbildungen  $K^n \rightarrow K^m$  gehörend betrachtet. Dabei war es aber eigentlich nur wichtig, dass wir in Definitions- und Wertebereich jeweils eine bestimmte Basis betrachten, in diesem Fall die Standard-Basis. Ganz analog können wir einer  $K$ -linearen Abbildung  $f : V \rightarrow V'$  eine Matrix zuordnen, wenn wir Basen  $B = (b_1, b_2, \dots, b_n)$  von  $V$  und  $B' = (b'_1, b'_2, \dots, b'_m)$  von  $V'$  fixieren. Es gibt dann nämlich eindeutig bestimmte Skalare  $a_{ij} \in K$ , sodass

$$f(b_j) = a_{1j}b'_1 + a_{2j}b'_2 + \dots + a_{mj}b'_m$$

für alle  $j \in \{1, 2, \dots, n\}$  gilt.

12.1. **Definition.** In der oben beschriebenen Situation heißt

**DEF**  
 $\text{Mat}_{B,B'}(f)$

$$\text{Mat}_{B,B'}(f) = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \in \text{Mat}(m \times n, K)$$

die *Matrix von  $f$  bezüglich der Basen  $B$  und  $B'$* .



Wie vorher auch enthält die  $j$ -te Spalte der Matrix die Koeffizienten des Bildes  $f(b_j)$  des  $j$ -ten Basisvektors in  $B$ , wenn es als Linearkombination der Basisvektoren in  $B'$  geschrieben wird.

12.2. **Beispiel.** Wir betrachten  $V = P_{<3}$ , den  $\mathbb{R}$ -Vektorraum der Polynomfunktionen vom Grad  $< 3$  und die lineare Abbildung  $D : V \rightarrow V, f \mapsto f'$ . Seien weiter  $B = (x \mapsto 1, x \mapsto x, x \mapsto x^2)$  und  $B' = (x \mapsto 1, x \mapsto x - 1, x \mapsto (x - 1)(x - 2))$  zwei Basen von  $V$ . Dann ist

**BSP**  
Matrix  
einer  
lin. Abb.

$$\begin{aligned} \text{Mat}_{B,B}(D) &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, & \text{Mat}_{B',B}(D) &= \begin{pmatrix} 0 & 1 & -3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \\ \text{Mat}_{B,B'}(D) &= \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} & \text{und} & \text{Mat}_{B',B'}(D) &= \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$



Wie wir sehen, kann ein und dieselbe lineare Abbildung durch viele verschiedene Matrizen beschrieben werden. Wie hängen diese miteinander zusammen? Dazu erst eine einfache Aussage über Verknüpfungen von linearen Abbildungen.

12.3. **Lemma.** Seien  $g : V \rightarrow V'$  und  $f : V' \rightarrow V''$  zwei  $K$ -lineare Abbildungen zwischen endlich-dimensionalen Vektorräumen. Seien weiter  $B$  eine Basis von  $V$ ,  $B'$  eine Basis von  $V'$  und  $B''$  eine Basis von  $V''$ . Dann gilt

**LEMMA**  
Matrix  
von  $f \circ g$

$$\text{Mat}_{B,B''}(f \circ g) = \text{Mat}_{B',B''}(f) \text{Mat}_{B,B'}(g).$$

*Beweis.* Das folgt aus der Definition der Matrixmultiplikation. □

Daraus ergibt sich sofort:

**12.4. Folgerung.** Sei  $f : V \rightarrow V'$  eine  $K$ -lineare Abbildung zwischen endlich-dimensionalen Vektorräumen. Seien  $B$  und  $\tilde{B}$  zwei Basen von  $V$  und  $B'$  und  $\tilde{B}'$  zwei Basen von  $V'$ . Dann ist

**FOLG**  
Basiswechsel

$$\text{Mat}_{\tilde{B}, \tilde{B}'}(f) = \text{Mat}_{B', \tilde{B}'}(\text{id}_{V'}) \text{Mat}_{B, B'}(f) \text{Mat}_{\tilde{B}, B}(\text{id}_V).$$

*Beweis.* Das folgt aus Lemma 12.3 und  $f = \text{id}_{V'} \circ f \circ \text{id}_V$ .  $\square$

Da  $\text{id}_V$  und  $\text{id}_{V'}$  Isomorphismen sind, sind die *Basiswechselmatrizen*  $\text{Mat}_{\tilde{B}, B}(\text{id}_V)$  und  $\text{Mat}_{B', \tilde{B}'}(\text{id}_{V'})$  invertierbar. Umgekehrt kann jede invertierbare Matrix als eine Basiswechselmatrix auftreten, wobei eine der beiden Basen beliebig vorgegeben werden kann.

**12.5. Lemma.** Sei  $K$  ein Körper, sei  $n \in \mathbb{N}$ , sei  $V$  ein  $K$ -Vektorraum mit Basis  $B = (b_1, b_2, \dots, b_n)$ . Sei weiter  $A \in \text{Mat}(n, K)$  invertierbar. Dann gibt es Basen  $B'$  und  $B''$  von  $V$ , sodass

**LEMMA**  
Basiswechsel-  
matrizen

$$A = \text{Mat}_{B, B'}(\text{id}_V) = \text{Mat}_{B'', B}(\text{id}_V)$$

ist.

*Beweis.* Sei  $A = (a_{ij})$  und  $B'' = (b''_1, b''_2, \dots, b''_n)$ . Die Aussage  $A = \text{Mat}_{B'', B}(\text{id}_V)$  bedeutet  $b''_j = a_{1j}b_1 + \dots + a_{nj}b_n$ . Wir definieren  $b''_j$  durch diese Gleichung für  $j \in \{1, 2, \dots, n\}$ ; dann gilt die gewünschte Aussage ( $B''$  ist eine Basis, weil  $A$  invertierbar ist: die  $b_i$  lassen sich als Linearkombinationen der  $b''_j$  ausdrücken, deren Koeffizienten die Einträge von  $A^{-1}$  sind).

Es gibt dann auch eine Basis  $B'$ , sodass  $A^{-1} = \text{Mat}_{B', B}(\text{id}_V)$  ist; damit folgt  $A = \text{Mat}_{B, B'}(\text{id}_V)$ , denn

$$\text{Mat}_{B, B'}(\text{id}_V) \text{Mat}_{B', B}(\text{id}_V) = \text{Mat}_{B', B'}(\text{id}_V) = I_n$$

nach Lemma 12.3.  $\square$

**12.6. Satz.** Sei  $K$  ein Körper und  $n \in \mathbb{N}$ . Die Menge der invertierbaren Matrizen in  $\text{Mat}(n, K)$  bildet eine Gruppe unter der Matrixmultiplikation.

**SATZ**  
Gruppe der  
invertierbaren  
Matrizen

*Beweis.* Die Matrixmultiplikation ist assoziativ, die (invertierbare) Einheitsmatrix  $I_n$  ist neutrales Element. Jede invertierbare Matrix hat per definitionem eine (selbst invertierbare) Inverse. Es bleibt zu zeigen, dass die Verknüpfung wohldefiniert ist, d.h., dass das Produkt zweier invertierbarer Matrizen wieder invertierbar ist. Das folgt aus

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I_n = (B^{-1}A^{-1})(AB);$$

die Inverse von  $AB$  ist also  $B^{-1}A^{-1}$ .  $\square$

**12.7. Definition.** Die Gruppe der invertierbaren Matrizen in  $\text{Mat}(n, K)$  heißt *allgemeine lineare Gruppe* und wird mit  $\text{GL}(n, K)$  bezeichnet.  $\diamond$

**DEF**  
 $\text{GL}(n, K)$

Die Abkürzung „GL“ kommt von englisch *general linear group*. Es ist auch die Notation  $\text{GL}_n(K)$  gebräuchlich.

- \* **12.8. Satz.** Sei  $K$  ein Körper, seien  $m, n \in \mathbb{N}$ , sei  $V$  ein  $n$ -dimensionaler und  $V'$  ein  $m$ -dimensionaler  $K$ -Vektorraum und sei  $f : V \rightarrow V'$  linear. Seien weiter  $B$  eine Basis von  $V$  und  $B'$  eine Basis von  $V'$  und  $A = \text{Mat}_{B, B'}(f)$ . Dann gilt: Die Menge der Matrizen von  $f$  bezüglich beliebiger Basen von  $V$  und  $V'$  ist genau

$$\{PAQ \mid P \in \text{GL}(m, K), Q \in \text{GL}(n, K)\}.$$

**SATZ**  
Matrizen  
derselben  
lin. Abb.

*Beweis.* Nach Folgerung 12.4 und der nachfolgenden Diskussion hat jede Matrix von  $f$  die Form  $PAQ$  mit invertierbaren Matrizen  $P$  und  $Q$ . Nach Lemma 12.5 gibt es zu beliebig vorgegebenen invertierbaren Matrizen  $P \in \text{GL}(m, K)$  und  $Q \in \text{GL}(n, K)$  Basen  $\tilde{B}$  von  $V$  und  $\tilde{B}'$  von  $V'$ , sodass  $P = \text{Mat}_{B', \tilde{B}'}(\text{id}_{V'})$  und  $Q = \text{Mat}_{\tilde{B}, B}(\text{id}_V)$ . Dann ist

$$PAQ = \text{Mat}_{B', \tilde{B}'}(\text{id}_{V'}) \text{Mat}_{B, B'}(f) \text{Mat}_{\tilde{B}, B}(\text{id}_V) = \text{Mat}_{\tilde{B}, \tilde{B}'}(f)$$

auch eine Matrix von  $f$ . □

- \* **12.9. Definition.** Seien  $K$  ein Körper,  $m, n \in \mathbb{N}$  und  $A, B \in \text{Mat}(m \times n, K)$ . Die Matrizen  $A$  und  $B$  heißen *äquivalent*, wenn es Matrizen  $P \in \text{GL}(m, K)$  und  $Q \in \text{GL}(n, K)$  gibt mit  $PAQ = B$ . ◇

**DEF**  
Äquivalenz  
von Matrizen

Zwei Matrizen in  $\text{Mat}(m \times n, K)$  sind also genau dann äquivalent, wenn sie dieselbe lineare Abbildung (aber evtl. bezüglich verschiedener Basen) repräsentieren.

- \* **12.10. Satz.** Seien  $K$  ein Körper,  $m, n \in \mathbb{N}$  und  $A, B \in \text{Mat}(m \times n, K)$ . Dann sind  $A$  und  $B$  äquivalent genau dann, wenn  $\text{rk}(A) = \text{rk}(B)$ . In diesem Fall sei  $r = \text{rk}(A)$ ; dann sind beide Matrizen äquivalent zur Matrix

**SATZ**  
Klassifikation  
von Matrizen  
bis auf  
Äquivalenz

$$M_r = \left( \begin{array}{c|c} I_r & \mathbf{0}_{r, n-r} \\ \hline \mathbf{0}_{m-r, r} & \mathbf{0}_{m-r, n-r} \end{array} \right).$$

Dabei steht  $\mathbf{0}_{k, l}$  für eine Nullmatrix mit  $k$  Zeilen und  $l$  Spalten.

*Beweis.* Sei  $r = \text{rk}(A)$ . Wir zeigen, dass  $A$  zu  $M_r$  äquivalent ist. Sei  $f : K^n \rightarrow K^m$  die zugehörige lineare Abbildung; sie hat Rang  $r$ , also ist  $\dim \ker(f) = n - r$ . Wir wählen eine Basis  $B = (b_1, \dots, b_n)$  von  $K^n$ , sodass  $(b_{r+1}, \dots, b_n)$  eine Basis von  $\ker(f)$  ist. Mit  $b'_i = f(b_i)$  für  $i \in \{1, 2, \dots, r\}$  ist dann  $(b'_1, \dots, b'_r)$  eine Basis von  $\text{im}(f)$  (vergleiche den Beweis von Satz 9.18). Wir ergänzen sie zu einer Basis  $B' = (b'_1, \dots, b'_m)$  von  $K^m$ . Dann ist  $\text{Mat}_{B, B'}(f)$  gerade  $M_r$ ;  $M_r$  ist damit äquivalent zu  $A$ .

Gilt auch  $\text{rk}(B) = r$ , dann ist  $B$  ebenfalls äquivalent zu  $M_r$ . Es folgt, dass  $A$  und  $B$  äquivalent sind: Es gibt  $P, P' \in \text{GL}(m, K)$  und  $Q, Q' \in \text{GL}(n, K)$  mit  $M_r = PAQ = P'BQ'$ . Dann ist  $B = (P'^{-1}P)A(QQ'^{-1})$ .

Umgekehrt gilt  $\text{rk}(B) = r = \text{rk}(A)$  für jede zu  $A$  äquivalente Matrix, denn der Rang einer Matrix ist gleich dem Rang jeder von ihr repräsentierten linearen Abbildung. □

Mit den Resultaten aus dem Kleingedruckten von Seite 81 ergibt sich aus Satz 12.10:

**Folgerung.** Jede Matrix  $A \in \text{Mat}(m \times n, K)$  lässt sich durch Zeilen- und Spaltenumformungen in die Matrix  $M_r$  mit  $r = \text{rk}(A)$  überführen.

**FOLG**  
Zeilen- und  
Spaltenumf.

Die Äquivalenz von Matrizen ist ein Beispiel einer Äquivalenzrelation. Eine Relation  $R$  zwischen Mengen  $X$  und  $Y$  ist formal eine Teilmenge  $R \subset X \times Y$ . Ist für  $x \in X$  und  $y \in Y$  das Paar  $(x, y)$  ein Element von  $R$ , dann sagt man,  $x$  und  $y$  stehen in der Relation  $R$  zueinander und schreibt auch  $x R y$  oder ähnlich. Im Fall  $X = Y$  spricht man auch von einer Relation auf  $X$ . Eine solche Relation heißt

- reflexiv, wenn  $\forall x \in X : x R x$ ,
- symmetrisch, wenn  $\forall x, y \in X : x R y \Rightarrow y R x$ , und
- transitiv, wenn  $\forall x, y, z \in X : (x R y \wedge y R z) \Rightarrow x R z$ .

Eine Relation auf  $X$ , die reflexiv, symmetrisch und transitiv ist, ist eine Äquivalenzrelation auf  $X$ . Beispiele sind die Gleichheitsrelation  $x = y$  (das ist die „feinste“ Äquivalenzrelation auf  $X$ ) oder auch die „Allrelation“  $R = X \times X$  (die „größte“ Äquivalenzrelation auf  $X$ ).

**Lemma.** Für Matrizen  $A, B \in \text{Mat}(m \times n, K)$  schreiben wir  $A \sim B$ , wenn  $A$  und  $B$  äquivalent sind, wenn es also  $P \in \text{GL}(m, K)$  und  $Q \in \text{GL}(n, K)$  gibt mit  $B = PAQ$ .

**LEMMA**  
Äquivalenz  
von Matrizen  
ist Äqu.rel.

Die Relation  $\sim$  ist eine Äquivalenzrelation auf  $\text{Mat}(m \times n, K)$ .

*Beweis.* Wir müssen die drei Eigenschaften nachprüfen.

- Reflexivität:  $A \sim A$ , denn man kann  $P = I_m$ ,  $Q = I_n$  wählen.
- Symmetrie: Es gelte  $A \sim B$ ; dann gibt es  $P \in \text{GL}(m, K)$  und  $Q \in \text{GL}(n, K)$  mit  $B = PAQ$ . Dann sind auch  $P^{-1} \in \text{GL}(m, K)$  und  $Q^{-1} \in \text{GL}(n, K)$  und es gilt  $B = P^{-1}A Q^{-1}$ , also  $B \sim A$ .
- Transitivität: Es gelte  $A \sim B$  und  $B \sim C$ . Dann gibt es  $P_1, P_2 \in \text{GL}(m, K)$  und  $Q_1, Q_2 \in \text{GL}(n, K)$  mit  $B = P_1A Q_1$  und  $C = P_2B Q_2$ . Es sind  $P_2P_1 \in \text{GL}(m, K)$  und  $Q_1Q_2 \in \text{GL}(n, K)$  und es gilt  $C = (P_2P_1)A(Q_1Q_2)$ , also ist  $A \sim C$ .  $\square$

Die wichtigste Eigenschaft einer Äquivalenzrelation auf einer Menge  $X$  ist, dass sie zu einer Einteilung von  $X$  in sogenannte Äquivalenzklassen führt. Ist  $\sim$  eine Äquivalenzrelation auf  $X$  und  $x \in X$ , dann schreiben wir  $[x]$  für die Menge  $\{y \in X \mid x \sim y\}$  der zu  $x$  äquivalenten Elemente von  $X$  und nennen  $[x]$  die Äquivalenzklasse von  $x$ . Jedes Element von  $[x]$  heißt ein Repräsentant der Äquivalenzklasse.

**Lemma.** Sei  $\sim$  eine Äquivalenzrelation auf einer Menge  $X$  und sei  $x \in X$ . Dann sind für  $y \in X$  die folgenden Aussagen äquivalent:

**LEMMA**  
Eigensch.  
Äqu.rel.

- (1)  $x \sim y$ .
- (2)  $y \in [x]$ .
- (3)  $[y] \cap [x] \neq \emptyset$ .
- (4)  $[y] = [x]$ .

Insbesondere sind zwei Äquivalenzklassen  $[x]$  und  $[y]$  entweder gleich oder disjunkt.

*Beweis.* Die Äquivalenz von (1) und (2) folgt aus der Definition von  $[x]$ .

„(2)  $\Rightarrow$  (3)“: Wegen der Reflexivität von  $\sim$  ist  $y \in [y]$ , also folgt aus  $y \in [x]$ , dass  $y \in [y] \cap [x]$ .

„(3)  $\Rightarrow$  (4)“: Sei  $z \in [y] \cap [x]$  und  $w \in [y]$ . Dann gilt  $y \sim w$ ,  $y \sim z$  und  $x \sim z$ ; mit Symmetrie und Transitivität von  $\sim$  folgt daraus  $x \sim w$ , also  $w \in [x]$ . Da  $w$  beliebig war, gilt  $[y] \subset [x]$ . Genauso erhalten wir  $[x] \subset [y]$ .

„(4)  $\Rightarrow$  (2)“: Aus  $y \in [y]$  und  $[y] = [x]$  folgt  $y \in [x]$ . □

Wir können die *Menge der Äquivalenzklassen*  $X/\sim = \{[x] \mid x \in X\}$  bilden. Dann gibt es eine natürliche (oder „kanonische“) surjektive Abbildung  $f : X \rightarrow X/\sim, x \mapsto [x]$ . Die Urbildmenge  $f^{-1}(\{[x]\})$  ist nach dem gerade bewiesenen Lemma genau  $[x]$ . Umgekehrt führt jede surjektive Abbildung  $f : X \rightarrow M$  zu einer Äquivalenzrelation auf  $X$  (man sagt auch,  $f$  induziert eine Äquivalenzrelation) durch  $x \sim y \iff f(x) = f(y)$ .

Die Aussage von Satz 12.10 bedeutet dann, dass die Äquivalenz von  $m \times n$ -Matrizen mit der durch  $\text{Mat}(m \times n, K) \rightarrow \{1, 2, \dots, \min\{m, n\}\}, A \mapsto \text{rk}(A)$ , induzierten Äquivalenzrelation übereinstimmt und dass  $M_r$  ein Repräsentant der durch  $\text{rk}(A) = r$  gegebenen Äquivalenzklasse ist.

13. DIE DETERMINANTE

In diesem Abschnitt führen wir die *Determinante* einer quadratischen Matrix ein. Das ist ein Skalar, der darüber Auskunft gibt, ob die Matrix invertierbar ist oder nicht. Wir definieren die Determinante rekursiv.

