# Vorlesung Lineare Algebra und Analytische Geometrie I

Andreas Knauf*

## Inhaltsverzeichnis



*Department Mathematik, Universität Erlangen-Nürnberg, Cauerstr. 11, 91058 Erlangen, Germany. e-mail: knauf@mi.uni-erlangen.de, web: www.mathematik.uni-erlangen.de/∼knauf

**Vorbemerkung:** Dieses Skript kann kein Lehrbuch ersetzen. Einige Lehrbücher zur Linearen Algebra sind im Literaturverzeichnis erwähnt.

**Danksagung** Ich danke Frau T. Dierkes und Frau A. Stroux, die viele Fehler im Manuskript fanden, und Frau I. Moch, die in detektivischer Arbeit meine Handschrift entzifferte und dieses Skript schrieb.

# 1 Einführung

Am Beginn des Mathematikstudiums stehen traditionell zwei Vorlesungen, die Analysis und die Lineare Algebra.

- Der zentrale Begriff der *Analysis* ist der des Grenzwertes. Durch Grenzwertbildung wird aus der Sekante die Tangente, aus der Summe das Integral. Differentiation und Integration führen zu Differential- und Integralgleichungen, mit deren Hilfe sich Naturvorgänge modellieren lassen.

- Die *Lineare Algebra*, mit der wir uns in dieser Vorlesung befassen, ist zunächst ein Teilgebiet der *Algebra*. Diese "ist die Lehre von den vier Grundrechenarten – Addition, Subtraktion, Multiplikation und Division – und der Auflösung der in diesem Zusammenhang entstehenden Gleichungen" (Erich Kähler, 1953).

  In der Linearen Algebra sind diese Gleichungen linear. Einfachstes Beispiel ist die Gleichung $ax = b$ mit der Unbekannten $x$ und der Lösung $x = b/a$.

Verallgemeinert werden Sie Methoden kennen lernen, mit denen sich die Lösungen eines Systems linearer Gleichungen mit mehreren Unbekannten finden lassen.

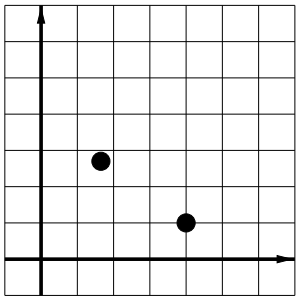
Eng verknüpft mit den linearen Gleichungssystemen sind die sog. *Vektorräume*. In unserem Beispiel ist die Lösung $x$ eine reelle Zahl oder, geometrisch gesehen, ein Punkt im eindimensionalen Vektorraum $\mathbb{R}$, der Zahlengerade. Ist der Koeffizient $a = 0$, dann erfüllt für $b = 0$ jedes $x$ unsere Gleichung, die Lösungsmenge ist also ganz $\mathbb{R}$, für $b \neq 0$ aber gibt es keine Lösung.

- Das zweite Thema dieser Vorlesung ist die *Analytische Geometrie*. In dieser wird unter anderem die Geometrie der Kegelschnitte (Ellipse, Parabel und Hyperbel) untersucht, die z.B. möglichen Bahnen von Himmelskörpern entsprechen. Dabei wird seit den Zeiten von Descartes (1596-1650) Gebrauch von den cartesischen Koordinaten in der Ebene gemacht. Diese gestatten es, geometrische Fragen algebraisch umzuformulieren und mit algebraischen Methoden zu lösen.

Im Lauf der Jahrhunderte wurde der Begriff des linearen oder auch Vektorraums aus der Intuition abstrahiert, die die Mathematiker über den dreidimensionalen Anschauungsraum oder die Zeichenebene entwickelt hatten.

In dieser Vorlesung wie in den meisten Mathematikvorlesungen wird umgekehrt bei einer axiomatischen Beschreibung der untersuchten Struktur, hier also des linearen Raumes, begonnen und die Anschauung danach in Sätzen und Beispielen entwickelt. Diese *deduktive* Vorgehensweise der Mathematik trägt in der Öffentlichkeit zu ihrem schlechten Ruf als abstrakt-unverständlicher Wissenschaft bei, hat aber ihren guten Grund. Sie dient der Entwicklung *neuer*, dem Untersuchungsobjekt angemessener Intuition. Ich möchte das am Beispiel der Zeichenebene (Tafel oder Papier) erläutern.

**1.1 Beispiel** Wenn wir ein Blatt Papier als Modell eines (zweidimensionalen) Raumes betrachten, abstrahieren wir zunächst von seiner endlichen Ausdehnung, stellen es uns nach allen Seiten fortgesetzt vor. Warum ist es zweidimensional? Nehmen wir an, das Papier ist kariert und ich zeichne einen Kreuzungspunkt als Nullpunkt aus. Dann kann ich jeden anderen *Kreuzungspunkt* durch Angabe *zweier* ganzer Zahlen eindeutig benennen;




durch Angabe zweier *reeller* Zahlen kann ich jeden *Punkt* der Ebene eindeutig benennen. Dagegen ist unser Anschauungsraum dreidimensional, denn wir benötigen für die Benennung seiner Punkte drei Koordinaten.

Den Punkten unseres Raumes ordnen wir jetzt Vektoren, also gerichtete Strecken, zu, die vom Nullpunkt ausgehen und im gegebenen Punkt enden. Einen solchen Vektor können wir mit seinem Endpunkt identifizieren. Es gibt also eine

$1:1$–Beziehung zwischen den Punkten des Raumes und den Vektoren (Diese Beziehung lässt sich aber erst dann herstellen, wenn man einen Nullpunkt gewählt hat).

Während man Punkte weder verlängern noch addieren kann, gilt dies für Vektoren. Wir können also damit die Elemente unseres Raumes der Vektoren mit beliebigen reellen Zahlen *multiplizieren* und wir können je zwei Elemente *addieren*.

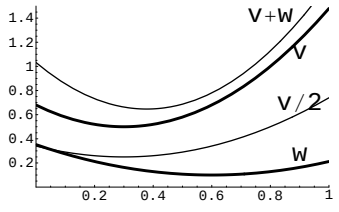
Wir können aber noch mehr tun:

- Wir können *Längen* $L$ von Vektoren messen; $L(a,b) := \sqrt{a^2+b^2}$ (Pythagoras).
- Wir können *Winkel* zwischen Vektoren messen; $(1,0)$ und $(0,1)$ beispielsweise sind rechtwinklig.
- Wir haben eine *Orientierung* in unserer Zeichenebene, können also sagen, dass $w$ aus $v$ durch eine Drehung im Uhrzeigersinn um einen Winkel $< \pi$ hervorgeht.
- Wir können *Flächen* geometrischer Figuren (bzw. *Volumina* im dreidimensionalen Raum) messen.

Kurz gesagt, unsere aus der Anschauung gewonnenen Räume haben eine große Zahl von Eigenschaften. Solange man sie als gegeben hinnimmt, ist es sinnlos, nach ihrer Abhängigkeit zu fragen.

Es hat sich aber im Verlauf der Mathematikgeschichte herausgestellt, dass der Raumbegriff auch in ganz anderen als geometrischen Zusammenhängen nützlich ist.

**1.2 Beispiel** Polynome des Grades $\leq 2$.
Wir betrachten Funktionen der Form $v(x) = ax^2 + bx + c$ mit reellen Koeffizienten $a, b, c$, also Polynome des Grades $\leq 2$. Multiplizieren wir eine solche Funktion $v$ mit einer reellen Zahl $r$, so erhalten wir die Funktion $rv$, $rv(x) = (ar)x^2 + (br)x + (cr)$,




also wieder ein Polynom vom Grad $\leq 2$. Zwei solche Polynome können wir addieren; mit $w(x) = dx^2 + ex + f$ ist $v + w$ von der Form $(v+w)(x) = (a+d)x^2 + (b+e)x + (c+f)$, also wieder ein entsprechendes Polynom.

Die Rechenoperationen sind so geartet, dass wir jedes dieser Polynome als einen Punkt im dreidimensionalen Raum auffassen können, wobei der Nullpunkt das Nullpolynom $p(x) \equiv 0$ ist.

**Frage:** Was ist die *Länge* eines so als Vektor aufgefassten Polynoms, was der *Winkel* zwischen zwei Polynomen etc.?

Erst einmal mutet diese Frage absurd an. Wenn man aber annimmt, dass man immer in einem linearen Raum Längen und Winkel messen kann, dann wird man diese Eigenschaft auch von unserem Polynomraum erwarten.

Tatsächlich ist es manchmal sinnvoll, die "Länge" $L$ eines Polynoms zu messen, z.B. nach der

**Definition:** $L(v) := \sqrt{\frac{1}{2}\int_{-1}^{1}(v(x))^2 dx}$.

Damit ergibt sich für das Polynom $v(x) = ax^2 + bx + c$

$$\begin{aligned} L^2(v) &= \tfrac{1}{2}\int_{-1}^{1}(ax^2+bx+c)^2 dx \\ &= \tfrac{1}{2}\int_{-1}^{1}(a^2x^4 + 2abx^3 + (2ac+b^2)x^2 + 2bcx + c^2)dx \\ L(v) &= \sqrt{\frac{a^2}{5} + \frac{2ac+b^2}{3} + c^2}. \end{aligned}$$

Warum sollte man diese Größe als "Länge" oder "Norm" von $v$ auffassen?

Nun, zunächst stellt sich heraus, dass diese Größe Eigenschaften besitzt, die wir von einer Länge erwarten:

- $L(v) \geq 0 \quad , \quad L(v) = 0 \quad$ nur wenn $\quad v \equiv 0 \quad$ ist
- $L(rv) = |r|L(v) \quad$ für jede reelle Zahl $\quad r$
- $L(v+w) \leq L(v) + L(w)$ (Dreiecksungleichung).

Wir können nun sagen, dass zwei solche Polynome $v_1$ und $v_2$ nahe benachbart sind, wenn ihr Abstand $L(v_1 - v_2)$ klein ist. Das wird dann der Fall sein, wenn die Differenzen $a_1 - a_2$, $b_1 - b_2$ und $c_1 - c_2$ ihrer Koeffizienten betragsmäßig klein sind.

Das dachte man sich auch schon vorher. Ein entscheidender Punkt ist aber der folgende: $L(v_1 - v_2)$ misst die mittlere quadratische Differenz der Polynome auf dem Intervall $[-1, 1]$. Durch Berechnung von $L$ bekommen wir eine *quantitative* Information.

Wenn wir am Verhalten der Polynome in einem größeren Intervall interessiert sind, z.B. $[-10, 10]$, dann verändert sich unser Abstandsbegriff, und wir setzen z.B.

$$L(v) := \sqrt{\tfrac{1}{20}\int_{-10}^{10}(v(x))^2 dx} = \sqrt{\frac{10000}{5}a^2 + \frac{100}{3}(2ac + b^2) + c^2}.$$

Der Koeffizient $a$ des quadratischen Terms hat jetzt ein viel größeres Gewicht als $b$ oder gar der Koeffizient $c$ des konstanten Terms.

Durch die Betrachtung von Räumen, die nur gewisse Strukturen mit dem Anschauungsraum teilen, wird offensichtlich unser Blick für diese Strukturen und ihre Zusammenhänge geschärft. Wir werden beispielsweise gezwungen, zu definieren, was wir abstrakt unter der Länge eines Vektors verstehen.

Natürlich machen es sich die Mathematiker insofern einfach, indem sie zunächst Räume mit ganz wenigen grundlegenden Strukturen untersuchen (Verlängerung um Faktor $r$, Addition von Vektoren). Entsprechend handelt die Lineare Algebra von diesen linearen oder auch Vektorräumen und ihren strukturerhaltenden Abbildungen.

Allerdings hat es sich herausgestellt, dass man den Vektorräumen statt der reellen Zahlen ebensogut die rationalen oder die komplexen Zahlen zugrundelegen kann. Ebenso wenig ist es notwendig, sich auf Vektorräume der Dimension 2 und 3 zu beschränken. Im Gegenteil werden unendlichdimensionale Vektorräume für die theoretische Behandlung vieler Fragen aus Naturwissenschaft und Technik benutzt.

In dieser Verallgemeinerung des Raumbegriffs bei gleichzeitiger Abstraktion von höheren Strukturen wie Länge, Winkel, Orientierung, Volumen etc. liegt aber auch eine Schwierigkeit, denn es ist gar nicht so leicht, von diesen in unserem Anschauungsraum vorkommenden Strukturen abzusehen.

# 2 Die Sprache der Mathematik

Die mathematischen Texte sind in einer Sprache abgefasst, die in ihrem Formalisierungsgrad zwischen den natürlichen Sprachen und etwa den Programmiersprachen steht. Basis dieser Sprache ist der Mengenbegriff und, damit eng verknüpft, der der Aussagen. Während in dieser Vorlesung sonst alles definiert werden wird, werde ich den Mengenbegriff nicht formal definieren, denn irgendwo muss man ja anfangen.

Die Objekte $x$, aus denen eine Menge $M$ besteht, werden ihre *Elemente* genannt. Man schreibt $x \in M$ falls $x$ Element von $M$, andernfalls $x \notin M$. Sind $x_1, x_2, \ldots, x_n$ Elemente von $M$, so schreibt man $x_1, \ldots, x_n \in M$. Es gibt eine ausgezeichnete Menge ohne Elemente, die *leere Menge* $\emptyset$. Manchmal gibt man Mengen durch Aufzählung ihrer Elemente in geschweiften Klammern an.

**2.1 Beispiele** 1. Die Menge der Wochentage, an denen diese Vorlesung stattfindet, ist

$$\{\text{Mittwoch}, \text{Freitag}\}.$$

2. $\mathbb{N}$ bezeichnet die Menge der *natürlichen Zahlen*, also $\mathbb{N} = \{1, 2, 3, \ldots\}$.[1] Will man die Null mit hinzunehmen, dann schreibt man $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$.

3. $\mathbb{Z}$ bezeichnet die Menge der *ganzen Zahlen*: $\mathbb{Z} = \{0, 1, -1, 2, -2, \ldots\}$.

Bei der Klammerschreibweise kommt es auf Reihenfolge und eventuelle Wiederholung nicht an:

**2.2 Beispiel** $\{\text{Mittwoch}, \text{Freitag}\} = \{\text{Freitag}, \text{Mittwoch}, \text{Freitag}\}$.

Gilt für alle Elemente $x \in M_1$, dass $x$ auch Element einer zweiten Menge $M_2$ ist, dann schreibt man $M_1 \subset M_2$ oder $M_2 \supset M_1$ (stilisiertes Kleinerzeichen) und nennt $M_1$ *Teilmenge von* $M_2$. $M_1$ heißt *echte Teilmenge von* $M_2$, wenn $M_1 \subset M_2$ und ein $x \in M_2$ mit $x \notin M_1$ existiert. Man schreibt dann $M_1 \subsetneq M_2$.

**2.3 Beispiel** $\mathbb{N} \subset \mathbb{Z}$ , $\mathbb{N} \neq \mathbb{Z}$ , denn $0 \in \mathbb{Z}$ , aber $0 \notin \mathbb{N}$.

[1] Mit den Punkten meint man, dass man sich die weiteren Elemente der Menge dazudenkt. Das ist natürlich bequem, aber nicht eindeutig ($\{1, 2, 3, \ldots\}$ könnte auch die Menge derjenigen natürlichen Zahlen bezeichnen, die nur durch sich und 1 ohne Rest teilbar sind. Dann würde die Liste durch 5 statt durch 4 fortgesetzt!). Für eine Definition der natürlichen Zahlen mittels der sog. *Peano-Axiome* siehe z.B. [3].

Die Elemente von Mengen dürfen selbst Mengen sein, auch wenn man hier vorsichtig sein muss und Konstrukte wie "die Menge der Mengen, die sich selbst nicht enthalten" vermeiden sollte (siehe Bsp. 2.15).

**2.4 Beispiel (Mengensysteme)** $\{\emptyset, \{\text{Mittwoch}\}, \{\text{Freitag}\}, \{\text{Mittwoch}, \text{Freitag}\}\}$.

Ist $S$ ein solches (nicht leeres) *Mengensystem*, dann kann man den *Durchschnitt* dieses Mengensystems als die Menge aller $x$ einführen, die in allen Mengen von $S$ enthalten sind. Diese Menge wird mit

$$\bigcap_{M \in S} M$$

bezeichnet. Für $S = \{M_1, \ldots, M_n\}$ schreibt man auch $M_1 \cap \ldots \cap M_n$.

**2.5 Beispiele**

1. Für $S := \{\mathbb{N}, \mathbb{Z}\}$ kann man statt $\bigcap_{M \in S} M = \mathbb{N}$ auch $\mathbb{N} \cap \mathbb{Z}$ schreiben.
2. Für $S = \{\text{Geraden } G \text{ durch den Nullpunkt der Ebene } E\}$ ist

$$\bigcap_{G \in S} G = \{0\}.$$

Gilt für zwei Mengen $M_1$ und $M_2$, dass ihr Durchschnitt die leere Menge ist, $M_1 \cap M_2 = \emptyset$, dann heißen $M_1$ und $M_2$ *disjunkt*.

**2.6 Beispiel (Disjunktheit)** $\mathbb{N} \cap \{\text{Mittwoch}, \text{Freitag}\} = \emptyset$.

Die *Vereinigung* $\bigcup_{M \in S} M$ der Mengen eines Mengensystems ist die Menge aller derjenigen Elemente, die zumindest zu einer Menge aus $S$ gehören.

**2.7 Beispiel** Wie im Beispiel 2.5.2. bezeichne $S$ die Menge der Geraden in der Ebene $E$. Es gilt dann

$$\bigcup_{G \in S} G = E.$$

Oft werden Mengen $M$ über Aussagen $A$ eingeführt, nach dem Schema

$$M := \{x \mid A(x)\},$$

also $M$ besteht aus den Elementen $x$, für die die Aussage $A$ wahr ist.

**2.8 Beispiel** $\{2\} = \{x \mid x \text{ ist prim und } x \text{ ist gerade}\}$.

Wie aus diesem Beispiel klar wird, können wir Vereinigung und Schnitt von Mengen über Aussagen formulieren:

$$\begin{aligned} M_1 \cap M_2 &= \{x \mid x \in M_1 \quad \text{und} \quad x \in M_2\}, \\ M_1 \cup M_2 &= \{x \mid x \in M_1 \quad \text{oder} \quad x \in M_2\}. \end{aligned}$$

Mengentheoretische Identitäten wie die *De Morgan-Regeln*

$$\begin{aligned} M \cap (N_1 \cup N_2) &= (M \cap N_1) \cup (M \cap N_2) \\ M \cup (N_1 \cap N_2) &= (M \cup N_1) \cap (M \cup N_2) \end{aligned}$$

lassen sich damit über Wahrheitstabellen verifizieren.

Soweit Mengen als Teilmengen definiert werden, ist es im Allgemeinen besser, die Obermenge auch direkt anzugeben.

**2.9 Beispiel (Definition von Mengen)** Die Menge $\mathbb{P}$ der *Primzahlen*:

$$\mathbb{P} := \big\{n \in \mathbb{N} \mid n \neq 1 \quad \text{und} \quad (k \in \mathbb{N} \text{ teilt } n) \Longrightarrow k \in \{1, n\}\big\}.$$

**Weitere Mengenoperationen**. Sind $M, N$ Mengen, dann

- $M \backslash N := \{x \in M \mid x \notin N\}$ (*Differenzmenge*)
- Speziell im Fall $N \subset M$ schreibt man dafür auch $M - N$.
- $M \Delta N := (M \backslash N) \cup (N \backslash M)$ (*Symmetrische Differenz*)
  Jedes Element $x \in M \Delta N$ ist entweder nur in $M$ oder nur in $N$.
- Die *Potenzmenge* $\mathcal{P}(M)$ (auch $2^M$ geschrieben) einer Menge $M$ ist durch

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

  definiert, also die Menge aller Teilmengen von $M$.

  **2.10 Beispiel** $\mathcal{P}(\{1, 2\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$.

- Das *kartesische Produkt* $M_1 \times M_2$ zweier Mengen ist die Menge aller geordneten Paare $(m_1, m_2)$ mit $m_1 \in M_1$, $m_2 \in M_2$. [2]

  Analog zu diesem zweifachen kartesischen Produkt kann man auch das $n$–fache kartesische Produkt $M_1 \times \ldots \times M_n$ definieren. Die Elemente $(x_1, \ldots, x_n)$ dieser Menge heißen dann *geordnete $n$–Tupel*, und man schreibt $M^n$ für das $n$-fache kartesische Produkt von $M$.

  **2.11 Beispiel** $\{1,2\} \times \{2,3\} = \{(1,2),(2,2),(1,3),(2,3)\}$

- Teilmengen $R \subset M \times N$ heißen *Relationen zwischen $M$ und $N$*.

  Eine Relation $R \subset M \times M$ heißt *Halbordnung*, wenn sie
  • *reflexiv* ist, d.h. für alle $x \in R$ auch $(x,x) \in R$ ist
  • *transitiv* ist, d.h. aus $(x,y) \in R$ und $(y,z) \in R$ auch $(x,z) \in R$ folgt, sie
  • *antisymmetrisch* ist, d.h. wenn mit $(x,y) \in R$ und $(y,x) \in R$ folgt, dass $x = y$ ist.
  Man schreibt dann $x \preceq y$ statt $(x,y) \in R$.

  Die Relation $R$ heißt *total*, wenn sich alle Elemente $x, y \in M$ vergleichen lassen, d.h. $(x,y) \in R$ oder $(y,x) \in R$ gilt. Eine totale Halbordnung auf $M$ heißt *Ordnung* (und $M$ *geordnet*).

  **2.12 Beispiele**
  1. Die Relation $R := \{(x,y) \in \mathbb{Z} \times \mathbb{Z} \mid y - x \in \mathbb{N}_0\}$ ordnet die ganzen Zahlen.
  2. Es sei $M := \mathbb{Z} \times \mathbb{Z}$, und für $x = (x_1, x_2)$, $y = (y_1, y_2) \in M$ gelte $x \preceq y$ falls $x_1 \preceq y_1$ und $x_2 \preceq y_2$. Dies ist eine Halbordnung, die aber nicht total ist. Genau alle Punkte im 1. Quadranten von $\mathbb{Z} \times \mathbb{Z}$ sind $\succeq (0,0)$, alle im 3. Quadranten sind $\preceq (0,0)$.

Eine *Abbildung* $f : M \to N$ von der Menge $M$ in die Menge $N$ "ordnet jedem Element $m \in M$ genau ein Element $f(m) \in N$ zu". Das ist natürlich keine formale Definition. Will man Abbildungen formal definieren, betrachtet man sie selbst als Mengen, und zwar heißt
eine Relation $f \subset M \times N$ *Abbildung von $M$ nach $N$*, wenn

1. $D(f) := \{m \in M \mid$ es gibt $n \in N$ mit $(m,n) \in f\} = M$ ist,

[2] Dabei kann man, wenn man will, $(m_1, m_2) := \{m_1, \{m_1, m_2\}\}$ definieren. Damit ist für $m_1 \neq m_2$ zwar $\{m_1, m_2\} = \{m_2, m_1\}$, aber $(m_2, m_1) = \{m_2, \{m_2, m_1\}\} \neq (m_1, m_2)$.

2. aus $(m, n_1) \in f$ und $(m, n_2) \in f$ immer $n_1 = n_2$ folgt.

Denn dann kann man jedem $m \in M$ genau ein Element $n \in N$ zuordnen, nämlich das mit $(m, n) \in f$.

$M$ heißt *Definitionsbereich*, $N$ *Bildbereich* von $f$.

$$W(f) := \{n \in N \mid \quad \text{es gibt} \quad m \in M \quad \text{mit} \quad (m, n) \in f\}$$

heißt *Wertebereich* von $f$. $f$ heißt

- *surjektiv*, wenn $W(f) = N$,
- *injektiv*, wenn aus $(m_1, n) \in f$ und $(m_2, n) \in f$ $m_1 = m_2$ folgt und
- *bijektiv*, wenn $f$ surjektiv und injektiv ist.

Man kann die als spezielle Relationen eingeführten Abbildungen dann wieder in gewohnter Manier als $f : M \to N, m \mapsto f(m)$ schreiben.

**2.13 Beispiele (Abbildungen)** 1. $f : \mathbb{R} \to \mathbb{R}$, $f(x) := x^2$ ist weder injektiv noch surjektiv.

2. $f : \mathbb{R} \to \mathbb{R}$, $f(x) := x^3$ ist bijektiv.

3. $f : \mathbb{R} \to \mathbb{R}$, $f(x) := x^3 - x$ ist nur surjektiv.

4. $f : \mathbb{R} \to \mathbb{R}$, $f(x) := e^x$ ist nur injektiv mit Wertebereich

$$W(f) = \mathbb{R}^+ := \{x \in \mathbb{R} \mid x > 0\}.$$

Will man betonen, dass man $f$ als Relation auffasst, so nennt man diese den *Graphen* von $f$.

- Wenn wir die Abbildung $f : M \to N$ auf eine Teilmenge $U \subset M$ anwenden, erhalten wir die Teilmenge

  $$f(U) := \{f(x) \mid x \in U\} \subset W(f)$$

  des Wertebereichs $W(f) = f(M)$. Es ist $f(M_1 \cup M_2) = f(M_1) \cup f(M_2)$ aber i.A. nur $f(M_1 \cap M_2) \subset f(M_1) \cap f(M_2)$.

- Zwei Abbildungen $f : M \to N$ und $g : N \to R$ kann man zur *Produktabbildung*
$$g \circ f : M \to R \quad , \quad m \mapsto g(f(m))$$
zusammenfügen (entgegen der Leserichtung wird erst $f$, dann $g$ angewandt!).

- Injektive Abbildungen $f : M \to N$ kann man auf ihrem Bild umkehren und erhält dann die *inverse Abbildung* $f^{-1} : f(M) \to M$.

- Bei nicht notwendig injektiven Abbildungen $f : M \to N$ bezeichnen wir für $V \subset N$ mit $f^{-1}(V)$ die *Menge*
$$f^{-1}(V) := \{m \in M \mid f(m) \in V\},$$
und für $n \in N$ setzen wir $f^{-1}(n) := \{m \in M \mid f(m) = n\}$ (identifizieren also hier einelementige Mengen mit ihren Elementen).

- Die *identische Abbildung* $\mathrm{Id} : M \to M$ ist durch $\mathrm{Id}(m) = m$ für $m \in M$ definiert (präziser: $\mathrm{Id}_M$). Bijektive Abbildungen $f : M \to N$ haben dann die Eigenschaften
$$f^{-1} \circ f = \mathrm{Id}_M \quad \text{und} \quad f \circ f^{-1} = \mathrm{Id}_N.$$

- Die *Einschränkung* oder *Restriktion* einer Abbildung $f : M \to N$ auf die Teilmenge $U \subset M$ ist die Abbildung
$$f|_U : U \to N \quad , \quad f|_U(x) := f(x).$$

Für $n \in \mathbb{N}$ hat eine Menge $M$ die *Mächtigkeit* $|M| = n$ (d.h. hat $n$ Elemente), wenn eine Bijektion $f : \{1, \ldots, n\} \to M$ existiert ($f$ nummeriert dann die Elemente). Allgemeiner haben die Mengen $M$ und $N$ die *gleiche Mächtigkeit* (geschrieben $|M| = |N|$), falls eine Bijektion $f : N \to M$ existiert.

**2.14 Beispiel** $|\mathbb{N}^2| = |\mathbb{N}|$, denn $f : \mathbb{N}^2 \to \mathbb{N}$, $(a, b) \mapsto \frac{1}{2}(a+b-1)(a+b-2)+b$ ist eine Bijektion (warum?).

Ein abschließendes Beispiel soll erläutern, dass der naive Mengenbegriff auf Grenzen stößt.

**2.15 Beispiel (Russellsche Antinomie)** Wir betrachten die "Menge" $M := \{x \mid x \text{ ist Menge}\}$ aller Mengen.

Weiter sei $N := \{x \in M \mid x \notin x\}$ die Menge aller Mengen, die sich nicht selbst enthalten.

*Frage:* Gilt $N \in N$?

1. *Antwort ja.* Dann ist $N \notin N$ nach Definition von $N$.

2. *Antwort nein.* Dann ist $N \in N$ nach Definition von $N$.

Wir werden offensichtlich auf einen Widerspruch, eben die Russellsche Antinomie, gestoßen, der zeigt, dass wir solche Selbstbezüge von Mengen tunlichst vermeiden sollten.[3]

Für unsere Zwecke ist eine *Aussage* $A$ ein Element der Menge $\{\text{wahr}, \text{falsch}\}$ und eine *Aussageform* eine Abbildung einer (oft nicht explizit angegebenen) Menge in diese zweielementige Menge.

Sind $A$, $B$ zwei Aussagen, so lassen sich mittels der *Junktoren* (d.h. "Verbinder") neue Aussagen herstellen:

| Junktor | Bedeutung | Zeichen |
|---|---|---|
| Negation | nicht $A$ | $\neg A$ |
| Konjugation | $A$ und $B$ | $A \wedge B$ |
| Adjunktion | $A$ oder $B$ | $A \vee B$ |
| Implikation | wenn $A$, dann $B$ | $A \Rightarrow B$ |
| Äquivalenz | $A$ genau dann, wenn $B$ | $A \Leftrightarrow B$ |

Die Wahrheitswerte der so zusammengesetzten Aussagen sind durch die folgende Tabelle festgelegt. (Eine Tabelle dieser Art heißt *Wahrheitstafel*).

| $A$ | $B$ | $\neg A$ | $A \wedge B$ | $A \vee B$ | $A \Rightarrow B$ | $A \Leftrightarrow B$ |
|---|---|---|---|---|---|---|
| $W$ | $W$ | $F$ | $W$ | $W$ | $W$ | $W$ |
| $W$ | $F$ | $F$ | $F$ | $W$ | $F$ | $F$ |
| $F$ | $W$ | $W$ | $F$ | $W$ | $W$ | $F$ |
| $F$ | $F$ | $W$ | $F$ | $F$ | $W$ | $W$ |

**2.16 Beispiele** 1. "$6$ ist eine Primzahl" ist eine Aussage und zwar eine falsche.

[3] Eine Umformulierung ist die folgende Frage: Es gibt zahlreiche Buchkataloge. Zwar sind Kataloge auch Bücher, aber die meisten von ihnen führen sich selbst nicht auf. Wir wollen einen Katalog aller Kataloge, die sich selbst nicht enthalten, erstellen. Müssen wir ihn selbst aufführen oder nicht?

2. $A : \mathbb{N} \to \{\text{wahr}, \text{falsch}\}$ , $A(x) := \begin{cases} \text{wahr} & , \ x \text{ ist Primzahl} \\ \text{falsch} & , \ x \text{ ist nicht Primzahl.} \end{cases}$

   $A$ ist eine Aussageform.

Aus Aussageformen kann man durch die so genannten *Quantoren* "für alle", abgekürzt mit $\wedge$ oder $\forall$, und "es existiert", abgekürzt mit $\vee$ oder $\exists$, Aussagen machen:

$$\begin{aligned} \big(\forall x \in M : A(x)\big) &:= \begin{cases} \text{wahr} & , \ \text{falsch} \notin A(M) \\ \text{falsch} & , \ \text{falsch} \in A(M) \end{cases} \\ \big(\exists x \in M : A(x)\big) &:= \begin{cases} \text{wahr} & , \ \text{wahr} \in A(M) \\ \text{falsch} & , \ \text{wahr} \notin A(M) \end{cases} . \end{aligned}$$

Es gilt

$$\big(\neg(\forall x \in M : A(x))\big) \ = \ \big(\exists x \in M : \neg A(x)\big)$$

oder kurz $\neg\forall = \exists\neg$ und umgekehrt $\neg\exists = \forall\neg$.

# 3 Gruppen

Nach dieser Einleitungsphase mit ihrem undefinierten Mengenbegriff werden nach Goethes "Faust" dem Geist spanische Stiefel angeschnürt und alles wird ordentlich definiert. Das ganz allgemeine Schema wird immer wieder darin bestehen,

- abstrakt auf Mengen Strukturen einzuführen (z.B. Addition zweier Elemente, Längenmessung etc.),
- dann zu schauen, ob Mengen mit derartigen Strukturen überhaupt existieren (z.B. ob eine sechs-elementige Menge ein so genannter Körper sein kann),
- um als nächstes Eigenschaften dieser Objekte zu untersuchen.
- Typischerweise werden zu guter Letzt strukturerhaltende Abbildungen zwischen solchen Mengen betrachtet.

Die erste so untersuchte Struktur wird die der *Gruppe* sein.

**3.1 Definition** *Eine* **Gruppe** *besteht aus einer Menge $G$ und einer Abbildung $G \times G \to G$, geschrieben $(a, b) \mapsto a \circ b$, mit den Eigenschaften*

*1.* $\forall\, a, b, c \in G : a \circ (b \circ c) = (a \circ b) \circ c$ ***(Assoziativität)***

2. $\exists\, e \in G \quad \forall\, a \in G : e \circ a = a$ *(Existenz eines* **neutralen Elements***)*

3. $\forall\, a \in G \quad \exists\, a' \in G : a' \circ a = e$ *(Existenz der* **inversen Elemente***)*

4. *Die Gruppe heißt* **abelsch** *oder* **kommutativ**, *wenn*

$$\forall\, a, b \in G : a \circ b = b \circ a \qquad \textit{(}\textbf{Kommutativgesetz}\textit{)}.$$

Wir können also eine Gruppe durch $(G, \circ)$, also das Paar (Menge, Verknüpfung) kennzeichnen, werden aber oft der Einfachheit halber nur $G$ schreiben.

**3.2 Beispiele (Gruppen)** 1. $(\mathbb{Z}, +)$, also die ganzen Zahlen mit additiver Verknüpfung. Die Null ist das neutrale Element, das inverse Element zu $a \in \mathbb{Z}$ schreibt man $-a$.

2. $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ und $(\mathbb{C}, +)$ sind genauso abelsche Gruppen wie $(\mathbb{Z}, +)$. Systematisch muss man nur zeigen, dass $(\mathbb{C}, +)$ eine Gruppe ist, denn $\mathbb{C} \supset \mathbb{R} \supset \mathbb{Q} \supset \mathbb{Z} \ni 0$, und die Teilmengen sind abgeschlossen unter der Addition und unter der Abbildung $x \mapsto -x$.

3. Ist $M$ eine beliebige nicht leere Menge, dann bildet die Menge der Bijektionen $\varphi : M \to M$ eine Gruppe $(G, \circ)$, wobei die Verknüpfung $\varphi \circ \psi$ von zwei bijektiven Abbildungen $\varphi, \psi \in G$ als die Produktabbildung definiert ist. Neutrales Element ist die identische Abbildung $\mathrm{Id}_M$, das zu $\varphi : M \to M$ inverse Element die inverse Abbildung $\varphi^{-1} : M \to M$. $(G, \circ)$ heißt *symmetrische Gruppe von* $M$. Die zur Menge

$$M := \{1, \ldots, n\} \equiv \{m \in \mathbb{N} \mid m \leq n\}$$

gehörige symmetrische Gruppe wird oft mit $\mathcal{S}_n$ bezeichnet. Sie ist also die Gruppe der Permutationen von $n$ Elementen.

Da jede Permutation $\varphi \in \mathcal{S}_n$ durch Angabe des geordneten $n$–Tupels der Bilder $(\varphi(1), \ldots, \varphi(n))$ fixiert ist, ist die Anzahl der Elemente von $\mathcal{S}_n$ gleich $n! := n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1$. Statt die Permutation $\varphi \in \mathcal{S}_n$ so zu schreiben, ist es allerdings üblich, noch die Urbilder mit anzugeben, d.h.

$$\left(\begin{smallmatrix} 1 & 2 & \ldots & n \\ \varphi(1) & \varphi(2) & \ldots & \varphi(n) \end{smallmatrix}\right).$$

Als Mengen sind $\mathcal{S}_1 = \{\mathrm{Id}\}$, $\mathcal{S}_2 = \{\mathrm{Id}, \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)\}$ und

$$\mathcal{S}_3 = \left\{\left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}\right)\right\}.$$

$\mathcal{S}_1$ und $\mathcal{S}_2$ sind abelsche Gruppen. Wegen

$$\alpha \circ \beta = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right) \neq \beta \circ \alpha = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}\right)$$

für $\alpha := \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right)$ und $\beta = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix}\right)$ ist dagegen $\mathcal{S}_3$ nicht abelsch.

4. $(\mathbb{R}^+, \cdot)$ ist eine abelsche Gruppe mit neutralem Element $1$.

**3.3 Definition** *Ist $G$ eine Gruppe und $M$ eine Menge, dann heißt eine Abbildung $\Phi : G \times M \to M$* **Gruppenwirkung** *(oder* **Operation***)* **von** $G$ **auf** $M$*, wenn stets*

$$\Phi(e, m) = m \quad \textit{und} \quad \Phi(g_2, \Phi(g_1, m)) = \Phi(g_2 \circ g_1, m)$$

*gilt.*

**3.4 Beispiel (Gruppenwirkung)** $M$ bezeichne die Menge der Ecken eines Quadrats $Q := [-1, 1] \times [-1, 1]$ und $G$ die Gruppe $\mathbb{Z}/4\mathbb{Z}$. Diese operiert durch Drehungen um Vielfache von $\pi/2$ auf $Q$ und permutiert dabei $M$.

Wegen der Assoziativität können wir bei beliebigen Verknüpfungen Klammern weglassen. Statt dem Zeichen "$\circ$" wird oft auch der Multiplikationspunkt benutzt, "$+$" allerdings nur bei abelschen Gruppen.

Es sollen jetzt weitere Eigenschaften von Gruppen aus den Axiomen gefolgert werden. Zunächst fällt ja auf, dass in den Axiomen relativ wenig gefordert wird, beispielsweise nicht die Eindeutigkeit des neutralen Elements oder die Beziehung $a \circ e = a$. Hier zeigt sich ein Sparsamkeitsprinzip der Mathematik. Was man ohnehin beweisen kann, soll man nicht auch noch fordern. Denn hier gilt:

**3.5 Satz** *1. Für jedes neutrale Element $e \in G$ gilt auch $a \circ e = a$ für alle $a \in G$. Aus $a' \circ a = e$ folgt auch $a \circ a' = e$.*

*2. Es gibt genau ein neutrales Element $e \in G$. Bereits aus $x \circ a = a$ für ein $a \in G$ folgt $x = e$.*

**Beweis:**

1. Zunächst der zweite Teil. Zu $a'$ gibt es nach Axiom 3 ein $a'' \in G$ mit $a'' \circ a' = e$. Damit ist

$$\begin{aligned} a \circ a' &\overset{2.}{=} e \circ (a \circ a') \overset{3.}{=} (a'' \circ a') \circ (a \circ a') \\ &\overset{1.}{=} a'' \circ (a' \circ a) \circ a' \overset{3.}{=} a'' \circ (e \circ a') \overset{2.}{=} a'' \circ a' \\ &\overset{3.}{=} e. \end{aligned}$$

Hierbei habe ich die Nummern der für die Umformung benutzten Axiome über das Gleichheitszeichen geschrieben.

Nun ist $a \circ e = a \circ (a' \circ a) = (a \circ a') \circ a = e \circ a = a$.

2. Ist $e' \in G$ ein weiteres neutrales Element, dann gilt insbesondere $e = e' \circ e = e'$, wobei das zweite Gleichheitszeichen aus dem ersten Teil des Satzes folgt.

   Mit $a' \circ a = a \circ a' = e$ folgt

$$x = x \circ e = x \circ (a \circ a') = (x \circ a) \circ a' = a \circ a' = e. \qquad \square$$

Nicht nur das neutrale Element $e \in G$, sondern auch das zu $a \in G$ inverse Element $a' \in G$ ist eindeutig bestimmt:

**3.6 Satz** *Zu $a \in G$ existiert genau ein inverses Element $a' \in G$.*

**Beweis:** Für $a'' \in G$ mit $a'' \circ a = a \circ a'' = e$ gilt

$$a'' = a'' \circ e = a'' \circ (a \circ a') = (a'' \circ a) \circ a' = e \circ a' = a'. \qquad \square$$

**3.7 Bemerkungen** 1. Mit dem Zeichen $\exists!$ (es existiert genau ein) könnte man schreiben: $\forall a \in G \;\; \exists! \, a' \in G : a' \circ a = e$.

2. Wegen Satz 3.6 ist es möglich, für das zu $a$ inverse Element $a^{-1}$ oder additiv $-a$ zu schreiben.

3. Ist die Gruppe nicht abelsch, empfiehlt es sich nicht, statt $a^{-1}$ das Symbol $\frac{1}{a}$ zu benutzen, weil sonst die Gefahr besteht, $\frac{b}{a}$ statt $a^{-1} \circ b$ zu schreiben, was auch $b \circ a^{-1}$ bedeuten könnte.
   Den Verknüpfungskringel werde ich aber oft weglassen.

**3.8 Satz** *Für $a, b \in G$ gilt $(a^{-1})^{-1} = a$ und $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$.*

**Beweis:**

$$\begin{aligned} (a^{-1})^{-1} &= (a^{-1})^{-1} e = (a^{-1})^{-1} a^{-1} a = ea = a. \\ (ab)^{-1} &= (ab)^{-1} e = (ab)^{-1} aa^{-1} = (ab)^{-1} a(bb^{-1}) a^{-1} \\ &= (ab)^{-1} (ab) b^{-1} a^{-1} = b^{-1} a^{-1} \end{aligned}$$

$\square$

Aus der Addition leitet sich die Subtraktion ab, aus der Multiplikation die Division. Allgemein gilt:

**3.9 Satz** $\forall a, b \in G \qquad \exists!\, x \in G : xa = b \quad$ *und* $\quad \exists!\, y \in G : ay = b.$

**Beweis:** Setze $x := ba^{-1}$. Dann ist $xa = ba^{-1}a = b$. Analog gilt für $y := a^{-1}b$ : $ay = aa^{-1}b = b$. Für jede Lösung $x'$ von $x'a = b$ gilt

$$x' = x'(aa^{-1}) = (x'a)a^{-1} = ba^{-1}$$

und analog für $y$. □

**3.10 Bemerkung** Ist die Gruppe nicht abelsch, dann brauchen die Lösungen $x$ und $y$ nicht übereinzustimmen. Z.B. gilt für $a, b \in \mathcal{S}_3$, $a := \left(\begin{smallmatrix}1&2&3\\3&2&1\end{smallmatrix}\right)$, $b := \left(\begin{smallmatrix}1&2&3\\2&3&1\end{smallmatrix}\right)$, dass $a^{-1} = a$ ist, und damit

$$x = ba^{-1} = \left(\begin{smallmatrix}1&2&3\\2&3&1\end{smallmatrix}\right) \circ \left(\begin{smallmatrix}1&2&3\\3&2&1\end{smallmatrix}\right) = \left(\begin{smallmatrix}1&2&3\\1&3&2\end{smallmatrix}\right) \neq y = a^{-1}b = \left(\begin{smallmatrix}1&2&3\\3&2&1\end{smallmatrix}\right) \circ \left(\begin{smallmatrix}1&2&3\\2&3&1\end{smallmatrix}\right) = \left(\begin{smallmatrix}1&2&3\\2&1&3\end{smallmatrix}\right).$$

Eine Teilmenge $U \subset G$ einer Gruppe $(G, \circ)$, die bez. der Verknüpfung $\circ$ eine Gruppe bildet, heißt *Untergruppe* von $G$. Praktisch muss man nur $e \in U$ und Abgeschlossenheit unter Multiplikation und Inversenbildung verifizieren. ($e \in U$ ist zu verifizieren, denn $\emptyset$ ist keine Gruppe!)

**3.11 Beispiel (Untergruppe)** Die Menge $2\mathbb{Z}$ der *geraden Zahlen* bildet eine Untergruppe von $(\mathbb{Z}, +)$. Allgemein wird für $n \in \mathbb{N}$ die Menge der durch $n$ ohne Rest teilbaren Zahlen mit

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

bezeichnet. $(n\mathbb{Z}, +)$ ist eine Gruppe, denn mit $nx$ und $ny$ sind auch $nx + ny = n(x + y)$ und $-nx = n(-x)$ in $n\mathbb{Z}$ und $0 = n0 \in n\mathbb{Z}$.

Allgemein heißen in der Mathematik die strukturverträglichen Abbildungen Homomorphismen. Im Zusammenhang mit der Gruppenstruktur ist also zu fordern:

**3.12 Definition** *Eine Abbildung $f : G_1 \to G_2$ von der Gruppe $G_1$ in die Gruppe $G_2$ heißt* **Homomorphismus**, *wenn*

$$\forall\, a, b \in G_1 \qquad f(ab) = f(a)f(b)$$

*gilt. Ein bijektiver Homomorphismus heißt* **Isomorphismus**.

**3.13 Bemerkungen** 1. Beachte, dass auf der linken Seite die Multiplikation in $G_1$, auf der rechten die Multiplikation in $G_2$ gemeint ist.

2. Es folgt $f(e_1) = f(e_1 \cdot e_1) = f(e_1) \cdot f(e_1)$ für das neutrale Element $e_1$ von $G_1$, also ist $f(e_1)$ das neutrale Element von $G_2$.

**3.14 Beispiele (Gruppenhomomorphismen)** 1. Für ein $n \in \mathbb{N}$ sei $f : \mathbb{Z} \to \mathbb{Z},\ f(x) := nx$. Dann gilt

$$f(a+b) = n(a+b) = na + nb = f(a) + f(b),$$

$f$ ist also ein Homomorphismus. Der Wertebereich $f(\mathbb{Z}) \subset \mathbb{Z}$ ist dann $f(\mathbb{Z}) = n\mathbb{Z}$.

2. Wir schränken nun den Bildbereich auf den Wertebereich ein, setzen also $f : \mathbb{Z} \to n\mathbb{Z},\ f(x) := nx$. Dieser Homomorphismus ist bijektiv, also ein Isomorphismus. Z.B. sind die additiven Gruppen $\mathbb{Z}$ der ganzen Zahlen und $2\mathbb{Z}$ der geraden Zahlen isomorph, als abstrakte Gruppen also nicht unterscheidbar.

3. Die Menge $G := \{-1, 1\}$ bildet eine zweielementige multiplikative Gruppe:

| $\cdot$ | 1 | -1 |
|---|---|---|
| 1 | 1 | -1 |
| -1 | -1 | 1 |

.

Eine Permutation aus $\mathcal{S}_n (n \geq 2)$ heißt *Transposition*, wenn sie zwei der Zahlen $1, \ldots, n$ vertauscht, die anderen aber elementweise festhält. Es lässt sich nun zeigen: Jede Permutation $\varphi \in \mathcal{S}_n$ lässt sich in der Form

$$\varphi = \alpha_1 \circ \ldots \circ \alpha_m$$

schreiben, wobei die $\alpha_i \in \mathcal{S}_n$ Transpositionen sind. Zwar ist ihre Zahl $m$ nicht durch $\varphi$ bestimmt, wohl aber $(-1)^m$, d.h. falls $m$ in einer Darstellung (un)gerade ist, dann in allen. Setze nun

$$\operatorname{sign} : \mathcal{S}_n \to \{-1, 1\} \quad , \quad \operatorname{sign}(\varphi) := \frac{\prod_{i<j}(\varphi(j) - \varphi(i))}{\prod_{i<j}(j-i)}. \tag{3.1}$$

Man zeigt, dass $\operatorname{sign}$ ein Gruppenhomomorphismus ist, indem man die Zuordnung $g \mapsto 1$, $u \mapsto -1$ und die Additionstabelle

| $+$ | g | u |
|---|---|---|
| g | g | u |
| u | u | g |

beachtet.

**3.15 Satz** *Der sogenannte* **Kern**

$$\ker(f) := f^{-1}(e_2) \subset G_1$$

*eines Gruppenhomomorphismus $f : G_1 \to G_2$ ist eine Untergruppe von $G_1$.*

**Beweis:** • Das neutrale Element $e_1$ von $G_1$ liegt in $\ker(f)$, denn $f(e_1) = e_2$.
• Weiter ist mit $a, b \in \ker(f)$ auch $ab \in \ker(f)$, denn

$$f(ab) = f(a)f(b) = e_2 \cdot e_2 = e_2, \quad \text{und}$$

• mit $a \in \ker(f)$ ist $a^{-1} \in \ker(f)$, denn $f(a^{-1}) = (f(a))^{-1} = e_2^{-1} = e_2$. □

**3.16 Beispiel** In den Beispielen für Gruppenhomomorphismus war in Beispiel 3.14.1. und 2. $\ker(f) = \{0\}$, die einelementige Gruppe.
3. Den Kern $\mathcal{A}_n := \ker(\text{sign})$ der in (3.1) definierten Signums-Funktion nennt man die *alternierende Gruppe* von $n$ Objekten. Ihre Elemente sind die *geraden Permutationen*. So ist

$$\mathcal{A}_2 = \{\text{Id}\} \quad , \quad \mathcal{A}_3 = \{\left(\begin{smallmatrix}1&2&3\\1&2&3\end{smallmatrix}\right), \left(\begin{smallmatrix}1&2&3\\2&3&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&2&3\\3&1&2\end{smallmatrix}\right)\}.$$

**3.17 Definition** *Eine Gruppe $G$ heißt zyklisch, wenn es ein $a \in G$ gibt, so dass sich jedes $x \in G$ in der Form $x = a^k$ $(k \in \mathbb{Z})$ darstellen läßt, mit $a^0 := e$ (neutrales Element) und $a^k := a \circ \ldots \circ a$, $a^{-k} := (a^{-1})^k$ für $k \in \mathbb{N}$.*
*$a$ heißt dann Erzeuger von $G$, formal: $G = \langle a \rangle$.*

**3.18 Beispiele (zyklische Gruppe)** 1. Die additive Gruppe $(\mathbb{Z}, +)$ ist zyklisch mit Erzeuger $a = 1$, denn jedes $x \in \mathbb{Z}$ lässt sich als positives oder negatives Vielfaches von 1 darstellen.

2. Die zyklische Gruppe $(n\mathbb{Z}, +)$, $n \in \mathbb{N}$, wird von $n$ erzeugt.

**3.19 Bemerkung** Jede zyklische Gruppe ist kommutativ. (Dies folgt direkt aus der Definition.)

Eine Gruppe heißt *endlich*, wenn sie als Menge endlich ist. Die Zahl $|G|$ ihrer Elemente heißt dann *Ordnung* der Gruppe.

**Frage:** Gibt es Gruppen beliebiger Ordnung $n \in \mathbb{N}$? Die Antwort heißt "ja". Wir werden solche Gruppen durch Nebenklassenbildung konstruieren.

**3.20 Definition** *Eine Relation $R \subset M \times M$ heißt* **Äquivalenzrelation**, *wenn für alle $x, y, z \in M$*

1. $(x, x) \in R$ *($R$ ist* **reflexiv***)*
2. $(x, y) \in R \Longrightarrow (y, x) \in R$ *($R$ ist* **symmetrisch***)*

*3.* $(x,y) \in R$ *und* $(y,z) \in R \implies (x,z) \in R$ *(*$R$ *ist* **transitiv***).*

*Statt* $(x,y) \in R$ *schreibt man dann* $x \sim y$*.*

**3.21 Beispiele (Äquivalenzrelationen)** 1. Die Gleichheits-Relation auf der Menge $M$ ist $R := \{(x,y) \in M \times M \mid x = y\}$.

2. $M = \mathbb{Z},\ n \in \mathbb{N}.\ \ x \sim y :\Longleftrightarrow x - y \in n\mathbb{Z}$.

   (Das Zeichen $:\Longleftrightarrow$ steht dabei für die Definition der linken Seite durch die rechte.)

**3.22 Definition** *Die von* $m \in M$ **erzeugte Äquivalenzklasse** $[m]$ *ist die Teilmenge* $[m] := \{n \in M \mid n \sim m\}$.
*Die Elemente einer Äquivalenzklasse heißen ihre* **Repräsentanten***.*

Zwei Äquivalenzklassen sind entweder gleich oder disjunkt; jedes Element von $M$ ist in genau einer Äquivalenzklasse enthalten.

**3.23 Beispiele (Fortsetzung von 3.21)** 1. Die Äquivalenzklassen der Gleichheits-Relation sind einelementig.

2. $\mathbb{Z} = \bigcup_{i=0}^{n-1} [i]$ mit $[i] = \{x \in \mathbb{Z} \mid x - i \in n\mathbb{Z}\}$.

Das letzte Beispiel lässt sich als Spezialfall einer gruppentheoretischen Konstruktion auffassen:

**3.24 Definition** *Es sei* $H \subset G$ *eine Untergruppe von* $G$ *und* $a \in G$*. Dann heißt*

$$\begin{array}{rcll} aH & := & \{ax \mid x \in H\} & \text{eine } \textbf{Linksnebenklasse}, \\ Ha & := & \{xa \mid x \in H\} & \text{eine } \textbf{Rechtsnebenklasse} \end{array}$$

*von* $H$ *in* $G$*.* $H$ *heißt* **Normalteiler** *von* $G$*, wenn für alle* $a \in G$ *gilt:* $aH = Ha$*.*

**3.25 Bemerkungen** 1. $a \in aH$, $a \in Ha$

2. $aH = bH \Longleftrightarrow b^{-1}a \in H$

   Denn insbesondere $ae = bh$ mit $h \in H \implies b^{-1}a \in H$.

3. Entweder $aH = bH$ oder $aH \cap bH = \emptyset$.

   Denn mit $ah_1 = bh_2$ für gewisse $h_1, h_2 \in H$ ist $b^{-1}a = h_2 h_1^{-1} \in H$.

4. Zugehörigkeit zur gleichen Links- (bzw. Rechts-)nebenklasse ist eine Äquivalenzrelation.

Kerne von Gruppenhomomorphismen $\varphi : G_1 \to G_2$ sind immer Normalteiler, denn

$$\varphi(ah) = \varphi(a)\varphi(h) = \varphi(a) = \varphi(h)\varphi(a) = \varphi(ha) \qquad (h \in \ker(\varphi)).$$

**3.26 Beispiele (Normalteiler)** 1. $\mathcal{A}_n$ ist Normalteiler von $\mathcal{S}_n$, denn $\mathcal{A}_n = \operatorname{sign}^{-1}(1)$ und $\operatorname{sign}$ ist ein Gruppenhomomorphismus.

2. Die Untergruppe $H := \{(\begin{smallmatrix}1&2&3\\1&2&3\end{smallmatrix}), (\begin{smallmatrix}1&2&3\\2&1&3\end{smallmatrix})\} \subset \mathcal{S}_3$ ist *kein* Normalteiler, denn für $a := (\begin{smallmatrix}1&2&3\\1&3&2\end{smallmatrix})$ ist

$$aH = \{(\begin{smallmatrix}1&2&3\\1&3&2\end{smallmatrix}), (\begin{smallmatrix}1&2&3\\3&1&2\end{smallmatrix})\} \neq Ha = \{(\begin{smallmatrix}1&2&3\\1&3&2\end{smallmatrix}), (\begin{smallmatrix}1&2&3\\2&3&1\end{smallmatrix})\}.$$

**3.27 Satz** *Es sei $H \subset G$ Normalteiler. Durch $aH \cdot bH := abH$ wird die Menge der Nebenklassen von $H$ zu einer Gruppe, der so genannten* **Faktorgruppe** *von $G$ nach $H$, in Zeichen $G/H$.*

**Beweis:**

- *Wohldefiniertheit:*

  $(aH = a'H$ und $bH = b'H) \Longrightarrow (a^{-1}a' \in H$ und $b^{-1}b' \in H) \Longrightarrow (ab)^{-1}a'b' \equiv b^{-1}(a^{-1}a')b' \in b^{-1}Hb' = b^{-1}b'H = H.$
  $\Longrightarrow abH = a'b'H$
  $\Longrightarrow aH \cdot bH = a'H \cdot b'H.$

- *Assoziativität:*

  $(aH \cdot bH) \cdot cH = abH \cdot cH = abcH = a(bc)H = aH \cdot (bc)H = aH \cdot (bH \cdot cH).$

- *Neutrales Element $eH$:*

  $eH \cdot aH = eaH = aH.$

- *Zu $aH$ inverses Element $a^{-1}H$:*

  $a^{-1}H \cdot aH = a^{-1}aH = eH = H.$

□

**3.28 Beispiele (Faktorgruppe)** 1. Wir betrachten den Normalteiler $H := n\mathbb{Z}$ von $G := \mathbb{Z}$. Die Nebenklassen bezeichnen wir mit

$$\bar{k} := k + n\mathbb{Z} \qquad (k \in \{0, \ldots, n-1\}).$$

Addition in der Faktorgruppe $\mathbb{Z}/n\mathbb{Z}$, z.B. für $n = 5$:

| $+$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ | $\bar{4}$ |
|---|---|---|---|---|---|
| $\bar{0}$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ | $\bar{4}$ |
| $\bar{1}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ | $\bar{4}$ | $\bar{0}$ |
| $\bar{2}$ | $\bar{2}$ | $\bar{3}$ | $\bar{4}$ | $\bar{0}$ | $\bar{1}$ |
| $\bar{3}$ | $\bar{3}$ | $\bar{4}$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ |
| $\bar{4}$ | $\bar{4}$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ |

2. Als abstrakte Gruppe ist $\mathcal{S}_n/\mathcal{A}_n \cong (\{-1, 1\}, \cdot)$ falls $n \neq 1$.

Das erste Beispiel zeigt, dass es endliche Gruppen jeder vorgegebenen Ordnung $n \in \mathbb{N}$ gibt.

# 4 Ringe und Körper

Ganze Zahlen können wir nicht nur addieren, sondern auch multiplizieren. Will man allerdings auch dividieren, so muss man vom Ring $(\mathbb{Z}, +, \cdot)$ der ganzen Zahlen zum Körper $(\mathbb{Q}, +, \cdot)$ der rationalen Zahlen übergehen.

Allgemein definieren wir:

**4.1 Definition** *Eine Menge $A$ mit zwei Operationen*

$$+ : A \times A \to A \quad \textit{und} \quad \cdot : A \times A \to A$$

*heißt* **Ring**, *falls gilt:*

1. $(A, +)$ *ist eine abelsche Gruppe (mit neutralem Element $0$ und zu $a$ inversem Element $-a$).*
2. *Assoziativität:* $(a \cdot b) \cdot c = a \cdot (b \cdot c) \qquad (a, b, c \in A).$
3. *Distributivität:*
   $a \cdot (b + c) = a \cdot b + a \cdot c$ *und* $(b + c) \cdot a = b \cdot a + c \cdot a \qquad (a, b, c \in A).$

4. *Falls zusätzlich gilt:*

$$\exists\, 1 \in A\ \forall a \in A : 1 \cdot a = a,$$

*heißt $A$* **Ring mit Einselement**[4]*.*

5. *Falls $\forall a, b \in A : a \cdot b = b \cdot a$, heißt $A$* **kommutativer Ring***.*

**4.2 Bemerkung** In Ausdrücken wie den rechten Seiten von 3. wird erst multipliziert, dann addiert.

Wir wollen nun sehen, wie man die rationalen Zahlen $\mathbb{Q}$, ausgehend von der Menge $\mathbb{Z}$ der ganzen Zahlen, konstruiert.

Zunächst erscheint es gar nicht notwendig, sich diese Arbeit zu machen, denn uns sind die rationalen Zahlen ja als die Brüche $p/q$ ganzer Zahlen $p$ und $q$ mit $q \neq 0$ vertraut. Bei einer formalen Definition ist es allerdings notwendig zu sagen, wann zwei solche Brüche die gleiche rationale Zahl bezeichnen. Daher führt man auf der Menge $\mathbb{Z} \times \mathbb{N}$ der Paare $(p, q)$ von Zähler und Nenner die Relation

$$(p_1, q_1) \sim (p_2, q_2) :\Longleftrightarrow p_1 q_2 = p_2 q_1$$

ein. Dies ist eine Äquivalenzrelation (nachprüfen!).

Als Repräsentanten der Äquivalenzklassen können wir z.B. die Paare $(p, q)$ mit teilerfremdem $p$ und $q$ nehmen; diese entsprechen den gekürzten Brüchen.

**4.3 Definition** *Die Menge der* **rationalen Zahlen** *ist durch $\mathbb{Q} := \mathbb{Z} \times \mathbb{N}/\sim$ gegeben, die Addition wird durch*

$$(p_1, q_1) + (p_2, q_2) := (p_1 q_2 + p_2 q_1, q_1 q_2)$$

*induziert und die Multiplikation durch*

$$(p_1, q_1) \cdot (p_2, q_2) := (p_1 p_2, q_1 q_2).$$

*Mit diesen Operationen wird $\mathbb{Q}$ zum Ring:*

**4.4 Beispiele (Ringe)** 1. $(\mathbb{Q}, +, \cdot)$ ist ein kommutativer Ring mit $1$, ebenso $\mathbb{R}$ und $\mathbb{C}$. Z.B. in [3] findet man weitere Informationen über rationale, reelle und komplexe Zahlen.

[4] Wir bezeichnen zwar das Einselement mit dem Symbol 1, es ist aber von $1 \in \mathbb{N}$ verschieden.

2. $(\mathbb{Z}, +, \cdot)$ ebenso, Division ist aber nicht möglich.

3. Für $n > 1$ ist $(n\mathbb{Z}, +, \cdot)$ ein kommutativer Ring ohne $1$.

4. Für $n \in \mathbb{N}$ können wir auf $(\mathbb{Z}/n\mathbb{Z}, +)$ eine Multiplikation einführen, indem wir die Repräsentanten der Nebenklassen miteinander multiplizieren und dann die Nebenklasse des Produktes als Ergebnis auffassen. Das Ergebnis ist unabhängig von der Wahl der Repräsentanten, denn für

$$a_2 = a_1 + rn \quad \text{und} \quad b_2 = b_1 + sn$$

ist

$$a_2 b_2 = a_1 b_1 + n(rb_1 + rsn + sa_1),$$

liegt also in der gleichen Nebenklasse wie $a_1 b_1$.

Damit wird $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ zu einem kommutativen Ring mit neutralem Element $\bar{0}$ der Addition und $\bar{1}$ der Multiplikation. Etwa für $n = 5$ ergibt sich folgende Multiplikationstabelle:

| $\cdot$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ | $\bar{4}$ |
|---|---|---|---|---|---|
| $\bar{0}$ | $\bar{0}$ | $\bar{0}$ | $\bar{0}$ | $\bar{0}$ | $\bar{0}$ |
| $\bar{1}$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ | $\bar{4}$ |
| $\bar{2}$ | $\bar{0}$ | $\bar{2}$ | $\bar{4}$ | $\bar{1}$ | $\bar{3}$ |
| $\bar{3}$ | $\bar{0}$ | $\bar{3}$ | $\bar{1}$ | $\bar{4}$ | $\bar{2}$ |
| $\bar{4}$ | $\bar{0}$ | $\bar{4}$ | $\bar{3}$ | $\bar{2}$ | $\bar{1}$ |

Diese Ringe heißen *Restklassenringe*.

5. Nicht kommutative Ringe werden wir im Zusammenhang mit linearen Abbildungen eines Raumes in sich kennen lernen.

**4.5 Satz** *In einem Ring $R$ gilt für alle $b \in R$ $0 \cdot b = b \cdot 0 = 0$.*

**Beweis:** Wegen $0 + 0 = 0$ gilt $0 \cdot b + 0 \cdot b = (0 + 0)b = 0b$, also $0 \cdot b = 0$. Analog für $b \cdot 0 = 0$. □

**4.6 Satz** *In einem Ring gilt $a \cdot (-b) = (-a)b = -(ab)$.*

**Beweis:** $ab + a(-b) = a(b + (-b)) = a \cdot 0 = 0$.

Die erste Identität ist dabei das Distributivgesetz, die zweite wurde im letzten Satz bewiesen. Analog ist $(-a)b = -(ab)$. □

Eine reelle Funktion $f : \mathbb{R} \to \mathbb{R}$ nennen wir (reelles) *Polynom*, wenn wir für ein $n \in \mathbb{N}$ und geeignete *Koeffizienten* $a_0, \ldots, a_n \in \mathbb{R}$ diese Funktion in der Form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \qquad (x \in \mathbb{R})$$

schreiben können. Nicht jede reelle Funktion lässt sich so schreiben (Gegenbeispiele sind $\cos$, $\ln$, $\tan$ etc).