\*

**13.1. Definition.** Sei  $K$  ein Körper. Für  $A = (a_{ij}) \in \text{Mat}(n, K)$  mit  $n > 0$  sei  $A_{ij} \in \text{Mat}(n - 1, K)$  (für  $i, j \in \{1, 2, \dots, n\}$ ) die Matrix, die aus  $A$  entsteht, wenn man die  $i$ -te Zeile und die  $j$ -te Spalte entfernt. Wir definieren die *Determinante* von  $A$ ,  $\det(A)$  rekursiv wie folgt:

**DEF**  
Determinante  
einer Matrix

- (1) Im Fall  $n = 0$  ist  $\det(A) = 1$ .
- (2) Im Fall  $n > 0$  ist

$$\det(A) = \sum_{j=1}^n (-1)^{j-1} a_{1j} \det(A_{1j}).$$

Für die Determinante ist auch folgende Schreibweise üblich:

$$\det((a_{ij})) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad \diamond$$

**13.2. Beispiele.** Für kleine positive Werte von  $n$  erhalten wir folgende Formeln:

**BSP**  
Determinante

$$\begin{aligned} \det((a)) &= a \\ \begin{vmatrix} a & b \\ c & d \end{vmatrix} &= ad - bc \\ \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= aei - afh + bfg - bdi + cdh - ceg \end{aligned}$$

Die Formel für die  $3 \times 3$ -Determinante lässt sich mit Hilfe der „Sarrus-Regel“ merken: Man schreibt die ersten beiden Spalten noch einmal hinter die Matrix und bildet die Summe der Produkte über die nach rechts fallenden Diagonalen minus die Summe der Produkte über die nach rechts steigenden Diagonalen.

$$\begin{vmatrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{vmatrix}$$

Für größere Determinanten gibt es allerdings keine solche Merkregel!



Welche Eigenschaften hat die Determinante?

**13.3. Satz.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $A \in \text{Mat}(n, K)$ .

**SATZ**  
Eigensch.  
der Det.

- (1)  $\det(A)$  ist linear als Funktion jeder Zeile von  $A$  (dabei werden die Einträge der übrigen Zeilen als fest angesehen).
- (2) Hat  $A$  zwei gleiche Zeilen, dann ist  $\det(A) = 0$ .
- (3)  $\det(I_n) = 1$ .

- (4) Geht  $A'$  aus  $A$  durch Vertauschen zweier Zeilen hervor, dann gilt  $\det(A') = -\det(A)$ .
- (5) Führt man eine elementare Zeilenumformung  $\mathbf{I}_i(\lambda)$  an  $A$  aus, dann multipliziert sich  $\det(A)$  mit  $\lambda$ .
- (6) Führt man eine elementare Zeilenumformung  $\mathbf{II}_{i,j}(\lambda)$  an  $A$  aus, dann ändert sich  $\det(A)$  nicht.
- (7) Es gilt  $\det(A) \neq 0 \iff \text{rk}(A) = n \iff A$  invertierbar.
- (8) Ist  $d : \text{Mat}(n, K) \rightarrow K$  eine Abbildung, die die Eigenschaften (1) und (2) hat, dann gilt  $d(A) = \det(A)d(I_n)$  für alle  $A \in \text{Mat}(n, K)$ . Insbesondere ist  $\det : \text{Mat}(n, K) \rightarrow K$  die einzige Abbildung, die (1), (2) und (3) erfüllt.

*Beweis.* Der Beweis der ersten drei Aussagen erfolgt durch Induktion über  $n$ . Die verbleibenden Aussagen folgen aus den ersten drei. Im Fall  $n = 0$  sind die ersten drei Aussagen trivial. Sei also  $n > 0$  und die Aussagen seien für kleinere Werte von  $n$  richtig.

- (1)  $\det(A)$  ist linear in der ersten Zeile von  $A$ , denn nach Definition ist  $\det(A)$  eine Linearkombination der Einträge der ersten Zeile, deren Koeffizienten nicht von der ersten Zeile abhängen. Sei  $k \in \{1, 2, \dots, n-1\}$ . Nach Induktionsannahme sind alle  $\det(A_{1j})$  linear in der  $k$ -ten Zeile von  $A_{1j}$  und damit linear in der  $(k+1)$ -ten Zeile von  $A$ .  $\det(A)$  ist somit eine Linearkombination von Abbildungen, die linear als Funktion der  $(k+1)$ -ten Zeile von  $A$  sind (mit Koeffizienten, die nicht von der  $(k+1)$ -ten Zeile abhängen) und somit ebenfalls linear in der  $(k+1)$ -ten Zeile von  $A$ .
- (2) Sei  $A$  eine Matrix, in der die  $k$ -te und die  $l$ -te Zeile übereinstimmen, wobei  $1 \leq k < l \leq n$ . Ist  $k > 1$ , dann stimmt in jeder Matrix  $A_{1j}$  die  $(k-1)$ -te mit der  $(l-1)$ -ten Zeile überein; nach Induktionsannahme gilt  $\det(A_{1j}) = 0$  für alle  $j$ , also auch  $\det(A) = 0$ . Es bleibt der Fall  $k = 1$  zu betrachten. Falls  $l > 2$  ist, dann vertauschen wir die  $l$ -te mit der zweiten Zeile. Nach Induktionsannahme (Teil (4)) bewirkt das einen Vorzeichenwechsel in allen  $\det(A_{1j})$ , ändert also nichts daran, ob  $\det(A) = 0$  ist oder nicht. Wir können also annehmen, dass die beiden ersten Zeilen von  $A$  gleich sind. Wir schreiben  $d_{jk} = d_{kj}$  für die Determinante der Matrix, die aus  $A$  durch Streichen der ersten beiden Zeilen und der Spalten  $j$  und  $k$  entsteht. Dann gilt (unter Beachtung von  $a_{2k} = a_{1k}$ )

$$\begin{aligned}
 \det(A) &= \sum_{j=1}^n (-1)^{j-1} a_{1j} \det(A_{1j}) \\
 &= \sum_{j=1}^n (-1)^{j-1} a_{1j} \left( \sum_{k=1}^{j-1} (-1)^{k-1} a_{2k} d_{jk} + \sum_{k=j+1}^n (-1)^k a_{2k} d_{jk} \right) \\
 &= \sum_{1 \leq k < j \leq n} (-1)^{j-k} a_{1j} a_{1k} d_{jk} + \sum_{1 \leq j < k \leq n} (-1)^{k-j-1} a_{1j} a_{1k} d_{jk} \\
 &= \sum_{1 \leq j < k \leq n} (-1)^{k-j} a_{1j} a_{1k} d_{jk} + \sum_{1 \leq j < k \leq n} (-1)^{k-j-1} a_{1j} a_{1k} d_{jk} \\
 &= \sum_{1 \leq j < k \leq n} ((-1)^{k-j} + (-1)^{k-j-1}) a_{1j} a_{1k} d_{jk} \\
 &= 0.
 \end{aligned}$$

(Wir haben in der ersten Summe  $j$  und  $k$  vertauscht und dabei ausgenutzt, dass  $a_{1k}a_{1j}d_{kj} = a_{1j}a_{1k}d_{jk}$  ist.)

- (3) Nach der rekursiven Definition ist  $\det(I_n) = 1 \cdot \det(I_{n-1}) = 1$ .
- (4) Wir schreiben  $d(v_1, \dots, v_n)$  für die Determinante der Matrix, deren Zeilen  $v_1, \dots, v_n \in K^n$  sind; für  $1 \leq i < j \leq n$  sei  $d_{ij}(v_i, v_j) = d(v_1, \dots, v_n)$ , wobei die  $v_k$  mit  $k \notin \{i, j\}$  fest gewählt sind. Dann gilt

$$\begin{aligned} 0 &\stackrel{(2)}{=} d_{ij}(v_i + v_j, v_i + v_j) \\ &\stackrel{(1)}{=} d_{ij}(v_i, v_i) + d_{ij}(v_i, v_j) + d_{ij}(v_j, v_i) + d_{ij}(v_j, v_j) \\ &\stackrel{(2)}{=} d_{ij}(v_i, v_j) + d_{ij}(v_j, v_i), \end{aligned}$$

also ist  $d_{ij}(v_j, v_i) = -d_{ij}(v_i, v_j)$ .

- (5) Das ist ein Spezialfall von Teil (1).
- (6) In der Notation des Beweises von Teil (4) haben wir

$$d_{ij}(v_i + \lambda v_j, v_j) \stackrel{(1)}{=} d_{ij}(v_i, v_j) + \lambda d_{ij}(v_j, v_j) \stackrel{(2)}{=} d_{ij}(v_i, v_j).$$

- (7) Aus den Teilen (4), (5) und (6) folgt, dass  $\det(A)$  genau dann null ist, wenn  $\det(A')$  null ist, wobei  $A'$  die reduzierte Zeilenstufenform von  $A$  ist. Gilt  $\text{rk}(A) = n$ , dann ist  $A$  invertierbar, und nach Lemma 11.19 ist  $A' = I_n$  und damit  $\det(A') = \det(I_n) = 1 \neq 0$  nach Teil (3). Gilt  $\text{rk}(A) < n$ , dann hat  $A'$  eine Null-Zeile und damit ist  $\det(A') = 0$  nach Teil (1). Die zweite Äquivalenz folgt daraus, dass eine lineare Abbildung  $K^n \rightarrow K^n$  genau dann surjektiv ist, wenn sie ein Isomorphismus ist, vgl. Folgerung 9.14.
- (8) Aus (1) und (2) folgen (4), (5) und (6). Daraus folgt, dass  $d(A) = d_0(A)d(I_n)$  ist, wobei  $d_0(A)$  nur von  $A$  und nicht von  $d$  abhängt —  $d_0(A) = 0$ , wenn die reduzierte Zeilenstufenform  $A'$  von  $A$  eine Null-Zeile hat; sonst ist  $d_0(A)$  der Faktor, der aus den elementaren Zeilenumformungen  $\mathbf{I}_i(\lambda)$  und  $\mathbf{III}_{i,j}$  herrührt, die man ausführt, um von  $A$  zu  $A' = I_n$  zu gelangen. Der Spezialfall  $d = \det$  liefert  $d_0(A) = \det(A)$  und damit die Behauptung.  $\square$

Wenn wir die Determinante einer  $n \times n$ -Matrix  $A$  als Funktion der  $n$  Zeilen von  $A$  betrachten, die selbst Vektoren in  $K^n$  sind, dann erhalten wir eine sogenannte *alternierende Multilinearform*. Das ist ein Spezialfall einer multilinearen Abbildung.

**Definition.** Sei  $K$  ein Körper und seien  $V_1, V_2, \dots, V_m$  und  $W$   $K$ -Vektorräume. Eine Abbildung  $f : V_1 \times V_2 \times \dots \times V_m \rightarrow W$  heißt *multilinear*, wenn  $f$  in jedem Argument  $K$ -linear ist, wenn also gilt

$$f(v_1, \dots, v_{i-1}, \lambda v_i, v_{i+1}, \dots, v_m) = \lambda f(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_m)$$

und

$$\begin{aligned} f(v_1, \dots, v_{i-1}, v_i + v'_i, v_{i+1}, \dots, v_m) \\ = f(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_m) + f(v_1, \dots, v_{i-1}, v'_i, v_{i+1}, \dots, v_m) \end{aligned}$$

für alle  $i \in \{1, 2, \dots, m\}$ ,  $v_j \in V_j$ ,  $\lambda \in K$ ,  $v'_i \in V_i$ . Ist  $W = K$ , dann heißt  $f$  auch eine *Multilinearform*.

Eine Multilinearform  $f : V^m = V \times V \times \dots \times V \rightarrow K$  heißt *alternierend*, wenn  $f(v_1, \dots, v_m) = 0$  ist, sobald  $v_i = v_j$  ist für gewisse  $1 \leq i < j \leq m$ .  $\diamond$

Aussagen (1) und (2) in Satz 13.3 besagen gerade, dass  $\det(A)$  eine alternierende Multilinearform der Zeilen von  $A$  ist. Da man einen  $K$ -Vektorraum  $V$  mit Basis  $(b_1, \dots, b_n)$  mit  $K^n$  identifizieren kann, hat Aussage (8) in Satz 13.3 die folgende Interpretation:

**DEF**  
multilineare  
Abbildung  
alternierende  
Multilinear-  
form

**Satz.** Sei  $V$  ein  $K$ -Vektorraum mit Basis  $(b_1, b_2, \dots, b_n)$ . Dann gibt es genau eine alternierende Multilinearform  $d : V^n \rightarrow K$  mit  $d(b_1, b_2, \dots, b_n) = 1$ .

**SATZ**  
Existenz u.  
Eindeutigkeit  
alternierender  
Multilinear-  
formen

Für praktische Zwecke wichtig sind die Aussagen in Satz 13.3, die zeigen, wie sich die Determinante unter elementaren Zeilenumformungen verhält. Das liefert ein praktisches Verfahren zur Berechnung auch größerer Determinanten.

13.4. **Beispiel.**

$$\begin{vmatrix} 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 3 & 3 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 6 & 6 \end{vmatrix} \\ = 2 \cdot \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 6 \end{vmatrix} = 2 \cdot 6 \cdot \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 12 \quad \clubsuit$$

**BSP**  
Determi-  
nanten-  
berechnung

Die Eindeutigkeitsaussage (8) in Satz 13.3 ist wichtig, weil sie weitere Eigenschaften der Determinante zur Folge hat.

\* 13.5. **Satz.** Sei  $K$  ein Körper, sei  $n > 0$  und  $A = (a_{ij}) \in \text{Mat}(n, K)$ . Mit der in Definition 13.1 eingeführten Schreibweise  $A_{ij}$  gilt für jedes  $i \in \{1, 2, \dots, n\}$ :

**SATZ**  
Entwicklung  
der Det.  
nach der  
 $i$ -ten Zeile

$$\det(A) = \sum_{j=1}^n (-1)^{j-i} a_{ij} \det(A_{ij}).$$

*Beweis.* Wie im Beweis von Satz 13.3 zeigt man, dass die rechte Seite die Eigenschaften (1), (2) und (3) hat. Wegen der Eindeutigkeit folgt, dass die rechte Seite gleich  $\det(A)$  sein muss.  $\square$

13.6. **Beispiel.** Die Berechnung der Determinante in Beispiel 13.4 lässt sich vereinfachen, indem man nach der zweiten Zeile entwickelt:

**BSP**

$$\begin{vmatrix} 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{vmatrix} = - \begin{vmatrix} -1 & 1 & -1 \\ 1 & 1 & 1 \\ 2 & 4 & 8 \end{vmatrix} = \dots = 12 \quad \clubsuit$$

\* 13.7. **Satz.** Sei  $K$  ein Körper, sei  $n \in \mathbb{N}$  und seien  $A, B \in \text{Mat}(n, K)$ . Dann gilt

**SATZ**  
Multiplika-  
tivität  
der Det.

$$\det(AB) = \det(A) \det(B).$$

Ist  $A$  invertierbar, dann ist  $\det(A^{-1}) = \det(A)^{-1}$ .

*Beweis.* Wir fixieren  $B$  und betrachten  $A$  als variabel. Sei  $d_B : \text{Mat}(n, K) \rightarrow K$ ,  $A \mapsto \det(AB)$ . Aus den Eigenschaften der Matrixmultiplikation folgt, dass die  $k$ -te Zeile von  $AB$  nur von der  $k$ -ten Zeile von  $A$  abhängt und zwar linear. Es folgt, dass  $d_B$  linear in den Zeilen von  $A$  ist. Ebenso gilt, dass aus der Gleichheit der  $k$ -ten und der  $l$ -ten Zeile von  $A$  die entsprechende Aussage für  $AB$  folgt. Damit erfüllt  $d_B$  auch die Eigenschaft (2) in Satz 13.3. Die Eindeutigkeitsaussage in Satz 13.3

liefert nun  $\det(AB) = d_B(A) = \det(A)d_B(I_n) = \det(A)\det(B)$ . Die letzte Aussage ergibt sich aus  $\det(A)\det(A^{-1}) = \det(I_n) = 1$ .  $\square$

\* 13.8. **Satz.** *Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $A \in \text{Mat}(n, K)$ . Dann gilt*

$$\det(A^\top) = \det(A).$$

**SATZ**  
Symmetrie  
der Det.

*Beweis.* Wir müssen zeigen, dass  $\det(A^\top)$  die Eigenschaften (1), (2) und (3) aus Satz 13.3 hat.  $\det(I_n^\top) = \det(I_n) = 1$  ist klar. Die beiden anderen Aussagen sind dazu äquivalent, dass  $\det(A)$  linear in den *Spalten* von  $A$  ist und verschwindet, wenn  $A$  zwei gleiche *Spalten* hat. Die erste Aussage folgt leicht mit Induktion aus der rekursiven Definition der Determinante, denn für festes  $k$  ist jeder Term in der Summe linear in der  $k$ -ten Spalte von  $A$  (entweder durch  $a_{1k}$  oder durch  $\det(A_{1j})$ ). Die zweite Aussage kann wie folgt gezeigt werden: Wenn  $A$  zwei gleiche Spalten hat, dann ist  $\text{rk}(A) < n$ , also  $\det(A) = 0$  nach Satz 13.3, Teil (7).  $\square$

Daraus folgt zum Beispiel, dass man auch *Spaltenumformungen* bei der Berechnung der Determinante verwenden kann, auch mit *Zeilenumformungen* gemischt. Ebenso ergibt sich eine Formel zur Entwicklung der Determinante nach einer Spalte.

\* 13.9. **Folgerung.** *Sei  $K$  ein Körper, sei  $n > 0$  und  $A = (a_{ij}) \in \text{Mat}(n, K)$ . Mit der in Definition 13.1 eingeführten Schreibweise  $A_{ij}$  gilt für jedes  $j \in \{1, 2, \dots, n\}$ :*

$$\det(A) = \sum_{i=1}^n (-1)^{j-i} a_{ij} \det(A_{ij}).$$

**FOLG**  
Entwicklung  
der Det.  
nach der  
 $j$ -ten Spalte

*Beweis.* Das folgt aus Satz 13.5, angewandt auf  $A^\top$  und aus  $\det(A^\top) = \det(A)$ .  $\square$

13.10. **Beispiel.** Eine Matrix  $A \in \text{Mat}(n, \mathbb{R})$  mit  $AA^\top = I_n$  heißt *orthogonal*. Was kann man über  $\det(A)$  sagen?

**BSP**

Es gilt

$$1 = \det(I_n) = \det(AA^\top) = \det(A)\det(A^\top) = \det(A)^2,$$

also ist  $\det(A) = \pm 1$ . ♣

13.11. **Beispiel.** Wir berechnen die Determinante aus Beispiel 13.4 noch einmal.

**BSP**  
Determinanten-  
berechnung

$$\begin{aligned} \begin{vmatrix} 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{vmatrix} &= - \begin{vmatrix} -1 & 1 & -1 \\ 1 & 1 & 1 \\ 2 & 4 & 8 \end{vmatrix} = - \begin{vmatrix} -1 & 1 & 0 \\ 1 & 1 & 0 \\ 2 & 4 & 6 \end{vmatrix} = -6 \cdot \begin{vmatrix} -1 & 1 \\ 1 & 1 \end{vmatrix} \\ &= -6((-1) \cdot 1 - 1 \cdot 1) = 12 \end{aligned}$$

(Entwicklung nach der zweiten Zeile, elementare Spaltenumformung  $\mathbf{II}_{3,1}(-1)$ , Entwicklung nach der dritten Spalte, Formel für  $2 \times 2$ -Determinante). ♣

Die Entwicklung der Determinante nach Zeilen und Spalten führt zu folgender „Formel“ für die Inverse einer Matrix.