Wir können ein Polynom durch die Folge $(a_0, \ldots, a_n)$ seiner Koeffizienten festlegen. Umgekehrt sind auch die Koeffizienten reeller Polynome eindeutig bestimmt:

**4.7 Satz** *Es bezeichne $R$ einen der Ringe $\mathbb{C}, \mathbb{R}, \mathbb{Q}$ oder $\mathbb{Z}$. Sind $a_i, b_i \in R$ und ist*

$$\sum_{i=0}^{n} a_i x^i = \sum_{i=0}^{n} b_i x^i \qquad (x \in R),$$

*dann ist $a_i = b_i \quad (i \in \{0, \ldots, n\})$.*

**Beweis:** Das Polynom $p(x) := \sum_{i=0}^{n} c_i x^i$ mit $c_i := a_i - b_i$ hat dann den konstanten Wert Null, und es ist zu zeigen, dass alle $c_i = 0$ sind. Andernfalls gibt es ein $m \in \{0, \ldots, n\}$ mit $c_m \neq 0$ und $c_k = 0$ für $k > m$. Wäre $m = 0$, dann wäre $p$ das konstante Polynom $p(x) = c_0 \neq 0$. Damit ist $p(x) = \sum_{i=0}^{m} c_i x^i$ und für $x \neq 0$ ist $p(x) = c_m x^m \left(1 + \sum_{i=0}^{m-1} \frac{c_i}{c_m} x^{i-m}\right)$. Für $|x| > 2m \cdot \max_{1 \leq i \leq m} |c_i / c_m|$ ist $\left|\sum_{i=0}^{m-1} \frac{c_i}{c_m} x^{i-m}\right| \leq \sum_{i=0}^{m-1} \left|\frac{c_i}{c_m}\right| \left|\frac{1}{x}\right| \leq \frac{1}{2}$, sodass $|p(x)| \geq |\frac{c_m x^m}{2}| > 0$. Widerspruch. □

Nichts hindert uns daran, analog über einem *beliebigen* Ring $R$ Folgen $(a_0, \ldots, a_n)$ von Koeffizienten $a_i \in R$ und die Funktion $x \mapsto \sum_{i=0}^{n} a_i x^i$ zu betrachten. Ist $f(x) = \sum_{i=0}^{m} a_i x^i$ und $g(x) = \sum_{i=0}^{n} b_i x^i$, dann können wir diese Polynome punktweise addieren und multiplizieren, indem wir

$$(f + g)(x) := f(x) + g(x) \quad , \quad (f \cdot g)(x) := f(x) \cdot g(x) \qquad (x \in R)$$

setzen. Es gilt dann

$$(f+g)(x) = \sum_{i=0}^{\max(m,n)} (a_i + b_i) x^i \quad \text{und} \quad (f \cdot g)(x) = \sum_{i=0}^{m+n} \left( \sum_{k=0}^{i} a_{i-k} b_k \right) x^i \quad (x \in R)$$

(wobei wir von der Möglichkeit Gebrauch gemacht haben, die Koeffizientenfolgen rechts mit Nullen aufzufüllen).

Satz 4.7 gilt aber nicht für beliebige Ringe:

**4.8 Beispiel (Polynomringe)** Wir betrachten im Restklassenring $R := \mathbb{Z}/5\mathbb{Z}$ das Polynom

$$p(x) := x^5 - x = x(x^4 - \overline{1}) = x(x^4 + \overline{4}).$$

Obwohl die Folge $(a_0, \ldots, a_5) = (\overline{0}, \overline{4}, \overline{0}, \overline{0}, \overline{0}, \overline{1})$ der Koeffizienten von $p$ ungleich der Nullfolge ist, ist die Funktion $p$ die Nullfunktion, denn es ist

$$\overline{1}^4 = \overline{2}^4 = \overline{3}^4 = \overline{4}^4 = \overline{1}.$$

Wegen dieser Uneindeutigkeit der Koeffizienten ist es sinnvoller, für beliebige Ringe $R$ Polynome über $R$ als Koeffizientenfolgen zu definieren, statt sie als Funktionen anzusehen.

Der Index $r$ des letzten von Null verschiedenen Koeffizienten $a_r \neq 0$ heißt *Grad* des Polynoms. Sind alle $a_i = 0$, dann setzen wir den Grad gleich $-\infty$. Damit ist der Grad des Summenpolynoms höchstens gleich dem Maximum der Grade der Summanden. Der Grad des Produktpolynoms ist höchstens gleich der Summe der Grade der Faktorpolynome.

**4.9 Satz** *Die Polynome über einem Ring $R$ mit punktweiser Addition und Multiplikation bilden einen kommutativen Ring. Besitzt $R$ ein Einselement, dann ist $p(x) \equiv 1$ Einselement des Polynomrings.*

Dieser Ring wird mit $R[x]$ bezeichnet. Analog kann man auch Ringe $R[x, y]$ von Polynomen zweier Variablen einführen etc.

**4.10 Definition** *Ein kommutativer Ring $K$ mit Einselement $1$ heißt* **Körper**, *wenn*

$$\forall a \in K\backslash\{0\} \qquad \exists b \in K : ba = 1 \quad \textit{(inverses Element der Multiplikation)}$$

*und*

$$1 \neq 0.$$

Das zu $a$ inverse Element $b$ muss eindeutig sein, denn $(K\backslash\{0\}, \cdot)$ bildet eine Gruppe. Wir schreiben daher $a^{-1}$ statt $b$.

**Der Körper $\mathbb{Q}$ der rationalen Zahlen:**
Der Ring $(\mathbb{Z}, +, \cdot)$ ist kein Körper, denn außer $-1$ und $1$ besitzt keine Zahl ein multiplikatives inverses Element. Wir haben aber aus $\mathbb{Z}$ den so genannten Quotientenkörper $\mathbb{Q}$ konstruiert (vgl. Definition 4.3). Repräsentanten der neutralen Elemente von Addition und Multiplikation sind $(0, 1)$ und $(1, 1)$, während

$-(p,q) = (-p,q)$ und $(1,1)/(p,q) = \mathrm{sign}(p) \cdot (q,p)$ gilt (Letzteres natürlich nur für $p \neq 0$). Die Relation $(p_1,q_1) \leq (p_2,q_2)$, falls $p_1 q_2 \leq p_2 q_1$, überträgt sich auf $\mathbb{Q}$.

**Der Körper $\mathbb{R}$ der reellen Zahlen:**
Analog bilden die reellen Zahlen den Körper $(\mathbb{R}, +, \cdot)$. Man konstruiert sie mithilfe der so genannten *Dedekind-Schnitte*, das sind Teilmengen $M \subset \mathbb{Q}$ mit $\emptyset \neq M \neq \mathbb{Q}$ und der Eigenschaft, dass aus $a \in M$, $b \in \mathbb{Q}$ mit $b \geq a$ auch $b \in M$ folgt. Außerdem wird gefordert, dass $M$ kein kleinstes Element besitzt. Als Menge definieren wir die reellen Zahlen durch

$$\mathbb{R} := \{M \subset \mathbb{Q} \mid M \text{ ist Dedekind} - \text{Schnitt}\}.$$

Die rationale Zahl $a \in \mathbb{Q}$ entspricht dabei der Menge

$$M := \{b \in \mathbb{Q} \mid b > a\} \in \mathbb{R}.$$

Die Addition bzw. Multiplikation der Dedekind-Schnitte erfolgt durch die entsprechende Operation auf den Mengen, d.h.

$$M_1 + M_2 := \{m_1 + m_2 \mid m_1 \in M_1,\ m_2 \in M_2\}$$

und

$$M_1 \cdot M_2 := \{m_1 \cdot m_2 \mid m_1 \in M_1,\ m_2 \in M_2\}.$$

Letztere Definition ist nur für Dedekind-Schnitte $M_1$ und $M_2$ sinnvoll, die nicht negative reelle Zahlen repräsentieren, die also als Mengen im neutralen Element $\{b \in \mathbb{Q} \mid b > 0\} \in \mathbb{R}$ der Addition enthalten sind. Andernfalls überträgt man das Produkt der Vorzeichen auf den Produktschnitt.

Man prüft für diese Operationen die Gültigkeit der Körperaxiome nach (siehe z.B. [1]). Auf $\mathbb{R}$ wird durch $M_1 \leq M_2 :\Longleftrightarrow M_1 \supset M_2$ eine Ordnungsrelation eingeführt.

Neben den Dedekind-Schnitten, die rationale Zahlen repräsentieren, finden wir weitere Zahlen, z.B. $\sqrt{2}$, repräsentiert durch den Dedekind-Schnitt

$$\{a \in \mathbb{Q} \mid a > 0, a^2 > 2\}.$$

**Der Körper $\mathbb{C}$ der komplexen Zahlen:**
Aus $\mathbb{R}$ konstruieren wir folgendermaßen den Körper $(\mathbb{C}, +, \cdot)$ der komplexen Zahlen. Als Menge ist

$$\mathbb{C} := \mathbb{R} \times \mathbb{R},$$

geometrisch entspricht einer komplexen Zahl also ein Punkt in der Ebene $\mathbb{R}^2$. Addition und Multiplikation von $a = (a_1, a_2)$ und $b = (b_1, b_2) \in \mathbb{C}$ sind durch

$$a + b := (a_1 + a_2, b_1 + b_2) \quad \text{und} \quad a \cdot b := (a_1 b_1 - a_2 b_2, a_1 b_2 + a_2 b_1)$$

gegeben. Damit bildet $\mathbb{C}$ eine additive Gruppe mit neutralem Element $(0, 0)$ und zu $a$ inversem Element $-a = (-a_1, -a_2)$. Das neutrale Element der Multiplikation ist $(1, 0)$. Das zu $(a_1, a_2) \in \mathbb{C} \setminus \{(0, 0)\}$ inverse Element ist

$$(a_1/(a_1^2 + a_2^2), -a_2/(a_1^2 + a_2^2)).$$

$a_1 = \mathrm{Re}(a)$ wird der *Realteil* von $a = (a_1, a_2) \in \mathbb{C}$ genannt, $a_2 = \mathrm{im}(a)$ der *Imaginärteil*. $|a| := \sqrt{a_1^2 + a_2^2}$ heißt der *Betrag* von $a$. Mit der Abkürzung $i := (0, 1)$ für die *imaginäre Einheit* ist $i^2 = -1$, und wir schreiben kurz $a_1 + ia_2$ statt $(a_1, a_2) \in \mathbb{C}$. Die Abbildung

$$\mathbb{C} \to \mathbb{C} \quad , \quad z = \mathrm{Re}(z) + i\ \mathrm{im}(z) \mapsto \overline{z} := \mathrm{Re}(z) - i\ \mathrm{im}(z)$$

heißt *Konjugation*. Sie ist ein sog. *Körperautomorphismus* von $\mathbb{C}$, d.h. es gilt immer $\overline{v + w} = \overline{v} + \overline{w}$, $\overline{v \cdot w} = \overline{v} \cdot \overline{w}$ und $\overline{1} = 1$. Außerdem ist $\overline{\overline{z}} = z$.
Warum sind komplexe Zahlen nützlich? Z.B. weil wir aus jeder Zahl die Wurzel ziehen können, oder allgemeiner, weil jedes nicht konstante komplexe Polynom eine Nullstelle besitzt. $\mathbb{C}$ ist aber auch dann nützlich, wenn man sich eigentlich nur für $\mathbb{R}$ interessiert.

**4.11 Beispiele (komplexe Zahlen)** 1. Für $a := 1 + i$ und $b := 3 - 2i$ ist

$$a + b = 4 - i \quad , \quad a \cdot b = 3 - 2(i^2) + i(3 - 2) = 5 + i \quad , \quad \overline{b} = 3 + 2i,$$

$|b|^2 = b\overline{b} = 13$ und $a/b = a\overline{b}/(b\overline{b}) = (1 + 5i)/13$.

2. Es gibt 3 komplexe dritte Wurzeln der 1: 1, $\frac{1}{2}(-1 + \sqrt{3}i)$ und $\frac{1}{2}(-1 - \sqrt{3}i)$.

3. Jede reelle kubische Gleichung lässt sich auf die Form

$$x^3 = 3px + 2q \tag{4.1}$$

mit geeigneten reellen Koeffizienten $p$ und $q$ bringen, indem man einen evtl. vorhandenen quadratischen Term durch eine Substitution der Form $x \mapsto x + k$ eliminiert. Die Lösung

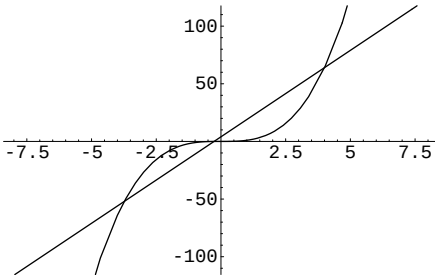
$$x = \sqrt[3]{q + \sqrt{q^2 - p^3}} + \sqrt[3]{q - \sqrt{q^2 - p^3}} \tag{4.2}$$

stammt von Cardano und findet sich in seinem 1545 erschienenen Buch 'Ars Magna'.

Im Gegensatz zu den quadratischen Gleichungen besitzt (4.1) immer eine reelle Lösung. Ist z.B.

$$x^3 = 15x + 4, \quad \text{also} \quad p = 5, q = 2,$$

dann überprüft man durch Einsetzen, dass $x_1 = 4$ eine Lösung ist. Polynomdivision ergibt die beiden anderen Lösungen:




$$(x^3 - 15x - 4)/(x - 4) = x^2 + 4x + 1, \quad \text{also} \quad x_{2/3} = -2 \pm \sqrt{3}.$$

Wie können wir alternativ die Formel (4.2) benutzen? Nun, diese lautet

$$x = \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}} = \sqrt[3]{2 + 11i} + \sqrt[3]{2 - 11i}.$$

Wie man durch Bilden der dritten Potenz überprüft, ist $2 \pm i$ eine dritte Wurzel von $2 \pm 11i$.[5] Dies entspricht der Lösung $x_1 = (2 + i) + (2 - i)$. Die beiden anderen Lösungen $x_2$ und $x_3$ erhält man, indem man analog die beiden anderen dritten Wurzeln von $2 + 11i$ einsetzt. Diese lauten (unter Verwendung des 2. Bsp.) $(2 + i) \cdot \frac{1}{2}(-1 - \sqrt{3}i) = -1 + \frac{1}{2}\sqrt{3} + (-\sqrt{3} - \frac{1}{2})i$ und $(2 + i) \cdot \frac{1}{2}(-1 + \sqrt{3}i) = -1 - \frac{1}{2}\sqrt{3} + (\sqrt{3} - \frac{1}{2})i$. Die Moral von der Geschicht'? Die Cardano-Formel liefert am Ende die drei reellen Lösungen, um diese aber zu berechnen, muss man 'ins Komplexe gehen'.

**Restklassenkörper:**
Wie wir durch Betrachtung der Multiplikationstabelle von $\mathbb{Z}/5\mathbb{Z}$ feststellen können, ist dieser Ring ein Körper, denn in jeder Spalte taucht die $\bar{1}$ auf, und $\bar{1} \neq \bar{0}$.

Das könnte uns glauben lassen, dass die Ringe $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ immer Körper sind. Das ist aber falsch: Für $\mathbb{Z}/4\mathbb{Z}$ haben wir die Multiplikationstabelle

[5] Wird das Zeichen $\pm$ in einer Gleichung benutzt, dann gilt diese für beide Vorzeichen. Taucht das Zeichen mehrmals auf, dann muss es an allen Stellen als $+$ bzw. an allen Stellen als $-$ gelesen werden.

| $\cdot$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ |
|---|---|---|---|---|
| $\bar{0}$ | $\bar{0}$ | $\bar{0}$ | $\bar{0}$ | $\bar{0}$ |
| $\bar{1}$ | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ |
| $\bar{2}$ | $\bar{0}$ | $\bar{2}$ | $\bar{0}$ | $\bar{2}$ |
| $\bar{3}$ | $\bar{0}$ | $\bar{3}$ | $\bar{2}$ | $\bar{1}$ |

Wir finden kein $b \in \mathbb{Z}/4\mathbb{Z}$ mit $b \cdot \bar{2} = \bar{1}$. Das Element $\bar{2}$ hat also kein inverses Element und $\bar{2} \neq \bar{0}$. Stattdessen hat $\mathbb{Z}/4\mathbb{Z}$ den so genannten Nullteiler $\bar{2}$, denn es gilt $\bar{2} \cdot \bar{2} = \bar{0}$.

**4.12 Definition** *Ist $R$ ein Ring und $a \in R\backslash\{0\}$, dann heißt $a$* **linker** *bzw.* **rechter Nullteiler**, *falls $b \in R\backslash\{0\}$ mit $ab = 0$ bzw. $ba = 0$ existiert.*

**4.13 Bemerkung** In einem kommutativen Ring fallen beide Begriffe natürlich zusammen.

**4.14 Satz** *Körper haben keine Nullteiler.*

**Beweis:** Es sei $a \in K\backslash\{0\}$ und $b \in K$. Aus $ab = 0$ folgt $b = a^{-1}ab = a^{-1}0 = 0$. □

Welche Restklassenringe sind nun Körper?

Die Antwort auf diese Frage ergibt sich aus folgender Beobachtung. Um ein Körper sein zu können, muss ein Ring $R$ ein Einselement $1$ besitzen. Für $n \in \mathbb{Z}$ können wir jetzt $n \times 1$ durch $n \times 1 := \underbrace{1 + 1 + \ldots + 1}_{n-\text{mal}}$ definieren. Wie das Beispiel des Ringes $\mathbb{Z}/n\mathbb{Z}$ zeigt, kann der Fall $n \times 1 = 0$ eintreten.

**4.15 Definition** *Die* **Charakteristik** *eines Körpers $K$ ist $0$, falls $n \times 1 \neq 0$ für alle $n \in \mathbb{N}$ gilt. Sonst ist sie das kleinste $n \in \mathbb{N}$ mit $n \times 1 = 0$.*

**4.16 Beispiele (Charakteristik)** 1. Die Charakteristiken von $\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ sind Null.

2. Die Charakteristik von $\mathbb{Z}/5\mathbb{Z}$ ist $5$.

**4.17 Satz** *Ist die Charakteristik $p$ eines Körpers $> 0$, dann ist $p$ eine Primzahl.*

**Beweis:** $p$ ist jedenfalls $\neq 1$, denn
$$\begin{array}{ccccccccc} \mathbb{N} & & K & & K & & K \\ \cup & & \cup & & \cup & & \cup \\ 1 & \times & 1 & = & 1 & \neq & 0 \end{array}.$$
Wäre $p > 1$ nicht prim, dann gäbe es $s, t \in \mathbb{N}$, $s, t > 1$ mit $p = st$. Damit wären $s \times 1$ und $t \times 1$ Nullteiler, denn $s \times 1 \neq 0 \neq t \times 1$ wegen $s, t < p$. □

**4.18 Korollar** *Ist $n \in \mathbb{N}$ nicht prim, dann ist $\mathbb{Z}/n\mathbb{Z}$ kein Körper.*

**4.19 Satz** *Der Restklassenring $\mathbb{Z}/n\mathbb{Z}$ ist genau dann ein Körper, wenn $n \in \mathbb{N}$ prim ist. $n$ ist dann die Charakteristik von $\mathbb{Z}/n\mathbb{Z}$.*

**Beweis:** Wir nehmen an, dass $n$ *prim* ist und müssen für jedes $\bar{a} \in (\mathbb{Z}/n\mathbb{Z})\backslash\{0\}$ die Existenz eines inversen Elementes $\bar{b} \in \mathbb{Z}/n\mathbb{Z}$ mit $\bar{b}\bar{a} = \bar{1}$ zeigen. Es reicht *hier* aber aus zu zeigen, dass $(\mathbb{Z}/n\mathbb{Z})$ nullteilerfrei ist. Denn dann ist für jedes $\bar{a} \in (\mathbb{Z}/n\mathbb{Z})\backslash\{\bar{0}\}$ die Abbildung

$$\mathbb{Z}/n\mathbb{Z} \mapsto \mathbb{Z}/n\mathbb{Z} \quad , \quad \bar{b} \mapsto \bar{b}\bar{a}$$

injektiv.

Wie sieht man diese Injektivität? Nun, wenn $\bar{b}_1 \neq \bar{b}_2$, aber $\bar{b}_1\bar{a} = \bar{b}_2\bar{a}$ wäre, dann wäre $\bar{c} := \bar{b}_1 - \bar{b}_2 \neq \bar{0}$, aber $\bar{c}\bar{a} = \bar{0}$, $\bar{a}$ und $\bar{c}$ wären also Nullteiler.

Es kann aber keine Zahlen $a, c \in \{1, \ldots, n-1\}$ geben, für die $ac$ ein Vielfaches von $n$ ist. Denn sonst müsste $a$ oder $c$ den Primfaktor $n$ enthalten.

Injektive Abbildungen einer endlichen Menge in sich sind aber auch surjektiv, also bijektiv. Es gibt also ein $\bar{b}$ mit $\bar{b} \cdot \bar{a} = \bar{1}$ und wir haben bewiesen, dass $\mathbb{Z}/n\mathbb{Z}$ ein Körper ist.

Dass $n$ auch die Charakteristik ist, sieht man daran, dass $n \times \bar{1} = \bar{n} = \bar{0}$ ist und für $m \in \mathbb{N}$, $m < n \quad \bar{m} \neq \bar{0}$ ist. □

# 5 Vektorräume

In diesem Kapitel beginnen wir mit dem eigentlichen Thema der Linearen Algebra, der Theorie der Vektorräume und der linearen Abbildungen zwischen Vektorräumen.

**5.1 Definition** *Eine abelsche Gruppe $(V, +)$ heißt* **Vektorraum** *oder* **linearer Raum über** *einem Körper $K$, wenn sie mit einer Abbildung $\cdot : K \times V \to V$, genannt* **Skalarmultiplikation**[6]*, versehen ist, für die folgende Axiome gelten:*

1. $(k_1k_2) \cdot v = k_1 \cdot (k_2 \cdot v) \qquad (k_1, k_2 \in K, v \in V)$
   *(Pseudoassoziativität).*

2. $(k_1 + k_2) \cdot v = k_1 \cdot v + k_2 \cdot v \qquad (k_1, k_2, k \in K, v, v_1, v_2 \in V)$
   $k \cdot (v_1 + v_2) = k \cdot v_1 + k \cdot v_2$
   *(Distributivität)*

[6] darf nicht mit dem *Skalarprodukt* zweier Vektoren verwechselt werden.

*3.* $1 \cdot v = v \qquad (1 \in K, v \in V)$.

*Die Elemente von* $V$ *heißen dann* **Vektoren**.

**5.2 Bemerkungen** 1. Man stelle sich die Skalarmultiplikation mit $k \in K$ als Streckung des Vektors $v$ um den Faktor $k$ vor. Für $K = \mathbb{R}$ ist das die richtige geometrische Interpretation.

2. **Pseudo**assoziativität deshalb, weil hier Objekte aus verschiedenen Mengen miteinander multipliziert werden.

3. Man beachte, dass beispielsweise bei $(k_1 + k_2) \cdot v = k_1 \cdot v + k_2 \cdot v$ links in $K$, rechts in $V$ addiert wird.

4. Würde man $1 \cdot v = v$ nicht verlangen, könnte man ohne Konflikt mit den anderen Axiomen die Skalarmultiplikation so definieren, dass alles auf $0 \in V$ abgebildet wird. Das wäre zu langweilig und wird durch Axiom 3. ausgeschlossen, falls die Gruppe $V$ aus mehr als dem Element $0$ besteht.

**5.3 Beispiele (Vektorräume)** 1. $V = (\{0\}, +)$, $K$ beliebig, $k \cdot 0 := 0$ für alle $k \in K$. Dieser *Vektorraum* heißt der *nulldimensionale Raum über* $K$. Nulldimensionale Räume über verschiedenen Körpern werden unterschieden.

2. $V := K$, Skalarmultiplikation := Körpermultiplikation. Dieser Vektorraum heißt der *eindimensionale arithmetische Vektorraum über* $K$.

3. Für $n \in \mathbb{N}$ ist auf der Menge $V := \underbrace{K \times \ldots \times K}_{n-\text{mal}}$, also dem $n$–fachen cartesischen Produkt von $K$, eine Gruppenstruktur durch

$$v + w := (v_1 + w_1, \ldots, v_n + w_n)$$

für $v := (v_1, \ldots, v_n) \in V$ und $w := (w_1, \ldots, w_n) \in V$ definiert.

Die Gruppe ist abelsch und hat das neutrale Element $0 := (0, \ldots, 0)$.

$(-v_1, \ldots, -v_n)$ ist zu $v$ invers.

Man kann jetzt eine Skalarmultiplikation $K \times V \to V$ durch

$$k \cdot v := (kv_1, \ldots, kv_n) \qquad \big(k \in K, v = (v_1, \ldots, v_n) \in V\big)$$

einführen. Der entstehende Vektorraum über $K$ wird mit $K^n$ bezeichnet und heißt *$n$–dimensionaler arithmetischer Vektorraum über $K$*. Beispielsweise ist der $\mathbb{R}^3$ unser Anschauungsraum.[7]

4. Ist $K$ ein Körper, $S$ eine beliebige, nicht leere Menge und $V := \{f : S \to K\}$, dann wird mit den Definitionen

$$(f+g)(s) := f(s) + g(s) \qquad (f, g \in V, s \in S)$$

und

$$(kf)(s) := k \underbrace{\cdot}_{\text{Mult. in } K} f(s) \qquad (f \in V, k \in K, s \in S)$$

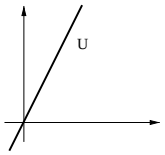
$(V, +, K, \cdot)$ zu einem Vektorraum über $K$.

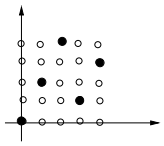
Dieser Vektorraum ist für viele praktische Zwecke zu groß. Beispielsweise wird man in der Analysis statt $V := \{f : \mathbb{R} \to \mathbb{R}\}$ die Teilmenge der stetigen (oder differenzierbaren, etc.) reellen Funktionen betrachten. Diese enthält die Nullfunktion und ist unter Addition und Skalarmultiplikation abgeschlossen.

5. Die Menge $V := K[x]$ der Polynome $a_n x^n + \ldots + a_0$ über einem Körper $K$ (aufgefasst als Koeffizientenfolgen $(a_0, \ldots, a_n)$) mit Addition der Koeffizienten und Skalarmultiplikation $k \cdot (a_0, \ldots, a_n) := (k \cdot a_0, \ldots, k \cdot a_n)$ bildet einen Vektorraum.

**5.4 Definition** *Eine Teilmenge $U \subset V$ eines $K$–Vektorraums $V$ heißt* **Untervektorraum** *oder* **Unterraum**, *wenn $U$ Untergruppe von $(V, +)$ und $k \cdot u \in U$ für $k \in K$, $u \in U$.*

**5.5 Beispiele (Untervektorräume)** 1. Die Gerade $U := \{(x, 2x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$ in der Ebene, und die Gerade $\{(x, 2x) \in (\mathbb{Z}/5\mathbb{Z})^2 \mid x \in \mathbb{Z}/5\mathbb{Z}\}$



[7] Hier zeigt sich ein typisches Dilemma der Mathematik. Eigentlich bedeutet $K^n$ ja zunächst nur eine *Menge*. Um die additive *Gruppe* zu bezeichnen, müsste man die Gruppenverknüpfung $+$ mitangeben, also $(K^n, +)$ schreiben. Der *Vektorraum* ist erst nach Angabe von Gruppe, Körper und Skalarmultiplikation fixiert, also durch das Vier-Tupel $(K^n, +, K, \cdot)$. Trotzdem schreibt man meistens kurz $K^n$ und denkt sich den Rest.

2. $U := C(\mathbb{R}, \mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ ist stetig}\}$ bildet einen Unterraum des $\mathbb{R}$–Vektorraums $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ aller Funktionen $f : \mathbb{R} \to \mathbb{R}$.

3. Für einen Körper $K$ und $n \in \mathbb{N}$ ist

$$V_n := \{p \mid p \text{ ist Polynom } \leq n\text{−ten Grades}\}$$

ein Unterraum des Vektorraumes $K[x]$ aller Polynome über $K$.

Für $K = \mathbb{R}$ ist

$$U := \left\{ v \in V_n \,\middle|\, \int_{-1}^{1} v(x)dx = 0 \right\}$$

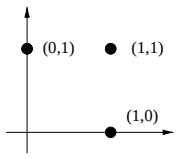
ein Unterraum, der die ungeraden Polynome (aber für $n > 1$ nicht nur diese) enthält.

**5.6 Bemerkung** Damit sind Untervektorräume von $K$–Vektorräumen selbst $K$–Vektorräume.

**5.7 Satz** *Der Durchschnitt beliebig vieler Untervektorräume eines Vektorraums ist selbst ein Unterraum.*

**Beweis:** Es sei $\mathcal{S}$ ein System von Unterräumen $U \subset V$ und $D := \bigcap_{U \in \mathcal{S}} U$. Dann sind für $a, b \in D$ auch $a, b \in U$ für alle $U \in \mathcal{S}$ und wegen der Unterraumeigenschaft der $U$ ist $a + b \in U$ und für $k \in K$ $\; ka \in U$. Also auch $a + b \in D$ und $ka \in D$. Wegen $0 \in U$ für $U \in \mathcal{S}$ ist $D$ ein Unterraum. □

Analog könnte man glauben, dass die Vereinigung von Unterräumen einen Unterraum ergibt.
Das ist aber falsch. Zwar sind $\mathbb{R} \times \{0\}$ und $\{0\} \times \mathbb{R}$ Unterräume des $\mathbb{R}^2$, aber ihre Vereinigung $V := \mathbb{R} \times \{0\} \cup \{0\} \times \mathbb{R}$, das Achsenkreuz, ist kein Unterraum, denn $(1, 0) \in V$, $(0, 1) \in V$, aber $(1, 0) + (0, 1) = (1, 1) \notin V$.




In einer gleich zu beschreibenden Weise spannen aber schon die Vektoren $(1, 0)$ und $(0, 1)$ den $\mathbb{R}^2$ auf.

**5.8 Definition** *Der von einer Teilmenge $M \subset V$* **aufgespannte** *oder* **erzeugte Untervektorraum** *ist*

$$\mathrm{span}(M) \equiv [M] := \bigcap \{U \subset V \mid U \text{ ist Unterraum von } V, M \subset U\}.$$

*$M$ heißt* **Erzeugendensystem** *von $V$, wenn $[M] = V$ ist.*

Nach dem gerade formulierten Satz ist $[M]$ tatsächlich ein Unterraum von $V$. Es ist $M \subset [M]$ und $M = [M]$ genau dann, wenn $M$ ein Unterraum ist. Es gilt $[\emptyset] = \{0\}$, denn die leere Menge ist in jedem Unterraum enthalten.

**5.9 Beispiel** $M := \{(1,0),(0,1)\} \subset \mathbb{R}^2$ spannt $[M] = \mathbb{R}^2$ auf.

Manchmal möchte man $[M]$ direkter als durch den Schnitt aller $M$ enthaltenden Unterräume aus $M$ gewinnen. So ist es im letzten Beispiel möglich, jeden Vektor $v \in [M] = \mathbb{R}^2$ als $v = c_1 \cdot (1,0) + c_2 \cdot (0,1)$ mit geeigneten $c_1, c_2 \in \mathbb{R}$ darzustellen. Allgemeiner definiert man

**5.10 Definition** *Es seien $v_1, \ldots, v_n \in V$ endlich viele Vektoren. Jeder Vektor der Form $\sum_{i=1}^n c_i v_i \in V$ mit $c_i \in K$, wird dann eine* **Linearkombination der Vektoren** *$v_1, \ldots, v_n$ genannt.*

*Ein Vektor heißt* **Linearkombination** *einer nicht leeren* **Menge** *$M \subset V$, wenn er Linearkombination endlich vieler Vektoren aus $M$ ist.*

**5.11 Satz** *Es sei $M \subset V$ nicht leer und $M^*$ die Menge der Linearkombinationen von $M$. Dann ist $M^* = [M]$.*

**Beweis:** $M^*$ ist ein Unterraum von $V$, denn $M^* \supset M \neq \emptyset$, und mit $v, w \in M^*$ und $k \in K$ ist auch

$$v + w \in M^* \quad , \quad kv \in M^*.$$

(Wir schreiben nämlich $v = \sum_{i=1}^n c_i v_i$, $w = \sum_{j=1}^m d_j w_j$ mit $v_i, w_j \in M$, $c_i, d_j \in K$. Dann sind $v + w = \sum_{i=1}^n c_i v_i + \sum_{j=1}^m d_j w_j$ ebenso wie $kv = \sum_{i=1}^n (k \cdot c_i) v_i$ wieder (endliche) Linearkombinationen von Vektoren aus $M$).

Der Unterraum $M^*$ enthält $M$, sodass $M^* \supset [M]$. Andererseits ist $[M]$ ein $M$ enthaltender Unterraum, enthält also alle Linearkombinationen von $M$. Daher gilt auch die umgekehrte Inklusion $[M] \supset M^*$. □

Unser Ausgangspunkt war die Feststellung, dass die Vereinigung von Unterräumen im Allgemeinen kein Unterraum ist. Es liegt nun nahe, statt der Vereinigung die *Summe* der Unterräume zu betrachten.

**5.12 Definition** *Es sei $\mathcal{S}$ ein Mengensystem von Untervektorräumen $U \in \mathcal{S}$ von $V$. Der* **Summenraum** *$\sum_{U \in \mathcal{S}} U$ ist durch*

$$\sum_{U \in \mathcal{S}} U := \left[ \bigcup_{U \in \mathcal{S}} U \right]$$

*definiert.*

*Ist $\mathcal{S} = \{U_1, \ldots, U_n\}$, dann schreibt man auch $U_1 + \ldots + U_n$.*

**5.13 Satz** *Der Summenraum $\sum_{U\in\mathcal{S}} U$ besteht genau aus den Vektoren $v \in V$, die sich in der Form*

$$v = \sum_{i=1}^{n} u_i \quad \textit{mit} \quad u_i \in U_i$$

*schreiben lassen, wobei die Unterräume $U_1, \ldots, U_n \in \mathcal{S}$ voneinander verschieden sind.*

**Beweis:** Da $M := \cup_{U\in\mathcal{S}} U \neq \emptyset$, lässt sich nach Satz 5.11 jeder Vektor $v \in \sum_{U\in\mathcal{S}} U$ als Linearkombination von $M$ darstellen. Also ist $v = \sum_{j=1}^{m} d_j \tilde{u}_j$, wobei wir durch Umordnung

$$\tilde{u}_j \in U_i \quad \text{für} \quad J(i) \leq j < J(i+1) \quad \text{und} \quad U_i \neq U_k \quad \text{für} \quad i \neq k$$

erreichen können. Dann ist wegen der Unterraumeigenschaft von $U_i$ aber auch $u_i := \sum_{j=J(i)}^{J(i+1)-1} c_j \tilde{u}_j \in U_i$. □

**Anwendung: Codes.** Daten werden als Folgen von Bits über Leitungen (sog. Kanäle) übertragen. Dabei passieren Übertragungsfehler. Aufgabe der sog. Kanalkodierung ist es, so weit möglich derartige Fehler zu erkennen und zu beheben.

Ein Bit können wir als die Menge $B := \{0, 1\}$ ansehen.

Da wir Bits addieren und multiplizieren wollen, fassen wir $B$ als den zweielementigen Körper auf.

Damit sind $n$ Bits mit dem arithmetischen Vektorraum $B^n$ gleichzusetzen. Es ist $|B^n| = 2^n$, wir können durch Identifikation der zu übertragenden Zeichen mit gewissen Vektoren aus $B^n$ also maximal $2^n$ Zeichen übertragen. Es ist aber ratsam, nicht alle Vektoren als Codewörter zu verwenden, will man eventuell auftretende Übertragungsfehler erkennen und beheben.

Der sog. *Hammingabstand*

$$d : B^n \times B^n \to \{0, \ldots, n\} \quad , \quad d(v, w) := n - \sum_{k=1}^{n} \delta(v_k, w_k)$$

definiert eine Metrik auf $B^n$, die translationsinvariant im Sinne $d(v + u, w + u) = d(v, w)$ ist. Um nun möglichst viele Übertragungsfehler beheben oder zumindest erkennen zu können, wird man versuchen den minimalen Hammingabstand der Codewörter möglichst groß zu halten.

Typischerweise wählt man die Vektoren, die Zeichen codieren, aus einem Unterraum.

**5.14 Definition** $C \subset B^n$ *heißt (linearer)* **Code**, *wenn* $C$ *ein Unterraum von* $B^n$ *ist. Ist* $k := \dim(C)$, *dann nennt man* $C$ *einen* $(n,k)$*–Code.*

**5.15 Beispiele (lineare Codes)** 1. $n = 3$, $C = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\} \subset B^3$ ist ein $k = 1$–dimensionaler Unterraum, es handelt sich also um einen $(3,1)$–Code. Das Zeichen $0$ wird durch $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \in C$ codiert, die $1$ durch $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in C$ (daher heißt $C$ auch *Wiederholungscode*).

Bei der Decodierung werden den Zeichen $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ die $1$ zugeordnet, den Zeichen $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ die $0$.

Damit kann ein Fehler bei der Übertragung erkannt und beseitigt werden, allerdings bei einer auf ein Drittel gesenkten Übertragungsrate.

2. $n$ beliebig, $C = \left\{ \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in B^n \mid b_n = \sum_{i=1}^{n-1} b_i \right\}$.

Hier dient das letzte Bit als sog. *Paritätsbit*, der Code heißt entsprechend *Paritätscode*.

Die Vektoren $b \in C$ haben die Eigenschaft, dass $\sum_{i=1}^{n} b_i = 0$ ist.

Ein einzelner Übertragungsfehler wird dadurch erkannt, dass der fehlerhaft übertragene Vektor die Paritätssumme $\sum_{i=1}^{n} b_i = 1$ besitzt. Zwar kann der Fehler nicht korrigiert werden (da im Allgemeinen unbekannt ist, an welcher der $i$ möglichen Positionen er auftrat). Man kann aber die Nachricht ein zweites Mal übertragen und auf mehr Glück hoffen.

Der Vorteil dieses *Paritätscodes* ist, dass er ein $(n, n-1)$–Code ist, also eine gute Übertragungsrate bewirkt.

# 6 Basis und Dimension

Die Vektoren $v_1 := (3,0)$, $v_2 := (0,2)$, $v_3 := (2,3) \in \mathbb{R}^2$ spannen den $\mathbb{R}^2$ auf, es gilt also

$$\mathbb{R}^2 = [\{v_1, v_2, v_3\}].$$

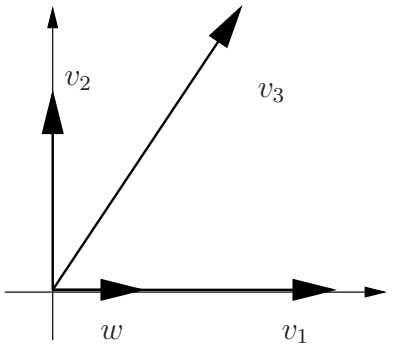
Wir können damit jeden Vektor $w = (w_1, w_2) \in \mathbb{R}^2$ als Linearkombination der $v_i$ darstellen. Diese Darstellung ist aber nie eindeutig. So gilt für

$$w := (1,0) \qquad w = \tfrac{1}{3} v_1,$$

aber auch

$$w = \tfrac{1}{2}v_3 - \tfrac{3}{4}v_2 \quad \text{etc.}$$

Der intuitive Grund für diese Uneindeutigkeit ist der, dass wir mit $v_1, v_2$ und $v_3$ zu viele Vektoren zur Linearkombination zur Verfügung haben, genauer gesagt, einen zu viel. Denn die "Dimension" des $\mathbb{R}^2$ ist ja zwei, es sollten also zwei dieser drei Vektoren ausreichen. Man versteht das Problem besser, wenn man feststellt, dass der Nullvektor $0 = (0, 0)$ die Darstellung




$$0 = -\tfrac{2}{3}v_1 - \tfrac{3}{2}v_2 + v_3$$

besitzt. Hat man $w$ als Linearkombination der $v_1, v_2, v_3$ dargestellt, kann man auf beiden Seiten ein Vielfaches des Nullvektors hinzuaddieren und erhält eine neue Darstellung von $w$ als Linearkombination der $v_1, v_2, v_3$. Z.B. ist

$$w = \tfrac{1}{3}v_1 + 0 = -\tfrac{1}{3}v_1 - \tfrac{3}{2}v_2 + v_3.$$

Es ist nun *praktisch*, Vektoren als Linearkombinationen fester Vektoren zu schreiben, denn dann überträgt sich Addition von Vektoren und Multiplikation mit Skalaren auf die Koeffizienten. Andererseits ist es oft *unpraktisch*, wenn die Darstellung als Linearkombination nicht eindeutig ist. Daher definiert man

**6.1 Definition** *Vektoren $v_1, \ldots, v_n \in V$ heißen* **linear unabhängig**, *wenn der Nullvektor nur die* **triviale Darstellung** $0 = 0 \cdot v_1 + \ldots + 0 \cdot v_n$ *zulässt, wenn aus $0 = \sum_{i=1}^{n} c_i v_i$, also $c_1 = \ldots = c_n = 0$ folgt.*

*Andernfalls heißen sie* **linear abhängig**.

**6.2 Beispiel** $\{v_1, v_2, v_3\}$ sind linear abhängig. $\{v_1, v_2\}$ sind linear unabhängig, ebenso wie $\{v_1, v_3\}, \{v_2, v_3\}, \{v_1\}, \{v_2\}$ und $\{v_3\}$.

**6.3 Definition** *Eine Teilmenge $M \subset V$ heißt* **linear unabhängig**, *wenn je endlich viele verschiedene Vektoren aus $M$ linear unabhängig sind, sonst* **linear abhängig**.

**6.4 Bemerkungen**

1. Insbesondere ist $M = \emptyset$ linear unabhängig.
2. Teilmengen linear unabhängiger Mengen sind linear unabhängig.
3. Obermengen linear abhängiger Mengen sind linear abhängig.

4. Ein einzelner Vektor $v \neq 0$ ist immer linear unabhängig.

Ist eine Menge von mindestens zwei Vektoren linear abhängig, so kann man mindestens einen durch die anderen ersetzen. So ist in unserem Beispiel

$$v_3 = \tfrac{2}{3}v_1 + \tfrac{3}{2}v_2.$$

**6.5 Satz** *Eine aus mindestens zwei Vektoren bestehende Menge $M \subset V$ ist genau dann linear abhängig, wenn ein Vektor $m \in M$ sich als Linearkombination $m = \sum_{i=1}^{n} c_i v_i$ mit $v_i \in M$, $v_i \neq m$ schreiben lässt.*

**Beweis:** • Wenn eine solche Beziehung besteht, ist der Nullvektor $0 = -m + \sum_{i=1}^{n} c_i v_i$, also eine nicht triviale Linearkombination aus Vektoren von $M$. Damit ist $M$ linear abhängig.
• Ist andererseits $M$ linear abhängig, dann ist $0 = \sum_{j=1}^{r} d_j m_j$ mit $m_j \in M$ und $d_j \in K$, wobei o.B.d.A. $d_r \neq 0$ angenommen werden darf. Damit ist

$$m_r = \sum_{j=1}^{r-1} c_j m_j \quad \text{mit} \quad c_j = -\frac{d_j}{d_r}.$$

□

**6.6 Definition** $M \subset V$ *heißt* **Basis von** $V$, *wenn* $[M] = V$ *und* $M$ *linear unabhängig.*

$M$ darf also nicht zu klein sein, denn sonst gilt nicht $[M] = V$ und nicht zu groß, denn sonst wäre $M$ linear abhängig.

Es ist also erst mal unklar, ob jeder Vektorraum eine Basis besitzt.

**6.7 Beispiele (Basen)** 1. $\emptyset$ ist Basis des Nullraums $V = \{0\}$.

2. Für jede Zahl $v \in V \backslash \{0\}$ des eindimensionalen arithmetischen Vektorraums $V := K$ über dem Körper $K$ ist $\{v\}$ eine Basis desselben, denn wir können ja jedes $w \in V$ als $\frac{w}{v}$–faches von $v$ darstellen.

3. Im einleitenden Beispiel ist $\{v_1, v_2\}$ eine Basis, aber auch $\{v_2, v_3\}$ und $\{v_1, v_3\}$, denn diese Vektoren sind linear unabhängig (und spannen den $\mathbb{R}^2$ auf). Allgemeiner betrachten wir zwei Vektoren $v = (v_1, v_2)$ und $w = (w_1, w_2)$ aus dem $\mathbb{R}^2$, und fragen, ob sie eine Basis bilden. Dazu müsste sich jeder Vektor

$b = (b_1, b_2) \in \mathbb{R}^2$ eindeutig als Linearkombination $b = x \cdot v + y \cdot w$ darstellen lassen, das Gleichungssystem

$$\begin{aligned} v_1 x + w_1 y &= b_1 \\ v_2 x + w_2 y &= b_2 \end{aligned}$$

müsste also für alle $b_1, b_2$ eindeutig lösbar sein. Subtrahieren wir das $w_1$-fache der zweiten Gleichung vom $w_2$-fachen der ersten, bzw. das $v_2$-fache der ersten Gleichung vom $v_1$-fachen der zweiten, dann erhalten wir

$$\begin{aligned} (v_1 w_2 - v_2 w_1) x &= w_2 b_1 - w_1 b_2 \\ (v_1 w_2 - v_2 w_1) y &= v_1 b_2 - v_2 b_1, \end{aligned}$$

was wir für den Fall $v_1 w_2 - v_2 w_1 \neq 0$ nach $x$ und $y$ auflösen können. Ist dagegen $v_1 w_2 - v_2 w_1 = 0$, dann kann das daran liegen, dass $v = w = 0$ ist. Andernfalls gilt aber

$$v_1 \cdot w = w_1 \cdot v \quad \text{und} \quad v_2 \cdot w = w_2 \cdot v,$$

wobei nicht alle Faktoren verschwinden. Auch in diesem Fall sind also $v$ und $w$ linear abhängig und bilden damit keine Basis.

4. Im arithmetischen Vektorraum $K^n$ bilden die Vektoren

$$\begin{aligned} e_1 &:= (1, 0, \ldots, 0) \\ e_2 &:= (0, 1, 0, \ldots, 0) \\ &\vdots \\ e_n &:= (0, \ldots, 0, 1) \end{aligned}$$

eine Basis, die so genannte *kanonische Basis*.

5. Im Vektorraum $\mathbb{R}[x]$ der reellen Polynome bilden die Vektoren $e_n, n \in \mathbb{N}_0$, $e_n(x) := x^n$, eine Basis.

Besitzt nun jeder Vektorraum $V$ eine Basis? Ja, aber der Beweis ist nicht elementar. Er benutzt das sog. *Zornsche Lemma*.

Die Grundidee ist, dass zunächst sicher Erzeugendensysteme von $V$ existieren (z.B. $V$ selbst). Diese sind i.A. nicht linear unabhängig, wir betrachten aber die Familie

$$\mathcal{S} := \{M \subset V \mid M \text{ linear unabhängig}\}$$

linear unabhängiger Mengen $M$ von Vektoren. Beginnend mit der (linear unabhängigen!) leeren Menge, nehmen wir zu unserer Menge immer neue Vektoren hinzu, bis bei Hinzunahme eines weiteren Vektors die lineare Unabhängigkeit verloren gehen würde.

Zur exakten Formulierung benötigt man den Begriff der Kette:

**6.8 Definition** *Es sei $\mathcal{S}$ ein nicht leeres Mengensystem und $\mathcal{K} \subset \mathcal{S}$, $\mathcal{K} \neq \emptyset$. Dann heißt $\mathcal{K}$ eine* **Kette***, falls aus $M_1, M_2 \in \mathcal{K}$ stets $M_1 \subset M_2$ oder $M_2 \subset M_1$ folgt. Eine Menge $M \in \mathcal{S}$ heißt* **maximales Element** *von $\mathcal{S}$, falls aus $N \in \mathcal{S}$ und $M \subset N$ stets $M = N$ folgt.*

**6.9 Beispiele** 1. Für $\mathcal{S} := \mathcal{P}(\{1,2,3\})$ ist die Teilmenge $\{\{3\},\{2,3\}\}$ eine Kette, $\{1,2,3\}$ maximal.
2. Das Mengensystem $\mathcal{S} := \{\{1,\ldots,n\} \mid n \in \mathbb{N}\}$ ist eine Kette ohne maximales Element.

**Zornsches Lemma**
Wenn für jede Kette $\mathcal{K}$ von $\mathcal{S}$ die Vereinigungsmenge $\bigcup_{k\in\mathcal{K}} k \in \mathcal{S}$, dann gibt es zu jeder Menge $A \in \mathcal{S}$ ein maximales Element $M$ von $\mathcal{S}$ mit $A \subset M$.

**6.10 Satz** *Jeder Vektorraum $V$ besitzt eine Basis $B$.*

**Beweis:** Die Voraussetzung des Zornschen Lemmas ist erfüllt, denn für eine Kette $\mathcal{K} \subset \mathcal{S}$ linear unabhängiger Teilmengen ist $W := \bigcup_{k\in\mathcal{K}} k$ linear unabhängig: eine *endliche* Teilmenge von $W$ ist auch schon in einem Element $k \in \mathcal{K}$ der Kette enthalten und damit linear unabhängig.

Nach dem Zornschen Lemma gibt es insbesondere zu $A := \emptyset \in \mathcal{S}$ ein maximales Element $B \in \mathcal{S}$, also eine linear unabhängige Menge, die in keiner größeren enthalten ist. $B$ muss eine Basis sein, denn würde $B$ nicht den gesamten Vektorraum aufspannen, dann gäbe es einen Vektor $v \in V - [B]$, und dieser wäre noch linear unabhängig. □

Wirklich wichtig wird für uns die Existenz einer Basis *endlich*-dimensionaler Räume sein. Später werden bei den interessierenden unendlichdimensionalen Vektorräumen Wege gefunden werden, "Linearkombinationen" unendlich vieler Vektoren zu definieren.

Mit der selben Beweismethode können wir den folgenden *Basisergänzungssatz* zeigen:

**6.11 Satz (Basisergänzungssatz)** *Es sei $A \subset V$ linear unabhängig. Dann gibt es eine Basis $B$ von $V$ mit $B \supset A$.*

**Beweis:** Nehme im Beweis statt der leeren Menge $A$ als Ausgangspunkt. □

Die informelle Bemerkung, dass Basen nicht zu klein und nicht zu groß sein dürfen, wird jetzt in einem Satz präzisiert.

**6.12 Satz** *Für eine Teilmenge $B \subset V$ eines Vektorraums $V$ sind folgende Aussagen paarweise äquivalent:*

1. *$B$ ist Basis von $V$.*
2. *$[B] = V$, aber für jede echte Teilmenge $C \subsetneq B$ ist $[C] \neq V$.*
3. *$B$ ist linear unabhängig, aber für $C \supsetneq B$ ist $C$ linear abhängig.*

*Ist $V$ nicht der Nullraum, dann ist außerdem äquivalent:*

4. *Jeder Vektor aus $V$ kann auf genau eine Weise als Linearkombination von $B$ dargestellt werden.*

**Beweis:** Wir schauen uns zunächst den Spezialfall $V = \{0\}$ an. $V$ hat nur die Teilmengen $V$ und $\emptyset$. Für $V$ ist keine der vier Aussagen wahr, für $\emptyset$ Aussagen 1.–3.

Also können wir im Folgenden $V \neq \{0\}$ voraussetzen und müssen die Äquivalenz aller vier Aussagen zeigen, das sind im Prinzip 12 Implikationen. Es reicht aber so viele Implikationen zu zeigen, dass man im gerichteten Graphen, dessen Ecken die Aussagen und dessen gerichtete Kanten die Implikationen sind, von jeder Stelle aus in Pfeilrichtung jede andere erreichen kann, z.B.

• $1. \Longrightarrow 4.$**:** Wir zeigen $\neg 4. \Longrightarrow \neg 1.$ Es gibt zwei mögliche Gründe dafür, dass 4. nicht erfüllt ist:

- Es lässt sich nicht jeder Vektor $v \in V$ als Linearkombination von $B$ darstellen. Dann ist $[B] \neq V$ , also $B$ keine Basis.

- Es lässt sich $v \in V$ als $v = \sum_{i=1}^{n} c_i b_i$ und $v = \sum_{i=1}^{n} d_i b_i$ mit $c_i, d_i \in K$ und $b_i \in B$ darstellen, wobei die $b_i$ voneinander verschieden und $c_j \neq d_j$ für mindestens ein $j \in \{1, \ldots, n\}$ ist. Dann ist $0 = \sum_{i=1}^{n} k_i b_i$ mit $k_i := c_i - d_i$ und $k_j \neq 0$. $B$ ist damit ebenfalls keine Basis.

• $4. \Longrightarrow 2.$**:** Da jeder Vektor als Linearkombination von $B$ dargestellt werden kann, ist $[B] = V$, denn wir haben gesehen, dass für $B \neq \emptyset$ $[B]$ aus den Linearkombinationen von $B$ besteht.

Es sei $C \subsetneq B$, aber $[C] = V$. Dann lässt sich ein Vektor $b \in B - C$ finden, und $b$ lässt sich als Linearkombination von $C$ darstellen. Andererseits ist $b = 1 \cdot b$

eine andere Darstellung von $b$. Widerspruch zu 4.!

• $2. \Longrightarrow 3.$**:** Wegen $[B] = V \neq \{0\}$ ist $B \neq \emptyset$.

Wir zeigen zunächst die lineare Unabhängigkeit von $B$.

1. Ist $B = \{v\}$, dann muss $v \neq 0$ sein und damit $B$ linear unabhängig.

2. Ist $B$ mindestens zweielementig, dann haben wir unter der Voraussetzung der linearen Abhängigkeit von $B$ die Existenz eines Vektors $m \in B$ bewiesen, der Linearkombination von $C := B - \{m\}$ ist. $C$ wäre echte Teilmenge von $B$, aber $B \subset [C]$ und damit $V = [B] \subset [C]$ im Widerspruch zur Voraussetzung.

Also ist $B$ linear unabhängig und wir müssen noch zeigen, dass jede echte Obermenge $C \supsetneq B$ linear abhängig ist. Es gibt also einen Vektor $c \in C - B$, und mit $c \in V = [B]$ lässt dieser sich als Linearkombination von $B$ darstellen. Nach Satz 6.5 ist damit $C$ linear abhängig.

• $3. \Longrightarrow 1.$**:** Da $B$ linear unabhängig ist, müssen wir nur noch die zweite Basiseigenschaft $[B] = V$ nachprüfen.

Mittels des Zornschen Lemmas hatten wir jedenfalls die Existenz einer Basis $B^* \supset B$ von $V$ gezeigt. Wäre $B^* \neq B$, dann wäre nach der Voraussetzung $B^*$ linear abhängig, also keine Basis. Damit ist $B^* = B$ eine Basis. □

Für alle vom Nullraum verschiedenen Vektorräume außer dem eindimensionalen arithmetischen Vektorraum $\mathbb{Z}/2\mathbb{Z}$ gibt es *verschiedene* Basen, denn ich kann ja z.B. alle Basiselemente mit einer von $0$ und $1$ verschiedenen Zahl aus $K$ multiplizieren.

In den Anwendungen der Linearen Algebra sind Basen rechnerische Hilfsmittel und ihre Wahl ist eine Frage der Bequemlichkeit. Es ist also eine natürlich auftretende Frage, ob ich eine vorgegebene Basis so abändern kann, dass sie bestimmte vorgegebene Vektoren enthält.

Wir schauen uns das Problem für Vektorräume mit endlicher Basis $B$ an, also $B = \{b_1, \ldots, b_n\}$. Wir vereinbaren hier und im Folgenden, dass in dieser Schreibweise für die Basisvektoren $b_i \neq b_j$ gilt, falls $i \neq j$.[8]

**6.13 Satz** *Es sei $B := \{b_1, \ldots, b_n\} \subset V$ eine Basis von $V$ und der Vektor $b \in V$ besitze die Darstellung $b = \sum_{i=1}^{n} k_i b_i$, $k_i \in K$. Ist für einen Index $j \in \{1, \ldots, n\}$ $k_j \neq 0$, dann ist auch $B' := \{b_1, \ldots, b_{j-1}, b, b_{j+1}, \ldots, b_n\}$ eine Basis von $V$.*

[8] Das folgt *nicht* aus der Tatsache, dass sonst schon $b_i$ und $b_j$ linear abhängig wären, denn die *Menge* $B$ könnte auch bei Mehrfachaufführung eines Elementes linear unabhängig sein!

Wir können also $b_j$ durch $b$ austauschen.

**Beweis:** O.B.d.A. $j = 1$, also $b_1 = \left(b - \sum_{j=2}^{n} k_j b_j\right)/k_1$.

Da $B$ eine Basis ist, lässt sich jeder Vektor $v \in V$ in der Form

$$v = \sum_{i=1}^{n} c_i b_i$$

darstellen oder auch in der Form

$$v = c_1/k_1 \cdot b + \sum_{i=2}^{n} \left(c_i - \frac{c_1 k_i}{k_1}\right) b_i.$$

Damit ist $v \in [B']$, also $[B'] = V$.

$B'$ ist auch linear unabhängig, denn aus

$$0 = cb + \sum_{i=2}^{n} c_i b_i$$

folgt mit $b = \sum_{i=1}^{n} n_i b_i$

$$0 = ck_1 b_1 + \sum_{i=2}^{n} (c_i + c \cdot k_i) b_i.$$

Wegen der linearen Unabhängigkeit von $B$ muss $ck_1 = 0 = c_i + k_i c \ (i = 2, \ldots, n)$ gelten, wir haben aber $k_1 \neq 0$ vorausgesetzt, sodass $c = 0$ und $c_i = 0 \ \ (i = 2, \ldots, n)$ folgt. Also ist $B'$ linear unabhängig. □

Statt einem Vektor der Basis können wir unter Umständen auch mehrere austauschen.

**Austauschsatz von Steinitz.** Es sei $B := \{b_1, \ldots, b_n\}$ eine Basis von $V$, und die Vektoren $a_1, \ldots, a_k \in V$ seien linear unabhängig. Dann gilt $k \leq n$, und für eine geeignete Nummerierung (=Permutation) $\pi \in \mathcal{S}_n$ ist auch

$$B' = \{a_1, \ldots, a_k, b_{\pi(k+1)}, \ldots, b_{\pi(n)}\}$$

eine Basis.

**Beweis:** Durch Induktion über $k$.

Für $k = 1$ reduziert sich der Satz auf den vorigen (Induktionsanfang).

*Induktionsschritt* $k-1 \to k$: Für eine geeignete Permutation $\sigma \in S_n$ ist $B'' := \{a_1, \ldots, a_{k-1}, b_{\sigma(k)}, \ldots, b_{\sigma(n)}\}$ eine Basis von $V$, und es gilt $k-1 \leq n$. Nun wäre für $k-1 = n$ schon $\{a_1, \ldots, a_{k-1}\} = B''$ eine Basis, was aber der linearen Unabhängigkeit von $\{a_1, \ldots, a_k\}$ widerspräche. Also $k \leq n$.

Nun besitzt $a_k$ eine Darstellung

$$a_k = \sum_{i=1}^{k-1} c_i a_i + \sum_{i=k}^{n} c_i b_{\sigma(i)} \quad \text{mit} \quad c_1, \ldots, c_n \in K$$

durch die Basis $B''$. Wegen der linearen Unabhängigkeit von $\{a_1, \ldots, a_k\}$ muss mindestens ein Koeffizient $c_j$, $j \geq k$, von Null verschieden sein.

Nach dem letzten Satz können wir damit $b_{\sigma(j)}$ gegen $a_k$ austauschen und erhalten nach einer Transposition $j \longleftrightarrow k$ die Basis $B'$. □

Der Austauschsatz führt uns zum Dimensionsbegriff. Sind nämlich $\{b_1, \ldots, b_n\}$ und $\{a_1, \ldots, a_k\}$ zwei Basen, dann ist $k \leq n$, aber aus Symmetriegründen auch $n \leq k$, also $k = n$. Besitzt andererseits ein Vektorraum eine unendliche Basis, dann müssen alle seine Basen aus unendlich vielen Elementen bestehen.

Wir können daher invariant definieren:

**6.14 Definition** *Besitzt ein Vektorraum $V$ eine endliche Basis $\{a_1, \ldots, a_n\}$, so wird $\dim(V) := n$ die* **Dimension** *von $V$ genannt. Sonst heißt $V$* **unendlichdimensional** *und man setzt $\dim(V) := \infty$.*

**6.15 Beispiele**

1. $\dim(\{0\}) = 0$.
2. Für die arithmetischen Vektorräume $K^n$ über $K$ ist $\dim(K^n) = n$.
3. Der Raum $V$ der Polynome über $\mathbb{R}$ ist unendlichdimensional: $\dim(V) = \infty$, denn die $e_n$, $n \in \mathbb{N}_0$, $e_n(x) = x^n$ bilden eine unendliche Basis.

**6.16 Bemerkung** Man kann den Austauschsatz auch auf den Fall unendlicher Basen ausdehnen und zeigen, dass je zwei Basen die gleiche Kardinalität $=:$ $\dim$ besitzen. Aber auch mit dieser Präzisierung ist der folgende Satz ohne die Voraussetzung $\dim(V) < \infty$ falsch.

**6.17 Satz** *Ist $\dim(V) < \infty$ und $U \subset V$ ein Unterraum, dann ist $\dim(U) \leq \dim(V)$. Aus $\dim(U) = \dim(V)$ folgt $U = V$.*

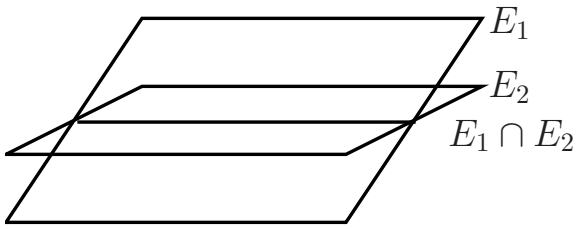
**Beweis:** Wir können eine Basis von $U$ zu einer Basis von $V$ erweitern, sodass die erste Aussage gilt. Ist $\dim(U) = \dim(V) =: n$, dann besitzt $U$ eine Basis $\{b_1, \ldots, b_n\}$, die aber auch eine Basis von $V$ sein muss. □

Betrachten wir einmal zwei Ebenen $E_1, E_2 \subset \mathbb{R}^3$. Es gilt $\dim(E_1) = \dim(E_2) = 2$ nach Definition dessen, was wir als eine Ebene betrachten.

Entweder ist nun $E_1 = E_2 (= E_1 \cap E_2)$ oder $E_1 \cap E_2$ ist eine Gerade, also $\dim(E_1 \cap E_2) = 1$.

In beiden Fällen gilt

$$\begin{array}{ccccccc} \dim E_1 & + & \dim E_2 & = & \dim(E_1 \cap E_2) & + & \dim(E_1 + E_2) \\ 2 & + & 2 & = & 2 & + & 2 \\ 2 & + & 2 & = & 1 & + & 3 \end{array}$$




Im zweiten Fall ist $\mathbb{R}^3 = E_1 + E_2$.

Allgemein wird diese Situation durch den Dimensionssatz erfasst:

**Dimensionssatz:** Sind $E, F \subset V$ Unterräume endlicher Dimension, dann gilt

$$\dim(E) + \dim(F) = \dim(E \cap F) + \dim(E + F).$$

**Beweis:** Zunächst ist nach dem vorigen Satz $\dim(E \cap F) < \infty$.

Es sei $\{a_1, \ldots, a_m\}$ eine Basis von $E \cap F$.

Diese lässt sich zu einer Basis

$$\{a_1, \ldots, a_m, e_1, \ldots, e_n\} \quad \text{von} \quad E$$

und zu einer Basis

$$\{a_1, \ldots, a_m, f_1, \ldots, f_r\} \quad \text{von} \quad F$$

erweitern. (Ist $\dim(E \cap F) = 0$, dann lässt man die $a$'s weg und setzt $m := 0$.)

Es soll jetzt gezeigt werden, dass

$$B := \{a_1, \ldots, a_m, e_1, \ldots, e_n, f_1, \ldots, f_r\}$$

eine Basis von $E + F$ ist.