\* 13.12. **Definition.** Seien  $K$  ein Körper,  $n > 0$  und  $A \in \text{Mat}(n, K)$ . Die Matrix  $\tilde{A} \in \text{Mat}(n, K)$ , deren Eintrag in der  $i$ -ten Zeile und  $j$ -ten Spalte durch  $(-1)^{i-j} \det(A_{ji})$  (nicht  $A_{ij}$ !) gegeben ist, heißt die *adjungierte Matrix* zu  $A$ .  $\diamond$

**DEF**  
Adjungierte  
Matrix

\* 13.13. **Satz.** Seien  $K$  ein Körper,  $n > 0$  und  $A \in \text{Mat}(n, K)$ . Dann gilt

$$A\tilde{A} = \tilde{A}A = \det(A)I_n.$$

**SATZ**  
Adjungierte  
Matrix

Ist  $A$  invertierbar, dann ist  $A^{-1} = \det(A)^{-1}\tilde{A}$ .

*Beweis.* Der Eintrag an der Stelle  $(i, k)$  im Produkt  $A\tilde{A}$  ist

$$\sum_{j=1}^n a_{ij}(-1)^{j-k} \det(A_{kj}).$$

Im Fall  $k = i$  ergibt das  $\det(A)$  nach dem Satz 13.5 über die Entwicklung der Determinante nach der  $i$ -ten Zeile. Im Fall  $k \neq i$  ergibt sich analog die Determinante der Matrix, die aus  $A$  entsteht, wenn man die  $k$ -te Zeile durch die  $i$ -te ersetzt. Da diese Matrix zwei gleiche Zeilen hat, ist ihre Determinante null. Das zeigt  $A\tilde{A} = \det(A)I_n$ . Die Aussage  $\tilde{A}A = \det(A)I_n$  sieht man analog unter Verwendung von Folgerung 13.9. Die letzte Aussage folgt durch Multiplikation mit  $\det(A)^{-1}A^{-1}$ .  $\square$

13.14. **Beispiel.** Für  $n = 2$  bekommen wir die Formel

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

**BSP**  
Inverse einer  
 $2 \times 2$ -Matrix



Wir werden uns jetzt mit einer Verallgemeinerung der Formeln für die Determinante wie in Beispiel 13.2 beschäftigen. Diese Formeln erhält man aus der rekursiven Definition der Determinante wie in Definition 13.1. Das Resultat ist eine Summe von Termen der Form  $\pm a_{1,\sigma(1)}a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$ , wobei die Spaltenindizes  $\sigma(1), \sigma(2), \dots, \sigma(n)$  paarweise verschieden sind (denn jede „verbrauchte“ Spalte wird in der weiteren Entwicklung entfernt). Die Abbildung  $\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$  ist demnach bijektiv, also eine Permutation. Wie man sich leicht überlegt, kommt auch jede Permutation in der Entwicklung der Determinante vor. Wir erinnern uns daran, dass die Permutationen von  $\{1, 2, \dots, n\}$  die Elemente der *symmetrischen Gruppe*  $S_n$  sind; die Verknüpfung in dieser Gruppe ist die Komposition von Abbildungen. Damit haben wir Folgendes gezeigt:

\* 13.15. **Satz.** Seien  $K$  ein Körper und  $n \in \mathbb{N}$ . Dann gibt es eine Abbildung  $\varepsilon : S_n \rightarrow \{-1, 1\}$ , sodass für alle  $A = (a_{ij}) \in \text{Mat}(n, K)$  gilt

**SATZ**  
Leibniz-  
Formel

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}.$$

Diese Formel hat  $\#S_n = n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$  Terme und taugt damit außer für sehr kleine Werte von  $n$  nicht zur praktischen Berechnung der Determinante! Sie ist aber nützlich für theoretische Überlegungen. Zum Beispiel folgt sofort, dass die Determinante einer Matrix mit ganzzahligen Einträgen eine ganze Zahl ist.

**13.16. Definition.**  $\varepsilon(\sigma) \in \{-1, 1\}$  heißt das *Signum* oder *Vorzeichen* der Permutation  $\sigma \in S_n$ .  $\sigma$  heißt *gerade*, wenn  $\varepsilon(\sigma) = 1$  und *ungerade*, wenn  $\varepsilon(\sigma) = -1$  ist.

**DEF**  
Signum einer  
Permutation  
◇

Um etwas über diese Vorzeichenfunktion herauszufinden, führen wir die zu  $\sigma$  gehörende Permutationsmatrix ein.

**13.17. Definition.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $\sigma \in S_n$ . Dann bezeichnen wir mit  $P(\sigma)$  die Matrix  $(\delta_{\sigma(i),j})_{1 \leq i,j \leq n} \in \text{Mat}(n, K)$  und nennen  $P(\sigma)$  die *zu  $\sigma$  gehörende Permutationsmatrix*.

**DEF**  
Permutations-  
matrix  
◇

Die Einträge von  $P(\sigma)$  sind 1 an Positionen der Form  $(i, \sigma(i))$  und sonst 0.

**13.18. Lemma.** Seien  $K$  ein Körper und  $n \in \mathbb{N}$ .

**LEMMA**  
Eigensch.  
Permutations-  
matrix

- (1) Für  $\sigma, \tau \in S_n$  gilt  $P(\sigma \circ \tau) = P(\tau)P(\sigma)$ .
- (2) Für  $\sigma \in S_n$  gilt  $\varepsilon(\sigma) = \det(P(\sigma))$ .
- (3) Für  $\sigma, \tau \in S_n$  gilt  $\varepsilon(\sigma \circ \tau) = \varepsilon(\sigma)\varepsilon(\tau)$ .
- (4) Ist  $\sigma$  eine Transposition (also eine Permutation, die zwei Elemente vertauscht und alle anderen nicht ändert), dann ist  $\varepsilon(\sigma) = -1$ .

*Beweis.*

- (1) Der Eintrag an der Stelle  $(i, k)$  von  $P(\tau)P(\sigma)$  ist

$$\sum_{j=1}^n \delta_{\tau(i),j} \delta_{\sigma(j),k} = \begin{cases} 1, & \text{falls } k = \sigma(j) \text{ und } j = \tau(i), \\ 0, & \text{sonst,} \end{cases}$$

also  $\delta_{\sigma(\tau(i)),k}$ , genau wie in  $P(\sigma \circ \tau)$ .

- (2) Der einzige von null verschiedene Term in der Formel für  $\det(P(\sigma))$  ist  $\varepsilon(\sigma)\delta_{\sigma(1),\sigma(1)} \cdots \delta_{\sigma(n),\sigma(n)} = \varepsilon(\sigma)$ .
- (3) Das folgt aus (1) und (2) und der Multiplikativität der Determinante.
- (4) In diesem Fall erhält man  $P(\sigma)$  aus  $I_n$  durch Vertauschen zweier Zeilen (oder Spalten), also ist  $\varepsilon(\sigma) = \det(P(\sigma)) = -\det(I_n) = -1$ .  $\square$

Da sich (wie man sich leicht überlegen kann) jede Permutation als Komposition von Transpositionen schreiben lässt, ist  $\varepsilon$  durch die Eigenschaften (3) und (4) in Lemma 13.18 eindeutig festgelegt.

Es gibt eine Art Formel für  $\varepsilon(\sigma)$ . Dazu eine kleine Definition:

**Definition.** Sei  $\sigma \in S_n$ . Ein Paar  $(i, j)$  mit  $1 \leq i < j \leq n$  heißt *Fehlstand* von  $\sigma$ , wenn  $\sigma(i) > \sigma(j)$  ist.

**DEF**  
Fehlstand  
◇

Dann gilt der folgende Satz.

**Satz.** Sei  $\sigma \in S_n$  und sei  $m$  die Anzahl der Fehlstände von  $\sigma$ . Dann ist  $\varepsilon(\sigma) = (-1)^m$ .

**SATZ**  
Signum und  
Fehlstände

*Beweis.* Sei  $\varepsilon'(\sigma)$  die durch  $(-1)^{\text{Anzahl Fehlstände von } \sigma}$  definierte Funktion. Die Transposition  $\tau$ , die  $k$  und  $l$  vertauscht (mit  $k < l$ ), hat genau  $m = 1 + 2(l - k - 1)$  Fehlstände (nämlich  $(k, l)$  sowie  $(k, j)$  und  $(j, l)$  für alle  $k < j < l$ ). Da  $m$  ungerade ist, ist  $\varepsilon'(\tau) = (-1)^m = -1 = \varepsilon(\tau)$ .

Außerdem gilt

$$\varepsilon'(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i},$$

denn das rechts stehende Produkt hat Betrag 1 (jeder Faktor im Nenner tritt bis aufs Vorzeichen auch im Zähler auf) und  $\sigma(j) - \sigma(i)$  ist negativ genau dann, wenn  $(i, j)$  ein Fehlstand von  $\sigma$  ist. Es folgt

$$\begin{aligned} \varepsilon'(\sigma \circ \tau) &= \prod_{1 \leq i < j \leq n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} \\ &= \prod_{1 \leq i < j \leq n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \prod_{1 \leq i < j \leq n} \frac{\tau(j) - \tau(i)}{j - i} \\ &= \prod_{1 \leq k < l \leq n} \frac{\sigma(l) - \sigma(k)}{l - k} \prod_{1 \leq i < j \leq n} \frac{\tau(j) - \tau(i)}{j - i} = \varepsilon'(\sigma)\varepsilon'(\tau). \end{aligned}$$

Die Funktion  $\varepsilon'$  hat also die Eigenschaften (3) und (4) aus Lemma 13.18 und muss daher mit  $\varepsilon$  übereinstimmen.  $\square$

Hier ist ein Beispiel für eine allgemeine Formel für eine spezielle Determinante:

**Satz.** Sei  $K$  ein Körper, sei  $n \in \mathbb{N}$ , seien  $a_1, a_2, \dots, a_n \in K$  und sei  $A$  folgende „Vandermonde-Matrix“ zu  $a_1, a_2, \dots, a_n$ :

**SATZ**  
Vandermonde-  
Determinante

$$A = (a_i^{j-1})_{1 \leq i, j \leq n} = \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{pmatrix} \in \text{Mat}(n, K).$$

Dann ist

$$\det(A) = \prod_{1 \leq i < j \leq n} (a_j - a_i).$$

*Beweis.* Durch Induktion nach  $n$ . Der Fall  $n = 0$  ist klar ( $\det(A) = 1$  und das Produkt ist leer). Sei also  $n > 0$  und die Behauptung für  $n - 1$  bewiesen. Wir subtrahieren die erste Zeile von den übrigen und entwickeln nach der ersten Spalte; das liefert

$$\det(A) = \begin{vmatrix} a_2 - a_1 & a_2^2 - a_1^2 & \cdots & a_2^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ a_n - a_1 & a_n^2 - a_1^2 & \cdots & a_n^{n-1} - a_1^{n-1} \end{vmatrix}.$$

Nun ist  $x^m - y^m = (x - y)(x^{m-1} + x^{m-2}y + \dots + xy^{m-2} + y^{m-1})$ . Wir können aus der ersten, ...,  $(n - 1)$ -ten Zeile also jeweils einen Faktor  $(a_2 - a_1), \dots, (a_n - a_1)$  herausziehen:

$$\det(A) = \prod_{j=2}^n (a_j - a_1) \cdot \begin{vmatrix} 1 & a_2 + a_1 & \cdots & a_2^{n-2} + a_1 a_2^{n-3} + \cdots + a_1^{n-2} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & a_n + a_1 & \cdots & a_n^{n-2} + a_1 a_n^{n-3} + \cdots + a_1^{n-2} \end{vmatrix}.$$

Durch Subtraktion von  $a_1$ -mal der vorletzten Spalte von der letzten, dann Subtraktion von  $a_1$ -mal der drittletzten Spalte von der vorletzten, ...,  $a_1$ -mal der ersten Spalte

von der zweiten erhält man aus der verbliebenen Matrix die Vandermonde-Matrix zu  $a_2, a_3, \dots, a_n$ . Die Behauptung folgt aus der Induktionsvoraussetzung.  $\square$

Die Vandermonde-Matrix ist genau die Matrix der linearen Abbildung

$$\phi : P_{<n} \longrightarrow \mathbb{R}^n, \quad f \longmapsto (f(a_1), f(a_2), \dots, f(a_n))$$

(wo  $P_{<n}$  der  $\mathbb{R}$ -Vektorraum der Polynomfunktionen vom Grad  $< n$  ist), die bei der Interpolation von gegebenen Werten durch Polynome eine Rolle spielt, bezüglich der Basis  $(x \mapsto 1, x \mapsto x, x \mapsto x^2, \dots, x \mapsto x^{n-1})$  von  $P_{<n}$  und der Standardbasis von  $\mathbb{R}^n$ , siehe Beispiele 8.27 und 9.15.

## 14. EIGENWERTE UND EIGENVEKTOREN

Im vorletzten Abschnitt haben wir einen Klassifikationssatz bewiesen (Satz 12.10). Man kann ihn so interpretieren, dass die einzige Eigenschaft, die lineare Abbildungen zwischen zwei gegebenen endlich-dimensionalen Vektorräumen voneinander unterscheidet, der *Rang* ist: Ist  $f : V \rightarrow W$  linear mit  $\text{rk}(f) = r$ , dann können wir Basen von  $V$  und  $W$  so wählen, dass  $f$  durch die Matrix  $M_r$  gegeben ist, und  $M_r$  hängt nur von  $r$  (und den Dimensionen von  $V$  und  $W$ ) ab.

Wir werden jetzt statt linearen Abbildungen zwischen verschiedenen Vektorräumen  $V$  und  $W$  *Endomorphismen*  $f : V \rightarrow V$  betrachten. Da wir nur einen Vektorraum haben, können wir auch nur *eine* Basis wählen. Wir haben also deutlich weniger Spielraum, was sich in einem erheblich schwierigeren Klassifikationsproblem niederschlägt.

Natürlich kann man lineare Abbildungen  $f : V \rightarrow V$  auch als Spezialfall von linearen Abbildungen  $V \rightarrow W$  betrachten, wo „zufällig“  $W = V$  ist. Dann wird man die Wahl von verschiedenen Basen von  $V$  auf der Quell- und der Zielseite zulassen. Zum Beispiel erhält man so die Basiswechselformen  $\text{Mat}_{B,B'}(\text{id}_V)$ . Auf der anderen Seite geht so aber die Information verloren, dass es sich wirklich auf beiden Seiten um *denselben* Vektorraum handelt und nicht um zwei Vektorräume, die zufällig isomorph sind (d.h., dieselbe Dimension haben). Für das Klassifikationsproblem, das wir in diesem Abschnitt (und dann weiter in der *Linearen Algebra II*) studieren wollen, ist es aber wesentlich, dass  $f$  als Endomorphismus von  $V$  betrachtet wird. Anderenfalls wäre eine Aussage der Form  $f(v) = \lambda v$  (siehe Definition 14.3 unten) nicht sinnvoll, bzw. sie würde sich nicht auf die  $f$  beschreibenden Matrizen übertragen.

Wir schreiben erst einmal auf, wie die Matrizen von  $f$  bezüglich verschiedener Basen von  $V$  miteinander zusammenhängen.

**14.1. Satz.** Sei  $V$  ein  $K$ -Vektorraum mit Basis  $B = (b_1, b_2, \dots, b_n)$  und sei  $f$  ein Endomorphismus von  $V$ . Sei weiter  $A = \text{Mat}_{B,B}(f) \in \text{Mat}(n, K)$  die Matrix von  $f$  bezüglich  $B$ . Dann ist

$$\{\text{Mat}_{B',B'}(f) \mid B' \text{ Basis von } V\} = \{PAP^{-1} \mid P \in \text{GL}(n, K)\}.$$

**SATZ**  
Matrizen  
eines Endo-  
morphimus

*Beweis.* Es ist  $\text{Mat}_{B',B'}(f) = \text{Mat}_{B',B'}(\text{id}_V) \text{Mat}_{B,B}(f) \text{Mat}_{B',B}(\text{id}_V) = PAP^{-1}$  mit  $P = \text{Mat}_{B',B}(\text{id}_V) \in \text{GL}(n, K)$ . Umgekehrt lässt sich jede Matrix  $P \in \text{GL}(n, K)$  in dieser Form schreiben (Folgerung 12.4).  $\square$

\* **14.2. Definition.** Sei  $K$  ein Körper und  $n \in \mathbb{N}$ . Zwei Matrizen  $A, A' \in \text{Mat}(n, K)$  heißen *ähnlich*, wenn es eine Matrix  $P \in \text{GL}(n, K)$  gibt mit  $A' = PAP^{-1}$ .  $\diamond$

**DEF**  
Ähnlichkeit  
von Matrizen

Ähnlich wie für die Äquivalenz von Matrizen zeigt man, dass die Ähnlichkeit von Matrizen eine Äquivalenzrelation ist.

Wenn  $A$  eine Matrix eines Endomorphismus  $f$  von  $V$  ist, dann sind die Matrizen von  $f$  bezüglich beliebiger Basen von  $V$  also gerade die zu  $A$  ähnlichen Matrizen.

Die Klassifikation von Matrizen bis auf Ähnlichkeit (und damit die Klassifikation der Endomorphismen endlich-dimensionaler Vektorräume) ist relativ kompliziert. Sie wird durch die *Jordan-Normalform* geleistet, die wir im nächsten Semester besprechen werden. Hier werden wir uns erst einmal auf die Diskussion einfacherer „Invarianten“ (also Daten, die nur von  $f$  und nicht von der Basis abhängen) beschränken.

Die Idee ist, den Endomorphismus  $f : V \rightarrow V$  mit anderen besonders einfachen Endomorphismen zu vergleichen. Die einfachsten Endomorphismen sind sicher die Multiplikationen mit einem Skalar  $\lambda \in K$ :  $v \mapsto \lambda v$ . Wir können uns fragen, ob es Elemente von  $V$  gibt, die sich unter  $f$  und dieser Abbildung gleich verhalten. Das führt auf folgende Definition.

- \* 14.3. **Definition.** Sei  $V$  ein  $K$ -Vektorraum und  $f \in \text{End}(V)$ . Ein Skalar  $\lambda \in K$  heißt *Eigenwert* von  $f$ , wenn es einen Vektor  $\mathbf{0} \neq v \in V$  gibt, sodass  $f(v) = \lambda v$  ist. Jeder solche Vektor heißt ein *Eigenvektor* von  $f$  zum Eigenwert  $\lambda$ .  $\diamond$  **DEF**  
Eigenwert  
Eigenvektor

Man beachte die Bedingung  $v \neq \mathbf{0}$ ! Ohne sie wäre die Definition sinnlos, weil dann jedes  $\lambda$  ein Eigenwert wäre (denn  $f(\mathbf{0}) = \mathbf{0} = \lambda \mathbf{0}$ ).

- \* 14.4. **Definition.** Sei  $V$  ein  $K$ -Vektorraum,  $\lambda \in K$  und  $f \in \text{End}(V)$ . Der Untervektorraum **DEF**  
Eigenraum

$$E_\lambda(f) = \{v \in V \mid f(v) = \lambda v\} = \ker(\lambda \text{id}_V - f)$$

von  $V$  heißt der  $\lambda$ -Eigenraum von  $f$ .

Die Dimension  $\dim E_\lambda(f)$  des  $\lambda$ -Eigenraums von  $f$  heißt die *geometrische Vielfachheit* des Eigenwerts  $\lambda$  von  $f$ .  $\diamond$

$E_\lambda(f)$  besteht also aus dem Nullvektor und den Eigenvektoren zum Eigenwert  $\lambda$ .  $\lambda$  ist genau dann ein Eigenwert von  $f$ , wenn  $E_\lambda(f) \neq \{\mathbf{0}\}$ , also die geometrische Vielfachheit positiv ist.