Nach einem Satz über den Summenraum lässt sich jeder Vektor $v \in E + F$ in der Form $v = e + f$ mit $e \in E$ und $f \in F$ darstellen. Diese Darstellung wird im Allgemeinen nicht eindeutig sein, aber jedenfalls lässt sich $e$ als Linearkombination der $\{a_1, \ldots, a_m, e_1, \ldots, e_n\}$ und $f$ als Linearkombination der $\{a_1, \ldots, a_m, f_1, \ldots, f_r\}$ darstellen.

Damit ist $v$ Linearkombination der Vektoren aus $B$, also $[B] = U + V$.

Wir müssen noch nachweisen, dass $B$ linear unabhängig ist. Es sei dazu

$$0 = \sum_{i=1}^{m} \lambda_i a_i + \sum_{i=1}^{n} \mu_i e_i + \sum_{i=1}^{r} \nu_i f_i,$$

also

$$E \ni \sum_{i=1}^{m} \lambda_i a_i + \sum_{i=1}^{n} \mu_i e_i = -\sum_{i=1}^{r} \nu_i f_i \in F.$$

Da der Vektor aus $E$ und aus $F$ ist, muss er aus $E \cap F$ sein, also $\mu_i = \nu_i = 0$. Da $\{a_1, \ldots, a_m\}$ eine Basis von $E \cap F$ ist, müssen dann aber auch die $\lambda_i = 0$ sein.

Damit ist $\dim(E \cap F) = m$, $\dim(E) = m + n$, $\dim(F) = m + r$ und $\dim(E + F) = m + n + r$, woraus die Aussage folgt. □

# 7 Koordinaten

Die arithmetischen Vektorräume haben den rechnerischen Vorteil, dass in ihnen Vektoraddition und Multiplikation mit Skalaren auf die entsprechenden Körperoperationen zurückgeführt sind.

Durch Einführung einer Basis $\{b_1, \ldots, b_n\}$ in einem $n$–dimensionalen $K$–Vektorraum $V$ können wir in $V$ rechnen wie im $K^n$. Sind nämlich $v, w \in V$ zwei Vektoren der Form $v = \sum_{i=1}^{n} v_i b_i$, $w = \sum_{i=1}^{n} w_i b_i$ mit $v_i, w_i \in K$, dann ist

$$v + w = \sum_{i=1}^{n} (v_i + w_i) b_i.$$

Ähnlich ist für $c \in K$

$$c \cdot v = \sum_{i=1}^{n} (c v_i) b_i.$$

Die eindeutig bestimmten Koeffizienten $v_1, \ldots, v_n$ der Darstellung von $v$ als Linearkombination der Basisvektoren werden die *Koordinaten* von $v$ bezüglich der Basis $B$ genannt, $(v_1, \ldots, v_n)$ der *Koordinatenvektor* von $v$ bez. $B$.

Die (geordnete) Basis $B$ vermittelt eine Abbildung

$$\Phi_B : K^n \to V \quad , \quad (v_1, \ldots, v_n) \mapsto \sum_{i=1}^{n} v_i b_i,$$

für die

$$\Phi_B(a+b) = \Phi_B(a) + \Phi_B(b) \qquad (a, b \in K^n)$$

und

$$\Phi_B(ca) = c\Phi_B(a) \qquad (a \in K^n, c \in K)$$

gilt. $\Phi_B$ ist damit eine sog. *lineare Abbildung*.

Nur der Nullvektor $0 \in K^n$ wird unter $\Phi_B$ auf den Nullvektor $0 \in V$ abgebildet, und wir können $\Phi_B$ umkehren:

$$\Phi_B^{-1}(v) = (v_1, \ldots, v_n) \in K^n \quad \text{für} \quad v = \sum_{i=1}^{n} v_i b_i.$$

Die lineare Abbildung $\Phi_B$ ist damit ein sog. *$K$–Vektorraum-Isomorphismus*, und wir können Rechnungen im einen Vektorraum auch im anderen durchführen und dann wieder zurücktransformieren, ohne das Ergebnis zu verändern. Insbesondere gilt:

**7.1 Satz** *Vektoren $a_1, \ldots, a_k \in V$ sind genau dann linear unabhängig, wenn ihre Koordinatenvektoren $\Phi_B^{-1}(a_1), \ldots, \Phi_B^{-1}(a_k) \in K^n$ linear unabhängig sind.*

Ist nun $c := \{c_1, \ldots, c_n\}$ eine zweite (geordnete) Basis von $V$, dann stellt sich die Frage, wie die Koordinatenvektoren $\Phi_B^{-1}(v)$ und $\Phi_C^{-1}(v)$ bez. der beiden Basen zusammenhängen.

Dazu schreiben wir die Vektoren $c_i$ der zweiten Basis als Linearkombinationen

$$c_k = \sum_{i=1}^{n} m_{i,k} b_i \quad , \quad (k = 1, \ldots, n)$$

der ersten Basis mit Koeffizienten $m_{i,k} \in K$. Ist nun $v = \sum_{k=1}^{n} w_k \cdot c_k$, also $\Phi_C^{-1}(v) = (w_1, \ldots, w_n)$, dann gilt

$$v = \sum_{k=1}^{n} w_k \cdot \left( \sum_{i=1}^{n} m_{i,k} b_i \right) = \sum_{i=1}^{n} \left( \sum_{k=1}^{n} m_{i,k} w_k \right) b_i,$$

also schließt man durch Koeffizientenvergleich

$$v_i = \sum_{k=1}^{n} m_{i,k} w_k \qquad (i = 1, \ldots, n).$$

Damit folgt der

**7.2 Satz** *Der* Basistransformation $c_k = \sum_{i=1}^{n} m_{i,k} b_i \quad (k = 1, \ldots, n)$ *entspricht die* Koordinatentransformation

$$v_i = \sum_{k=1}^{n} m_{i,k} w_k \qquad (i = 1, \ldots, n).$$

Die in den Transformationsformeln auftretenden Zahlen $m_{i,k}, \;\; i, k \in \{1, \ldots, n\}$ bilden zusammengefasst eine quadratische Matrix der Größe $n$:

**7.3 Definition** • *Für $m, n \in \mathbb{N}$ ist eine $m \times n$–***Matrix** $A$ *mit* **Einträgen** $a_{ik}$ *aus einem Ring $R$ ein "rechteckiges Schema"*

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

*von Elementen $a_{ik} \in R$. Für Puristen: $A$ ist eine Abbildung*

$$A : \{1, \ldots, m\} \times \{1, \ldots, n\} \to R \quad , \quad (i, k) \mapsto (A)_{ik} := a_{ik}.$$

- *Für $i \in \{1, \ldots, m\}$ ist das $n$–Tupel $(a_{i1}, \ldots, a_{i,n})$ die $i$–***te Zeile von** $A$,
- *für $k \in \{1, \ldots, n\}$ das $m$–Tupel $(a_{1k}, \ldots, a_{mk})$ die $k$–***te Spalte von** $A$.
- *Für $m = n$ heißt $A$* **quadratische Matrix der Größe** $m$.
- *Die Menge der $m \times n$–Matrizen mit Einträgen aus $R$ bezeichnet man mit $\mathrm{Mat}(m \times n, R)$, für $m = n$ kurz mit $\mathrm{Mat}(n, R)$.*
- *Die* **Transponierte** $A^t$ *der Matrix $A \in \mathrm{Mat}(m \times n, R)$ ist die Matrix*

$$A^t \in \mathrm{Mat}(n \times m, R) \quad \textit{mit} \quad (A^t)_{i,k} := (A)_{k,i}.$$

- *Ist $R$ ein Ring mit $1$, dann bezeichnet*

$$\mathbb{1}_n \in \mathrm{Mat}(n, R) \quad , \quad (\mathbb{1}_n)_{i,k} := \begin{cases} 1 & , i = k \\ 0 & , i \neq k \end{cases}$$

*die* **Einheitsmatrix**.

Offensichtlich ist diese Menge isomorph zu $R^{m\times n}$:

$$A \mapsto (a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{m1}, \ldots, a_{mn})$$

ist ein solcher Isomorphismus.

Kehren wir zurück zu den Koordinatentransformationen. Diese ergeben eine Abbildung $\hat{M} : K^n \to K^n$,

$$\hat{M}((w_1, \ldots, w_n)) := (v_1, \ldots, v_n) = \left( \sum_{k=1}^{n} m_{1k} w_k, \ldots, \sum_{k=1}^{n} m_{nk} w_k \right),$$

und nach dem letzten Satz ist

$$\hat{M} = \Phi_B^{-1} \circ \Phi_C.$$

Um die in dieser Beziehung beteiligten Mengen und Abbildungen zu verdeutlichen, kann man das Diagramm

$$\begin{array}{ccc} K^n & \xrightarrow{\hat{M}} & K^n \\ \Phi_C \searrow & & \swarrow \Phi_B \\ & V & \end{array}$$

hinmalen. Man sagt dann, dass das Diagramm *kommutiert*, was heißt, zusammengesetzte Abbildungen mit gleichem Anfangs- und Endpunkt sind gleich.

**7.4 Beispiel** $V := \{\text{Polynome über } \mathbb{R} \text{ vom Grad} \leq 2\}$.
Basen $B := \{b_1, b_2, b_3\}$ mit $b_1(x) = 1, b_2(x) = x, b_3(x) = x^2$
und $C := \{c_1, c_2, c_3\}$ mit $c_1 = b_1$, $c_2 = b_2$ aber $c_3(x) = \frac{3x^2-1}{2}$ (dies sind die ersten drei sog. Legendre-Polynome).
$\Phi_B : K^3 \to V \quad , \quad \Phi_B(v_1, v_2, v_3) := v_1 + v_2 x + v_3 x^2$
$\Phi_C : K^3 \to V \quad , \quad \Phi_C(w_1, w_2, w_3) := \left(w_1 - \frac{w_3}{2}\right) + w_2 x + \frac{3}{2} w_3 x^2$
$\hat{M} : K^3 \to K^3 \qquad \hat{M}(w_1, w_2, w_3) = (w_1 - \frac{w_3}{2}, w_2, \frac{3}{2} w_3)$.

Hier ist also $M = \begin{pmatrix} 1 & 0 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 3/2 \end{pmatrix}$.

Wir können umgekehrt fragen, unter welchen Bedingungen eine $n \times n$–Matrix $M$ mit Einträgen aus $K$ einen Basiswechsel bewirkt, d.h., wann die Linearkombinationen

$$c_k := \sum_{i=1}^{n} m_{i,k} b_i \in V \quad , \quad (k = 1, \ldots, n)$$

der Vektoren $b_1, \ldots, b_n$ der Basis $B$ eine neue Basis $C = \{c_1, \ldots, c_n\}$ von $V$ ergibt.

Nach dem eingangs bewiesenen Satz können wir das nachprüfen, indem wir feststellen, wann die Koordinatenvektoren

$$\Phi_B^{-1}(c_1), \ldots, \Phi_B^{-1}(c_n) \in K^n$$

linear unabhängig sind.

Nun ist aber für $k \in 1, \ldots, n$

$$\begin{aligned} \Phi_B^{-1}(c_k) &= \Phi_B^{-1}\left(\sum_{i=1}^{n} m_{ik} b_i\right) = \sum_{i=1}^{n} m_{ik} \Phi_B^{-1}(b_i) \\ &= \sum_{i=1}^{n} m_{ik} c_i = (m_{1k}, \ldots, m_{nk}), \end{aligned}$$

also die $k$–te Spalte der Matrix $M$. Um die Verwirrung wahlweise zu verkleinern oder zu vergrößern, schreibt man dieses $n$–Tupel als *Spaltenvektor*:

$$\Phi_B^{-1}(c_k) = \begin{pmatrix} m_{1k} \\ \vdots \\ m_{nk} \end{pmatrix},$$

also in der Form, wie es auch in $M$ vorkommt. Wir müssen also die Frage beantworten, ob die Spalten der Transformationsmatrix $M$, aufgefasst als Vektoren im $K^n$, linear unabhängig sind.

Dazu schreiben wir die Matrix $M$ in der Form $M = (s_1, \ldots, s_n)$, wobei $s_k = \begin{pmatrix} m_{1k} \\ \vdots \\ m_{nk} \end{pmatrix}$ die $k$–te Spalte bezeichnet. Die Frage ist, ob die Gleichung

$$0 = \sum_{k=1}^{n} \lambda_k s_k$$

nur durch die Koeffizienten $\lambda_k = 0$ erfüllt ist.

Die Antwort auf diese Frage bleibt unverändert, wenn ich statt der Matrix $M$ die Matrix betrachte, die durch *elementare Spaltenumformungen*, d.h.

- Austausch zweier Spalten

- Multiplikation einer Spalte mit einem Faktor $k \neq 0$
- Addition des $l \in K$–fachen einer Spalte zu einer anderen Spalte

aus $M$ hervorgeht.

Ist nämlich $M' = (s'_1, \ldots, s'_n)$ die neue Matrix, dann ist in diesen Fällen

- für den Austausch der $i$–ten mit der $k > i$–ten Spalte

$$M' = (s_1, \ldots, s_k, \ldots, s_i, \ldots, s_n),$$

- Für die Multiplikation der $i$–ten Spalte

$$M' = (s_1, \ldots, ls_i, \ldots s_n),$$

- Addition des $l$–fachen der $k$–ten Spalte zur $i$–ten

$$M' = (s_1, \ldots, s_i + ls_k, \ldots s_n),$$

und eine nicht triviale Darstellung der $0$ geht in eine solche über.

Man definiert elementare Spaltenumformungen für nicht quadratische Matrizen in gleicher Weise. *Elementare Zeilenumformungen* werden dadurch definiert, dass man das Wort "Spalte" in der obigen Definition durch "Zeile" ersetzt.

Die Idee bei den elementaren Umformungen ist, $M$ so umzuformen, dass man der transformierten Matrix die lineare (Un-)Abhängigkeit direkt ansieht.

Allgemeiner gilt:

**7.5 Satz** *Die Maximalzahl linear unabhängiger Zeilenvektoren einer Matrix* $M \in \mathrm{Mat}(m \times n, k)$ *wird durch elementare Zeilenumformungen nicht geändert.*

**Beweis:** Diese Maximalzahl ist die Dimension der von den Zeilenvektoren $z_1, \ldots, z_m$ aufgespannten Unterraums $[\{z_1, \ldots, z_m\}] \subset K^n$. Diese ist offensichtlich invariant unter Unnummerierung der Zeilen und Multiplikation mit von Null verschiedenen Faktoren. Der Übergang zu $z'_1 := z_1, \ldots, z'_i = z'_i + lz_k, \ldots z'_m = z_m$ erniedrigt die Dimension nicht, denn die $z_i$ sind wegen $z_i = z'_i - lz'_k$ Linearkombinationen der $z'_i$. □

Das Ziel besteht nun darin, $M$ in Zeilenstufenform überzuführen, weil man dieser die Maximalzahl linear unabhängiger Zeilen direkt ansieht.

**7.6 Definition** $M \in \mathrm{Mat}(m \times n, K)$ *ist in* **Zeilenstufenform**, *wenn*

1. *für* $i > k$ *immer* $M_{ik} = 0$ *ist und*

2. *aus* $M_{i,1} = M_{i,2} = \ldots = M_{i,r} = 0$ *für einen Zeilenindex* $i \in \{1, \ldots, m-1\}$ *und einen Spaltenindex* $r \in \{1, \ldots, n-1\}$ *folgt:*

$$M_{i+1,1} = M_{i+1,2} = \ldots = M_{i+1,r+1} = 0.$$

Mit anderen Worten ist $M$ genau dann in Zeilenstufenform, wenn die Zahl der führenden Nullen der $(i+1)$-ten Zeile immer mindestens um eins größer ist als die der $i$-ten Zeile.

$M$ sieht dann etwa so aus:

$$M = \begin{pmatrix} \blacksquare & * & * & * & * & * & * & \ldots & * \\ 0 & 0 & \blacksquare & * & * & * & * & \ldots & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & \ldots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & * & \ldots & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ r\text{te Zeile} \\ \\ \\ \\ \end{matrix}.$$

Die ■ bezeichnen von Null verschiedene Matrixeinträge, die $*$ irgendwelche Zahlen.

Die Maximalzahl linear unabhängiger Zeilen entspricht der Zahl der *Pivotelemente* ■, denn einerseits sind die nur mit Nullen gefüllten Zeilen immer linear abhängig, andererseits existiert für $(\lambda_1, \ldots, \lambda_r) \neq (0, \ldots, 0)$ ein kleinster Index $l \in \{1, \ldots, r\}$ mit $\lambda_l \neq 0$, und an der Stelle des Pivotelements der $l$–ten Zeile tritt in $\sum_{i=1}^{r} \lambda_i z_i$ eine von Null verschiedene Zahl auf.

**7.7 Beispiel (Zeilenstufenform)** $\left(\begin{smallmatrix} 1 & 2 & 1 \\ 0 & 0 & 3 \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 1 & 0 & 1 \\ 0 & 0 & 3 \end{smallmatrix}\right)$ sind in Zeilenstufenform, $\left(\begin{smallmatrix} 0 & 0 & 3 \\ 1 & 2 & 1 \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 7 & 8 \end{smallmatrix}\right)$ nicht.

Die Überführung einer beliebigen Matrix $M$ in Zeilenstufenform geht folgendermaßen vor sich: Ist in der ersten Spalte ein Eintrag $\neq 0$, so kann man die entsprechende Zeile durch Vertauschung mit der ersten Zeile an die oberste Position bringen. Danach addiert man Vielfache der ersten Zeile zu den folgenden, sodass überall sonst in der ersten Spalte nur noch Nullen stehen. Man wendet darauf das Verfahren auf die Matrix an, die entsteht, wenn man die erste Zeile streicht etc.

Dieses Verfahren wird *Gaußverfahren* genannt.

**7.8 Beispiel (Gaußverfahren)** $M = \begin{pmatrix} 1 & 3 & -4 & 3 \\ 3 & 9 & -2 & -11 \\ 4 & 12 & -6 & -6 \\ 2 & 6 & 2 & -10 \end{pmatrix}$

$$\begin{array}{cccc|l} 1 & 3 & -4 & 3 & \\ 0 & 0 & 10 & -20 & z_2 - 3z_1 \\ 0 & 0 & 10 & -18 & z_3 - 4z_1 \\ 0 & 0 & 10 & -16 & z_3 - 2z_1 \\ \hline 1 & 3 & -4 & 3 & \\ 0 & 0 & 10 & -20 & \\ 0 & 0 & 0 & 2 & z_3 - z_2 \\ 0 & 0 & 0 & 4 & z_4 - z_2 \\ \hline 1 & 3 & -4 & 3 & \\ 0 & 0 & 10 & -20 & \\ 0 & 0 & 0 & 2 & \\ 0 & 0 & 0 & 0 & z_4 - 2z_3 \end{array}$$

**7.9 Definition** • *Die Maximalzahl linear unabhängiger Zeilen einer Matrix* $M \in \mathrm{Mat}(m \times n, K)$ *mit Koeffizienten aus einem Körper* $K$ *heißt der* **Zeilenrang von** $M$*.*
• *Die Maximalzahl linear unabhängiger Spalten von* $M$ *heißt der* **Spaltenrang von** $M$*.*

Im letzten Beispiel war der Zeilenrang von $M$ gleich $3$, die Zeilenvektoren waren also nicht alle linear unabhängig.

Es wird sich herausstellen, dass der *Zeilen*rang von $M$ gleich dem *Spalten*rang von $M$ ist. Man schreibt dann nur noch

$$\mathrm{rang}(M).$$

# 8 Lineare Abbildungen

sind die strukturerhaltenden Abbildungen von Vektorräumen. Die beiden vorkommenden Strukturen sind die der abelschen Gruppe der Vektoren und der Multiplikation von Vektoren mit Körperelementen. Entsprechend definiert man

**8.1 Definition** *Eine Abbildung* $f : V \to W$ *zwischen* $K$*–Vektorräumen* $V, W$ *heißt* **linear***, wenn für alle* $v, v_1, v_2 \in V$ *und* $c \in K$ *gilt:*

$$f(v_1 + v_2) = f(v_1) + f(v_2) \quad \textit{und} \quad f(cv) = cf(v).$$

**8.2 Beispiele (lineare Abbildungen)** 1. $f : V \to \{0\} \subset W$ heißt die *Nullabbildung* . Man findet also immer lineare Abbildungen zwischen Vektorräumen über dem gleichen Körper $K$.

2. Die Multiplikation mit einem Skalar $c \in K$: $f : V \to V$, $x \mapsto cx$.

3. Sind $S \subset T$, $S \neq \emptyset$ zwei Mengen und $V := \mathrm{Abb}(T, K)$, $W := \mathrm{Abb}(S, K)$, dann ist $f : V \to W$, $v \mapsto v|_S$ eine lineare Abbildung, die *Restriktion auf* $S$.

4. Ist $V := C^\infty(\mathbb{R}, \mathbb{R})$ der Vektorraum der unendlich oft differenzierbaren reellen Funktionen, dann ist $f : V \to V$, $v \mapsto \frac{d}{dx}v$ linear. Diese lineare Abbildung ist jedoch nicht injektiv, $\ker(f)$ besteht aus den konstanten Funktionen.

   Sie ist surjektiv, denn $w \in C^\infty(\mathbb{R}, \mathbb{R})$, $w(x) := \int_0^x v(y)\,dy$ ist Urbild von $v \in C^\infty(\mathbb{R}, \mathbb{R})$.

Ein besonders wichtiges Beispiel ist mit Matrizen verbunden. Für beliebige Ringe $R$ definieren wir:

**8.3 Definition** • *Die* **Summe** $A + B$ *zweier Matrizen* $A, B \in \mathrm{Mat}(m \times n, R)$ *ist die Matrix* $A + B \in \mathrm{Mat}(m \times n, R)$ *mit den Einträgen*

$$(A + B)_{ik} := (A)_{ik} + (B)_{ik}.$$

- *Das* $k \in R$**–fache** $k \cdot A$ *einer Matrix* $A \in \mathrm{Mat}(m \times n, R)$ *ist die Matrix* $k \cdot A \in \mathrm{Mat}(m \times n, R)$ *mit den Einträgen* $(kA)_{ik} := k \cdot (A)_{ik}$.

- *Ist* $A \in \mathrm{Mat}(m \times n, R)$ *und* $B \in \mathrm{Mat}(n \times r, R)$*, dann ist das* **Matrixprodukt** $A \cdot B$ *die Matrix* $A \cdot B \in \mathrm{Mat}(m \times r, R)$ *mit den Einträgen*

$$(AB)_{i,s} := \sum_{k=1}^{n} (A)_{ik}(B)_{ks}.$$

**8.4 Beispiel** $\left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 5 & 6 \\ 7 & 8 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 6 & 8 \\ 10 & 12 \end{smallmatrix}\right)$, $2 \cdot \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 & 4 \\ 6 & 8 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)\left(\begin{smallmatrix} 5 & 6 \\ 7 & 8 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 19 & 22 \\ 43 & 50 \end{smallmatrix}\right)$.

$$(a_1, \ldots, a_n)\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \left(\sum_{i=1}^{n} a_i b_i\right) \text{ , aber } \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}(a_1, \ldots, a_n) = \begin{pmatrix} b_1a_1 & \ldots & b_1a_n \\ \vdots & & \vdots \\ b_na_1 & \ldots & b_na_n \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix},$$

aber

$$(0,\dots,1,\dots,0)\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = (a_{k1},\dots,a_{kn}),$$

falls die $1$ an der $k$ten Stelle steht.

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1+\dots+a_{1n}x_n \\ \vdots \\ a_{m1}x_1+\dots+a_{mn}x_n \end{pmatrix}.$$

Mit diesen Definitionen wird

• $\mathrm{Mat}(n,R)$ mit Matrizenaddition und Matrizenmultiplikation zu einem Ring. Dieser ist i.A. nicht kommutativ. Z.B. ist

$$\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) \quad \text{aber} \quad \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right).$$

• für einen Körper $K$ $\quad \mathrm{Mat}(m\times n,K)$ mit Matrizenaddition und Skalarmultiplikation zu einem Vektorraum über $K$.

Wir haben schon festgestellt, dass dieser Vektorraum isomorph zum arithmetischen Vektorraum $K^{m\cdot n}$ ist. Die (lexikalisch geordnete) Familie von Matrizen

$$E_{i,j} := \begin{pmatrix} 0 & \vdots & 0 \\ \cdots & 1 & \cdots \\ 0 & \vdots & 0 \end{pmatrix} \in \mathrm{Mat}(m\times n,K) \qquad (1\le i\le m, 1\le j\le n), \quad (8.1)$$

bei denen die einzige Eins im Kreuzungspunkt der $i$-ten Zeile mit der $k$-ten Spalte steht, bildet eine kanonische Basis von $\mathrm{Mat}(m\times n,K)$; insbesondere ist

$$\dim\big(\mathrm{Mat}(m\times n,K)\big) = m\,n.$$

Fassen wir andererseits Vektoren $v=(v_1,\dots,v_n)$ im $K^n$ als Spaltenmatrizen $\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathrm{Mat}(n\times 1,K)$ auf, dann gilt:

**8.5 Satz** *Ist $A\in\mathrm{Mat}(m\times n,K)$ eine $m\times n$–Matrix mit Einträgen aus dem Körper $K$, dann ist die Abbildung $f: K^n\to K^m$, $v\mapsto A\cdot v$ linear.*

**Beweis:**

1. Für $v\in K^n$ und $k\in K$ gilt

$$f(k\cdot v) = A(k\cdot v) = k\cdot Av = k\cdot f(v),$$

denn für $i = 1, \ldots, m$ ist

$$(A(k \cdot v))_{i,1} = \sum_{l=1}^{n} (A)_{il}(k \cdot v_l) = k \cdot \sum_{l=1}^{n} (A)_{il} v_l = k(Av)_i.$$ [9]

2. Für $v, w \in K^n$ ist

$$f(v + w) = A(v + w) = Av + Aw = f(v) + f(w),$$

denn für $i = 1, \ldots, m$ ist

$$\begin{aligned}(A(v+w))_i &= \sum_{l=1}^{n} (A)_{il}(v+w)_l \\ &= \sum_{l=1}^{n} (A)_{il}(v_l + w_l) = \sum_{l=1}^{n} (A)_{il} v_l + \sum_{l=1}^{n} (A)_{il} w_l \\ &= (Av)_i + (Aw)_i.\end{aligned}$$

□

**8.6 Beispiel (Drehung)** $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, $K = \mathbb{R}$. $f(v) = f\left(\left(\begin{smallmatrix} v_1 \\ v_2 \end{smallmatrix}\right)\right) = \left(\begin{smallmatrix} -v_2 \\ v_1 \end{smallmatrix}\right)$.

$f$ entspricht geometrisch einer Drehung im Gegenuhrzeigersinn um $\pi/2 = 90°$, denn die "Länge" $\sqrt{v_1^2 + v_2^2}$ des Vektors $v$ ist gleich der "Länge" $\sqrt{w_1^2 + w_2^2} = \sqrt{(-v_2)^2 + v_1^2}$ des Vektors $w = f(v)$ (Pythagoras!), und eine doppelte Anwendung von $f$ spiegelt den Vektor am Nullpunkt (= Drehung um $\pi$): $f(f(v)) = A(Av) = (AA)v$, aber $AA = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, also $f(f(v)) = -v$.

Da eine lineare Abbildung $f : V \to W$ wegen der Regel $f(v_1+v_2) = f(v_1)+f(v_2)$ insbesondere ein Gruppenhomomorphismus der abelschen Gruppen $V$ und $W$ ist, gilt $f(0) = 0$ und $f(-v) = -f(v)$ $(v \in V)$.

Eine Abbildung $f : V \to W$ ist erst dann gegeben, wenn wir sie an jedem Punkt aus $V$ definiert haben. Es ist aber offensichtlich nicht jede solche Abbildung linear.

**8.7 Beispiele** 1. $V = W = \mathbb{R}$. Für $a, b \in \mathbb{R}$ sei $f(v) := av+b$. Diese Abbildung ist für $b \neq 0$ nicht linear, denn $f(0) = b \neq 0$.

[9] $Av$ und $A(k \cdot v)$ sind $m \times 1$-Matrizen, also Spaltenvektoren der Länge $m$. Manchmal schreibt man bei Matrizen aus $\mathrm{Mat}(m \times 1, R)$ und $\mathrm{Mat}(1 \times n, R)$ den Index 1 nicht.

2. $V = W = \mathbb{R}^2$.

$$f(v) := \begin{cases} 0 & , \quad v = 0 \\ \frac{\sqrt{v_1^2+v_2^2}}{\max(|v_1|,|v_2|)} v & , \quad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \neq 0 \end{cases} .$$

Zwar gilt $f(kv) = k(v)$ für alle $k \in \mathbb{R}$ und $v \in V$, denn für $k = 0$ bzw. $v = 0$ ist das offensichtlich und für $k \neq 0$, $v \neq 0$ ist

$$\begin{aligned} f(kv) &= \frac{kv\sqrt{(kv_1)^2 + (kv_2)^2}}{\max(|kv_1|, |kv_2|)} = kv \cdot \frac{|k|\sqrt{v_1^2 + v_2^2}}{|k|\max(|v_1|, |v_2|)} \\ &= kv\frac{\sqrt{v_1^2 + v_2^2}}{\max(|v_1|, |v_2|)} = kf(v). \end{aligned}$$

Aber im Allgemeinen ist $f(v_1 + v_2) \neq f(v_1) + f(v_2)$:

$$\begin{aligned} f \quad & (\tbinom{1}{0}) = \tbinom{1}{0} \quad , \quad f(\tbinom{0}{1}) = \tbinom{0}{1} \quad , \\ f \quad & (\tbinom{1}{0} + \tbinom{0}{1}) = f(\tbinom{1}{1}) = \sqrt{2}\tbinom{1}{1} \neq f(\tbinom{1}{0}) + f(\tbinom{0}{1}) . \end{aligned}$$

3. Die komponentenweise Komplex-Konjugation

$$f : \mathbb{C}^n \to \mathbb{C}^n \quad , \quad f((v_1, \ldots, v_n)) := (\overline{v_1}, \ldots, \overline{v_n})$$

auf dem $n$-dimensionalen arithmetischen Vektorraum $\mathbb{C}^n$ ist nicht linear, denn für gilt zwar

$$f(v + w) = \overline{v + w} = \overline{v} + \overline{w} = f(v) + f(w) \qquad (v, w \in \mathbb{C}^n),$$

aber z.B. für $k := \sqrt{-1} \in \mathbb{C}$ ist

$$f(k \cdot v) = \overline{k}f(v) = -kf(v),$$

was für $v \neq 0$ ungleich $kf(v)$ ist.

Wenn wir eine *lineare* Abbildung $f : V \to W$ definieren wollen, dann reicht es aus, die Bilder einer Basis von $V$ anzugeben.

**8.8 Satz** *Es sei $B \subset V$ eine Basis des $K$–Vektorraums $V$ und $\tilde{\varphi} : B \to W$ eine Abbildung in einen $K$–Vektorraum $W$.*

*Dann existiert genau eine lineare Abbildung $\varphi : V \to W$ mit $\varphi|_B = \tilde{\varphi}$.*

**Beweis:** • *Eindeutigkeit:* Jeder Vektor $v \in V$ besitzt eine eindeutige Basisdarstellung

$$v = \sum_{i=1}^{n} \lambda_i b_i \quad \text{mit} \quad b_1, \ldots, b_n \in B \quad \text{(voneinander verschieden)} \quad .$$

Wegen der Linearität von $\varphi$ muss dann

$$\varphi(v) \;=\; \varphi\left(\sum_{i=1}^{n} \lambda_i b_i\right) = \sum_{i=1}^{n} \varphi(\lambda_i b_i) = \sum_{i=1}^{n} \lambda_i \varphi(b_i) = \sum_{i=1}^{n} \lambda_i \tilde{\varphi}(b_i)$$

gelten. Der Bildvektor $\varphi(v)$ ist also schon durch die Koordinate von $v$ und die Bilder $\tilde{\varphi}(b_i)$ der Basisvektoren bestimmt.