- 14.5. **Beispiel.**  $E_0(f)$  ist gerade der Kern von  $f$ . Null ist also genau dann ein Eigenwert von  $f$ , wenn  $f$  nicht injektiv ist. Ist  $V$  endlich-dimensional, dann ist das auch dazu äquivalent, dass  $f$  kein Isomorphismus ist.  $\clubsuit$  **BSP**  
Eigenwert  
null

- 14.6. **Beispiel.** Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $(x, y) \mapsto (y, x)$ . Dann hat  $f$  die Eigenwerte 1 und  $-1$ , denn  $v_1 = (1, 1) \in E_1(f)$  und  $v_{-1} = (1, -1) \in E_{-1}(f)$ . Man sieht leicht, dass beide Eigenräume eindimensional sind.  $\clubsuit$  **BSP**  
Eigenwerte  
Eigenräume

- 14.7. **Beispiel.** Sei  $f : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ ,  $h \mapsto h'$ . Dann hat  $f$  jedes  $\lambda \in \mathbb{R}$  als Eigenwert, und es gilt  $E_\lambda(f) = \langle x \mapsto e^{\lambda x} \rangle$ . (Beweis wie in Beispiel 11.11.)  $\clubsuit$  **BSP**  
Eigenwerte  
Eigenräume

Wir werden sehen, dass die Situation von Beispiel 14.6 für Endomorphismen von endlich-dimensionalen Vektorräumen recht typisch ist. Wir zeigen erst einmal, dass es nicht zu viele Eigenwerte geben kann.

- \* 14.8. **Satz.** Sei  $V$  ein  $K$ -Vektorraum und  $f \in \text{End}(V)$ . Seien  $\lambda_1, \lambda_2, \dots, \lambda_m \in K$  paarweise verschieden und für  $j \in \{1, 2, \dots, m\}$  sei  $v_j \in V$  ein Eigenvektor von  $f$  zum Eigenwert  $\lambda_j$ . Dann sind  $v_1, v_2, \dots, v_m$  linear unabhängig. **SATZ**  
Lin. Unabh.  
von Eigenvektoren

*Beweis.* Seien  $\alpha_1, \alpha_2, \dots, \alpha_m \in K$  mit  $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m = \mathbf{0}$ . Wir müssen zeigen, dass alle  $\alpha_j = 0$  sind. Dazu verwenden wir Induktion über  $m$ . Der Fall

$m = 0$  ist klar (null Vektoren sind stets linear unabhängig). Sei also  $m > 0$  und die Behauptung für  $m - 1$  schon bewiesen. Es ist

$$\begin{aligned} \mathbf{0} &= \lambda_m \mathbf{0} - f(\mathbf{0}) \\ &= \lambda_m(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m) - f(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m) \\ &= \lambda_m \alpha_1 v_1 + \lambda_m \alpha_2 v_2 + \dots + \lambda_m \alpha_m v_m - \alpha_1 \lambda_1 v_1 - \alpha_2 \lambda_2 v_2 - \dots - \alpha_m \lambda_m v_m \\ &= \alpha_1(\lambda_m - \lambda_1)v_1 + \alpha_2(\lambda_m - \lambda_2)v_2 + \dots + \alpha_{m-1}(\lambda_m - \lambda_{m-1})v_{m-1}. \end{aligned}$$

Aus der Induktionsannahme folgt

$$\alpha_1(\lambda_m - \lambda_1) = \alpha_2(\lambda_m - \lambda_2) = \dots = \alpha_{m-1}(\lambda_m - \lambda_{m-1}) = 0.$$

Weil  $\lambda_m \neq \lambda_1, \lambda_2, \dots, \lambda_{m-1}$  ist, ergibt sich daraus  $\alpha_1 = \alpha_2 = \dots = \alpha_{m-1} = 0$ . Die ursprüngliche Gleichung reduziert sich also auf  $\alpha_m v_m = \mathbf{0}$ . Weil  $v_m \neq \mathbf{0}$  ist, folgt auch  $\alpha_m = 0$ .  $\square$

**14.9. Folgerung.** Sei  $V$  ein  $K$ -Vektorraum und  $f \in \text{End}(V)$ . Seien  $\lambda_1, \lambda_2, \dots, \lambda_m$  paarweise verschiedene Elemente von  $K$  und für  $j \in \{1, 2, \dots, m\}$  sei  $n_j = \dim E_{\lambda_j}(f)$  und  $(b_{j1}, b_{j2}, \dots, b_{jn_j})$  eine Basis von  $E_{\lambda_j}(f)$ . Dann sind die  $b_{ji}$  (mit  $j \in \{1, 2, \dots, m\}$  und  $i \in \{1, 2, \dots, n_j\}$ ) linear unabhängig. Insbesondere gilt

$$\dim E_{\lambda_1}(f) + \dim E_{\lambda_2}(f) + \dots + \dim E_{\lambda_m}(f) \leq \dim V.$$

**FOLG**  
Dimension  
von Eigen-  
räumen

Ist  $V$  endlich-dimensional, dann kann  $f$  also höchstens  $\dim V$  Eigenwerte haben.

*Beweis.* Sei

$$\sum_{j=1}^m \sum_{i=1}^{n_j} \alpha_{ji} b_{ji} = \mathbf{0}$$

mit  $\alpha_{ji} \in K$ . Sei  $v_j = \sum_{i=1}^{n_j} \alpha_{ji} b_{ji} \in E_{\lambda_j}(f)$ , dann gilt  $v_1 + v_2 + \dots + v_m = \mathbf{0}$ . Aus Satz 14.8 folgt dann  $v_1 = v_2 = \dots = v_m = \mathbf{0}$ , denn eventuell vorkommende Vektoren  $\neq \mathbf{0}$  müssten linear unabhängig sein und könnten sich also nicht zum Nullvektor addieren. Da  $(b_{j1}, b_{j2}, \dots, b_{jn_j})$  eine Basis von  $E_{\lambda_j}(f)$  ist, folgt dann aus  $v_j = \mathbf{0}$  auch  $\alpha_{ji} = 0$  für alle  $i \in \{1, 2, \dots, n_j\}$ . Da das für jedes  $j \in \{1, 2, \dots, m\}$  gilt, sind alle  $\alpha_{ji} = 0$ . Das zeigt die Behauptung. Die letzte Aussage folgt aus

$$\begin{aligned} \dim V &\geq \#\{b_{ji} \mid j \in \{1, 2, \dots, m\}, i \in \{1, 2, \dots, n_j\}\} \\ &= n_1 + n_2 + \dots + n_m \\ &= \dim E_{\lambda_1}(f) + \dim E_{\lambda_2}(f) + \dots + \dim E_{\lambda_m}(f). \end{aligned} \quad \square$$

Wie können wir die Eigenwerte (und dann die Eigenräume) finden? Dazu wählen wir eine Basis und bestimmen die Matrix  $A$  von  $f$  bezüglich dieser Basis. Wir übertragen die Begriffe Eigenwert usw. auf Matrizen.

**14.10. Definition.** Sei  $K$  ein Körper und  $A \in \text{Mat}(n, K)$ . Sei  $\lambda \in K$ . Dann heißt  $\lambda$  ein *Eigenwert* von  $A$ , wenn es einen Spaltenvektor  $\mathbf{0} \neq \mathbf{x} \in K^n$  gibt mit  $A\mathbf{x} = \lambda\mathbf{x}$ . In diesem Fall heißt  $\mathbf{x}$  ein *Eigenvektor* von  $A$  zum Eigenwert  $\lambda$ . Der Untervektorraum

$$E_\lambda(A) = \{\mathbf{x} \in K^n \mid A\mathbf{x} = \lambda\mathbf{x}\} = \ker(\lambda I_n - A)$$

heißt der *Eigenraum* von  $A$  zum Eigenwert  $\lambda$ ; seine Dimension heißt die *geometrische Vielfachheit* des Eigenwerts  $\lambda$  von  $A$ .  $\diamond$

Die Eigenwerte von  $A$  und ihre geometrischen Vielfachheiten entsprechen dann denen von  $f$ .

**DEF**  
Eigenwert  
etc. für  
Matrizen

Der Schlüssel zur Bestimmung der Eigenwerte ist folgende einfache Beobachtung.

**14.11. Lemma.** Sei  $K$  ein Körper,  $\lambda \in K$  und  $A \in \text{Mat}(n, K)$ .  $\lambda$  ist genau dann ein Eigenwert von  $A$ , wenn  $\det(\lambda I_n - A) = 0$  ist. Die geometrische Vielfachheit des Eigenwerts  $\lambda$  ist  $\dim \ker(\lambda I_n - A) = n - \text{rk}(\lambda I_n - A)$ .

**LEMMA**  
Charakterisierung von Eigenwerten

*Beweis.* Wir haben folgende Kette von Äquivalenzen:

$$\begin{aligned} \lambda \text{ ist Eigenwert von } A &\iff E_\lambda(A) \neq \{\mathbf{0}\} \\ &\iff \ker(\lambda I_n - A) \neq \{\mathbf{0}\} \\ &\iff \det(\lambda I_n - A) = 0 \end{aligned}$$

Die letzte Aussage folgt aus der Definition der geometrischen Vielfachheit. □

**14.12. Beispiel.** Wir betrachten

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \text{Mat}(2, \mathbb{R}).$$

**BSP**  
Bestimmung der Eigenwerte

Dann ist (für  $\lambda \in \mathbb{R}$ )

$$\det(\lambda I_2 - A) = \begin{vmatrix} \lambda & -1 \\ -1 & \lambda \end{vmatrix} = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1).$$

Das verschwindet genau für  $\lambda = 1$  und  $\lambda = -1$ , also sind das die Eigenwerte von  $A$ . Wir können Basen der Eigenräume  $E_\lambda(A)$  mit dem Zeilenstufenform-Algorithmus, angewandt auf  $\lambda I_2 - A$ , berechnen. Für  $\lambda = 1$  haben wir

$$\lambda I_2 - A = I_2 - A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix};$$

das liefert die Basis  $(1, 1)$  für  $E_1(A)$ . Für  $\lambda = -1$  sieht es so aus:

$$\lambda I_2 - A = -I_2 - A = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix};$$

die Basis ist  $(-1, 1)$ . ♣

An diesem Beispiel sieht man, dass die Determinante, deren Verschwinden anzeigt, dass  $\lambda$  ein Eigenwert ist, ein *Polynom* in  $\lambda$  (mit Koeffizienten in  $K$ ) ist. Wir müssen daher etwas ausholen und ein wenig über Polynome sprechen.

**Exkurs: Polynome.**

**14.13. Definition.** Sei  $K$  ein Körper. Ein *Polynom* in der Variablen (oder *Unbestimmten*)  $X$  über  $K$  ist ein Ausdruck der Form

**DEF**  
Polynom

$$p = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

mit  $n \in \mathbb{N}$  und  $a_0, a_1, \dots, a_n \in K$ .  $a_j$  heißt der *j-te Koeffizient* von  $p$  oder der *Koeffizient von  $X^j$*  in  $p$ . Wir setzen  $a_j = 0$  für  $j > n$ . Ist  $a_n \neq 0$ , dann hat das Polynom *Grad  $n$* :  $\text{deg}(p) = n$  (englisch „degree“). Ist  $a_n = 1$ , dann heißt  $p$  *normiert*. Sind alle  $a_j = 0$ , dann ist  $p$  das *Nullpolynom*; sein Grad ist  $\text{deg}(\mathbf{0}) = -\infty$ . Ist  $n = 0$ , dann heißt  $p$  *konstant* (d.h.,  $p = \mathbf{0}$  oder  $\text{deg}(p) = 0$ ). Wir schreiben  $K[X]$  für die Menge der Polynome in  $X$  über  $K$ .

Sei  $q = b_m X^m + \dots + b_1 X + b_0$ . Dann sind  $p$  und  $q$  genau dann gleich, wenn ihre Koeffizienten übereinstimmen:  $a_j = b_j$  für alle  $j \in \mathbb{N}$  (mit der Konvention  $a_j = 0$  für  $j > n$  und  $b_j = 0$  für  $j > m$ ). Die *Summe* von  $p$  und  $q$  ist

$$p + q = \sum_{j=0}^{\max\{m,n\}} (a_j + b_j) X^j;$$

es gilt  $\deg(p + q) \leq \max\{\deg(p), \deg(q)\}$ , und das *Produkt* von  $p$  und  $q$  ist

$$pq = \sum_{k=0}^{m+n} \left( \sum_{i,j:i+j=k} a_i b_j \right) X^k.$$

Es gilt  $\deg(pq) = \deg(p) + \deg(q)$ . Wir identifizieren  $K$  mit der Teilmenge der konstanten Polynome:  $K \subset K[X]$ . Die Menge  $K[X]$  wird mit der eben definierten Addition und Multiplikation ein kommutativer Ring, dem *Polynomring* in  $X$  über  $K$ ; die Einschränkung der Multiplikation auf  $K \times K[X]$  macht  $K[X]$  zu einem unendlich-dimensionalen  $K$ -Vektorraum mit Basis  $(1, X, X^2, X^3, \dots)$ .  $\diamond$

Wenn es Sie stört, dass in der Definition von einem „Ausdruck der Form ...“ gesprochen wird, ohne dass gesagt wird, was das eigentlich „ist“, dann lesen Sie hier weiter.

Formal kann man die Definition auf stabile FüÙe stellen, indem man setzt

$$K[X] = \{(a_n)_{n \in \mathbb{N}} \in K^{\mathbb{N}} \mid \exists N \in \mathbb{N} \forall n > N : a_n = 0\}.$$

Das sind also die endlichen Folgen von Elementen von  $K$ , in dem Sinne, dass alle bis auf endlich viele Folgenglieder null sind. Man definiert weiter  $X = (0, 1, 0, 0, 0, \dots) \in K[X]$  und die Abbildung  $i : K \rightarrow K[X]$ ,  $a \mapsto (a, 0, 0, 0, \dots)$ . Die Addition in  $K[X]$  wird komponentenweise definiert, die Multiplikation mit  $X$  durch

$$X \cdot (a_0, a_1, a_2, \dots) = (0, a_0, a_1, a_2, \dots)$$

und die mit  $i(a)$  durch

$$i(a) \cdot (a_0, a_1, a_2, \dots) = (aa_0, aa_1, aa_2, \dots).$$

Dann ist

$$(a_0, a_1, \dots, a_n, 0, 0, 0, \dots) = i(a_0) + i(a_1)X + i(a_2)X^2 + \dots + i(a_n)X^n,$$

und die Multiplikation damit wird so definiert, dass das Assoziativ- und das Distributivgesetz gelten. Mittels der Abbildung  $i$  wird  $K$  mit seinem Bild in  $K[X]$  identifiziert; man schreibt also einfach  $a$  statt  $i(a)$ . Die Ringaxiome muss man dann noch nachprüfen. Die  $K$ -Vektorraum-Struktur von  $K[X]$  ist einfach die als Untervektorraum von  $K^{\mathbb{N}}$ .

Das funktioniert auch dann noch, wenn man die Endlichkeitsbedingung in der Definition weglässt. Man erhält dann den Ring  $K[[X]]$  der *formalen Potenzreihen* in  $X$  über  $K$ . Für  $K = \mathbb{R}$  oder  $\mathbb{C}$  spielen diese Potenzreihen eine wichtige Rolle in der Analysis (bzw. Funktionentheorie).

In Polynome kann man einsetzen:

**14.14. Definition.** Sei  $K$  ein Körper und  $p = a_n X^n + \dots + a_1 X + a_0 \in K[X]$  ein Polynom. Für  $\lambda \in K$  ist der *Wert von  $p$  bei  $\lambda$*  gegeben durch

$$p(\lambda) = a_n \lambda^n + \dots + a_1 \lambda + a_0.$$

$\lambda$  heißt eine *Nullstelle* von  $p$ , wenn  $p(\lambda) = 0$  ist.  $\diamond$

**DEF**  
Werte und  
Nullstellen  
von  
Polynomen

Ein Polynom  $p \in K[X]$  führt also zu einer *Polynomfunktion*  $K \rightarrow K$ ,  $\lambda \mapsto p(\lambda)$ . Die Abbildung  $K[X] \rightarrow \text{Abb}(K, K)$ , die einem Polynom die zugehörige Polynomfunktion zuordnet, ist injektiv, wenn der Körper  $K$  unendlich ist. Das ergibt sich aus dem folgenden Satz.

**14.15. Satz.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $x_1, x_2, \dots, x_n \in K$  paarweise verschieden. Seien weiter  $y_1, y_2, \dots, y_n \in K$ . Dann gibt es ein eindeutig bestimmtes Polynom  $p \in K[X]$  mit  $\deg(p) < n$ , sodass  $p(x_j) = y_j$  ist für alle  $j \in \{1, 2, \dots, n\}$ .

**SATZ**  
Eindeutigkeit  
von  
Polynomen

*Beweis.* Die Existenz folgt mit der Lagrangeschen Interpolationsformel wie in Beispiel 9.15. Damit ist die lineare Abbildung

$$\phi : \{p \in K[X] \mid \deg(p) < n\} \longrightarrow K^n, \quad p \longmapsto (p(x_1), p(x_2), \dots, p(x_n))$$

surjektiv. Da die beiden beteiligten Vektorräume dieselbe Dimension  $n$  haben (der Vektorraum der Polynome vom Grad  $< n$  hat Basis  $(1, X, X^2, \dots, X^{n-1})$ ), folgt aus der Surjektivität die Bijektivität.  $\square$

**14.16. Folgerung.** Ein Polynom  $p \in K[X]$  mit  $\deg(p) = n \in \mathbb{N}$  kann nicht mehr als  $n$  Nullstellen in  $K$  haben.

**FOLG**  
Nullstellen  
von  
Polynomen

*Beweis.* Angenommen,  $p$  hat  $n + 1$  Nullstellen  $a_1, a_2, \dots, a_{n+1} \in K$ . Dann muss  $p$  das eindeutig bestimmte Polynom von Grad  $< n + 1$  sein, dass  $p(a_j) = 0$  erfüllt für alle  $j \in \{1, 2, \dots, n + 1\}$ . Das Nullpolynom hat aber diese Eigenschaft, also muss  $p = \mathbf{0}$  sein. Das ist aber ein Widerspruch zur Voraussetzung  $\deg(p) = n$ .  $\square$

**14.17. Folgerung.** Ist  $K$  ein unendlicher Körper, dann ist die Abbildung  $K[X] \rightarrow \text{Abb}(K, K)$ , die einem Polynom die zugehörige Polynomfunktion zuordnet, injektiv.

**FOLG**  
Polynome  
und  
Polynom-  
funktionen

Ein Polynom  $p \in K[X]$  ist dann also durch seine Werte  $p(\lambda)$  für  $\lambda \in K$  eindeutig bestimmt. Wir können also zum Beispiel den Vektorraum  $P$  der Polynomfunktionen mit  $\mathbb{R}[X]$  identifizieren.