• *Existenz* ist durch die obige Definition gesichert und

• für $w = \sum_{i=1}^{n} \mu_i b_i$, $k \in K$ gilt

$$\varphi(v + w) = \sum_{i=1}^{n} (\lambda_i + \mu_i)\tilde{\varphi}(b_i) = \varphi(v) + \varphi(v)$$

und

$$\varphi(kv) = \sum_{i=1}^{n} k\lambda_i \tilde{\varphi}(b_i) = k\varphi(v).$$

$\varphi$ ist also *linear*. □

**8.9 Satz** *Ist $\varphi : V \to W$ linear, dann*

- *ist für $M \subset V$ $\varphi([M]) = [\varphi(M)]$,*
- *ist mit $U \subset V$ auch $\varphi(U) \subset W$ ein Unterraum und $\dim(\varphi(U)) \leq \dim(U)$.*

**Beweis:** • Wegen $[\emptyset] = \{0\}$ ist der Fall $M = \emptyset$ offensichtlich.
• Ist $M \neq \emptyset$, dann ist das Bild

$$\varphi\left(\sum_{i=1}^{n} \lambda_i m_i\right) = \sum_{i=1}^{n} \lambda_i \varphi(m_i)$$

von einer Linearkombination der $m_i \in M$ eine Linearkombination von $\varphi(M)$. Liest man diese Gleichung von rechts nach links, dann sieht man, dass man

umgekehrt jede Linearkombination von $\varphi(M)$ so darstellen kann. Wir wissen aber, dass $[N]$ für $N \neq \emptyset$ genau aus allen Linearkombinationen von $N$ besteht, also $\varphi([M]) = [\varphi(M)]$.
• Um die zweite Aussage zu beweisen, bemerken wir, dass für den Unterraum $U$ gilt, dass $[U] = U$, also auch $\varphi(U) = \varphi([U]) = [\varphi(U)]$ ein Unterraum ist (denn aus $[X] = X$ folgt umgekehrt, dass $X \subset W$ ein Unterraum ist).
• Ist $B \subset U$ eine Basis von $U$, dann enthält wegen $\varphi(U) = \varphi([B]) = [\varphi(B)]$ die Menge $\varphi(B)$ eine Basis von $\varphi(U)$. Also $\dim(\varphi(U)) \leq |B| = \dim(U)$. □

Insbesondere ist das Bild $\varphi(U) \subset V$ ein Unterraum und man definiert

**8.10 Definition** *Das* **Bild** *einer linearen Abbildung* $\varphi : U \to V$ *ist*

$$\mathrm{im}(\varphi) := \varphi(U) \subset V,$$

*ihr* **Rang**

$$\mathrm{rang}(\varphi) := \dim\big(\mathrm{im}(\varphi)\big).$$

Wir haben schon für Matrizen $A \in \mathrm{Mat}(m \times n, K)$ den Rangbegriff definiert. Es war der Spaltenrang (Zeilenrang) die Maximalzahl linear unabhängiger Spaltenvektoren (Zeilenvektoren) von $A$. Betrachtet man die der Matrix $A$ zugeordnete lineare Abbildung $f : K^n \to K^m$, $x \mapsto Ax$, dann erkennt man, dass die Basisvektoren $e_1, \ldots, e_n$ von $K^n$, die ja als Spaltenvektoren geschrieben die Form

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{–te Stelle}$$

haben, auf die Vektoren

$$f(e_i) = \begin{pmatrix} \sum_{l=1}^n (A)_{1l}(e_i)_l \\ \vdots \\ \sum_{l=1}^n (A)_{ml}(e_i)_l \end{pmatrix} = \begin{pmatrix} (A)_{1i} \\ \vdots \\ (A)_{mi} \end{pmatrix}$$

abgebildet werden.

Es gilt also: Die Spaltenvektoren von $A$ sind die Bilder der Vektoren der kanonischen Basis.

Das Bild $f(K^n) \subset K^m$ besitzt also eine Basis, die von einer Maximalzahl linear unabhängiger Spaltenvektoren gebildet wird.

In diesem Sinn stimmen also die Begriffe des (Spalten-)Rangs einer Matrix und der von ihr vermittelten linearen Abbildung überein:

$$\operatorname{rang}(f) = \operatorname{rang}(A).$$

Neben $\varphi(V)$ ist auch $\varphi^{-1}(0) = \ker(\varphi)$ ein UnterraumAllgemein gilt

**8.11 Satz** *Ist $Y \subset W$ ein Unterraum, dann ist für eine lineare Abbildung $\varphi : V \to W$ auch $\varphi^{-1}(Y) \subset V$ ein Unterraum.*

**Beweis:** Mit $x_1, x_2 \in X := \varphi^{-1}(Y)$ und $k \in K$ folgt

$$\varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2) \in Y \quad \text{und} \quad \varphi(kx_1) = k\varphi(x_1) \in Y,$$

denn $\varphi(x_i) \in Y$, und $Y$ ist ein Unterraum. □

Insbesondere ist $\operatorname{Kern}(\varphi)$ ein Unterraum von $V$, denn $\{0\}$ ist ein Unterraum von $W$.

**8.12 Definition** *Der* **Defekt** *einer linearen Abbildung $\varphi : V \to W$ ist*

$$\operatorname{def}(\varphi) := \dim\big(\ker(\varphi)\big).$$

**8.13 Satz** *Ist $\dim(V) < \infty$, dann ist für jede lineare Abbildung $\varphi : V \to W$*

$$\operatorname{def}(\varphi) + \operatorname{rang}(\varphi) = \dim(V).$$

**Beweis:** Nach dem Basisergänzungssatz kann man eine Basis $B' = \{b'_1, \ldots, b'_n\}$ von $\ker(\varphi)$ zu einer Basis $B = B' \cup B''$, $B'' = \{b''_1, \ldots, b''_m\}$ von $V$ ergänzen. Es gilt $\varphi(B') = \{0\}$ und $[\varphi(B'')] = [\varphi(B)] = \varphi([B]) = \varphi(V)$ nach Satz 8.9.

Als nächstes zeigen wir, dass die Vektoren $\varphi(b''_1), \ldots, \varphi(b''_m) \in W$ linear unabhängig sind, also

$$0 = \sum_{i=1}^{M} \lambda_i \varphi(b''_i) = \varphi\left(\sum_{i=1}^{m} \lambda_i b''_i\right)$$

nur für $\lambda_1 = \ldots \lambda_m = 0$ gilt. $\sum_{i\in I}^{m} \lambda_i b''_i \in \ker(\varphi)$, es gilt also

$$\sum_{i=1}^{m} \lambda_i b''_i = \sum_{l=1}^{n} \mu_l b_l,$$

denn $B'$ ist eine Basis von $\ker(\varphi)$. Da aber $B' \dot\cup B'' = B$ eine Basis von $V$ ist, müssen $\lambda_i = \mu_l = 0$ gelten.

Da die $\varphi(b''_1), \ldots, \varphi(b''_m) \in W$ linear unabhängig sind, bilden sie wegen $[\varphi(B'')] = \varphi(V)$ eine Basis von $\varphi(V)$, es ist also

$$\dim(\mathrm{Im}(\varphi)) = |B''| = m.$$

Damit ergibt sich mit $\mathrm{def}(\varphi) = \dim(\ker(\varphi)) = |B'| = n$ der Satz

$$\dim(V) = |B| = |B'| + |B''| = \mathrm{def}(\varphi) + \mathrm{rang}(\varphi).$$

□

Wir haben im Fall arithmetischer Vektorräume gesehen, dass Matrizen lineare Abbildungen definieren.

Andererseits wissen wir, dass wir einen $K$–Vektorraum $V$ mit $m := \dim(V) < \infty$ und einer (geordneten) Basis $B = (b_1, \ldots, b_m)$ mittels $\varphi_B : K^m \mapsto V$, $(x_1, \ldots, x_m) \mapsto \sum_{i=1}^m x_i b_i$ mit dem arithmetischen $K$–Vektorraum gleicher Dimension in Verbindung bringen können. $\Phi_B$ ist linear und umkehrbar. Ist nun $f : V \to W$ eine lineare Abbildung in einen $K$–Vektorraum $W$ der Dimension $n := \dim(W) < \infty$ mit Basis $C = (c_1, \ldots, c_n)$, dann ist $\psi := \Phi_C^{-1} \circ f \circ \Phi_B : K^m \to K^n$ eine Abbildung zwischen den arithmetischen Vektorräumen, die als Verkettung linearer Abbildungen linear ist.

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle \Phi_B}\uparrow & & \uparrow{\scriptstyle \Phi_C} \\ K^m & \xrightarrow[\psi]{} & K^n \ . \end{array}$$

Eine solche lineare Abbildung $\psi : K^m \to K^n$ lässt sich, wie wir wissen, durch Angabe der Bilder $\psi(e_l)$, $l = 1, \ldots, m$, der kanonischen Basisvektoren $e_l$ des $K^m$ definieren.

Wir bezeichnen mit $(A)_{il} := (\psi(e_l))_i$, $i = 1, \ldots, n$ deren $i$–te Komponente. Dann ist für einen beliebigen Vektor $v = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \in K^m$ wegen $v = \sum_{l=1}^m v_l e_l$ die $i$–te Komponente seines Bildes durch

$$\begin{aligned} (\psi(v))_i &= \left( \sum_{l=1}^m \psi(e_l) v_l \right)_i = \sum_{l=1}^n (\psi(e_l))_i v_l \\ &= \sum_{l=1}^n (A)_{il} v_l = (Av)_i \end{aligned}$$

gegeben, also $\psi(v) = Av$.

Wir können also in natürlicher Weise *alle* linearen Abbildungen $\psi : K^m \to K^n$ durch Matrixmultiplikation darstellen.

**8.14 Definition** $A \in \mathrm{Mat}(n \times m, K)$ *heißt die* **Matrixdarstellung** *von* $f : V \to W$ *bezüglich der Basen* $B$ *von* $V$ *und* $C$ *von* $W$*.*

**8.15 Notation** $M_C^B(f)$ bezeichnet die Matrixdarstellung von $f : V \to W$ bez. der Basen $B$ von $V$ und $C$ von $W$. Im Fall eines Endomorphismus, also $V = W$, und nur eine Basis $B = C$ schreiben wir kurz $M_B(f)$ statt $M_B^B(f)$.

Insbesondere ist für Basen $B = (b_1, \ldots, b_n)$ und $C = (c_1, \ldots, c_n)$ von $V$

$$T := M_C^B(\mathrm{Id}_V)$$

die sog. *Transformationsmatrix* für den Basiswechsel von $B$ auf $C$. Ist also $v = \sum_{i=1}^n x_i b_i$, dann ist $v = \sum_{i=1}^n y_i c_i$ mit

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = T \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Damit ist $b_i = \sum_{j=1}^n (T)_{ij} c_j$ bzw. $c_i = \sum_{j=1}^n (T^{-1})_{ij} b_j$.

Die darstellenden Matrizen verändern sich bei Basiswechsel damit gemäß

$$M_C(\varphi) = T M_B(\varphi) T^{-1}. \tag{8.2}$$

**8.16 Beispiel** $V = \{p \in \mathbb{R}[x] \mid p \text{ ist Polynom vom Grad} \leq n\}$

1. $f : V \to V \quad , \quad p \mapsto p'$.

   Bezüglich der Basis $B = \{e_0, \ldots, e_n\}$ von $V$, $e_i(x) = x^i$, ist die darstellende Matrix von $f$ von der Form $M_B(f) = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 2 & 0 & \ldots & 0 \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & n-1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & n \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}$, denn

$$f(e_0) = 0 \quad \text{und} \quad f(e_i) = i e_{i-1} \quad (i \in \mathbb{N}).$$

2. Ist $f$ auf dem gleichen Vektorraum für $n = 2$ von der Form

$$f(p)(x) := p(x+1) \qquad (x \in \mathbb{R}),$$

dann ist die darstellende Matrix $M_B(f) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$, denn für alle $x \in \mathbb{R}$ gilt

$$\begin{array}{rcll} f(e_0)(x) & = & 1 & = e_0(x) \\ f(e_1)(x) & = & x+1 & = e_0(x) + e_1(x) \\ f(e_2)(x) & = & (x+1)^2 & = e_0(x) + 2e_1(x) + e_2(x). \end{array}$$

Es gibt einfache Kriterien dafür, ob eine lineare Abbildung injektiv bzw. surjektiv ist:

**8.17 Satz** *Für eine lineare Abbildung $\varphi : V \to W$ sind folgende Aussagen gleichwertig:*

1. *$\varphi$ ist surjektiv.*
2. *$[\varphi(B)] = W$ für eine Basis $B$ von $V$.*

*Falls $\dim(W) < \infty$, ist außerdem gleichwertig:*

3. *$\operatorname{rang}(\varphi) = \dim(W)$.*

**Beweis:** $1. \iff 2.:\ [\varphi(B)] = \varphi([B]) = \varphi(V)$
$1. \implies 3.:\ \operatorname{rang}(\varphi) = \dim(\varphi(V)) = \dim(W)$ für $\varphi(V) = W$. Das gilt immer.
$3. \implies 1.$ für $\dim(W) < \infty$: Für diesen Fall folgt nach Satz 6.17 aus $\dim(U) = \dim(W)$ für einen Unterraum $U$ schon $U = W$. □

**8.18 Satz** *Für eine lineare Abbildung $\varphi : V \to W$ sind folgende Aussagen paarweise gleichwertig:*

1. *$\varphi$ ist injektiv.*
2. *$\ker(\varphi) = \{0\}$.*
3. *$\operatorname{def}(\varphi) = 0$.*
4. *Für eine Basis $B$ ist $\varphi\,|_B$ injektiv und $\varphi(B)$ linear unabhängig.*

*Falls* $\dim(V) < \infty$*, ist außerdem gleichwertig:*

*5.* $\operatorname{rang}(\varphi) = \dim(V)$.

**Beweis:** • $1. \implies 2.:\ \ker(\varphi) = \varphi^{-1}(\{0\})$. Wegen der Linearität gilt immer $0 \in \ker(\varphi)$, wegen der Injektivität Gleichheit.

• $2. \iff 3.:\ 0 = \operatorname{def}(\varphi) = \dim(\ker(\varphi)) = \dim(\{0\}) = 0.$

• $2. \implies 4.$: Wäre $\varphi\,|_B$ nicht injektiv, also $\varphi(b_1) = \varphi(b_2)$ für $b_1 \neq b_2 \in B$, dann wäre $b_1 - b_2 \in \ker(\varphi)\backslash\{0\}$.
Wäre $\varphi(B)$ linear abhängig, dann gäbe es eine nicht triviale Linearkombination

$$W \ni 0 = \sum_{i=1}^{n} \lambda_i \varphi(b_i) = \varphi\left(\sum_{i=1}^{n} \lambda_i b_i\right).$$

Wegen der linearen Unabhängigkeit der Basis wäre dann aber

$$0 \neq \sum_{i=1}^{n} \lambda_i b_i \in \ker(\varphi).$$

• $4. \implies 1.$: Wäre $\varphi(v_1) = \varphi(v_2)$ für $v_1 \neq v_2 \in V$, dann auch $\varphi(v) = 0$ für $v := v_1 - v_2 \neq 0$. Mit $v = \sum_{i=1}^{n} \lambda_i b_i$ ist $0 = \varphi(v) = \sum_{i=1}^{n} \lambda_i \varphi(b_i)$, also wäre für $\varphi|_B$ injektiv das Bild $\varphi(B)$ der Basis linear abhängig.

• $3. \iff 5.:\ \operatorname{def}(\varphi) = 0 \iff \operatorname{rang}(\varphi) \equiv \dim(V) - \operatorname{def}(\varphi) = \dim(V).$ □

# 9 Lineare Gleichungssysteme und Invertierung von Matrizen

**9.1 Definition** *Ein Gleichungssystem der Form*

$$\begin{array}{c} \sum_{k=1}^{n} a_{1k} x_k = b_1 \\ \vdots \\ \sum_{k=1}^{n} a_{mk} x_k = b_m \end{array}$$

*mit Koeffizienten* $a_{ik}, b_i$ *aus einem Körper* $K$ *nennt man ein System von* $m$ **linearen Gleichungen** *mit* $n$ **Unbekannten** $x_k$*. Für* $b_1 = \ldots = b_m = 0$ *heißt das Gleichungssystem* **homogen**, *sonst* **inhomogen**.

Offensichtlich lässt sich ein solches Gleichungssystem unter der Verwendung der so genannten Koeffizientenmatrix $A \in \mathrm{Mat}(m \times n, K)$, $(A)_{i,k} := a_{i,k}$ und des Vektors $b := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m$ kurz in der Form

$$Ax = b$$

schreiben. Die Unbekannten wurden dabei zu dem Vektor $x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ zusammengefasst.

**9.2 Definition** *Die* **Lösungsmenge** $L$ *des linearen Gleichungssystems* $Ax = b$ *ist*

$$L := \{x \in K^n \mid Ax = b\}.$$

Die Matrix $A$ definiert eine lineare Abbildung

$$\varphi : K^n \to K^m \quad , \quad \varphi(x) := Ax.$$

Offensichtlich ist für das homogene Gleichungssystem $Ax = 0$ die Lösungsmenge

$$L = \ker(\varphi),$$

sodass für das homogene Gleichungssystem immer die Lösung $x = 0$ existiert und die Lösungsmenge $L \subset K^n$ einen Unterraum der Dimension

$$\mathrm{def}(\varphi) = \dim(K^n) - \mathrm{rang}(\varphi) = n - \mathrm{rang}(A)$$

bildet.

Für ein inhomogenes Gleichungssystem $Ax = b$ existiert genau dann eine Lösung, wenn $b \in \varphi(K^n)$ gilt. Das muss natürlich nicht der Fall sein. Andernfalls ist $L = \emptyset$.

Wie findet man nun heraus, ob ein inhomogenes Gleichungssystem überhaupt eine Lösung besitzt? Dazu schreiben wir $A$ in der Form $A = (s_1, \ldots, s_n)$ mit den Spaltenvektoren $s_k := \begin{pmatrix} a_{1,k} \\ \vdots \\ a_{m,k} \end{pmatrix} \in K^m$. Damit ist $Ax = \sum_{k=1}^{n} s_k x_k$ eine Linearkombination der Spaltenvektoren, und es muss sich $b$ als eine solche Linearkombination darstellen lassen.

**9.3 Definition** *Die* **erweiterte Koeffizientenmatrix** $A_{\mathrm{erw}} \in \mathrm{Mat}(m \times (n+1), K)$ *des linearen Gleichungssystems* $Ax = b$ *ist* $A_{\mathrm{erw}} := (s_1, \ldots, s_n, b)$.

Damit ergibt sich:

**9.4 Satz** *Das lineare Gleichungssystem $Ax = b$ besitzt genau dann eine Lösung, wenn* $\operatorname{rang}(A_{\text{erw}}) = \operatorname{rang}(A)$.

**Beweis:** Es gilt immer $\operatorname{rang}(A_{\text{erw}}) \geq \operatorname{rang}(A)$. Bei Gleichheit lässt sich die letzte Spalte $b$ von $A_{\text{erw}}$ als Linearkombination eines maximalen Systems linear unabhängiger Spaltenvektoren $s_k$ von $A$ darstellen (siehe Satz 6.5). Die (mit Nullen ergänzten) Koeffizienten $x_k$ der Spalten $s_k$ ergeben dann eine Lösung. Entsprechend umgekehrt. □

Für das inhomogene lineare Gleichungssystem ist die Lösungsmenge $L \subset K^n$ kein Unterraum mehr, denn wegen $A \cdot 0 = 0 \neq b$ ist $0 \notin L$. Es gilt aber:

**9.5 Satz** *Ist $L_0$ die Lösungsmenge des homogenen Gleichungssystems $Ax = 0$ und $x^* \in K^n$ eine Lösung des inhomogenen Gleichungssystems $Ax = b$, dann ist dessen Lösungsmenge $L_b$ von der Form $L_b = \{x + x^* \mid x \in L_0\}$.*

**Beweis:** • Ist $y = x + x^*$, $x \in L_0$, dann ist $Ay = Ax + Ax^* = 0 + b = b$.
• Ist $y \in L_b$, dann ist $x := y - x^* \in L_0$, denn $Ax = Ay - Ax^* = b - b = 0$. □

Für $b \neq 0$ ist die Lösungsmenge $L_b \subset K^n$ wie schon bemerkt kein Unterraum mehr, aber immerhin ein so genannter affiner Unterraum.[10]

**9.6 Definition** *Eine Teilmenge $L \subset V$ eines $K$–Vektorraumes $V$ heißt* **affiner Unterraum** *von $V$, falls $L = \emptyset$, oder falls es ein $l \in V$ und einen Unterraum $U \subset V$ gibt, sodass*

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

*Falls $L = \emptyset$ setzen wir* $\dim(L) := -1$, *sonst* $\dim(L) := \dim(U)$.

In unserem Fall ist $V = K^n$, $U = L_0$ und $l = x^*$, falls eine Lösung $x^*$ des inhomogenen Gleichungsystems existiert.

Jeder Unterraum $U \subset V$ ist auch ein affiner Unterraum, denn $U = 0 + U$.

Beispiele affiner Unterräume des $\mathbb{R}^n$ sind Punkte, Geraden und Ebenen, die den Ursprung nicht enthalten müssen. Die Dimensionen dieser affinen Unterräume sind $0, 1$ bzw. $2$.

[10] "This is in accordance with the principle that in mathematics a *red herring* does not have to be either red or a herring" (M. Hirsch; Im Englischen bedeutet red herring ein Ablenkungsmanöver).

In der Darstellung $L = l + U$ eines nicht leeren affinen Unterraumes $L$ ist der Unterraum $U$ durch $L$ eindeutig bestimmt, $l$ dagegen im Allgemeinen nicht, denn

$$U = \{l_1 - l_2 \mid l_1, l_2 \in L\},$$

während für jedes $l' \in L$ auch $L = l' + U$ gilt. Andererseits *muss* $l \in L$ sein.

Zurück zum linearen Gleichungssystem. Wie berechnet man nun seine Lösungsmenge?

Dazu schreiben wir die erweiterte Koeffizientenmatrix $A_{\text{erw}}$ auf, wobei wir zweckmäßigerweise $b$ durch einen Strich abtrennen:

$$A_{\text{erw}} = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right)$$

Die Lösungsmenge $L$ ändert sich nun nicht, wenn wir die elementaren *Zeilen*umformungen auf $A_{\text{erw}}$ anwenden. Denn

1. Bei Vertauschen von Zeilen vertauscht man nur die entsprechenden Gleichungen.
2. Für $c \neq 0$ ist $\sum_{k=1}^{n} a_{i,k}x_k = b_i$ genau dann, wenn $\sum_{k=1}^{n}(ca_{i,k})x_k = cb_i$.
   Man kann also mit von Null verschiedenen Konstanten $c$ multiplizieren.
3. Auch die Addition verschiedener Zeilen lässt sich rückgängig machen und produziert damit keine neuen Lösungen.

Man geht jetzt so vor:

1. Man bringt $A$ durch Anwendung von elementaren Zeilenumformungen auf $A_{\text{erw}}$ auf Zeilenstufenform und erhält:

$$A'_{\text{erw}} = \left(\begin{array}{ccccc|c} a'_{1,1} & \dots & a'_{1,q} & \dots & a'_{1,n} & b'_1 \\ 0 & \ddots & \vdots & & \vdots & \vdots \\ \vdots & \ddots & a'_{p,q} & \dots & a'_{p,n} & b'_p \\ 0 & \dots & 0 & \dots & 0 & b'_{p+1} \\ \vdots & & \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 0 & b'_m \end{array}\right).$$

   Dieser Schritt heißt *Vorwärtselimination*.

2. Ist ein Koeffizient $b_l$, $p < l \leq m$, ungleich Null, dann besitzt das Gleichungssystem keine Lösung, denn schon die der $l$ten Zeile von $A'_{\text{erw}}$ entsprechende Gleichung $\sum_{k=1}^{n} 0 \cdot x_k = b_l \neq 0$ besitzt keine Lösung $x$.

3. Sonst erzeugt man durch elementare Zeilenumformungen auch *über* den Pivotelementen Nullen und erhält

$$A''_{\text{erw}} = \left(\begin{array}{ccccccc|c} a''_{1,1} & 0 & \dots & 0 & a''_{1,q+1} & \dots & a''_{1,n} & b''_1 \\ 0 & \ddots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & a''_{p,q} & a''_{p,q+1} & \dots & a''_{p,n} & b''_p \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \end{array}\right).$$

   Dieser Schritt heißt *Rückwärtssubstitution*.

4. Gibt es $p$ Pivotelemente, dann findet man die Lösungsmenge folgendermaßen:

   Man nennt $x_k$ *freie Variable*, falls in der $k$–ten Spalte von $A''$ kein Pivotelement steht. Es gibt also neben den $x_k$ mit $q < k \leq n$ weitere $q - p$ freie Variable.

   Die Werte der freien Variablen kann man frei aus dem $K$ wählen. Die Werte der restlichen $p$ *Pivot-Variablen* $x_k$ erhält man in Abhängigkeit von den Koeffizienten $b''_l$ und den freien Variablen, indem man nach ihnen auflöst.

**9.7 Beispiel (Lösungsmenge eines linearen Gleichungssystems)**

$$A'_{\text{erw}} = \left(\begin{array}{ccccc|c} 1 & -2 & 3 & 4 & 2 & 2 \\ 0 & 0 & 2 & 1 & -4 & 3 \\ 0 & 0 & 0 & -1 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Rückwärtssubstitution:

$$\begin{array}{rrrrr|c|l}
1 & -2 & 3 & 0 & 14 & 6 & z_1 + 4z_5 \\
0 & 0 & 2 & 0 & -1 & 4 & z_2 + z_3 \\
0 & 0 & 0 & -1 & 3 & 1 & \\
0 & 0 & 0 & 0 & 0 & 0 & \\
\hline
1 & -2 & 0 & 0 & 15\frac{1}{2} & 0 & z_1 - \frac{3}{2}z_2 \\
0 & 0 & 2 & 0 & -1 & 4 & \\
0 & 0 & 0 & -1 & 3 & 1 & \\
0 & 0 & 0 & 0 & 0 & 0 & \\
\hline
 & x_2 & & & x_5 & & 
\end{array}$$

Es sind also die Variablen $x_2$ und $x_5$ frei, und

$$A''_{\text{erw}} = \left(\begin{array}{rrrrr|c} 1 & -2 & 0 & 0 & 15\frac{1}{2} & 0 \\ 0 & 0 & 2 & 0 & -1 & 4 \\ 0 & 0 & 0 & -1 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Die Pivot-Variablen besitzen damit die Werte

$$x_1 = 2x_2 - \frac{31}{2}x_5 \quad , \quad x_3 = \tfrac{1}{2}(x_5 + 4) \quad \text{und} \quad x_4 = -(-3x_5 + 1),$$

sodaß sich die Lösungsmenge

$$L = \left\{ \left( \begin{array}{c} 2x_2 - \frac{31}{2}x_5 \\ x_2 \\ \frac{1}{2}x_5 + 2 \\ 3x_5 - 1 \\ x_5 \end{array} \right) \middle| \; x_2, x_5 \in K \right\}$$

ergibt.

Manchmal möchte man mehrere lineare Gleichungssysteme

$$Ax = b_1, \ldots, Ax = b_r$$

mit $b_1, \ldots, b_r \in K^m$ lösen (worunter natürlich nicht die Bestimmung *gemeinsamer* Lösungen zu verstehen ist!). Dann kann man das Verfahren der Vorwärtselimination und Rückwärtssubstitution auf die um $b_1, \ldots, b_r$ erweiterte Matrix $A$ anwenden. Der Rechenaufwand verringert sich entsprechend.

**9.8 Beispiel** 1. Gleichungssystem: $2x - 3y = 2\ ,\ \ 4x - 8y = 3$,
2. Gleichungssystem: $2x - 3y = -5$, $4x - 8y = 1$.

Also ist die Matrix $A := \left(\begin{smallmatrix} 2 & -3 \\ 4 & -8 \end{smallmatrix}\right)$, und die Vektoren $b_1 := \left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right)$, $b_2 := \left(\begin{smallmatrix} -5 \\ 1 \end{smallmatrix}\right)$.

$$\begin{array}{cc|cc} 2 & -3 & 2 & -5 \\ 4 & -8 & 3 & 1 \\ \hline 2 & -3 & 2 & -5 \\ 0 & -2 & -1 & 11 \\ \hline 2 & 0 & 7/2 & -43/2 \\ 0 & 1 & 1/2 & -11/2 \\ \hline 1 & 0 & 7/4 & -43/4 \\ 0 & 1 & 1/2 & -11/2 \end{array}$$

Also ist die Lösungsmenge der ersten Gleichung

$$L_1 \equiv \{x \mid Ax = b_1\} = \left\{ \begin{pmatrix} 7/4 \\ 1/2 \end{pmatrix} \right\}$$

und die der zweiten

$$L_2 \equiv \{x \mid Ax = b_2\} = \left\{ \begin{pmatrix} -43/4 \\ -11/2 \end{pmatrix} \right\}.$$

Betrachten wir den Spezialfall einer $n \times n$–Matrix $A$ und der Vektoren $b_1 := e_1, \ldots, b_n := e_n$, wobei die $e_i$ die kanonischen Basisvektoren der $\mathbb{R}^n$ sind.

Wir beginnen also mit

$$\left( \begin{array}{c|ccc} & 1 & \ldots & 0 \\ A & \vdots & & \vdots \\ & 0 & \ldots & 1 \end{array} \right),$$

rechts vom Strich steht, wenn man so will, die Einheitsmatrix $\mathbb{1}_n$. Diese hat den Rang $n$, der durch die auf beiden Seiten durchgeführten elementaren Zeilenumformungen nicht verändert wird.

Wenn also am Ende der Vorwärtselimination $A$ in Zeilenstufenform $A'$ übergeführt wurde, dann muss Rang $A$, also die Zahl der von Null verschiedenen Zeilen von $A'$, gleich $n$ sein, also alle Zeilen ungleich Null sein, damit man die Gleichungen $Ax = e_1, \ldots, Ax = e_n$ lösen kann. Eine Zeilenstufenmatrix der Größe $n \times n$ vom Rang $n$ ist aber eine *obere Dreiecksmatrix* (d.h. unterhalb

der Diagonale stehen nur Nullen), deren Diagonalelemente $(A)'_{i,i}$ alle ungleich Null sind. Dann können wir bei der Rückwärtssubstitution eine Matrix $A''$ mit den *gleichen* Diagonalelementen erreichen, die auch oberhalb der Diagonale nur Nullen aufweist. Durch zeilenweise Division durch die Diagonalelemente erreicht man die Einheitsmatrix.

Unser Schema ist also:

| $A$ | $\mathbb{1}_n$ |
|---|---|
| $A'$ | $*$ |
| $\mathbb{1}$ | $C$ |

Es ist nun mit $C = (c_1, \ldots, c_n)$ $Ac_i = e_i,\ i = 1, \ldots, n$.

Wenn wir diese $n$ Vektorgleichungen zusammenfassen, sehen wir, dass $AC = (e_1, \ldots, e_n) = \mathbb{1}_n$ ist.

Damit ist $A^{-1} := C$ die inverse Matrix zu $A$. Wir haben also ein Verfahren hergeleitet, mit dem die inverse Matrix $A^{-1}$ einer quadratischen Matrix $A$ berechnet werden kann, soweit diese überhaupt existiert.

**9.9 Beispiel (Bestimmung der inversen Matrix)** $A = \left(\begin{smallmatrix} 1 & 0 & -1 \\ 3 & 1 & -3 \\ 1 & 2 & -2 \end{smallmatrix}\right)$

| | | | | | | |
|---|---|---|---|---|---|---|
| $1$ | $0$ | $-1$ | $1$ | $0$ | $0$ | |
| $3$ | $1$ | $-3$ | $0$ | $1$ | $0$ | |
| $1$ | $2$ | $-2$ | $0$ | $0$ | $1$ | |
| $1$ | $0$ | $-1$ | $1$ | $0$ | $0$ | |
| $0$ | $1$ | $0$ | $-3$ | $1$ | $0$ | $z_2 - 3z_1$ |
| $0$ | $2$ | $-1$ | $-1$ | $0$ | $1$ | $z_3 - z_1$ |
| $1$ | $0$ | $-1$ | $1$ | $0$ | $0$ | |
| $0$ | $1$ | $0$ | $-3$ | $1$ | $0$ | |
| $0$ | $0$ | $-1$ | $5$ | $-2$ | $1$ | $z_3 - 2z_1$ |
| $1$ | $0$ | $0$ | $-4$ | $2$ | $-1$ | $z_1 - z_3$ |
| $0$ | $1$ | $0$ | $-3$ | $1$ | $0$ | |
| $0$ | $0$ | $1$ | $-5$ | $2$ | $-1$ | $-z_3$ |

Damit ist

$$A^{-1} = \begin{pmatrix} -4 & 2 & -1 \\ -3 & 1 & 0 \\ -5 & 2 & -1 \end{pmatrix}.$$

Offensichtlich ist $A^{-1}$ diejenige Matrix, die die zu $\varphi$ inverse lineare Abbildung $\varphi^{-1} : K^n \to K^n$ erzeugt.

# 10 Isomorphismen und Endomorphismen

**10.1 Definition** *Eine lineare Abbildung $\varphi : V \to W$ heißt auch* **Vektorraumhomomorphismus**. *Sie heißt*
• **Isomorphismus**, *wenn $\varphi$ bijektiv ist,*
• **Endomorphismus**, *wenn $W = V$ ist und*
• **Automorphismus**, *wenn $\varphi$ ein bijektiver Endomorphismus ist.*

**10.2 Satz** *Für eine lineare Abbildung $\varphi : V \to W$ sind folgende Aussagen gleichwertig:*

1. *$\varphi$ ist ein Isomorphismus.*
2. *Ist $B$ eine Basis von $V$, dann ist $\varphi|_B$ injektiv und $\varphi(B) \subset W$ eine Basis.*
3. *Besitzen $V$ und $W$ endliche Dimension, dann ist außerdem gleichwertig:* $\dim(V) = \operatorname{rang}(\varphi) = \dim(W)$.

**Beweis:** • Wir wissen aus Satz 8.18, dass $\varphi$ genau dann injektiv ist, wenn $\varphi|_B$ injektiv und $\varphi(B)$ linear unabhängig. Weiter wurde gezeigt, dass $\varphi$ genau dann surjektiv ist, wenn $[\varphi(B)] = W$. Zusammengenommen ergibt das die Äquivalenz von 1. und 2.
• Im endlichdimensionalen Fall war $\varphi$ außerdem nach Satz 8.18 genau dann injektiv, wenn $\operatorname{rang}(\varphi) = \dim(V)$ und nach Satz 8.17 genau dann surjektiv, wenn $\operatorname{rang}(\varphi) = \dim(W)$. Das ergibt die Äquivalenz der Aussagen 1. und 3. □

Das Kriterium 3. lässt sich wieder leicht rechnerisch verifizieren: Die Forderung 3. lässt sich überhaupt nur dann erfüllen, wenn $\dim(V) = \dim(W)$, also die darstellende Matrix $A = \operatorname{Mat}_C^B(\varphi)$ von $\varphi$ bez. einer Basis $B$ von $V$ und $C$ von $W$ quadratisch ist. Dann muss zusätzlich $\operatorname{rang}(A)$ maximal sein.

**10.3 Definition** *Zwei Vektorräume $V, W$ heißen* **isomorph**, *wenn ein Isomorphismus $\varphi : V \to W$ existiert.*

**10.4 Satz** *Zwei $K$–Vektorräume $V$ und $W$ sind genau dann isomorph, wenn* $\dim(V) = \dim(W)$.

**10.5 Bemerkung** Hier ist Dimension im Sinn von Kardinalität = **Mächtigkeit** der Basen verstanden!

**Beweis:** Ist $\varphi : V \to W$ ein Isomorphismus, dann ist für eine Basis $B$ von $V$ nach dem letzten Satz $\varphi|_B$ injektiv und $\varphi(B)$ eine Basis von $W$. Damit ist $\varphi|_B : B \to \varphi(B)$ eine Bijektion zwischen Basen von $V$ und $W$, sodass $\dim(V) = \dim(W)$.

Existieren umgekehrt Basen $B$ von $V$ und $C$ von $W$ gleicher Kardinalität $|B| = |C|$, dann bedeutet das die Existenz einer Bijektion $\tilde{\varphi} : B \to C$, die wir linear zu einer linearen Abbildung $\varphi : V \to W$ fortsetzen können. Da diese auf $B$ mit $\tilde{\varphi}$ übereinstimmt, ist $\varphi$ nach dem eingangs bewiesenen Satz ein Isomorphismus. □

**Achtung:** Ist $\dim(V) = \infty$, dann existieren Isomorphismen $\varphi : V \to W \subset V$ auf echte Teilräume von $W$.

**10.6 Beispiel (Vektorraumhomomorphismen)** Der Homomorphismus $\varphi : V \to V$ auf $V = C^\infty(\mathbb{R}, \mathbb{R})$ mit $\varphi(v)(x) := \int_0^x v(y)\,dy$ ist injektiv, denn $\ker(\varphi) = \{0\}$, aber $W := \varphi(V) \subset V$ ist ein echter Teilraum:

$$W = \{v \in V \mid v(0) = 0\}.$$

Ist $\varphi : V \to W$ ein Isomorphismus, dann existiert die Umkehrabbildung $\varphi^{-1} : W \to V$, denn $\varphi$ ist ja bijektiv.

Auch $\varphi^{-1} : W \to V$ ist *linear*, denn für $w, w_1, w_2 \in W$ und $c \in K$ gilt

$$\begin{aligned} \varphi^{-1}(w_1 + w_2) &= \varphi^{-1}\left(\varphi \circ \varphi^{-1}(w_1) + \varphi \circ \varphi^{-1}(w_2)\right) \\ &= \varphi^{-1}\left(\varphi\left(\varphi^{-1}(w_1) + \varphi^{-1}(w_2)\right)\right) \\ &= \varphi^{-1}(w_1) + \varphi^{-1}(w_2) \end{aligned}$$

und

$$\varphi^{-1}(cw) = \varphi^{-1}(c\varphi \circ \varphi^{-1}(w)) = \varphi^{-1}(\varphi(c\varphi^{-1}(w))) = c\varphi^{-1}(w).$$

Damit ergibt sich

**10.7 Satz** *Isomorphie von Vektorräumen ist eine Äquivalenzrelation.*

**Beweis:**

1. *Reflexivität:* $\mathrm{Id}_V : V \to V$ ist ein Isomorphismus $\Longrightarrow V \sim V$.
2. *Symmetrie:* Mit $\varphi : V \to W$ ist wie eben gezeigt $\varphi^{-1} : W \to V$ ein Isomorphismus, also $V \sim W \Longrightarrow W \sim V$.

3. *Transitivität:* Mit $\varphi : V \to W$, $\psi : W \to X$ ist auch $\psi \circ \varphi : V \to X$ ein Isomorphismus, denn

$$\psi \circ \varphi(v_1 + v_2) = \psi(\varphi(v_1) + \varphi(v_2)) = \psi \circ \varphi(v_1) + \varphi \circ \psi(v_2)$$

und

$$\psi \circ \varphi(cv) = \psi(c\varphi(v)) = c \cdot \psi \circ \varphi(v).$$

Außerdem war am Anfang der Vorlesung gezeigt worden, dass die Verknüpfung von Bijektionen eine Bijektion ergibt. □

Nach dem vorletzten Satz sind die Äquivalenzklassen isomorpher $K$–Vektorräume durch die Dimension charakterisiert. Insbesondere sind alle $K$–Vektorräume $V$ der Dimension $n := \dim(V) < \infty$ zum $K^n$ isomorph.

**10.8 Definition** *Für $K$–Vektorräume $V, W$ sei*

$$L(V, W) := \{\varphi : V \to W \mid \varphi \text{ ist linear}\}$$

*und*

$$L(V) := L(V, V).$$

**10.9 Satz** *Mit der Addition linearer Abbildungen $\varphi, \psi \in L(V, W)$*

$$(\varphi + \psi)(v) := \varphi(v) + \psi(v) \qquad (v \in V)$$

*und Skalarmultiplikation mit $c \in K$*

$$(c\varphi)(v) := c \cdot \varphi(v) \qquad (v \in V)$$

*wird $L(V, W)$ zu einem $K$–Vektorraum. Sind $V$ und $W$ endlichdimensional, dann ist*

$$\dim\big(L(V, W)\big) = \dim(V) \cdot \dim(W).$$

**Beweis:** Setze $\rho := \varphi + \psi$. Dann ist

$$\begin{aligned} \rho(v_1 + v_2) &= \varphi(v_1 + v_2) + \psi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) + \psi(v_2) + \psi(v_2) \\ &= \rho(v_1) + \rho(v_2), \\ \rho(cv) &= \varphi(cv) + \psi(cv) = c\varphi(v) + c\psi(v) \\ &= c\rho(v). \end{aligned}$$

Setze $\mu := c\varphi$. Dann ist

$$\begin{aligned}\mu(v_1+v_2) &= c\cdot\varphi(v_1+v_2) = c\cdot\varphi(v_1)+c\cdot\varphi(v_2) = \mu(v_1)+\mu(v_2)\\ \mu(kv) &= c\cdot\varphi(kv) = c\cdot k\cdot\varphi(v) = k\cdot\mu(v).\end{aligned}$$

Also ist $L(V,W)$ unter Addition und Multiplikation mit Skalaren abgeschlossen. Die Vektorraumaxiome lassen sich nach dem gleichen Schema überprüfen.

Zum Nachrechnen der Dimensionsformel sei $B := \{b_1,\ldots,b_m\}$ eine Basis von $V$ und $C := \{c_1,\ldots,c_n\}$ eine Basis von $W$. Die Abbildungen

$$\varphi_{j,i} : V \to W \qquad (i\in\{1,\ldots,m\}, j\in\{1,\ldots,n\})$$

mit

$$\varphi_{j,i}\left(\sum_{k=1}^{m}\lambda_k b_k\right) := \lambda_i c_j$$

sind linear und als Vektoren $\varphi_{j,i} \in L(V,W)$ linear unabhängig. Denn ist für Koeffizienten $\mu_{j,i} \in K$

$$\psi := \sum_{i=1}^{m}\sum_{j=1}^{n}\mu_{j,i}\cdot\varphi_{j,i} = 0$$

die Nullabbildung, dann ist insbesondere $\psi(b_k) = 0$, also

$$\psi(b_k) = \sum_{j=1}^{n}\mu_{j,k}c_j = 0 \qquad (k\in\{1,\ldots,m\}),$$

was wegen der linearen Unabhängigkeit der Basisvektoren $c_j$ das Verschwinden aller Koeffizienten $\mu_{j,k}$ impliziert.

Die $\varphi_{j,i} \in L(V,W)$ bilden auch eine Basis, denn für $\psi \in L(V,W)$ lässt sich jeder Bildvektor $\psi(b_i)$ der Basisvektoren als Linearkombination

$$\psi(b_i) = \sum_{j=1}^{n}\mu_{j,i}c_j \in W$$

der Basisvektoren $c_j$ darstellen. Daher gilt $\psi = \sum_{i=1}^{m}\sum_{j=1}^{n}\mu_{j,i}\varphi_{j,i}$. Also ist

$$\begin{aligned}\dim\big(L(V,W)\big) &= \big|\{\varphi_{i,j} \mid i\in\{1,\ldots,m\}, j\in\{1,\ldots,n\}\}\big|\\ &= m\,n = \dim(V)\dim(W).\end{aligned}$$

□

Im Fall der Vektorräume $V = K^m$, $W = K^n$ werden die Basiselemente $\varphi_{j,i} \in L(K^m, K^n)$ bez. der kanonischen Basen von $V$ und $W$ durch die in (8.1) definierten $n \times m$ Matrizen $E_{j,i} \in \mathrm{Mat}(n \times m, K)$ dargestellt, sodaß

$$L(K^m, K^n) \cong \mathrm{Mat}(n \times m, K)$$

ein Vektorraum-Isomorphismus ist, wenn man die Matrizen wie üblich addiert und mit Körperelementen multipliziert.

Matrizen kann man auch miteinander multiplizieren, wenn die Dimensionen stimmen. Auf der Ebene $K$-linearer Abbildungen $\varphi \in L(U, V)$ und $\psi \in L(V, W)$ entspricht dies der Verknüpfung

$$\psi \circ \varphi \in L(U, W).$$

Insbesondere können wir Endomorphismen $\varphi, \psi \in L(V)$ hintereinander ausführen und erhalten $\varphi \circ \psi \in L(V)$. Mit dieser Multiplikation wird $L(V)$ zu einem Ring, dem so genannten *Endomorphismenring* von $V$, dessen $1$–Element die identische Abbildung $\mathrm{Id}_V$ ist. Statt $L(V)$ schreibt man auch $\mathrm{End}(V)$.

**10.10 Bemerkung** Allerdings ist dieser Ring *nicht kommutativ*, falls $\dim(V) > 1$. Denn dann gibt es zwei linear unabhängige Vektoren $b_1, b_2 \in V$, die man zu einer Basis $B$ von $V$ ergänzen kann, mit

$$\varphi\left(\sum_i \lambda_i b_i\right) := (\lambda_1 + \lambda_2)b_1 + \lambda_2 b_2 \quad \text{und} \quad \psi\left(\sum_i \lambda_i b_i\right) := \lambda_1 b_1 + (\lambda_1 + \lambda_2)b_2$$

wird

$$\psi\left(\varphi\left(\sum_i \lambda_i b_i\right)\right) = \psi((\lambda_1 + \lambda_1)b_1 + \lambda_2 b_2) = (\lambda_1 + \lambda_2)b_1 + (\lambda_1 + 2\lambda_2)b_2$$

und

$$\varphi\left(\psi\left(\sum_i \lambda_i b_i\right)\right) = \varphi(\lambda_1 b_1 + (\lambda_1 + \lambda_2)b_2) = (2\lambda_1 + \lambda_2)b_1 + (\lambda_1 + \lambda_2)b_2,$$

also

$$(\varphi \circ \psi - \psi \circ \varphi)\left(\sum_i \lambda_i b_i\right) = \lambda_1 b_1 - \lambda_2 b_2 \neq 0.$$

**10.11 Definition** *Die Menge* $\mathrm{GL}(V)$ *der Automorphismen* $\varphi \in L(V)$ *wird als* **Allgemeine Lineare Gruppe** *des Vektorraums* $V$ *bezeichnet.*

Da die Verknüpfung $\psi \circ \varphi$ von Automorphismen ebenso wie $\varphi^{-1}$ ein Automorphismus ist, handelt es sich tatsächlich um eine Gruppe (die für $\dim(V) > 1$ nicht kommutativ ist).

Im Fall $V = K^n$ schreibt man oft

$$\mathrm{GL}(n, K) := \mathrm{GL}(K^n).$$

Wir hatten gesehen, dass endlichdimensionale Vektorräume $V$ über $K$ zu $K^{\dim(V)}$ isomorph sind. Daher brauchen wir nur $\mathrm{GL}(n, K)$ zu untersuchen.

Jeden Endomorphismus $\varphi : K^n \to K^n$ können wir als Matrix $A$ (mit $\varphi(x) = Ax$) auffassen. Die Verknüpfung von Endomorphismen geht dabei in die Matrizenmultiplikation über. Daher fasst man $\mathrm{GL}(n, K)$ als Gruppe von $n \times n$–Matrizen auf: $\mathrm{GL}(n, K) \subset \mathrm{Mat}(n, K)$.

**10.12 Beispiele (Allgemeine Lineare Gruppe)**

1. Für Dimension $n = 1$ ist $\mathrm{GL}(1, K) = K^* := \{k \in K \mid k \neq 0\}$.

2. $n = 2$. $A \in \mathrm{GL}(2, K) \iff \mathrm{rang}(A) = 2$. Also muss für $A = \left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right)$ der Vektor $\binom{a_{11}}{a_{21}}$ und der Vektor $\binom{a_{12}}{a_{22}}$ linear unabhängig sein, d.h. $\lambda\binom{a_{11}}{a_{21}} + \mu\binom{a_{12}}{a_{22}} = 0$ nur für $\lambda = \mu = 0$. In Verallgemeinerung von Bsp. 6.7.3 ist das genau dann der Fall, wenn die *Determinante*

   $$\det(A) := a_{11}a_{22} - a_{12}a_{21}$$

   von $A$ ungleich Null ist.

   Denn bei linearer Abhängigkeit ist o.B.d.A. $\lambda \neq 0$ (sonst Austausch der beiden Vektoren), also $\binom{a_{11}}{a_{21}} = -\frac{\mu}{\lambda}\binom{a_{12}}{a_{22}}$ und $a_{11}a_{22} - a_{12}a_{21} = \left(-\frac{\mu}{\lambda}a_{12}\right)a_{22} - \left(-\frac{\mu}{\lambda}a_{22}\right)a_{12} = 0$.

   Umgekehrt ist für $\det(A) = 0$ entweder $\binom{a_{11}}{a_{21}} = \binom{0}{0}$, was offensichtlich lineare Abhängigkeit impliziert oder o.B.d.A. $a_{21} \neq 0$, also $a_{12} = \frac{a_{11}a_{22}}{a_{21}}$, also $\binom{a_{12}}{a_{22}} = \mu\binom{a_{11}}{a_{21}}$ mit $\mu := \frac{a_{22}}{a_{21}}$. Auch in diesem Fall sind die beiden Vektoren linear abhängig.

3. Für $n = 2$ Dimensionen und den Körper $K = \mathbb{Z}/2\mathbb{Z}$ ist

   $$\mathrm{GL}(2, \mathbb{Z}/2\mathbb{Z}) = \left\{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\right\},$$

   während $\mathrm{Mat}(2, \mathbb{Z}/2\mathbb{Z})$ aus $2^4 = 16$ Elementen besteht.

   Für alle $M \in \mathrm{GL}(2, \mathbb{Z}/2\mathbb{Z})$ gilt $\det(M) = 1$.

# 11 Determinante und Spur

Im letzten Kapitel hatten wir die Determinante $\det(A)$ von $2 \times 2$– Matrizen $A$ als $\det \left( \begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix} \right) = a_{11}a_{22} - a_{12}a_{21}$ eingeführt. Damit war $\det : \mathrm{Mat}(2, K) \to K$ sicher *nicht* linear, aber *bilinear* in den Spaltenvektoren.

**11.1 Definition** *Ist* $\Phi : \underbrace{V \times \ldots \times V}_{n-\text{mal}} \to K$ *eine Abbildung des* $n$*–fachen cartesischen Produktes eines* $K$*-Vektorraums in den Körper* $K$*, dann heißt* $\Phi$ *eine* $n$**–fache Linearform**, *wenn für alle* $i \in \{1, \ldots, n\}$, $v_l, w_l \in V$ *und* $c \in K$ *gilt:*

$$\begin{aligned} &\Phi(v_1, \ldots, v_{i-1}, v_i + w_i, v_{i+1}, \ldots, v_n) = \\ &\Phi(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) + \Phi(v_1, \ldots, v_{i-1}, w_i, v_{i+1}, \ldots, v_n) \end{aligned}$$

*und*

$$\Phi(v_1, \ldots, v_{i-1}, cv_i, v_{i+1}, \ldots, v_n) = c\Phi(v_1, \ldots, v_n).$$

Man verlangt also von einer $n$–fachen Multilinearform, dass sie in jedem Argument linear ist, unabhängig von den (festen) Werten der anderen $n - 1$ Argumente.

**11.2 Beispiel** Mit der Spaltendarstellung $M = (s_1, s_2)$ von $M \in \mathrm{Mat}(2, K)$ gilt

$\det(as_1 + bt_1, s_2) = a \det(s_1, s_2) + b \det(t_1, s_2)$ und $\det(s_2, s_1) = -\det(s_1, s_2)$.

Gleichzeitig ist (nach Bsp. 10.12.2) $\det(s_1, s_2)$ genau dann $0$, wenn $s_1$ und $s_2$ linear abhängig sind. Die letzteren beiden Eigenschaften, die Antisymmetrie und das Kriterium für lineare Abhängigkeit der Spaltenvektoren, hängen zusammen.

**11.3 Definition** *Eine* $n = \dim(V)$*–fache Multilinearform* $\Phi$ *von* $V$ *heißt* **Determinantenform**, *wenn*

1. *für linear abhängige Vektoren* $v_1, \ldots, v_n \in V$ *gilt:* $\Phi(v_1, \ldots, v_n) = 0$,
2. *es umgekehrt Vektoren* $v_1, \ldots, v_n \in V$ *mit* $\Phi(v_1, \ldots, v_n) \neq 0$ *gibt.*

**11.4 Lemma** *Ist* $\pi \in \mathcal{S}_n$ *eine Permutation, dann ist* $\Phi(v_{\pi(1)}, \ldots, v_{\pi(n)}) = \mathrm{sign}(\pi) \cdot \Phi(v_1, \ldots, v_n)$, *falls* $\Phi$ *eine Determinantenform ist.*

**Beweis:** Da $\mathrm{sign} : \mathcal{S}_n \to \{-1, 1\}$ ein Gruppenhomomorphismus ist (siehe Beispiel 3.14) und sich jede Permutation als Produkt von Transpositionen schreiben lässt, reicht es aus, die Formel für eine *Transposition* $\pi$ zu zeigen. Dann ist $\mathrm{sign}(\pi) = -1$.

Nun ist

$$\begin{aligned}&\Phi(v_1, \ldots, v_i, \ldots, v_k + cv_i, \ldots, v_n) = \\ &\Phi(v_1, \ldots, v_i, \ldots, v_k, \ldots, v_n) + c\Phi(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) = \\ &\Phi(v_1, \ldots, v_i, \ldots, v_k, \ldots, v_n)\end{aligned}$$

wegen der Multilinearität von $\Phi$ und der Eigenschaft, dass $\Phi$ verschwindet, wenn seine Argumente linear abhängig sind.

Daher ist

$$\begin{aligned}&\Phi(v_1, \ldots, v_i, \ldots, v_k, \ldots, v_n) \\ &\quad = \Phi(v_1, \ldots, v_i, \ldots, v_i + v_k, \ldots, v_n) \\ &\quad = \Phi(v_1, \ldots, v_i - (v_i + v_k), \ldots, v_i + v_k, \ldots, v_n) \\ &\quad = -\Phi(v_1, \ldots, v_k, \ldots, v_i + v_k, \ldots, v_n) \\ &\quad = -\Phi(v_1, \ldots, v_k, \ldots, v_i, \ldots, v_n).\end{aligned}$$

□

In der Definition der Determinantenform wird nur verlangt, dass diese auf *einer* Basis ungleich Null ist. Tatsächlich ist sie auf *jeder* Basis von Null verschieden, sodass man mit ihr auf lineare Unabhängigkeit testen kann.

**11.5 Lemma** *Sind die Vektoren $b_1, \ldots, b_n \in V$ linear unabhängig, dann ist die Determinantenform $\Phi(b_1, \ldots, b_n) \neq 0$.*

**Beweis:** In Teil 2. der Definition 11.3 wird verlangt, dass es (linear unabhängige) Vektoren $v_1, \ldots, v_n$ gibt, für die $\Phi(v_1, \ldots, v_n) \neq 0$ ist. Da die $b$'s eine Basis bilden, können wir die $v$'s in der Form

$$v_i = \sum_{k=1}^{n} a_{i,k} b_k \quad \text{mit} \quad a_{i,k} \in K$$

schreiben.

Wegen der Multilinearität können wir

$$\begin{aligned}\Phi(v_1,\ldots,v_n) &= \Phi\left(\sum_{k_1=1}^{n} a_{1,k_1}b_{k_1},\ldots,\sum_{k_n=1}^{n} a_{n,k_n}b_{k_n}\right)\\ &= \sum_{k_1,\ldots,k_n=1}^{n} a_{1,k_1}\cdot\ldots\cdot a_{n,k_n}\Phi(b_{k_1},\ldots,b_{k_n})\end{aligned}$$

schreiben. Tatsächlich muss man nur über Index–$n$–Tupel $(k_1,\ldots,k_n)$ summieren, die paarweise verschieden sind, weil die anderen Terme wegen linearer Abhängigkeit sowieso verschwinden. Also hat man nur Terme von der Form $\Phi(b_{\pi(1)},\ldots,b_{\pi(n)}) = \operatorname{sign}(\pi)\Phi(b_1,\ldots,b_n) \neq 0$. Damit ist also

$$\Phi(v_1,\ldots,v_n) = \left(\sum_{\pi\in\mathcal{S}_n}\left(\prod_{i=1}^{n} a_{i,\pi(i)}\right)\cdot\operatorname{sign}(\pi)\right)\cdot\Phi(b_1,\ldots,b_n).$$

Da die linke Seite der Gleichung ungleich Null ist, folgt die Behauptung. □

**11.6 Lemma** *Ist $c \in K\backslash\{0\}$, dann ist für $v_i = \sum_{k=1}^{n} a_{i,k}b_k$*

$$\Psi(v_1,\ldots,v_n) := c\cdot\sum_{\pi\in\mathcal{S}_n}\operatorname{sign}(\pi)\prod_{i=1}^{n} a_{i,\pi(i)}$$

*eine Determinantenform.*

**Beweis:** • Offensichtlich ist $\Psi$ $n$–linear und

$$\begin{aligned}\Psi(b_1,\ldots,b_n) &= c\cdot\sum_{\pi\in\mathcal{S}(n)}\operatorname{sign}(\pi)\prod_{i=1}^{n}\delta_{i,\pi(i)}\\ &= c\cdot\operatorname{sign}(\mathrm{Id}) = c \neq 0.\end{aligned}$$

• Es muss also nur noch nachgeprüft werden, dass für linear abhängige $v_1,\ldots,v_n$

$$\Psi(v_1,\ldots,v_n) = 0$$

ist.
- Ist $n = 1$, dann muss $v_1 = 0$ sein, also auch $\Psi(v_1) = 0$.
- Ist $n > 1$, dann ist o.B.d.A. $v_1 = \sum_{i=2}^{n} c_i v_i$, also

$$\Psi(v_1,\ldots,v_n) = \sum_{i=2}^{n} c_i\Psi(v_i,v_2,\ldots,v_i,\ldots,v_n).$$

Dass die Faktoren $\Psi(v_i, v_2, \ldots, v_i, \ldots, v_n) = 0$ sind, folgt aus der Antisymmetrie $\Psi(w_1, \ldots, w_i, \ldots, w_k, \ldots, w_n) = -\Psi(w_1, \ldots, w_k, \ldots, w_i, \ldots, w_n)$ unter Transpositionen. □

Es gibt also in jedem endlichdimensionalen Vektorraum Determinantenformen, und diese unterscheiden sich nur durch eine Konstante $c \in K\backslash\{0\}$.

**11.7 Satz** *Sind $\Phi_1, \Phi_2$ Determinantenformen des Vektorraums $V$, dann ist der Quotient*

$$c := \frac{\Phi_1(b_1, \ldots, b_n)}{\Phi_2(b_1, \ldots, b_n)}$$

*unabhängig von der Wahl der geordneten Basis $(b_1, \ldots, b_n)$ von $V$, und für $v_1, \ldots, v_n \in V$ ist*

$$\Phi_1(v_1, \ldots, v_n) = c\Phi_2(v_1, \ldots, v_n).$$

**Beweis:**

$$\begin{aligned}
\Phi_1(v_1, \ldots, v_n) &= \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi) \left( \prod_{i=1}^{n} a_{i,\pi(i)} \right) \Phi_1(b_1, \ldots, b_n) \\
&= c \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi) \left( \prod_{i=1}^{n} a_{i,\pi(i)} \right) \Phi_2(b_1, \ldots, b_n) \\
&= c\Phi_2(v_1, \ldots, v_n),
\end{aligned}$$

falls $v_i = \sum_{k=1}^{n} a_{i,k} b_k$. Ist $(v_1, \ldots, v_n)$ eine zweite Basis, dann können wir diese Gleichung wegen $\Phi_2(v_1, \ldots, v_n) \neq 0$ nach $c$ auflösen. Damit folgt die Unabhängigkeit des Koeffizienten $c$ von der Wahl der Basis. □

Man soll sich die Determinantenform als eine Messvorschrift für das Volumen des von $v_1, \ldots, v_n$ aufgespannten Quaders vorstellen. Die *Volumeneinheit* ist definitionsabhängig, das *Verhältnis* von Volumina nicht.

Es ist also möglich festzustellen, um welchen *Faktor* ein Endomorphismus $\varphi : V \to V$ das Volumen verändert:

**11.8 Definition** *Ist $n := \dim(V) < \infty$ und $\varphi \in L(V)$, dann ist die* **Determinante** *von $\varphi$ durch*

$$\det(\varphi) := \frac{\Psi(\varphi(b_1), \ldots, \varphi(b_n))}{\Psi(b_1, \ldots, b_n)}$$

*definiert, wobei $(b_1, \ldots, b_n)$ eine Basis von $V$ und $\Psi$ eine Determinantenform ist.*

Natürlich müssen wir zeigen, dass $\det(\varphi)$ nicht von der Wahl der Basis und der Determinantenform abhängt.
• Sind $\Psi_1$ und $\Psi_2$ zwei Determinantenformen auf $V$ und $c := \frac{\Psi_2(b_1,\ldots,b_n)}{\Psi_1(b_1,\ldots,b_n)}$, dann ist $\Psi_2(\varphi(b_1),\ldots,\varphi(b_n)) = c\Psi_1(\varphi(b_1),\ldots,\varphi(b_n))$, also

$$\frac{\Psi_2(\varphi(b_1),\ldots,\varphi(b_n))}{\Psi_2(b_1,\ldots,b_n)} = \frac{c}{c}\frac{\Psi_1(\varphi(b_1),\ldots,\varphi(b_n))}{\Psi_1(b_1,\ldots,b_n)} = \det(\varphi).$$

• Auch von der Wahl der Basis hängt die Definition nicht ab, denn für eine *zweite Basis* $(c_1,\ldots,c_n)$ mit $c_i = \sum_{k=1}^n a_{i,k}b_k$ ist

$$\begin{aligned}\Psi(\varphi(c_1),\ldots,\varphi(c_n)) &= \Psi\left(\sum_{k=1}^n a_{1,k}\varphi(b_k),\ldots,\sum_{k=1}^n a_{n,k}\varphi(b_k)\right)\\ &= \left[\sum_{\pi\in\mathcal{S}_n}\operatorname{sign}(\pi)\left(\prod_{i=1}^n a_{i,\pi(i)}\right)\right]\Psi(\varphi(b_1),\ldots,\varphi(b_n))\end{aligned}$$

und

$$\Psi(c_1,\ldots,c_n) = \left[\sum_{\pi\in\mathcal{S}_n}\operatorname{sign}(\pi)\left(\prod_{i=1}^n a_{i,\pi(i)}\right)\right]\Psi(b_1,\ldots,b_n),$$

sodass bei der Quotientenbildung der (von Null verschiedene!) Vorfaktor in eckigen Klammern wegfällt.

**11.9 Satz** *1.* $\det(\varphi)$ *ist genau dann* $\neq 0$*, wenn* $\varphi$ *ein Automorphismus ist.*

*2. Für* $\varphi_1,\varphi_2 \in L(V)$ *gilt die* **Determinantenproduktformel**

$$\det(\varphi_1\circ\varphi_2) = \det(\varphi_1)\det(\varphi_2)$$

**Beweis:**

1. Eine Determinantenform $\Phi(v_1,\ldots,v_n)$ ist genau dann $\neq 0$, wenn $v_1,\ldots,v_n$ eine Basis bilden.

   Andererseits ist $\varphi \in L(V)$ genau dann ein Automorphismus, wenn $\varphi(b_1),\ldots,\varphi(b_n)$ wieder eine Basis von $V$ ist.

2. Ist $\varphi_2$ ein Automorphismus, dann ist

$$\begin{aligned}\det(\varphi_1 \circ \varphi_2) &= \frac{\Psi(\varphi_1 \circ \varphi_2(b_1), \ldots, \varphi_1 \circ \varphi_2(b_n))}{\Psi(b_1, \ldots, b_n)} \\ &= \frac{\Psi(\varphi_1 \circ \varphi_2(b_1), \ldots, \varphi_1 \circ \varphi_2(b_n))}{\Psi(\varphi_2(b_1), \ldots, \varphi_2(b_n))} \cdot \frac{\Psi(\varphi_2(b_1), \ldots, \varphi_2(b_n))}{\Psi(b_1, \ldots, b_n)} \\ &= \det(\varphi_1) \cdot \det(\varphi_2).\end{aligned}$$

Ist $\varphi_2$ kein Automorphismus, also $\det\varphi_2 = 0$, dann ist auch $\varphi_1 \circ \varphi_2$ kein Automorphismus, also auch $\det(\varphi_1 \circ \varphi_2) = 0$. □

**11.10 Korollar** *Für* $\dim(V) < \infty$ *und* $\varphi \in \mathrm{GL}(V)$ *gilt*

1. $\det(\mathrm{Id}_V) = 1$.
2. $\det(\varphi^{-1}) = 1/\det(\varphi)$.