*Beweis.* Wir schreiben  $\Phi$  für die Abbildung  $K[X] \rightarrow \text{Abb}(K, K)$ .  $\Phi$  ist linear, also genügt es zu zeigen, dass  $\ker(\Phi) = \{\mathbf{0}\}$  ist. Sei also  $p \in \ker(\Phi)$ . Dann ist  $p(\lambda) = 0$  für alle  $\lambda \in K$ , also hat (da  $K$  unendlich ist) das Polynom  $p$  unendlich viele Nullstellen in  $K$ . Nach Folgerung 14.16 muss  $p$  das Nullpolynom sein.  $\square$

Für endliche Körper  $K$  ist die Aussage falsch: Ist  $\#K = q < \infty$ , dann ist  $\dim_K \text{Abb}(K, K) = q$ , denn eine Abbildung  $f : K \rightarrow K$  ist durch die  $q$  Werte  $f(a)$  für  $a \in K$  eindeutig bestimmt. Auf der anderen Seite ist  $\dim_K K[X] = \infty$ , und damit kann es keine injektive lineare Abbildung  $K[X] \rightarrow \text{Abb}(K, K)$  geben.

Der Kern von  $\Phi$  besteht in diesem Fall aus allen Polynomen, die alle Elemente von  $K$  als Nullstellen haben. Man kann zeigen, dass

$$\prod_{a \in K} (X - a) = X^q - X$$

ist; der Kern besteht demnach genau aus den Vielfachen von  $X^q - X$ .

So wie man ganze Zahlen mit Rest durcheinander dividieren kann, gibt es auch für Polynome eine Division mit Rest („Polynomdivision“).

**14.18. Satz.** Sei  $K$  ein Körper und seien  $f, g \in K[X]$  mit  $g$  normiert. Dann gibt es eindeutig bestimmte Polynome  $q$  („Quotient“) und  $r$  („Rest“) in  $K[X]$ , sodass  $f = qg + r$  und  $\deg(r) < \deg(g)$ .

**SATZ**  
Polynom-  
division

*Beweis.* Wir beweisen zunächst die Existenz. Sei  $\deg(g) = m$ , also

$$g = X^m + b_{m-1}X^{m-1} + \dots + b_1X + b_0.$$

Wir betrachten  $g$  als fest und führen den Beweis durch Induktion über  $\deg(f)$ . Ist  $\deg(f) < m$ , dann erfüllen  $q = 0$  und  $r = f$  die Bedingungen. Wir können also annehmen, dass  $n = \deg(f) \geq m$  ist; die Existenzaussage sei für  $\deg(f) < n$  bereits bewiesen. Es ist  $f = a_nX^n + a_{n-1}X^{n-1} + \dots + a_0$ , also ist

$$\tilde{f} = f - a_nX^{n-m}g = (a_{n-1} - a_nb_{m-1})X^{n-1} + \dots;$$

der Grad von  $\tilde{f}$  ist damit kleiner als  $n$ . Nach Induktionsvoraussetzung gibt es  $\tilde{q}, r \in K[X]$  mit  $\tilde{f} = \tilde{q}g + r$  und  $\deg(r) < m$ . Wir setzen  $q = a_nX^{n-m} + \tilde{q}$ ; dann ist

$$f = a_nX^{n-m}g + \tilde{f} = a_nX^{n-m}g + \tilde{q}g + r = qg + r$$

wie gewünscht.

Zur Eindeutigkeit: Seien  $q_1, q_2, r_1, r_2 \in K[X]$  mit  $f = q_1g + r_1 = q_2g + r_2$  und  $\deg(r_1), \deg(r_2) < \deg(g)$ . Es folgt

$$(q_1 - q_2)g = r_2 - r_1.$$

Die rechte Seite hat Grad  $< \deg(g)$ . Wäre  $q_1 \neq q_2$ , dann hätte die linke Seite Grad  $\deg(q_1 - q_2) + \deg(g) \geq \deg(g)$ , ein Widerspruch. Also ist  $q_1 = q_2$  und damit auch  $r_1 = r_2$ . □

Der Beweis übersetzt sich direkt in den üblichen Algorithmus zur Polynomdivision: Man subtrahiert geeignete Vielfache von  $g$  solange von  $f$ , bis man ein Polynom zurückbehält, dessen Grad kleiner als der von  $g$  ist.

**14.19. Folgerung.** Seien  $K$  ein Körper,  $p \in K[X]$  und  $\lambda \in K$ . Wir schreiben

$$p = q(X - \lambda) + r$$

**FOLG**  
Nullstellen

wie in Satz 14.18. Dann ist  $r = p(\lambda)$  konstant. Insbesondere ist  $\lambda$  genau dann eine Nullstelle von  $p$ , wenn  $r = 0$  ist.

*Beweis.*  $r$  ist konstant, da  $\deg(r) < 1 = \deg(X - \lambda)$ . Außerdem gilt

$$p(\lambda) = q(\lambda)(\lambda - \lambda) + r = r. \quad \square$$

Ist  $\lambda$  eine Nullstelle von  $p$ , dann ist demnach  $p = (X - \lambda)q$  mit einem Polynom  $q \in K[X]$ . Ist  $\lambda$  auch eine Nullstelle von  $q$ , dann ist  $p = (X - \lambda)^2\tilde{q}$  und so fort. Das führt zu folgender Definition.

**14.20. Definition.** Seien  $K$  ein Körper,  $0 \neq p \in K[X]$  ein Polynom und  $\lambda \in K$ . Die *Vielfachheit* der Nullstelle  $\lambda$  von  $p$  ist die größte Zahl  $n \in \mathbb{N}$ , sodass man  $p$  in der Form  $p = (X - \lambda)^nq$  schreiben kann mit einem Polynom  $q \in K[X]$ . In diesem Fall ist  $q(\lambda) \neq 0$ . ◇

**DEF**  
Vielfachheit  
einer  
Nullstelle

$\lambda$  ist also genau dann eine Nullstelle, wenn die Vielfachheit von  $\lambda$  als Nullstelle positiv ist.

14.21. **Beispiele.** Wir betrachten  $K = \mathbb{R}$ . Für  $p = X^3 - X^2 - X + 1 \in \mathbb{R}[X]$  gilt

**BSP**  
Vielfachheiten

$$p = (X - 1)^2(X + 1),$$

also hat  $p$  die Nullstellen 1 (mit Vielfachheit 2: eine „doppelte“ Nullstelle) und  $-1$  (mit Vielfachheit 1: eine „einfache“ Nullstelle).

Für  $q = X^3 + X^2 + X + 1 \in \mathbb{R}[X]$  gilt dagegen

$$q = (X + 1)(X^2 + 1),$$

also hat  $q$  nur die (einfache) Nullstelle  $-1$  in  $\mathbb{R}$ , denn der zweite Faktor  $X^2 + 1$  nimmt nur positive Werte an und hat daher keine reelle Nullstelle. Wenn wir  $q$  aber als Polynom in  $\mathbb{C}[X]$  betrachten, dann haben wir

$$q = (X + 1)(X + i)(X - i);$$

$q$  hat also die drei komplexen (einfachen) Nullstellen  $-1$ ,  $i$  und  $-i$ .



Zurück zu Eigenwerten und Eigenräumen. Wir haben gesehen, dass der Ausdruck

$$\det(\lambda I_n - A)$$

darüber entscheidet, ob  $\lambda$  ein Eigenwert von  $A$  ist oder nicht. Ist  $A = (a_{ij})$ , dann hat diese Determinante die folgende Form:

$$\begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & -a_{23} & \cdots & -a_{2n} \\ -a_{31} & -a_{32} & \lambda - a_{33} & \cdots & -a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & -a_{n3} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

Wenn wir das in die Leibniz-Formel einsetzen, dann bekommen wir

$$\begin{aligned} & (\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}) + \text{Terme mit } \leq n - 2 \text{ Faktoren } \lambda - a_{jj} \\ & = \lambda^n - (a_{11} + \dots + a_{nn})\lambda^{n-1} + \dots + (-1)^n \det(A). \end{aligned}$$

Das hat die Form  $p(\lambda)$  mit einem normierten Polynom  $p \in K[X]$  vom Grad  $n$ .

\* 14.22. **Definition.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $A \in \text{Mat}(n, K)$ . Das Polynom  $\det(XI_n - A) \in K[X]$  heißt das *charakteristische Polynom* von  $A$ . ◇

**DEF**  
Charakteristisches Polynom

Wir haben gesehen, dass die Eigenwerte von  $A$  genau die Nullstellen des charakteristischen Polynoms von  $A$  sind.

14.23. **Beispiel.** Was sind die Eigenwerte der „Telefonmatrix“

**BSP**  
Eigenwerte

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in \text{Mat}(3, \mathbb{R}) ?$$

Wir bestimmen das charakteristische Polynom:

$$\begin{aligned} \det(XI_3 - A) &= \begin{vmatrix} X-1 & -2 & -3 \\ -4 & X-5 & -6 \\ -7 & -8 & X-9 \end{vmatrix} \\ &= (X-1)(X-5)(X-9) - 2 \cdot 6 \cdot 7 - 3 \cdot 4 \cdot 8 \\ &\quad - (X-1) \cdot 6 \cdot 8 - (X-5) \cdot 3 \cdot 7 - (X-9) \cdot 2 \cdot 4 \\ &= X^3 - 15X^2 + 59X - 45 - 84 - 96 \\ &\quad - 48X + 48 - 21X + 105 - 8X + 72 \\ &= X^3 - 15X^2 - 18X = X(X^2 - 15X - 18) \end{aligned}$$

Ein Eigenwert ist  $\lambda_1 = 0$  ( $A$  ist also nicht invertierbar), die anderen beiden finden wir mit Hilfe der Lösungsformel für quadratische Gleichungen:

$$\lambda_2 = \frac{15 + 3\sqrt{33}}{2} \quad \text{und} \quad \lambda_3 = \frac{15 - 3\sqrt{33}}{2}. \quad \clubsuit$$

Wir können die Definitionen von Determinante und charakteristischem Polynom auch auf Endomorphismen übertragen.

**14.24. Definition.** Seien  $K$  ein Körper,  $V$  ein endlich-dimensionaler  $K$ -Vektorraum und  $f \in \text{End}(V)$ . Sei  $B$  eine beliebige Basis von  $V$  und  $A = \text{Mat}_{B,B}(f)$ . Dann ist die *Determinante* von  $f$  definiert als  $\det(f) = \det(A)$  und das *charakteristische Polynom* von  $f$  ist das charakteristische Polynom von  $A$ .  $\diamond$

**DEF**  
Determinante,  
char. Pol.  
von Endo-  
morphis-  
men

Die Definition ist sinnvoll, weil sie nicht von der Wahl der Basis  $B$  abhängt: Ist  $A' = \text{Mat}_{B',B'}(f)$  mit einer anderen Basis  $B'$  von  $V$ , dann gibt es eine Matrix  $P \in \text{GL}(n, K)$  (wenn  $\dim V = n$ ), sodass  $A' = PAP^{-1}$ . Dann ist

$$\begin{aligned} \det(A') &= \det(PAP^{-1}) = \det(P) \det(A) \det(P^{-1}) \\ &= \det(A) \det(PP^{-1}) = \det(A) \det(I_n) = \det(A). \end{aligned}$$

Es ist auch  $P(XI_n - A)P^{-1} = XPI_nP^{-1} - PAP^{-1} = XI_n - A'$ , und die gleiche Rechnung wie eben zeigt, dass  $A$  und  $A'$  dasselbe charakteristische Polynom haben.

\* **14.25. Definition.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$ ,  $A \in \text{Mat}(n, K)$  und  $\lambda \in K$ . Die *algebraische Vielfachheit* von  $\lambda$  als Eigenwert von  $A$  ist die Vielfachheit von  $\lambda$  als Nullstelle des charakteristischen Polynoms von  $A$ . Entsprechend definieren wir die *algebraische Vielfachheit* von  $\lambda$  als Eigenwert eines Endomorphismus  $f$  eines endlich-dimensionalen  $K$ -Vektorraums  $V$ .  $\diamond$

**DEF**  
algebraische  
Vielfachheit

Wir haben jetzt also zwei Vielfachheiten von Eigenwerten definiert, die geometrische und die algebraische. In welcher Beziehung stehen sie zueinander? Wir wissen bisher Folgendes:

$$\text{geom. Vielfachheit} > 0 \iff \text{Eigenwert} \iff \text{alg. Vielfachheit} > 0$$

Müssen die beiden Vielfachheiten immer gleich sein?

14.26. **Beispiel.** Sei

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \in \text{Mat}(2, K).$$

Das charakteristische Polynom von  $A$  ist  $(X - \lambda)^2$ , also hat  $\lambda$  die algebraische Vielfachheit 2. Auf der anderen Seite ist  $E_\lambda(A) = \langle (1, 0) \rangle$  (denn  $\text{rk}(\lambda I_2 - A) = 1$ ), also hat  $\lambda$  die geometrische Vielfachheit 1. ♣

**BSP**  
alg.  $\neq$  geom.  
Vielfachheit

Die Vielfachheiten können also verschieden sein. Eine Beziehung gilt jedoch.

\* 14.27. **Satz.** Seien  $K$  ein Körper,  $V$  ein endlich-dimensionaler  $K$ -Vektorraum,  $f \in \text{End}(V)$  und  $\lambda \in K$ . Dann ist die geometrische Vielfachheit von  $\lambda$  als Eigenwert von  $f$  nicht größer als seine algebraische Vielfachheit.

**SATZ**  
geom.  $\leq$  alg.  
Vielfachheit

Die analoge Aussage gilt dann natürlich auch für Matrizen  $A \in \text{Mat}(n, K)$ .

*Beweis.* Sei  $m = \dim E_\lambda(f)$  die geometrische Vielfachheit, sei  $n = \dim V$  und sei  $(b_1, b_2, \dots, b_m)$  eine Basis von  $E_\lambda(f)$ . Wir können diese Basis zu einer Basis  $B = (b_1, b_2, \dots, b_n)$  von  $V$  erweitern. Dann ist

$$A = \text{Mat}_{B,B}(f) = \left( \begin{array}{c|c} \lambda I_m & D \\ \hline \mathbf{0}_{n-m,m} & C \end{array} \right)$$

mit Matrizen  $D \in \text{Mat}(m \times (n - m), K)$  und  $C \in \text{Mat}(n - m, K)$ , denn für  $j \in \{1, 2, \dots, m\}$  ist  $f(b_j) = \lambda b_j$ ; in der  $j$ -ten Spalte von  $A$  kommt also das  $\lambda$ -fache des  $j$ -ten Standard-Basisvektors zu stehen. Das charakteristische Polynom von  $f$  ist dann

$$\det(XI_n - A) = \det \left( \begin{array}{c|c} (X - \lambda)I_m & -D \\ \hline \mathbf{0}_{n-m,m} & XI_{n-m} - C \end{array} \right) = (X - \lambda)^m \det(XI_{n-m} - C).$$

Das zeigt, dass die Vielfachheit von  $\lambda$  als Nullstelle des charakteristischen Polynoms mindestens  $m$  ist.  $\square$

Wir haben folgende Aussage verwendet, deren Beweis eine Übungsaufgabe ist: Sind  $A \in \text{Mat}(m, K)$ ,  $B \in \text{Mat}(m \times n, K)$  und  $C \in \text{Mat}(n, K)$ , dann ist

$$\det \left( \begin{array}{c|c} A & B \\ \hline \mathbf{0}_{n,m} & C \end{array} \right) = \det(A) \det(C).$$

## 15. DIAGONALISIERBARKEIT

Wir haben in Folgerung 14.9 gesehen, dass die Summe der geometrischen Vielfachheiten der Eigenwerte eines Endomorphismus  $f$  eines  $n$ -dimensionalen Vektorraums (oder einer  $n \times n$ -Matrix) höchstens  $n$  ist. Das macht den Fall interessant, in dem diese Schranke erreicht wird. Wir formulieren zunächst eine Definition und werden dann sehen, was sie mit dieser Frage zu tun hat.

\*

**15.1. Definition.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$ . Eine Matrix  $A = (a_{ij}) \in \text{Mat}(n, K)$  ist eine *Diagonalmatrix* oder *diagonal*, wenn  $a_{ij} = 0$  ist für alle  $i, j \in \{1, 2, \dots, n\}$  mit  $i \neq j$ . Ist  $a_{ii} = d_i$  für  $i \in \{1, 2, \dots, n\}$ , dann schreiben wir  $\text{diag}(d_1, d_2, \dots, d_n)$  für  $A$ :

**DEF**  
Diagonal-  
matrix

$$\text{diag}(d_1, d_2, \dots, d_n) = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}. \quad \diamond$$

**15.2. Lemma.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$ ,  $V$  ein  $n$ -dimensionaler  $K$ -Vektorraum und  $f \in \text{End}(V)$ . Dann sind die folgenden Aussagen äquivalent:

**LEMMA**  
Diagonal-  
isierbarkeit

- (1)  $V$  hat eine Basis  $B$ , die aus Eigenvektoren von  $f$  besteht.
- (2) Die Summe der geometrischen Vielfachheiten der Eigenwerte von  $f$  ist  $n$ .
- (3) Sei  $A = \text{Mat}_{B', B'}(f)$  die Matrix von  $f$  bezüglich einer beliebigen Basis  $B'$  von  $V$ . Dann ist  $A$  ähnlich zu einer Diagonalmatrix.

*Beweis.* Wir beweisen zuerst die Äquivalenz der ersten beiden Aussagen.

„(1)  $\Rightarrow$  (2)“: Seien  $\lambda_1, \lambda_2, \dots, \lambda_k \in K$  die (paarweise verschiedenen) Eigenwerte von  $f$ , sei  $n_j = \dim E_{\lambda_j}(f)$  die geometrische Vielfachheit von  $\lambda_j$  und sei  $m_j$  die Anzahl der Basisvektoren in  $B$ , die Eigenvektoren zum Eigenwert  $\lambda_j$  sind. Dann haben wir jeweils  $m_j$  linear unabhängige Vektoren in  $E_{\lambda_j}(f)$ , also ist  $m_j \leq n_j$ . Auf der anderen Seite ist

$$n = \#B = m_1 + m_2 + \dots + m_k \leq n_1 + n_2 + \dots + n_k \leq n,$$

also haben wir Gleichheit; insbesondere ist  $n_1 + n_2 + \dots + n_k = n$ .

„(2)  $\Rightarrow$  (1)“: Wir behalten die Bezeichnungen bei. Sei  $B_j$  eine Basis von  $E_{\lambda_j}(f)$ . Nach Folgerung 14.9 bilden die  $B_j$  zusammen eine linear unabhängige Familie  $B$  in  $V$ . Wegen

$$\#B = \#B_1 + \#B_2 + \dots + \#B_k = n_1 + n_2 + \dots + n_k = n = \dim V$$

ist  $B$  dann eine Basis von  $V$ , die nach Konstruktion aus Eigenvektoren von  $f$  besteht.

„(1)  $\Rightarrow$  (3)“: Sei  $B = (b_1, b_2, \dots, b_n)$  und sei  $\alpha_j$  der Eigenwert von  $f$ , sodass  $f(b_j) = \alpha_j b_j$ . Dann ist

$$\text{Mat}_{B, B}(f) = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n)$$

eine Diagonalmatrix, und nach Satz 14.1 ist  $A$  zu ihr ähnlich.

„(3)  $\Rightarrow$  (1)“: Ist  $A$  ähnlich zu  $D = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n)$ , dann gibt es nach Satz 14.1 eine Basis  $B = (b_1, b_2, \dots, b_n)$  von  $V$ , sodass  $D = \text{Mat}_{B, B}(f)$  ist. Daran liest man ab, dass  $f(b_j) = \alpha_j b_j$  ist (und  $b_j \neq \mathbf{0}$ ), also besteht  $B$  aus Eigenvektoren von  $f$ .  $\square$

\* **15.3. Definition.** Sei  $f$  ein Endomorphismus eines endlich-dimensionalen  $K$ -Vektorraums  $V$ . Dann heißt  $f$  *diagonalisierbar*, wenn  $V$  eine Basis hat, die aus Eigenvektoren von  $f$  besteht.

**DEF**  
diagonalisierbar

Eine Matrix  $A \in \text{Mat}(n, K)$  heißt *diagonalisierbar*, wenn sie zu einer Diagonalmatrix ähnlich ist, wenn es also eine Diagonalmatrix  $D \in \text{Mat}(n, K)$  und eine invertierbare Matrix  $P \in \text{GL}(n, K)$  gibt mit  $PAP^{-1} = D$ .  $\diamond$

Aus dem Lemma ergibt sich, dass ein Endomorphismus genau dann diagonalisierbar ist, wenn die zugeordnete Matrix (bezüglich irgendeiner Basis) diagonalisierbar ist.

Aus dem Beweis ergibt sich auch, dass die Einträge auf der Diagonalen der Diagonalmatrix gerade die Eigenwerte sind; sie kommen so oft vor, wie es ihrer geometrischen Vielfachheit entspricht.

Da die geometrische Vielfachheit eines Eigenwerts höchstens so groß ist wie seine algebraische Vielfachheit, ist eine *notwendige* Bedingung für die Diagonalisierbarkeit, dass die Summe der *algebraischen* Vielfachheiten der Eigenwerte  $n$  ist. Das ist eine Eigenschaft des charakteristischen Polynoms.

**15.4. Definition.** Sei  $K$  ein Körper und  $p \in K[X]$  ein normiertes Polynom vom Grad  $n$ . Wir sagen,  $p$  *zerfällt in Linearfaktoren* über  $K$ , wenn es  $\alpha_1, \alpha_2, \dots, \alpha_n \in K$  gibt, sodass

**DEF**  
Zerlegung in Linearfaktoren

$$p = \prod_{j=1}^n (X - \alpha_j) = (X - \alpha_1)(X - \alpha_2) \cdots (X - \alpha_n)$$

ist.  $\diamond$

Das bedeutet also, dass die Summe der Vielfachheiten der Nullstellen von  $p$  in  $K$  gleich dem Grad  $n$  von  $p$  ist.

\* **15.5. Folgerung.** Seien  $K$  ein Körper und  $f$  ein Endomorphismus eines endlich-dimensionalen  $K$ -Vektorraums  $V$ . Dann sind die folgenden Aussagen äquivalent:

**FOLG**  
Charakterisierung von Diagonalisierbarkeit

- (1)  $f$  ist diagonalisierbar.
- (2) Das charakteristische Polynom von  $f$  zerfällt in Linearfaktoren über  $K$  und für jeden Eigenwert von  $f$  stimmt die geometrische mit der algebraischen Vielfachheit überein.

*Insbesondere gilt: Hat  $f$  genau  $n = \dim V$  verschiedene Eigenwerte, dann ist  $f$  diagonalisierbar.*

**Wichtig:** Die Umkehrung des letzten Satzes gilt nicht. Zum Beispiel sind die Einheitsmatrix oder die Nullmatrix diagonalisierbar (weil schon diagonal), sie haben aber jeweils nur einen Eigenwert (nämlich 1 bzw. 0).

*Beweis.* Seien  $\lambda_1, \lambda_2, \dots, \lambda_k \in K$  die verschiedenen Eigenwerte von  $f$  und seien  $n_j$  bzw.  $m_j$  ihre geometrischen bzw. algebraischen Vielfachheiten. Sei  $n = \dim V$  und  $p$  das charakteristische Polynom von  $f$ . Dann gilt wegen  $n_j \leq m_j$  (Satz 14.27) und  $m_1 + m_2 + \dots + m_k \leq \deg(p) = n$ :

$$\begin{aligned} f \text{ diagonalisierbar} &\iff n_1 + n_2 + \dots + n_k = n \\ &\iff m_1 + m_2 + \dots + m_k = n \quad \text{und} \quad \forall j : m_j = n_j \\ &\iff p \text{ zerfällt in Linearfaktoren über } K \text{ und } \forall j : m_j = n_j. \end{aligned}$$

Hat  $f$   $n$  verschiedene Eigenwerte, dann ist  $k = n$  und  $m_j = 1$ . Dann muss auch die geometrische Vielfachheit  $n_j = 1$  sein, also ist  $n_1 + n_2 + \dots + n_k = n$ .  $\square$

**15.6. Beispiel.** Die Telefonmatrix aus Beispiel 14.23 ist diagonalisierbar, weil sie die drei verschiedenen Eigenwerte  $0$ ,  $\frac{1}{2}(15 + 3\sqrt{33})$  und  $\frac{1}{2}(15 - 3\sqrt{33})$  hat.  $\clubsuit$

**BSP**  
Telefonmatrix  
ist diagona-  
lisierbar  
**BSP**  
Nicht  
diagonalisi-  
erbare  
Matrix

**15.7. Beispiel.** Eine Matrix muss nicht diagonalisierbar sein, wenn ihr charakteristisches Polynom in Linearfaktoren zerfällt. Das hatten wir (siehe Beispiel 14.26) an Hand der Matrix

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

gesehen, deren charakteristisches Polynom  $(X - \lambda)^2$  in Linearfaktoren zerfällt, für die aber die geometrische Vielfachheit von  $\lambda$  (nämlich 1) kleiner ist als die algebraische Vielfachheit (nämlich 2).  $\clubsuit$

Die Gleichheit der geometrischen und algebraischen Vielfachheit ist also eine wesentliche Bedingung. Die Bedingung, dass das charakteristische Polynom in Linearfaktoren zerfällt, können wir hingegen erfüllen, wenn unser Körper „groß genug“ ist. Wir erinnern uns daran (Satz 4.4), dass der Körper  $\mathbb{C}$  der komplexen Zahlen *algebraisch abgeschlossen* ist. Das bedeutet, dass jedes nicht-konstante Polynom  $p \in \mathbb{C}[X]$  eine Nullstelle in  $\mathbb{C}$  hat. Daraus folgt, dass jedes normierte Polynom über  $\mathbb{C}$  in Linearfaktoren zerfällt:

**15.8. Folgerung.** Sei  $p \in \mathbb{C}[X]$  normiert. Dann zerfällt  $p$  in Linearfaktoren über  $\mathbb{C}$ .

**FOLG**  
Faktorisierung  
von  
Polynomen  
über  $\mathbb{C}$

*Beweis.* Wir beweisen die Aussage durch Induktion über den Grad von  $p$ . Im Fall  $\deg(p) = 0$  ist  $p = 1$  und damit gleich dem leeren Produkt (anders ausgedrückt,  $p$  hat genau  $\deg(p) = 0$  Nullstellen in  $\mathbb{C}$ ), also zerfällt  $p$  trivialerweise in Linearfaktoren. Die Aussage gelte für Polynome vom Grad  $n$ , und  $p$  habe Grad  $n + 1$ . Nach dem Fundamentalsatz der Algebra 4.4 hat  $p$  eine Nullstelle  $\alpha_1 \in \mathbb{C}$ . Dann ist  $p = (X - \alpha_1)q$  mit  $q \in \mathbb{C}[X]$  und  $\deg(q) = n$ . Nach der Induktionsvoraussetzung zerfällt  $q$  in Linearfaktoren:

$$q = (X - \alpha_2)(X - \alpha_3) \cdots (X - \alpha_{n+1}),$$

also gilt das auch für  $p$ :

$$p = (X - \alpha_1)q = (X - \alpha_1)(X - \alpha_2) \cdots (X - \alpha_{n+1}). \quad \square$$

Für den Beweis haben wir nur verwendet, dass  $\mathbb{C}$  algebraisch abgeschlossen ist; die Aussage gilt also entsprechend für jeden algebraisch abgeschlossenen Körper. Außerdem kann man ganz allgemein zeigen, dass es zu jedem Körper  $K$  einen algebraisch abgeschlossenen Körper  $\bar{K}$  gibt, der  $K$  als Teilkörper enthält (also sodass die Addition und Multiplikation in  $K$  die Einschränkungen derjenigen von  $\bar{K}$  sind). Durch Übergang von  $K$  zu  $\bar{K}$  kann man dann also immer erreichen, dass das charakteristische Polynom einer Matrix (oder eines Endomorphismus) in Linearfaktoren zerfällt. Die Bedingung an die Gleichheit der geometrischen und algebraischen Vielfachheiten der Eigenwerte ist also die eigentlich entscheidende für die Diagonalisierbarkeit.

Wir erinnern uns daran, dass das charakteristische Polynom  $p$  einer Matrix  $A = (a_{ij}) \in \text{Mat}(n, K)$  die Form

$$p = X^n - (a_{11} + a_{22} + \dots + a_{nn})X^{n-1} + \dots + (-1)^n \det(A)$$

hat; der Koeffizient von  $X^0$  ergibt sich dabei aus

$$p(0) = \det(0 \cdot I_n - A) = \det(-A) = (-1)^n \det(A).$$

Wenn  $p$  in Linearfaktoren zerfällt:

$$\begin{aligned} p &= (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n) \\ &= X^n - (\lambda_1 + \lambda_2 + \dots + \lambda_n)X^{n-1} + \dots + (-1)^n \lambda_1 \lambda_2 \cdots \lambda_n, \end{aligned}$$

dann sehen wir durch Vergleich der beiden Darstellungen von  $p$ , dass wir Summe und Produkt der Eigenwerte einfach von der Matrix ablesen können. Bevor wir das als Lemma formulieren, ist hier noch eine Definition:

**15.9. Definition.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $A = (a_{ij}) \in \text{Mat}(n, K)$ . Die *Spur* von  $A$  ist

**DEF**  
Spur einer  
Matrix

$$\text{Tr}(A) = \sum_{i=1}^n a_{ii} = a_{11} + a_{22} + \dots + a_{nn}.$$

(„Tr“ von englisch *trace*.)

◇

**15.10. Lemma.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $A \in \text{Mat}(n, K)$ . Das charakteristische Polynom von  $A$  zerfällt in Linearfaktoren. Dann ist  $\text{Tr}(A)$  die Summe und  $\det(A)$  das Produkt der Eigenwerte von  $A$ , jeweils entsprechend ihrer algebraischen Vielfachheit gezählt.

**LEMMA**  
Spur und  
Determinante  
durch  
Eigenwerte

Da die charakteristischen Polynome ähnlicher Matrizen gleich sind, folgt

$$\text{Tr}(PAP^{-1}) = \text{Tr}(A).$$

Die Spur erfüllt aber sogar noch eine etwas stärkere Aussage.

\* **15.11. Satz.** Die Spur ist eine  $K$ -lineare Abbildung  $\text{Mat}(n, K) \rightarrow K$ . Für alle  $A \in \text{Mat}(n, K)$  gilt  $\text{Tr}(A^\top) = \text{Tr}(A)$ .

**SATZ**  
Eigensch.  
der Spur

Sind  $A \in \text{Mat}(m \times n, K)$  und  $B \in \text{Mat}(n \times m, K)$ , dann gilt

$$\text{Tr}(AB) = \text{Tr}(BA).$$

Man beachte, dass  $AB$  eine  $m \times m$ -Matrix und  $BA$  eine  $n \times n$ -Matrix ist.

*Beweis.* Die erste Aussage ist klar (die Spur ist eine Linearkombination der Matrix-Einträge), die zweite ebenfalls, da  $A^\top$  und  $A$  dieselben Diagonaleinträge haben. Für die dritte Aussage sei  $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$  und  $B = (b_{kl})_{1 \leq k \leq n, 1 \leq l \leq m}$ . Der Diagonaleintrag in der  $i$ -ten Zeile und Spalte von  $C = AB$  ist

$$c_{ii} = \sum_{j=1}^n a_{ij} b_{ji} \quad (i \in \{1, 2, \dots, m\})$$

und der Diagonaleintrag in der  $j$ -ten Zeile und Spalte von  $C' = BA$  ist

$$c'_{jj} = \sum_{i=1}^m b_{ji} a_{ij} \quad (j \in \{1, 2, \dots, n\}).$$

Es folgt

$$\begin{aligned} \operatorname{Tr}(AB) &= \operatorname{Tr}(C) = \sum_{i=1}^m c_{ii} = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ji} \\ &= \sum_{j=1}^n \sum_{i=1}^m b_{ji} a_{ij} = \sum_{j=1}^n c'_{jj} = \operatorname{Tr}(C') = \operatorname{Tr}(BA). \quad \square \end{aligned}$$

Die Gleichung  $\operatorname{Tr}(A) = \operatorname{Tr}(PAP^{-1})$  folgt daraus:

$$\operatorname{Tr}(PAP^{-1}) = \operatorname{Tr}(P(AP^{-1})) = \operatorname{Tr}((AP^{-1})P) = \operatorname{Tr}(A(P^{-1}P)) = \operatorname{Tr}(A).$$

Analog zu Definition 14.24 können wir daher auch die Spur eines Endomorphismus definieren.

**15.12. Definition.** Seien  $K$  ein Körper,  $V$  ein endlich-dimensionaler  $K$ -Vektorraum und  $f \in \operatorname{End}(V)$ . Sei  $B$  eine beliebige Basis von  $V$  und  $A = \operatorname{Mat}_{B,B}(f)$ . Dann ist die *Spur* von  $f$  definiert als  $\operatorname{Tr}(f) = \operatorname{Tr}(A)$ .  $\diamond$

**DEF**  
Spur  
von Endomorphismen

Die Aussage, dass Spur und Determinante die Summe und das Produkt der Eigenwerte sind (mit algebraischer Vielfachheit gezählt), gilt dann entsprechend auch für Endomorphismen.

Zum Abschluss dieses Abschnitts wollen wir uns noch überlegen, wie man, wenn die Matrix  $A \in \operatorname{Mat}(n, K)$  diagonalisierbar ist, eine Matrix  $P$  findet, die  $A$  diagonalisiert, also sodass  $PAP^{-1} = D$  eine Diagonalmatrix ist.

**15.13. Lemma.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$  und  $A \in \operatorname{Mat}(n, K)$  eine diagonalisierbare Matrix. Sei  $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$  eine Basis von  $K^n$ , die aus Eigenvektoren von  $A$  besteht, mit  $A\mathbf{b}_j = \lambda_j \mathbf{b}_j$  (wir betrachten  $\mathbf{b}_j$  als Spaltenvektor). Sei  $Q \in \operatorname{GL}(n, K)$  die Matrix, deren  $j$ -te Spalte  $\mathbf{b}_j$  ist für  $j \in \{1, 2, \dots, n\}$ . Dann ist

$$Q^{-1}AQ = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

eine Diagonalmatrix.

**LEMMA**  
Diagonalisierung

Man kann dann also  $P = Q^{-1}$  nehmen.

*Beweis.* Sei  $\mathbf{e}_j$  der  $j$ -te Standard-Basisvektor von  $K^n$  als Spaltenvektor. Dann gilt  $Q\mathbf{e}_j = \mathbf{b}_j$  und damit auch  $Q^{-1}\mathbf{b}_j = \mathbf{e}_j$ . Es folgt

$$Q^{-1}AQ\mathbf{e}_j = Q^{-1}A\mathbf{b}_j = \lambda_j Q^{-1}\mathbf{b}_j = \lambda_j \mathbf{e}_j.$$

Das zeigt, dass die  $j$ -te Spalte von  $Q^{-1}AQ$  gerade  $\lambda_j \mathbf{e}_j$  ist, also ist

$$Q^{-1}AQ = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n). \quad \square$$

**15.14. Beispiel.** Sei  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \operatorname{Mat}(2, \mathbb{R})$ . Dann hat  $A$  die beiden Eigenwerte 1 und  $-1$  und ist daher diagonalisierbar. Wir hatten in Beispiel 14.12 Basen der beiden Eigenräume gefunden:

$$E_1(A) = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle \quad \text{und} \quad E_{-1}(A) = \left\langle \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\rangle.$$

Eine geeignete Matrix  $Q$  ist demnach  $Q = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ ; mit  $P = Q^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$  ist dann  $PAP^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ .  $\clubsuit$

**BSP**  
Diagonalisieren einer Matrix

## 16. BILINEARFORMEN

Wir unterbrechen die Untersuchung der Struktur von Endomorphismen an dieser Stelle; sie wird in der *Linearen Algebra II* fortgesetzt; dann werden wir auch genauer sehen, wie Endomorphismen aussehen, die nicht diagonalisierbar sind. Wir wenden uns jetzt erst einmal einem anderen Thema zu.

\*

**16.1. Definition.** Seien  $K$  ein Körper und  $V_1, V_2, W$  drei  $K$ -Vektorräume. Eine Abbildung  $\beta : V_1 \times V_2 \rightarrow W$  heißt *(K-)bilinear*, wenn  $\beta$  in jedem der beiden Argumente  $K$ -linear ist, also wenn für alle  $v_1, v'_1 \in V_1, v_2, v'_2 \in V_2$  und  $\lambda \in K$  gilt

$$\begin{aligned}\beta(v_1 + v'_1, v_2) &= \beta(v_1, v_2) + \beta(v'_1, v_2), & \beta(\lambda v_1, v_2) &= \lambda \beta(v_1, v_2) \\ \beta(v_1, v_2 + v'_2) &= \beta(v_1, v_2) + \beta(v_1, v'_2), & \beta(v_1, \lambda v_2) &= \lambda \beta(v_1, v_2).\end{aligned}$$

Ist  $W = K$ , dann heißt  $\beta$  eine *(K-)Bilinearform*. Gilt außerdem  $V_1 = V_2 = V$ , dann heißt  $\beta$  eine *(K-)Bilinearform auf  $V$* .

Ist  $\beta : V \times V \rightarrow K$  eine Bilinearform auf  $V$ , dann heißt  $\beta$  *symmetrisch*, wenn für alle  $v_1, v_2 \in V$  gilt, dass  $\beta(v_2, v_1) = \beta(v_1, v_2)$  ist.  $\beta$  heißt *alternierend*, wenn für alle  $v \in V$  gilt, dass  $\beta(v, v) = 0$  ist.  $\diamond$

Ist  $\beta : V \times V \rightarrow K$  eine alternierende Bilinearform, dann gilt  $\beta(v_2, v_1) = -\beta(v_1, v_2)$ . Das sieht man so:

$$\begin{aligned}0 &= \beta(v_1 + v_2, v_1 + v_2) \\ &= \beta(v_1, v_1) + \beta(v_1, v_2) + \beta(v_2, v_1) + \beta(v_2, v_2) \\ &= \beta(v_1, v_2) + \beta(v_2, v_1).\end{aligned}$$

Umgekehrt folgt aus  $\beta(v_2, v_1) = -\beta(v_1, v_2)$  für alle  $v_1, v_2 \in V$  die Gleichung  $\beta(v, v) = -\beta(v, v)$ , also  $2\beta(v, v) = 0$  für alle  $v \in V$ . Kann man in  $K$  durch 2 teilen (im Körper  $\mathbb{F}_2$  mit zwei Elementen ist  $2 = 0$ , dort geht das nicht, aber sonst praktisch immer), dann folgt, dass  $\beta$  alternierend ist. Im Normalfall sind alternierende Bilinearformen also dasselbe wie schief-symmetrische.

Bilineare Abbildungen treten häufig in Gestalt einer Multiplikationsabbildung auf.

**16.2. Beispiele.** Die Matrixmultiplikation

$$\text{Mat}(l \times m, K) \times \text{Mat}(m \times n, K) \longrightarrow \text{Mat}(l \times n, K), \quad (A, B) \longmapsto AB$$

ist eine bilineare Abbildung (das folgt aus den Rechenregeln für Matrizen). Genauso ist die Multiplikation von Polynomen

$$K[X] \times K[X] \longrightarrow K[X], \quad (p, q) \longmapsto pq$$

eine bilineare Abbildung.

Das *Standard-Skalarprodukt*

$$K^n \times K^n \longrightarrow K, \quad ((x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)) \longmapsto x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

ist eine symmetrische Bilinearform auf  $K^n$ .

Die Abbildung

$$K^2 \times K^2 \longrightarrow K, \quad ((x_1, x_2), (y_1, y_2)) \longmapsto x_1 y_2 - x_2 y_1$$

ist eine alternierende Bilinearform auf  $K^2$ .

Die *Spurform*

$$\text{Mat}(m \times n, K) \times \text{Mat}(m \times n, K) \longrightarrow K, \quad (A, B) \longmapsto \text{Tr}(A^\top B) = \text{Tr}(AB^\top)$$

**DEF**  
bilineare Abb.  
Bilinearform

**BSP**  
bilineare Abb.

ist eine symmetrische Bilinearform auf  $\text{Mat}(m \times n, K)$ . ♣

Allgemein gilt (leichte Übung): Ist  $\beta : V_1 \times V_2 \rightarrow W$  bilinear und  $f : W \rightarrow W'$  linear, dann ist  $f \circ \beta$  bilinear.

Auch Bilinearformen auf endlich-dimensionalen Vektorräumen lassen sich durch Matrizen darstellen.

**16.3. Definition.** Sei  $V$  ein endlich-dimensionaler  $K$ -Vektorraum mit  $n = \dim V$  und sei  $\beta : V \times V \rightarrow K$  eine Bilinearform auf  $V$ . Sei weiter  $B = (b_1, b_2, \dots, b_n)$  eine Basis von  $V$ . Dann heißt

**DEF**  
Matrix  
einer  
Bilinearform

$$\text{Mat}_B(\beta) = (\beta(b_i, b_j))_{1 \leq i, j \leq n} = \begin{pmatrix} \beta(b_1, b_1) & \beta(b_1, b_2) & \cdots & \beta(b_1, b_n) \\ \beta(b_2, b_1) & \beta(b_2, b_2) & \cdots & \beta(b_2, b_n) \\ \vdots & \vdots & \ddots & \vdots \\ \beta(b_n, b_1) & \beta(b_n, b_2) & \cdots & \beta(b_n, b_n) \end{pmatrix}$$

die Matrix von  $\beta$  bezüglich  $B$ . ◇

Sind  $v = x_1 b_1 + x_2 b_2 + \dots + x_n b_n$  und  $v' = y_1 b_1 + y_2 b_2 + \dots + y_n b_n$  zwei Elemente von  $V$ , dann ist  $\beta(v, v') = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \beta(b_i, b_j)$ , was sich in folgende Matrixmultiplikation übersetzen lässt (rechts steht eine  $1 \times 1$ -Matrix, die wir mit ihrem einzigen Eintrag identifizieren):

$$\beta(v, v') = (x_1, x_2, \dots, x_n) \text{Mat}_B(\beta) (y_1, y_2, \dots, y_n)^\top.$$

**16.4. Lemma.** Seien  $K$  ein Körper,  $n \in \mathbb{N}$ ,  $V$  ein  $n$ -dimensionaler  $K$ -Vektorraum mit Basen  $B$  und  $B'$  und  $\beta : V \times V \rightarrow K$  eine Bilinearform. Seien  $A = \text{Mat}_B(\beta)$  und  $A' = \text{Mat}_{B'}(\beta)$ .

**LEMMA**  
Matrizen  
einer  
Bilinearform

- (1)  $\beta$  ist symmetrisch  $\iff A^\top = A$  (d.h.  $A$  ist symmetrisch).
- (2) Sei  $P = \text{Mat}_{B', B}(\text{id}_V)$  die Basiswechselmatrix. Dann ist

$$A' = P^\top A P.$$

*Beweis.* Sei  $B = (b_1, b_2, \dots, b_n)$  und  $B' = (b'_1, b'_2, \dots, b'_n)$ .

- (1) „ $\Rightarrow$ “: Ist  $\beta$  symmetrisch, dann ist  $\beta(b_i, b_j) = \beta(b_j, b_i)$ ; das bedeutet gerade  $A^\top = A$ .
- „ $\Leftarrow$ “: Sei  $A^\top = A$ . Dann gilt für Spaltenvektoren  $\mathbf{x}, \mathbf{y} \in K^n$ :

$$\mathbf{x}^\top A \mathbf{y} = (\mathbf{x}^\top A \mathbf{y})^\top = \mathbf{y}^\top A^\top \mathbf{x} = \mathbf{y}^\top A \mathbf{x}.$$

(Beachte:  $\mathbf{x}^\top A \mathbf{y}$  ist eine  $1 \times 1$ -Matrix und damit gleich ihrer Transponierten.) Daraus folgt  $\beta(v, v') = \beta(v', v)$  für alle  $v, v' \in V$ .

- (2) Wir stellen die „neuen“ Basisvektoren  $b'_j$  als Linearkombination der „alten“ Basisvektoren  $b_i$  dar:  $b'_j = \sum_{i=1}^n \lambda_{ij} b_i$  mit  $\lambda_{ij} \in K$ . Dann ist  $P = (\lambda_{ij})_{1 \leq i, j \leq n}$  (denn in der  $j$ -ten Spalte stehen die Koeffizienten  $\lambda_{1j}, \dots, \lambda_{nj}$  der Darstellung von  $b'_j = \text{id}_V(b'_j)$  als Linearkombination der Basis  $B$ ). Aus der Bilinearität von  $\beta$  folgt

$$\beta(b'_k, b'_l) = \beta\left(\sum_{i=1}^n \lambda_{ik} b_i, \sum_{j=1}^n \lambda_{jl} b_j\right) = \sum_{i=1}^n \sum_{j=1}^n \lambda_{ik} \beta(b_i, b_j) \lambda_{jl};$$

das ist gerade der Eintrag in der  $k$ -ten Zeile und  $l$ -ten Spalte von  $P^\top A P$ . Da links der Eintrag in der  $k$ -ten Zeile und  $l$ -ten Spalte von  $A'$  steht, folgt die Behauptung. □

Wir betrachten im Folgenden den Fall  $K = \mathbb{R}$ . Dann können wir zwischen positiven und negativen Elementen von  $\mathbb{R}$  unterscheiden. Das führt zu folgender Definition.

\* **16.5. Definition.** Sei  $V$  ein  $\mathbb{R}$ -Vektorraum und  $\beta : V \times V \rightarrow \mathbb{R}$  eine symmetrische Bilinearform auf  $V$ .

**DEF**  
positiv/  
negativ  
definit

- (1)  $\beta$  heißt *positiv semidefinit*, wenn  $\beta(v, v) \geq 0$  ist für alle  $v \in V$ .
- (2)  $\beta$  heißt *positiv definit*, wenn  $\beta(v, v) > 0$  ist für alle  $\mathbf{0} \neq v \in V$ .
- (3)  $\beta$  heißt *negativ semidefinit*, wenn  $\beta(v, v) \leq 0$  ist für alle  $v \in V$ .
- (4)  $\beta$  heißt *negativ definit*, wenn  $\beta(v, v) < 0$  ist für alle  $\mathbf{0} \neq v \in V$ .
- (5)  $\beta$  heißt *indefinit*, wenn es  $v, v' \in V$  gibt mit  $\beta(v, v) > 0$  und  $\beta(v', v') < 0$ .

Seien  $n \in \mathbb{N}$  und  $A \in \text{Mat}(n, \mathbb{R})$  symmetrisch, also  $A^\top = A$ .

- (1)  $A$  heißt *positiv semidefinit*, wenn  $\mathbf{x}^\top A \mathbf{x} \geq 0$  ist für alle Spaltenvektoren  $\mathbf{x} \in \mathbb{R}^n$ .
- (2)  $A$  heißt *positiv definit*, wenn  $\mathbf{x}^\top A \mathbf{x} > 0$  ist für alle Spaltenvektoren  $\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$ .
- (3)  $A$  heißt *negativ semidefinit*, wenn  $\mathbf{x}^\top A \mathbf{x} \leq 0$  ist für alle Spaltenvektoren  $\mathbf{x} \in \mathbb{R}^n$ .
- (4)  $A$  heißt *negativ definit*, wenn  $\mathbf{x}^\top A \mathbf{x} < 0$  ist für alle Spaltenvektoren  $\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$ .
- (5)  $A$  heißt *indefinit*, wenn es Spaltenvektoren  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  gibt mit  $\mathbf{x}^\top A \mathbf{x} > 0$  und  $\mathbf{y}^\top A \mathbf{y} < 0$ .  $\diamond$

Daraus folgt im Fall  $\dim V < \infty$ , dass  $\beta$  genau dann positiv/negativ (semi-)definit bzw. indefinit ist, wenn das für  $\text{Mat}_B(\beta)$  mit irgendeiner Basis  $B$  von  $V$  gilt.

**16.6. Beispiele.** Das Standard-Skalarprodukt auf  $\mathbb{R}^n$  ist positiv definit, denn  $x_1^2 + x_2^2 + \dots + x_n^2 > 0$ , wenn nicht alle  $x_j$  null sind.

**BSP**  
positiv  
definite  
Bilinear-  
formen

Die Spurform auf  $\text{Mat}(m \times n, \mathbb{R})$  ist ebenfalls positiv definit, denn für eine Matrix  $A = (a_{ij}) \in \text{Mat}(m \times n, \mathbb{R})$  ist

$$\text{Tr}(A^\top A) = \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2.$$

(Wenn man  $\text{Mat}(m \times n, \mathbb{R})$  mit  $\mathbb{R}^{mn}$  in der üblichen Weise identifiziert, dann ist die Spurform einfach das Standard-Skalarprodukt.)  $\clubsuit$

Die Matrix einer positiv definiten symmetrischen Bilinearform kann auch negative Einträge haben und eine symmetrische Matrix mit lauter positiven Einträgen braucht nicht positiv definit zu sein.

**16.7. Beispiele.** Seien

**BSP**

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Dann ist  $A$  positiv definit, denn

$$\begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} = 2x^2 - 2xy + 2y^2 = 2(x - \frac{1}{2}y)^2 + \frac{3}{2}y^2,$$

und  $B$  ist nicht positiv semidefinit, denn

$$(1 \quad -1) B \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1^2 + 4 \cdot 1 \cdot (-1) + (-1)^2 = -2. \quad \clubsuit$$

Tatsächlich ist  $B$  indefinit, denn  $\mathbf{e}_1^\top B \mathbf{e}_1 = 1$ .

Unser Ziel wird es sein, ein relativ einfaches Kriterium herzuleiten, mit dem man entscheiden kann, ob eine symmetrische Matrix positiv (oder negativ) definit ist. Dies geschieht im Hinblick auf Anwendungen in der Analysis II (dort wird es um Kriterien gehen, wann eine Funktion mehrerer Variabler ein Maximum oder Minimum hat).

Dafür werden wir folgende Aussage verwenden, die wir allerdings jetzt noch nicht beweisen werden. Das werden wir dann in der *Linearen Algebra II* nachholen. Zuerst noch eine Definition.

**16.8. Definition.** Sei  $n \in \mathbb{N}$ . Eine Matrix  $A \in \text{Mat}(n, \mathbb{R})$  heißt *orthogonal*, wenn  $A^\top A = I_n$  ist. Wir schreiben  $O(n)$  für die Menge der orthogonalen Matrizen in  $\text{Mat}(n, \mathbb{R})$ . **DEF**  
orthogonale  
Matrix  $\diamond$

Dann ist insbesondere  $A$  invertierbar (mit  $A^{-1} = A^\top$ ). Man prüft ohne große Schwierigkeiten nach, dass  $O(n)$  eine Gruppe (mit der Matrixmultiplikation als Verknüpfung) ist.

**16.9. Satz.** Ist  $A \in \text{Mat}(n, \mathbb{R})$  symmetrisch, gilt also  $A^\top = A$ , dann ist  $A$  (über  $\mathbb{R}$ ) orthogonal diagonalisierbar: Es gibt eine orthogonale Matrix  $P \in O(n)$ , sodass  $P^\top A P = P^{-1} A P = D$  eine Diagonalmatrix ist. **SATZ**  
Spektral-  
satz

Daraus folgt leicht:

**16.10. Lemma.** Sei  $A \in \text{Mat}(n, \mathbb{R})$  symmetrisch. Dann gilt:

- (1)  $A$  ist genau dann positiv definit, wenn alle Eigenwerte von  $A$  positiv sind.
  - (2)  $A$  ist genau dann negativ definit, wenn alle Eigenwerte von  $A$  negativ sind.
  - (3)  $A$  ist genau dann positiv semidefinit, wenn  $A$  keinen negativen Eigenwert hat.
  - (4)  $A$  ist genau dann negativ semidefinit, wenn  $A$  keinen positiven Eigenwert hat.
  - (5)  $A$  ist genau dann indefinit, wenn  $A$  positive und negative Eigenwerte hat.
- LEMMA**  
Definitheit  
über Eigen-  
werte

*Beweis.* Nach Satz 16.9 gibt es  $P \in O(n)$ , sodass

$$P^\top A P = P^{-1} A P = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

eine Diagonalmatrix ist; ihre Diagonaleinträge sind gerade die Eigenwerte von  $A$ . Nach Lemma 16.4 ist  $D$  die Matrix der  $A$  entsprechenden symmetrischen Bilinearform auf  $\mathbb{R}^n$  bezüglich einer anderen Basis (gegeben durch die Spalten von  $P$ ), also ist  $A$  genau dann positiv definit, wenn  $D$  positiv definit ist. Für einen Spaltenvektor  $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$  gilt

$$\mathbf{x}^\top D \mathbf{x} = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_n x_n^2.$$

Sind alle  $\lambda_j > 0$ , dann ist das positiv für alle  $\mathbf{x} \neq \mathbf{0}$ , also ist  $D$  (und damit  $A$ ) positiv definit. Ist hingegen  $\lambda_j \leq 0$  für ein  $j$ , dann ist  $\mathbf{x}^\top D \mathbf{x} \leq 0$  für  $\mathbf{x} = \mathbf{e}_j$ , und

$D$  (und damit  $A$ ) ist nicht positiv definit. Die anderen Aussagen sieht man auf die gleiche Weise.  $\square$

Das Definitheitskriterium wird mit Hilfe von Determinanten geeigneter Untermatrizen formuliert, sogenannten Minoren.

**16.11. Definition.** Seien  $K$  ein Körper,  $m, n \in \mathbb{N}$ ,  $A = (a_{ij}) \in \text{Mat}(m \times n, K)$  und  $0 \leq r \leq \min\{m, n\}$ . Eine  $r \times r$ -Untermatrix von  $A$  ist eine Matrix der Form  $(a_{i_k, j_l})_{1 \leq k, l \leq r}$ , wobei  $1 \leq i_1 < i_2 < \dots < i_r \leq m$  und  $1 \leq j_1 < j_2 < \dots < j_r \leq n$ . Man wählt also  $r$  Zeilen und  $r$  Spalten von  $A$  aus und bildet die Matrix aus den Einträgen in diesen Zeilen und Spalten.

**DEF**  
Untermatrix  
Minor  
Hauptminor

Ein  $r$ -Minor von  $A$  ist die Determinante einer  $r \times r$ -Untermatrix von  $A$ . Ist  $m = n$  und  $A$  eine symmetrische Matrix, dann ist ein  $r$ -Hauptminor von  $A$  ein  $r$ -Minor von  $A$ , sodass in der obigen Notation  $i_1 = j_1, i_2 = j_2, \dots, i_r = j_r$  gilt (man wählt also dieselben Zeilen- und Spaltenindizes aus). Der führende  $r$ -Hauptminor von  $A$  ist die Determinante der Untermatrix  $(a_{ij})_{1 \leq i, j \leq r}$ , die aus den ersten  $r$  Zeilen und Spalten von  $A$  gebildet wird.  $\diamond$

Minoren sind manchmal nützlich, um den Rang einer Matrix zu beschreiben.

**16.12. Lemma.** Seien  $K$  ein Körper,  $m, n \in \mathbb{N}$ ,  $A \in \text{Mat}(m \times n, K)$  und sei  $1 \leq r \leq \min\{m, n\}$ . Dann sind folgende Aussagen äquivalent:

**LEMMA**  
Rang über  
Minoren

- (1)  $\text{rk}(A) < r$ .
- (2) Alle  $r$ -Minoren von  $A$  verschwinden.

*Beweis.* „(1)  $\Rightarrow$  (2)“: Sei  $A'$  eine  $r \times r$ -Untermatrix von  $A$ . Da je  $r$  Spalten von  $A$  linear abhängig sind, gilt das auch für die Spalten von  $A'$ , also ist  $\det A' = 0$ .

„(2)  $\Rightarrow$  (1)“: Wir nehmen an, dass  $\text{rk}(A) \geq r$  ist und zeigen, dass es einen nicht verschwindenden  $r$ -Minor gibt. Nach Voraussetzung gibt es  $r$  linear unabhängige Spalten in  $A$ ; sei  $B$  die  $m \times r$ -Matrix, die aus diesen  $r$  Spalten besteht. Dann ist  $\text{rk}(B) = r$ , also hat  $B$  auch  $r$  linear unabhängige Zeilen. Sei  $A'$  die Matrix, die aus diesen  $r$  Zeilen von  $B$  besteht; dann ist  $A'$  eine  $r \times r$ -Untermatrix von  $A$ . Außerdem ist  $\text{rk}(A') = r$ , also ist der  $r$ -Minor  $\det(A')$  von  $A$  nicht null.  $\square$

Mit Hilfe der Minoren lassen sich auch die weiteren Koeffizienten des charakteristischen Polynoms ausdrücken. Wenn wir „ $r$ -Hauptminor“ auch für nicht unbedingt symmetrische quadratische Matrizen wie oben definieren, dann gilt für eine Matrix  $A \in \text{Mat}(n, K)$  mit charakteristischem Polynom  $p \in K[X]$ :

$$p = \sum_{k=0}^n (-1)^k s_k(A) X^{n-k},$$

wobei  $s_k(A)$  die Summe der  $k$ -Hauptminoren von  $A$  ist. Für  $k = 1$  ist das gerade die Spur von  $A$ , denn die 1-Hauptminoren sind genau die Einträge auf der Diagonalen; für  $k = n$  ist das die Determinante von  $A$  (der einzige  $n$ -Hauptminor). Eine Möglichkeit das einzusehen besteht darin, die Multilinearität der Determinante als Funktion (z.B.) der Zeilen einer Matrix zu verwenden (vergleiche das Kleingedruckte auf Seite 95). Für eine Teilmenge  $T$  von  $\{1, 2, \dots, n\}$  sei  $A_T$  die  $n \times n$ -Matrix, deren  $j$ -te Zeile für  $j \in T$  mit der  $j$ -ten Zeile von  $A$  und für  $j \notin T$  mit der  $j$ -ten Zeile von  $I_n$  übereinstimmt. Dann ist

$$p = \det(XI_n - A) = \sum_{T \subset \{1, 2, \dots, n\}} \det(-A_T) X^{n-\#T} = \sum_{T \subset \{1, 2, \dots, n\}} (-1)^{\#T} \det(-A_T) X^{n-\#T}$$

und  $\det(A_T)$  ist gerade der  $\#T$ -Minor von  $A$ , der zu den Zeilen und Spalten mit Nummern in  $T$  gehört (wie man durch Entwicklung nach den anderen Zeilen sieht).

Wir wollen die Minoren jetzt aber benutzen, um nachzuweisen, dass eine symmetrische Matrix positiv (oder negativ) definit ist.

\*

**16.13. Satz.** Seien  $n \in \mathbb{N}$  und  $A = (a_{ij}) \in \text{Mat}(n, \mathbb{R})$  eine symmetrische Matrix. Für  $r \in \{1, 2, \dots, n\}$  sei  $d_r(A) = \det(a_{ij})_{1 \leq i, j \leq r}$  der führende  $r$ -Hauptminor von  $A$ . Dann gilt:

**SATZ**  
Determinantenkriterium für positiv definit

- (1)  $A$  ist positiv definit  $\iff d_r(A) > 0$  für alle  $r \in \{1, 2, \dots, n\}$ .
- (2)  $A$  ist negativ definit  $\iff (-1)^r d_r(A) > 0$  für alle  $r \in \{1, 2, \dots, n\}$ .

Die Bedingung für „negativ definit“ heißt also  $d_1(A) < 0, d_2(A) > 0, d_3(A) < 0$  usw.: Die führenden Hauptminoren alternieren im Vorzeichen. Man merkt sich das am besten an den Vorzeichen der führenden Hauptminoren von  $-I_n$ .

*Beweis.* Wir beweisen zunächst Aussage (1). Die Richtung „ $\implies$ “ folgt aus Lemma 16.10, denn mit  $A$  sind auch die Matrizen  $A_r = (a_{ij})_{1 \leq i, j \leq r}$  positiv definit (siehe Lemma 16.16 unten), und eine positiv definite Matrix hat positive Determinante (denn die ist das Produkt der (positiven) Eigenwerte). Die Richtung „ $\impliedby$ “ zeigen wir durch Induktion über  $n$ . Der Fall  $n = 0$  (oder  $n = 1$ ) ist klar. Für den Schritt von  $n$  auf  $n + 1$  sei  $A \in \text{Mat}(n + 1, \mathbb{R})$  symmetrisch mit positiven führenden Hauptminoren  $d_r(A)$  für alle  $r \in \{1, 2, \dots, n + 1\}$ . Das gilt dann entsprechend auch für die Matrix  $A_n \in \text{Mat}(n, \mathbb{R})$  (denn  $d_r(A_n) = d_r(A)$  für  $r \leq n$ ). Nach Induktionsvoraussetzung ist  $A_n$  positiv definit. Das heißt, dass für Spaltenvektoren  $\mathbf{0} \neq \mathbf{x} \in \langle \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \rangle \subset \mathbb{R}^{n+1}$  stets  $\mathbf{x}^\top A \mathbf{x} > 0$  ist. Wir zeigen jetzt, dass  $A$  höchstens einen negativen Eigenwert haben kann: Nach Satz 16.9 gibt es  $P \in O(n + 1)$  mit

$$P^\top A P = P^{-1} A P = \text{diag}(\lambda_1, \dots, \lambda_{n+1})$$

diagonal. Wären wenigstens zwei Eigenwerte negativ, etwa  $\lambda_i$  und  $\lambda_j$ , mit zugehörigen Eigenvektoren  $\mathbf{y}_i = P \mathbf{e}_i$  und  $\mathbf{y}_j = P \mathbf{e}_j$  (als Spaltenvektoren), dann hätten wir für  $(0, 0) \neq (\alpha, \beta) \in \mathbb{R}^2$

$$(\alpha \mathbf{y}_i + \beta \mathbf{y}_j)^\top A (\alpha \mathbf{y}_i + \beta \mathbf{y}_j) = (\alpha \mathbf{e}_i + \beta \mathbf{e}_j)^\top P^\top A P (\alpha \mathbf{e}_i + \beta \mathbf{e}_j) = \lambda_i \alpha^2 + \lambda_j \beta^2 < 0.$$

Da die  $n + 2$  Vektoren  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n, \mathbf{y}_i, \mathbf{y}_j \in \mathbb{R}^{n+1}$  nicht linear unabhängig sein können, gibt es  $(0, 0) \neq (\alpha, \beta) \in \mathbb{R}^2$  mit  $\mathbf{0} \neq \mathbf{x} = \alpha \mathbf{y}_i + \beta \mathbf{y}_j \in \langle \mathbf{e}_1, \dots, \mathbf{e}_n \rangle$ . Dann müsste aber sowohl  $\mathbf{x}^\top A \mathbf{x} < 0$  als auch  $\mathbf{x}^\top A \mathbf{x} > 0$  gelten, ein Widerspruch. Es kann also keine zwei negativen Eigenwerte geben. Da das Produkt aller Eigenwerte  $d_{n+1}(A) = \det(A)$  positiv ist, kann es auch nicht genau einen negativen Eigenwert geben (und natürlich kann null kein Eigenwert sein), also sind alle Eigenwerte von  $A$  positiv; nach Lemma 16.10 ist  $A$  also positiv definit.

Aussage (2) folgt aus Aussage (1):  $A$  ist genau dann negativ definit, wenn  $-A$  positiv definit ist, und für die führenden Hauptminoren gilt  $d_r(-A) = (-1)^r d_r(A)$ . □

16.14. **Beispiele.** Wir betrachten wieder

**BSP**

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Die führenden Hauptminoren von  $A$  sind  $d_1(A) = 2$ ,  $d_2(A) = 2^2 - 1^2 = 3$ , was bestätigt, dass  $A$  positiv definit ist. Hingegen sind die führenden Hauptminoren von  $B$  gegeben durch  $d_1(B) = 1$  und  $d_2(B) = 1^2 - 2^2 = -3$ , was bestätigt, dass  $B$  nicht positiv definit ist (und auch nicht negativ definit, denn dafür haben beide Minoren das falsche Vorzeichen). ♣

Ist die symmetrische Matrix  $A \in \text{Mat}(n, \mathbb{R})$  nur positiv semidefinit, dann folgt wie im Beweis von „ $\Rightarrow$ “, dass die führenden Hauptminoren von  $A$  alle  $\geq 0$  sein müssen. Die Umkehrung gilt dann aber im Allgemeinen nicht.

16.15. **Beispiel.** Die Matrix  $A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$  hat nichtnegative führende Hauptminoren (beide sind null), ist aber nicht positiv semidefinit (denn  $\mathbf{e}_2^\top A \mathbf{e}_2 = -1$ ). ♣

**BSP**

Hier ist noch das im Beweis von Satz 16.13 versprochene Lemma:

16.16. **Lemma.** Sei  $A = (a_{ij}) \in \text{Mat}(n, \mathbb{R})$  eine positiv definite symmetrische Matrix. Dann ist für  $r \in \{1, 2, \dots, n\}$  die Untermatrix  $A' = (a_{ij})_{1 \leq i, j \leq r}$  von  $A$  ebenfalls positiv definit.

**LEMMA**  
Untermatrizen erben positive Definitheit

*Beweis.* Sei  $\mathbf{0} \neq \mathbf{x}' = (x_1, \dots, x_r) \in \mathbb{R}^r$ . Wir müssen zeigen, dass  $(\mathbf{x}')^\top A' \mathbf{x}' > 0$  ist. Sei  $\mathbf{x} = (x_1, x_2, \dots, x_r, 0, \dots, 0) \in \mathbb{R}^n$  (wir fügen also  $n - r$  Nullen an). Dann ist  $\mathbf{x} \neq \mathbf{0}$ , also nach Voraussetzung  $\mathbf{x}^\top A \mathbf{x} > 0$ . Es genügt also zu zeigen, dass  $(\mathbf{x}')^\top A' \mathbf{x}' = \mathbf{x}^\top A \mathbf{x}$  ist. Mit  $x_j = 0$  für  $j \in \{r + 1, r + 2, \dots, n\}$  gilt

$$\mathbf{x}^\top A \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \sum_{i=1}^r \sum_{j=1}^r a_{ij} x_i x_j = (\mathbf{x}')^\top A' \mathbf{x}'. \quad \square$$

Es gibt auch ein Determinanten-Kriterium für positive (oder negative) Semidefinitheit. Es lautet wie folgt.

**Satz.** Seien  $n \in \mathbb{N}$  und  $A = (a_{ij}) \in \text{Mat}(n, \mathbb{R})$  eine symmetrische Matrix. Dann gilt:

**SATZ**  
Determinantenkriterium für semidefinit

- (1)  $A$  ist positiv semidefinit  $\iff d \geq 0$  für alle Hauptminoren  $d$  von  $A$ .
- (2)  $A$  ist negativ semidefinit  $\iff \forall r \in \{1, 2, \dots, n\} : (-1)^r d \geq 0$  für alle  $r$ -Hauptminoren  $d$  von  $A$ .
- (3)  $A$  ist indefinit  $\iff$  es gibt einen  $2r$ -Hauptminor  $d < 0$  von  $A$ , oder es gibt einen  $(2r + 1)$ -Hauptminor  $d > 0$  und einen  $(2r' + 1)$ -Hauptminor  $d' < 0$  von  $A$ .

*Beweis.* Aussage (3) folgt formal-logisch aus (1) und (2) ( $A$  ist genau dann indefinit, wenn  $A$  weder positiv noch negativ semidefinit ist). Aussage (2) folgt aus (1) durch Anwendung von (1) auf  $-A$ . Es genügt also, die erste Aussage zu zeigen. Die Richtung „ $\Rightarrow$ “ ist wieder klar: Jede Haupt-Untermatrix von  $A$  ist positiv semidefinit, hat also nichtnegative Eigenwerte und damit nichtnegative Determinante.

Zum Beweis von „ $\Leftarrow$ “ nehmen wir an, dass alle Hauptminoren von  $A$  nichtnegativ sind. Wir bemerken zunächst, dass aus Lemma 16.10 folgt, dass eine symmetrische Matrix mit nicht verschwindender Determinante positiv oder negativ definit oder indefinit sein muss. Sei  $K = \ker(A) \subset \mathbb{R}^n$  und  $k = \dim K$ . Wir wählen eine Basis  $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k)$  von  $K$ . Wir können diese Basis durch Hinzunahme von  $n - k$  Standard-Basisvektoren  $\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \dots, \mathbf{e}_{j_{n-k}}$  (mit  $1 \leq j_1 < j_2 < \dots < j_{n-k} \leq n$ ) zu einer Basis von  $\mathbb{R}^n$  ergänzen

(Basisergänzungssatz 8.14). Sei  $V = \langle \mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \dots, \mathbf{e}_{j_{n-k}} \rangle$ ; dann ist  $V \cap K = \{\mathbf{0}\}$  und jeder Vektor  $\mathbf{x} \in \mathbb{R}^n$  kann eindeutig geschrieben werden als  $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1$  mit  $\mathbf{x}_0 \in K$  und  $\mathbf{x}_1 \in V$ . Sei  $\beta : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  die symmetrische Bilinearform, deren Matrix bezüglich der Standard-Basis  $A$  ist. Dann gilt  $\beta(\mathbf{x}, \mathbf{x}_0) = \beta(\mathbf{x}_0, \mathbf{x}) = 0$  für alle  $\mathbf{x} \in \mathbb{R}^n$  und  $\mathbf{x}_0 \in K$ . Für  $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1$  wie oben gilt also  $\beta(\mathbf{x}, \mathbf{x}) = \beta(\mathbf{x}_1, \mathbf{x}_1)$ . Sei  $A'$  die  $(n-k) \times (n-k)$ -Untermatrix von  $A$  zu den Zeilen- und Spaltenindizes  $j_1, j_2, \dots, j_{n-k}$ . Dann ist  $A'$  eine Matrix der Bilinearform  $\beta' = \beta|_{V \times V}$ , und nach der obigen Überlegung ist  $A$  genau dann positiv semidefinit, wenn das für  $A'$  gilt. Außerdem ist  $\ker(A') = \{\mathbf{0}\}$  (wegen  $K \cap V = \{\mathbf{0}\}$ ), also ist  $\det(A') \neq 0$ . Damit ist  $A'$  positiv oder negativ definit oder indefinit. Wie im Beweis von Satz 16.13 zeigt man induktiv, dass  $A'$  keine zwei negativen Eigenwerte haben kann. Wegen  $\det(A') > 0$  (hier verwenden wir die Voraussetzung) müssen alle Eigenwerte von  $A'$  positiv sein. Damit ist  $A'$  positiv definit, also ist  $A$  positiv semidefinit.  $\square$

Dieses Kriterium ist sehr viel weniger nützlich als Satz 16.13: Es gibt  $2^n$  Hauptminoren, aber nur  $n$  führende Hauptminoren. Der Aufwand dafür, *alle* Hauptminoren zu testen, wird also schon für relativ kleine  $n$  zu groß, um praktikabel zu sein. Zum Glück gibt es bessere Möglichkeiten. Wir werden in der *Linearen Algebra II* genauer darauf eingehen.

## 17. VOLUMINA

Zum Abschluss der Vorlesung wollen wir noch die geometrische Bedeutung der Determinante untersuchen. Wir betrachten den  $\mathbb{R}^n$  und definieren erst einmal den Begriff der (positiven oder negativen) Orientierung einer Basis.

\* **17.1. Definition.** Sei  $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$  eine Basis des  $\mathbb{R}^n$  und  $A = (\mathbf{b}_1 \mid \dots \mid \mathbf{b}_n)$  die Matrix, deren Spalten die Basisvektoren sind. Wir sagen, die Basis sei *positiv orientiert*, wenn  $\det(A) > 0$  ist, und *negativ orientiert*, wenn  $\det(A) < 0$  ist.  $\diamond$  **DEF** Orientierung einer Basis

Die Standard-Basis ist positiv orientiert. Im Fall  $n = 2$  ist eine Basis positiv orientiert, wenn der gegen den Uhrzeigersinn gemessene Winkel vom ersten zum zweiten Basisvektor kleiner ist als  $\pi$  ( $= 180^\circ$ ).

Der Vergleich der Orientierung einer Basis und ihres Bildes führt zum Begriff der orientierungserhaltenden bzw. -umkehrenden linearen Abbildung.

**17.2. Definition.** Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  ein Automorphismus (also ein invertierbarer Endomorphismus). Dann heißt  $f$  *orientierungserhaltend*, wenn  $f$  positiv orientierte Basen auf positiv orientierte Basen abbildet, und *orientierungsumkehrend*, wenn  $f$  positiv orientierte Basen auf negativ orientierte Basen abbildet.  $\diamond$  **DEF** orientierungserhaltend, -umkehrend

Man sieht leicht, dass  $f$  genau dann orientierungserhaltend (-umkehrend) ist, wenn  $\det(f) > 0$  ( $< 0$ ) ist.

Wir wollen jetzt das Volumen von „verzerrten Würfeln“ betrachten. In der Ebene  $\mathbb{R}^2$  sind das Parallelogramme. Allgemeiner definieren wir:

**17.3. Definition.** Ein *Parallelotop* im  $\mathbb{R}^n$  ist die Menge **DEF** Parallelotop

$$P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = \{t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \dots + t_n\mathbf{x}_n \mid 0 \leq t_1, t_2, \dots, t_n \leq 1\} \subset \mathbb{R}^n$$

für ein  $n$ -Tupel  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$  von Vektoren im  $\mathbb{R}^n$ . Das Parallelotop heißt *ausgeartet*, wenn die es aufspannenden Vektoren  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  linear abhängig sind (dann ist  $P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$  im echten Untervektorraum  $\langle \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \rangle$  von  $\mathbb{R}^n$  enthalten).  $\diamond$

Wir wollen jetzt untersuchen, wie man das „orientierte Volumen“ solcher Parallelotope definieren kann. Es sollte folgende Eigenschaften haben:

- (1)  $\text{vol } P(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = 1$  (der  $n$ -dimensionale Einheitswürfel hat Volumen 1).
- (2)  $\text{vol } P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$  ist positiv (bzw. negativ), wenn  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$  eine positiv (bzw. negativ) orientierte Basis von  $\mathbb{R}^n$  ist.
- (3)  $\text{vol } P(\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \lambda\mathbf{x}_j, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n) = \lambda \text{vol } P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$  für  $\lambda \in \mathbb{R}$ .
- (4)  $\text{vol } P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = 0$ , wenn  $P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$  ausgeartet ist.
- (5)  $\text{vol } P(\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \mathbf{x}_j + \mathbf{x}'_j, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n)$   
 $= \text{vol } P(\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \mathbf{x}_j, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n)$   
 $+ \text{vol } P(\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \mathbf{x}'_j, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n)$ .

Diese letzte Eigenschaft kann man sich plausibel machen, wenn man an die Formel „Grundfläche mal Höhe“ denkt: Die Höhe von  $\mathbf{x}_j + \mathbf{x}'_j$  über der „Grundfläche“, die durch das von den übrigen Vektoren aufgespannte Parallelotop gegeben ist, ist die Summe der Höhen von  $\mathbf{x}_j$  und  $\mathbf{x}'_j$ .

**17.4. Satz.** Die einzige Abbildung  $\text{vol}$  von der Menge der Parallelotope im  $\mathbb{R}^n$  nach  $\mathbb{R}$ , die die obigen Eigenschaften hat, ist

$$\text{vol } P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = \det(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n),$$

wobei  $\det(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$  für die Determinante der Matrix steht, deren Spalten die Vektoren  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  sind.

**SATZ**  
Determinante  
ist Volumen

*Beweis.* Die Determinante hat jedenfalls die geforderten Eigenschaften (Satz 13.3 und Satz 13.8, sowie Definition 17.1). Aus der Eindeutigkeitsaussage in Satz 13.3, zusammen mit Satz 13.8, der besagt, dass die analoge Aussage auch für Spalten statt Zeilen gilt, zeigt, dass die Determinante die einzige Abbildung ist, die die Eigenschaften (1), (3), (4) und (5) hat.  $\square$

Man kann also mit Hilfe der Determinante Volumina messen.

**17.5. Beispiel.** Die Fläche des Dreiecks mit den Eckpunkten  $(x_1, y_1)$ ,  $(x_2, y_2)$  und  $(x_3, y_3)$  ist

$$\frac{1}{2} \left| \det \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{pmatrix} \right|.$$

**BSP**  
Fläche eines  
Dreiecks

Wir können das Dreieck so verschieben, dass die erste Ecke im Ursprung zu liegen kommt. Dann ist die gesuchte Fläche die Hälfte der Fläche des von  $(x_2 - x_1, y_2 - y_1)$  und  $(x_3 - x_1, y_3 - y_1)$  aufgespannten Parallelogramms. Die orientierte Fläche dieses Parallelogramms ist

$$\det \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix}.$$

Die obige Determinante lässt sich durch die Spaltenoperationen  $\mathbf{II}_{2,1}(-1)$  und  $\mathbf{II}_{3,1}(-1)$  und nachfolgender Entwicklung nach der dritten Zeile auf diese Form bringen. Durch den Absolutbetrag erhalten wir die Fläche statt der orientierten Fläche.  $\clubsuit$

Aus der Multiplikativität der Determinante folgt eine Interpretation der Determinante eines Endomorphismus, die für Anwendungen in der Analysis (z.B. die Transformationsformel für mehrdimensionale Integrale) relevant ist.

\* **17.6. Satz.** Sei  $f \in \text{End}(\mathbb{R}^n)$  und seien  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \in \mathbb{R}^n$ . Dann gilt

$$\text{vol } f(P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)) = \det(f) \text{vol } P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n).$$

**SATZ**  
Determinante  
ist Skalierung  
des Volumens

Die Determinante eines Endomorphismus gibt also an, mit welchem Faktor das Volumen bei seiner Anwendung multipliziert wird.

*Beweis.* Sei  $X$  die Matrix, deren Spalten die Vektoren  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  sind, und sei  $A$  die Matrix von  $f$  bezüglich der Standard-Basis von  $\mathbb{R}^n$ . Dann gilt

$$\begin{aligned} \text{vol } f(P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)) &= \det(f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_n)) \\ &= \det(A\mathbf{x}_1, A\mathbf{x}_2, \dots, A\mathbf{x}_n) \\ &= \det(AX) = \det(A) \det(X) \\ &= \det(f) \text{vol } P(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n). \end{aligned} \quad \square$$