**Beweis:**

1. $\det(\mathrm{Id}_V) \neq 0$ und $\det(\mathrm{Id}_V) \cdot \det(\mathrm{Id}_V) = \det(\mathrm{Id}_V)$.
2. $\det(\varphi) \cdot \det(\varphi^{-1}) = \det(\varphi \circ \varphi^{-1}) = \det(\mathrm{Id}_V) = 1$. □

Wie berechnet man nun Determinanten von Endomorphismen $\varphi : V \to V$?

Im Grunde wissen wir das schon. Denn ist $A := \mathrm{Mat}_B(\varphi)$ die darstellende Matrix von $\varphi$ bez. einer Basis $B = (b_1, \ldots, b_n)$ von $V$, dann ist für eine Determinantenform $\Psi$, wie schon gezeigt,

$$\Psi\big(\varphi(b_1), \ldots, \varphi(b_n)\big) = \left(\sum_{\sigma \in \mathcal{S}_n} \mathrm{sign}(\pi) \prod_{i=1}^{n} a_{i,\pi(i)}\right) \Psi(b_1, \ldots, b_n),$$

sodass mit

$$\begin{aligned}\det(A) &:= \sum_{\pi \in \mathcal{S}_n} \mathrm{sign}(\pi) \prod_{i=1}^{n} a_{i,\pi(i)} \qquad (11.1) \\ \det(\varphi) &= \det(A)\end{aligned}$$

ist.

**11.11 Beispiel** Für eine Matrix $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \in \mathrm{Mat}(n, K)$ gilt

$n = 1$: $\det(A) = a_{11} \equiv A$,

$n = 2$: $\det(A) = a_{11}a_{22} - a_{12}a_{21}$,

$n = 3$: $\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$
$\qquad -a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$,

entsprechend den geraden Permutationen $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{smallmatrix}\right)$ aus $\mathcal{A}_3$ mit positiven Vorzeichen und den ungeraden Permutationen
$\left(\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}\right)$ mit negativen Vorzeichen.

Eine Merkregel für die Determinante einer $3 \times 3$–Matrix $A$ ist die

**Regel von Sarrus:** Man schreibt die beiden ersten Spaltenvektoren noch einmal rechts neben $A$.

$$\begin{array}{ccccccccc} a_{11} & & a_{12} & & a_{13} & & a_{11} & & a_{12} \\ & \diagdown & & \times & & \times & & \diagup & \\ a_{21} & & a_{22} & & a_{23} & & a_{21} & & a_{22} \\ & \diagup & & \times & & \times & & \diagdown & \\ a_{31} & & a_{32} & & a_{33} & & a_{31} & & a_{33} \end{array}$$

Danach addiert man die Produkte der mit einer Nebendiagonale ($\diagdown$) verbundenen Koeffizienten und subtrahiert davon die Produkte bezüglich der Hauptdiagonalen ($\diagup$).

**Vorsicht:** Die analoge Regel, angewandt auf $4 \times 4$–Matrizen, muss im Allgemeinen zum falschen Ergebnis führen, denn statt $8 = 2 \times 4$ Summanden haben wir hier $|\mathcal{S}_4| = 4! = 24$ Summanden.

**11.12 Satz** *Für die Determinante einer quadratischen Matrix $A \in \mathrm{Mat}(n, K)$ gelten die folgenden Rechenregeln:*

1. *Ihr Wert wird durch eine Transposition nicht geändert:*

$$\det(A^t) = \det(A).$$

*Ist $A$ von der Form $A = (s_1, \dots, s_i, \dots, s_k, \dots, s_n)$, dann ist*

2. $\det(s_1, \dots, s_i, \dots, s_k + s_i, \dots, s_n) = \det(A)$ *(Summation)*
3. $\det(s_1, \dots, s_k, \dots, s_i, \dots, s_n) = -\det(A)$ *(Austausch)*
4. $\det(s_1, \dots, c \cdot s_i, \dots, s_k, \dots, s_n) = c \cdot \det(A)$, $c \in K$ *(Multiplikation)*

5. $\det(s_1, \ldots, s_i, \ldots, s_i, \ldots, s_n) = 0$ *(Gleichheit)*

6. $\det(c \cdot A) = c^n \cdot \det(A)$

7. $\det(A^{-1}) = (\det(A))^{-1}$ *falls* $\det(A) \neq 0$.

8. *Ist* $B \in \mathrm{Mat}(n, K)$ *eine zweite quadratische Matrix, dann ist*

$$\det(AB) = \det(BA) = \det(A)\det(B).$$

9. *Ist* $T \in \mathrm{Mat}(n, K)$ *regulär, dann gilt*

$$\det(TAT^{-1}) = \det(A).$$

10. $\det(\mathbb{1}_n) = 1$.

**Beweis:** 1.

$$\begin{aligned}
\det(A^t) &= \sum_{\pi \in \mathcal{S}_n} \mathrm{sign}(\pi) \prod_{i=1}^{n} a_{\pi(i),i} = \sum_{\pi \in \mathcal{S}_n} \mathrm{sign}(\pi) \prod_{j=1}^{n} a_{j,\pi^{-1}(j)} \\
&\qquad \text{mit der Bezeichnung } j := \pi(i), \text{ also } i = \pi^{-1}(j) \\
&= \sum_{\pi \in \mathcal{S}_n} \mathrm{sign}(\pi^{-1}) \prod_{j=1}^{n} a_{j,\pi^{-1}(j)}, \quad \text{denn} \quad \mathrm{sign}(\pi) = \mathrm{sign}(\pi^{-1}) \\
&= \sum_{\sigma \in \mathcal{S}_n} \mathrm{sign}(\sigma) \prod_{j=1}^{n} a_{j,\sigma(j)} \quad \text{mit} \quad \sigma := \pi^{-1} \\
&= \det(A).
\end{aligned}$$

Bei den Rechenregeln 2.–5. ist zu bedenken, dass die Spaltenvektoren $s_1, \ldots, s_n$ von $A$ die Koordinatenvektoren der Basisvektoren $b_1, \ldots, b_n$ bez. der Abbildung $\varphi$ sind. Daher haben wir diese Relationen schon gezeigt.

6. folgt aus 4., 8. ist die Matrixversion des Determinantenproduktes, und 7.,9. und 10. folgen aus diesem. □

Die Zahl der Summanden der Determinante einer $n \times n$–Matrix ist $n!$, wächst also mit $n$ schnell an. Um den Rechenaufwand klein zu halten, bietet sich z.B. an, $A$ durch elementare Umformungen in Zeilenstufenform überzuführen. Bei zwei der drei elementaren Umformungen ändert sich zwar der Wert der Determinante, aber eben in überschaubarer Weise:

1. Jede Vertauschung zweier Zeilen bzw. Spalten bringt eine Vorzeichenumkehr, $k$ Vertauschungen also einen Faktor $(-1)^k$.

2. Multiplikation einer Zeile oder Spalte mit $c \neq 0$ bringt nach Teil 1. und 6. des Satzes 11.12 einen Faktor $c$.

3. Addition das $c$-fachen einer Zeile (oder Spalte) zu einer anderen ändert den Wert der Determinante nicht.

   Ist $A'$ in Zeilenstufenform, dann ist

   $$\det(A') = \prod_{i=1}^{n} (A')_{i,i},$$

   denn für alle von der Identität verschiedenen Permutationen $\pi$ gibt es ein $i \in \{1, \ldots, n\}$ mit

   $$\pi(i) < i \quad \text{, also} \quad (A')_{i,\pi(i)} = 0.$$

Damit ergibt sich

$$\det(A) = (-1)^k \cdot \det(A') = (-1)^k \cdot \prod_{i=1}^{n} (A')_{i,i}.$$

**11.13 Beispiel (Determinanten-Berechnung)**

$$
\begin{array}{rrrrrl}
 & 1 & 3 & 4 & 0 & \\
A := & 2 & 5 & 7 & 1 & \\
 & -1 & 2 & -3 & 0 & \\
 & 0 & 0 & 1 & 4 & \\
\hline
 & 1 & 3 & 4 & 0 & \\
 & 0 & -1 & -1 & 1 & z_2 - 2z_1 \\
 & 0 & 5 & 1 & 0 & z_3 + z_1 \\
 & 0 & 0 & 1 & 4 & \\
\hline
 & 1 & 3 & 4 & 0 & \\
 & 0 & -1 & -1 & 1 & \\
 & 0 & 0 & -4 & 5 & z_3 + 5z_1 \\
 & 0 & 0 & 1 & 4 & \\
\hline
 & 1 & 3 & 4 & 0 & \\
A' := & 0 & -1 & -1 & 1 & \\
 & 0 & 0 & -4 & 5 & \\
 & 0 & 0 & 0 & \frac{21}{4} & z_4 + \frac{1}{4} z_5
\end{array}
$$

$$\det(A) = \det(A') = 1 \cdot (-1) \cdot (-4) \cdot \tfrac{21}{4} = 21.$$

Der folgende *Entwicklungssatz von Laplace* ermöglicht die rekursive Berechnung von $\det(A)$ für $A \in \mathrm{Mat}(n, K)$, $n > 1$, durch Zurückführung auf Determinanten von $(n-1) \times (n-1)$–Matrizen.

Für die Indizes $i, j \in \{1, \ldots, n\}$ sei dazu $A^{(ij)} \in \mathrm{Mat}(n-1, K)$ die Matrix, die durch Streichen der $i$–ten Zeile und der $j$–ten Spalte aus $A$ entsteht.

Benennen wir mit $\chi_l \in \mathcal{S}_n$, $l = 1, \ldots, n$ diejenige Permutation, die durch

$$\chi_l(k) := \begin{cases} k & , \ k < l \\ k+1 & , \ l \leq k < n \\ l & , \ k = n \end{cases}$$

definiert ist, dann gilt die Beziehung

$$(A^{(ij)})_{k,l} = (A)_{\chi_i(k), \chi_j(l)} \qquad (1 \leq k, l \leq n-1). \tag{11.2}$$

$A^{(ij)}$ heißt das *algebraische Komplement* des Matrix-Eintrages $a_{ij} := (A)_{ij}$.

**11.14 Satz (Laplace)**

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \cdot \det(A^{(ij)}) \qquad (i = 1, \ldots, n).$$

**Beweis:** In der Zyklenschreibweise von Permutationen ist $\chi_l = (l, l+1, \ldots, n)$, also

$$\mathrm{sign}(\chi_l) = (-1)^{n-l}.$$

Setzen wir $\tau := \chi_j^{-1} \circ \pi \circ \chi_i$ für eine Permutation $\pi \in \mathcal{S}_n$ mit $\pi(i) = j$, dann gilt

$$\tau(n) = \chi_j^{-1} \circ \pi(i) = \chi_j^{-1}(j) = n,$$

und

$$\mathrm{sign}(\tau) = \mathrm{sign}(\chi_j^{-1})\mathrm{sign}(\pi)\mathrm{sign}(\chi_i) = (-1)^{i+j}\mathrm{sign}(\pi).$$

Daher ist nach Definition (11.1) der Determinante und nach (11.2)

$$\begin{aligned}
\det(A) &= \sum_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi) \prod_{l=1}^{n} (A)_{l,\pi(l)} = \sum_{j=1}^{n} \sum_{\substack{\pi\in\mathcal{S}_n \\ \pi(i)=j}} \operatorname{sign}(\pi) \prod_{l=1}^{n} (A)_{l,\pi(l)} \\
&= \sum_{j=1}^{n} \sum_{\substack{\tau\in\mathcal{S}_n \\ \tau(n)=n}} (-1)^{i+j} \operatorname{sign}(\tau) \prod_{l=1}^{n} (A)_{l,\chi_j\circ\tau\circ\chi_i^{-1}(l)} \\
&= \sum_{j=1}^{n} (-1)^{i+j} \sum_{\substack{\tau\in\mathcal{S}_n \\ \tau(n)=n}} \operatorname{sign}(\tau) \prod_{l=1}^{n} (A)_{\chi_i(l),\chi_j(\tau(l))} \\
&= \sum_{j=1}^{n} (-1)^{i+j} \sum_{\substack{\tau\in\mathcal{S}_n \\ \tau(n)=n}} \operatorname{sign}(\tau) \cdot a_{ij} \cdot \prod_{l=1}^{n-1} \left(A^{(ij)}\right)_{l,\tau(l)} \\
&= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \sum_{\tilde{\tau}\in\mathcal{S}_{n-1}} \operatorname{sign}(\tilde{\tau}) \prod_{l=1}^{n-1} \left(A^{(ij)}\right)_{l,\tilde{\tau}(l)} \\
&= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det\left(A^{(ij)}\right).
\end{aligned}$$

□

**11.15 Beispiel (Determinanten-Berechnung)** $A := \begin{pmatrix} 0&1&2&3 \\ 1&0&1&2 \\ 2&1&0&1 \\ 3&2&1&0 \end{pmatrix}$. Entwicklung nach der ersten Zeile ergibt

$$\det(A) = 0 \cdot \det\left(A^{(11)}\right) - 1 \cdot \det\left(A^{(12)}\right) + 2 \cdot \det\left(A^{(13)}\right) - 3 \cdot \det\left(A^{(14)}\right)$$

mit

$$\begin{aligned}
\det\left(A^{(11)}\right) &= \det\begin{pmatrix} 0&1&2 \\ 1&0&1 \\ 2&1&0 \end{pmatrix} = 4 \\
\det\left(A^{(12)}\right) &= \det\begin{pmatrix} 1&1&2 \\ 2&0&1 \\ 3&1&0 \end{pmatrix} = 6 \\
\det\left(A^{(13)}\right) &= \det\begin{pmatrix} 1&0&2 \\ 2&1&1 \\ 3&2&0 \end{pmatrix} = 0 \\
\det\left(A^{(14)}\right) &= \det\begin{pmatrix} 1&0&1 \\ 2&1&0 \\ 3&2&1 \end{pmatrix} = 2.
\end{aligned}$$

Also ist

$$\det(A) = -6 - 6 = -12.$$

Wegen $\det(A^t) = \det(A)$ kann man $\det(A)$ analog nach der $j$–ten Spalte statt nach der $i$–ten Zeile entwickeln.

**11.16 Satz (Cramersche Regel)** *Ist $A \in \mathrm{Mat}(n, K)$ regulär und $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in K^n$, dann besitzt das lineare Gleichungssystem*

$$Ax = b$$

*die (eindeutige) Lösung $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ der Form*

$$x_i = \frac{\det(s_1, \ldots, s_{i-1}, b, s_{i+1}, \ldots, s_n)}{\det(A)} \qquad (i = 1, \ldots, n),$$

*wobei $s_k$ die $k$–te Spalte der Matrix $A$ bezeichnet.*

**Beweis:** Zunächst besitzt das lineare Gleichungssystem eine eindeutige Lösung, denn die quadratische Matrix $A$ ist regulär.

Einsetzen von $x$ ergibt unter Verwendung des Laplaceschen Satzes

$$\begin{aligned} (Ax)_k &= \sum_{j=1}^{n} a_{kj} \frac{\sum_{l=1}^{n} (-1)^{l+j} b_l \cdot \det\left(A^{(lj)}\right)}{\det(A)} \\ &= \det(A)^{-1} \sum_{l=1}^{n} b_l \left( \sum_{j=1}^{n} a_{kj} (-1)^{l+j} \det\left(A^{(lj)}\right) \right). \end{aligned}$$

Nun ist

$$\sum_{j=1}^{n} a_{kj} (-1)^{l+j} \det\left(A^{(lj)}\right) = \delta_{kl} \cdot \det(A),$$

also $(Ax)_k = b_k$, denn für

• $k = l$ ist dies die linke Seite die Entwicklung von $\det(A)$ nach der $k$–ten Spalte, und für

• $k \neq l$ die Entwicklung von $\det(A')$, wobei $A'$ aus $A$ durch Ersetzen von $s_l$ durch $s_k$ hervorgeht. Damit ist aber $\mathrm{rang}(A') < n$, also $\det(A') = 0$. □

**Frage:** Warum interessiert man sich überhaupt für die Determinante eines Endomorphismus?

**Antwort:** Zum einen kann man mit ihr entscheiden, ob $\varphi$ ein Automorphismus ist (nämlich genau dann, wenn $\det(\varphi) \neq 0$). Dafür haben wir zwar eine einfachere Methode, das Gaußverfahren, kennen gelernt. Wichtig ist aber nicht nur, ob $\det\varphi \neq 0$ oder $= 0$ ist, sondern der genaue Wert. Wir können diesen Wert als den Faktor interpretieren, um den das Volumen unter $\varphi$ vergrößert wird.

Nach Satz 11.12.9 wird der Wert der Determinante durch die Transformation $A \mapsto TAT^{-1}$ nicht geändert. Diese Transformation entspricht aber einem Basiswechsel.

Denn wir können $A$ immer als darstellende Matrix $M_B(\varphi)$ eines Endomorphismus $\varphi$ bezüglich einer Basis $B = (b_1, \ldots, b_n)$ auffassen (z.B. $\varphi(x) := Ax$ und $b_i := e_i$ kanonischer Basisvektor des $K^n$).

Ist $C = (c_1, \ldots, c_n)$ eine zweite Basis, dann gilt für die Transformationsmatrix

$$T := M_C^B(\mathrm{Id}_V)$$

für den Basiswechsel von $B$ auf $C$ nach (8.2)

$$M_C(\varphi) = TAT^{-1},$$

und wir haben uns noch einmal der Tatsache vergewissert, dass die Determinante eines Endomorphismus nicht von der Wahl der Basis abhängt.

**Frage:** Welche weiteren Invarianten eines Endomorphismus gibt es neben der Determinante?

**11.17 Definition** *Die* **Spur** *einer quadratischen Matrix* $A \in \mathrm{Mat}(n, K)$ *ist*

$$\mathrm{Spur}(A) := \sum_{i=1}^{n} (A)_{i,i}.$$

**11.18 Satz** *1.* $\mathrm{Spur}(AB) = \mathrm{Spur}(BA)$

*2.* $\mathrm{Spur}(TAT^{-1}) = \mathrm{Spur}(A)$ *für eine reguläre Matrix* $T \in \mathrm{Mat}(n, K)$.

**Beweis:**

1.

$$\begin{aligned} \mathrm{Spur}(AB) &= \sum_{i=1}^{n} (AB)_{i,i} = \sum_{i=1}^{n} \sum_{k=1}^{n} (A)_{i_k} (B)_{k_i} \\ &= \sum_{k=1}^{n} \sum_{i=1}^{n} (B)_{k_i} (A)_{i_k} = \sum_{k=1}^{n} (BA)_{k,k} = \mathrm{Spur}(BA) \end{aligned}$$

2. Setze $B := AT$. Dann ist

$$\mathrm{Spur}(TAT^{-1}) = \mathrm{Spur}(AT^{-1}T) = \mathrm{Spur}(A). \qquad \square$$

Die Aussage 2. bedeutet wieder, dass die Spur von Basistransformationen unberührt bleibt. Wir können also definieren:

**11.19 Definition** *Die* **Spur** *eines Endomorphismus $\varphi : V \to V$ eines $n$–dimensionalen Vektorraums $V$ ist*

$$\mathrm{Spur}(\varphi) := \mathrm{Spur}\big(\mathrm{Mat}_B(\varphi)\big),$$

*wobei $B$ eine beliebige Basis von $V$ ist.*

**Folgerung:** Für $k \in \mathbb{N}$ gilt damit auch

$$\mathrm{Spur}(A^k) = \mathrm{Spur}(TA^kT^{-1}) = \mathrm{Spur}\big((TAT^{-1})^k\big),$$

sodass wir weitere invariante Größen gewinnen.

Allerdings lässt sich für $k > n$

$$\mathrm{Spur}(A^k) \quad \text{aus den Zahlen} \quad \mathrm{Spur}(A), \ldots, \mathrm{Spur}(A^n)$$

berechnen, und auch die Determinante $\det(A)$ ist ein Polynom in diesen Größen.

**11.20 Beispiel (Spur und Determinante)** • $n = 2$: Es ist für $A := \left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right)$

$$\begin{aligned} \mathrm{Spur}(A) &= a_{11} + a_{22}, \\ \mathrm{Spur}(A^2) &= \mathrm{Spur}\begin{pmatrix} a_{11}^2+a_{12}a_{21} & a_{11}a_{12}+a_{12}a_{22} \\ a_{21}a_{11}+a_{22}a_{21} & a_{21}a_{12}+a_{22}^2 \end{pmatrix} = a_{11}^2 + a_{22}^2 + 2a_{12}a_{21}\ , \quad \text{und} \\ \det(A) &= a_{11}a_{22} - a_{12}a_{22}, \end{aligned}$$

also

$$\det(A) = \tfrac{1}{2}\big((\mathrm{Spur}A)^2 - \mathrm{Spur}(A^2)\big).$$

• $n = 3$: $\qquad \det(A) = \tfrac{1}{6}\big((\mathrm{Spur}A)^3 - 3(\mathrm{Spur}A)(\mathrm{Spur}A^2) + 2\mathrm{Spur}(A^3)\big).$

**11.21 Bemerkung** Während im englischen Determinante *determinant* heißt, ist das Wort für Spur *trace*. Daher findet man oft $\mathrm{tr}(A)$ für die $\mathrm{Spur}$ von $A$.

**Anwendungen der Determinanten**

1. **Dreipunkteformel des Kreises**
   Ein Kreis in der Ebene mit Mittelpunkt $(x_0, y_0)$ und Radius $r > 0$ ist die Menge

$$K := \{(x, y) \in \mathbb{R}^2 \mid (x - x_0)^2 + (y - y_0)^2 = r^2\}.$$

   Man kann diesen Kreis also auch in der Form

$$K = \{(x, y) \in \mathbb{R}^2 \mid a_1(x^2 + y^2) + a_2 x + a_3 y + 1 = 0\} \tag{11.3}$$

   mit geeigneten Koeffizienten $a_1, a_2, a_3 \in \mathbb{R}$ schreiben.

   Wir nehmen nun an, dass wir drei voneinander verschiedene Punkte $(x_i, y_i)$ $(i = 1, 2, 3)$ auf einem Kreis kennen, und wir wollen die Koeffizienten $a_j$ bestimmen. Einsetzen der Punkte in die Kreisgleichung in (11.3) liefert das Ergebnis nach Lösung des linearen Gleichungssystems.

   Direkter ist die folgende Überlegung. Nehmen wir noch einen vierten Punkt $(x, y)$ des Kreises hinzu, dann müssen die vier linearen Bestimmungsgleichungen für die Koeffizienten $a_1, a_2, a_3$ linear abhängig sein, sodass

$$K = \{(x, y) \in \mathbb{R}^2 \mid \det(A(x, y)) = 0\}$$

   mit

$$A(x, y) = \begin{pmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{pmatrix}.$$

   **11.22 Beispiel** $K$ enthält die Punkte $(x_1, y_1) = (0, 0)$, $(x_2, y_2) = (1, 0)$ und $(x_3, y_3) = (0, 1)$, also

$$A(x, y) = \begin{pmatrix} x^2 + y^2 & x & y & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$$

   und

$$\det A(x, y) = x^2 + y^2 - x - y = 0.$$

2. **Interpolation mit Polynomen**
   Gegeben sind $n$ Punkte $(x_i, y_i) \in \mathbb{R}^2$ mit $x_1 < \ldots < x_n$; gesucht ist das (eindeutige) Polynom $p(x) = \sum_{k=0}^{n-1} a_k x^k$ höchstens $(n-1)$–ten Grades, dessen Graph durch diese Punkte geht, d.h. $p(x_i) = y_i \quad (i = 1, \ldots, n)$.

Die $n$ linearen Gleichungen

$$\sum_{k=0}^{n-1} x_i^k a_k = y_i \qquad (i = 1, \ldots, n)$$

in den $n$ Unbekannten $a_k$ können um die Gleichung $p(x) = y$ ergänzt werden, die dann wegen linearer Abhängigkeit eine verschwindende Determinante

$$\det\big(A(x,y)\big) = 0 \tag{11.4}$$

mit

$$A(x,y) := \begin{pmatrix} 1 & x & x^2 & \ldots & x^{n-1} & -y \\ 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} & -y_1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} & -y_n \end{pmatrix}$$

ergibt. Die Entwicklung nach der ersten Zeile liefert den Koeffizienten $(-1)^n\, D$ mit der so genannten *Vandermonde-Determinante*

$$D := \det \begin{pmatrix} 1 & x_1 & \ldots & x_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \ldots & x_n^{n-1} \end{pmatrix}.$$

Es gilt $D = \prod_{k>i}(x_k - x_i)$ (versuchen Sie diese Formel zu beweisen!). Daher ist $D \neq 0$ und das interpolierende Polynom existiert und ist eindeutig.

**11.23 Beispiel (interpolierendes Polynom)** Die drei Punkte $(x_1, y_1) = (0, 0)$, $(x_2, y_2) = (1, 1)$ und $(x_3, y_3) = (2, 0)$ sollen durch ein Polynom $p$ interpoliert werden. Entwicklung nach der ersten Spalte ergibt

$$\det \begin{pmatrix} 1 & x & x^2 & -y \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & -1 \\ 1 & 2 & 4 & 0 \end{pmatrix} = -\det \begin{pmatrix} x & x^2 & -y \\ 1 & 1 & -1 \\ 2 & 4 & 0 \end{pmatrix} = -4x + 2x^2 + 2y,$$

also ist mit (11.4)

$$p(x) = -x^2 + 2x.$$

# 12 Eigenwerte und Eigenvektoren

Wir haben die Determinante eines Endomorphismus $\varphi$ eines $n$–dimensionalen Vektorraumes $V$ ausgerechnet, indem wir die Determinante der darstellenden Matrix $\mathrm{Mat}_B(\varphi)$ bezüglich einer beliebigen Basis $B = (b_1, \ldots, b_n)$ berechneten.

Das charakteristische Polynom liefert uns weitere basisunabhängige Informationen über den Endomorphismus $\varphi$.

**12.1 Definition** • *Ist $\varphi \in L(V)$ Endomorphismus des endlichdimensionalen $K$–Vektorraumes $V$, dann heißt $\chi_\varphi \in K[x]$ mit*

$$\chi_\varphi(x) := \det(x\,\mathrm{Id}_V - \varphi) \qquad (x \in K)$$

*das* **charakteristische Polynom** *von $\varphi$.*
• *Ähnlich heißt $x \mapsto \chi_A(x) := \det(x\,\mathbb{1}_n - A)$ das charakteristische Polynom von $A \in \mathrm{Mat}(n, K)$.*

Es gilt $\chi_\varphi = \chi_A$, wenn $A$ darstellende Matrix von $\varphi$ ist.

Es handelt sich bei $\chi_\varphi$ wirklich um ein Polynom, und zwar vom Grad $n$, mit Leitkoeffizient $1$ und mit konstantem Koeffizienten $(-1)^n\det(\varphi)$, wie man aus der Definition der Determinante abliest.[11]

**12.2 Beispiele (charakteristisches Polynom)** 1. Ist $A$ eine *Diagonalmatrix*, d.h. $(A)_{ij} = 0$ für $i \neq j$, dann brauchen wir nur die Diagonalelemente $a_i := (A)_{ii}$, $i = 1, \ldots, n$ zu notieren, und schreiben $A = \mathrm{diag}(a_1, \ldots, a_n) \in \mathrm{Mat}(n, K)$.

Es ist dann

$$\chi_A(x) = \prod_{i=1}^{n}(x - a_i).$$

Die gleiche Formel gilt für eine obere Dreiecksmatrix.

2. Ist $A := \left(\begin{smallmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{smallmatrix}\right) \in \mathrm{Mat}(2, \mathbb{R})$, d.h. $A$ ist Drehmatrix für eine Drehung um den Winkel $\varphi \in (0, 2\pi)$, dann ist

$$\begin{aligned} \chi_A(x) &= \det\left(\begin{smallmatrix} x-\cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & x-\cos(\varphi) \end{smallmatrix}\right) \\ &= (x - \cos(\varphi))^2 + \sin^2(\varphi) \\ &= \big[x - (\cos(\varphi) + i\sin(\varphi))\big]\,\big[x - (\cos(\varphi) - i\sin(\varphi))\big]. \end{aligned}$$

3. Für die obere Dreiecksmatrix $A := \left(\begin{smallmatrix} a & 1 \\ 0 & a \end{smallmatrix}\right) \in \mathrm{Mat}(2, \mathbb{R})$ und $a \in \mathbb{R}$ ist

$$\chi_A(x) = \det\left(\begin{smallmatrix} x-a & -1 \\ 0 & x-a \end{smallmatrix}\right) = (x - a)^2.$$

[11] Oft wird das Polynom $x \mapsto \det(\varphi - x\,\mathrm{Id}_V)$ als das charakteristische Polynom bezeichnet. Dieses unterscheidet sich von $\chi_\varphi$ um den Faktor $(-1)^n$.

In Beispiel 1. entsprechen die Nullstellen von $\chi_A$ gerade den Diagonaleinträgen $a_1, \ldots, a_n$. Das mag zunächst nicht verwundern, ist aber doch bemerkenswert. Denn nach Basiswechsel ist die transformierte Matrix $TAT^{-1}$ im Allgemeinen nicht mehr diagonal, und man sieht ihr die Nullstellen des charakteristischen Polynoms nicht mehr an.

Damit stellt sich die Frage nach der basisunabhängigen Bedeutung der Nullstellen.

Für eine Nullstelle $\lambda \in K$ des charakteristischen Polynoms eines beliebigen Endomorphismus $\varphi \in L(V)$ gilt $\det(\lambda \operatorname{Id}_V - \varphi) = 0$ oder gleichbedeutend

$$\operatorname{def}(\lambda \operatorname{Id}_V - \varphi) \geq 1.$$

Im Sinn der folgenden Definition ist damit $\lambda$ Eigenwert von $\varphi$.

**12.3 Definition** *Ist $V$ ein $K$-Vektorraum, $\varphi \in L(V)$, und*
• *es gelte für einen vom Nullvektor verschiedenen Vektor $v \in V$ und ein $\lambda \in K$*

$$\varphi(v) = \lambda\, v.$$

*$v$ heißt dann* **Eigenvektor** *von $\varphi$ zum* **Eigenwert** *$\lambda$.*
• *Das* **Spektrum** *von $\varphi$ ist die Menge*

$$\operatorname{Spec}(\varphi) := \{\lambda \in K \mid \varphi - \lambda \operatorname{Id}_V \textit{ ist kein Isomorphismus}\}.$$

• *Ist $\dim(V) < \infty$, dann heißt $\varphi$* **diagonalisierbar**, *wenn es eine Basis von $V$ gibt, die aus Eigenvektoren von $\varphi$ besteht.*

Analoge Definitionen werden für quadratische Matrizen verwandt.

**12.4 Satz** *Ist $V$ ein $K$-Vektorraum, $\varphi \in L(V)$*
• *und $\lambda \in K$ Eigenwert von $\varphi$, dann gilt $\lambda \in \operatorname{Spec}(\varphi)$.*
• *Ist $\dim(V) < \infty$, dann gilt*

$$\operatorname{Spec}(\varphi) = \{\lambda \in K \mid \lambda \textit{ ist Eigenwert von } \varphi\} = \{\lambda \in K \mid \chi_\varphi(\lambda) = 0\}.$$

**Beweis:** • Für einen Eigenvektor $v$ zum Eigenwert $\lambda$ gilt

$$(\varphi - \lambda \operatorname{Id}_V)v = 0$$

also $v \in \ker(\varphi - \lambda \operatorname{Id}_V)$. Damit ist $\lambda \in \operatorname{Spec}(\varphi)$.
• Ist ein Endomorphismus $\rho \in L(V)$ eines endlichdimensionalen Vektorraums

nicht bijektiv, dann ist er auch nicht injektiv, d.h. $\mathrm{def}(\rho) > 0$. Anwendung auf $\rho := \varphi - \lambda \operatorname{Id}_V$ zeigt, dass jeder Punkt des Spektrums auch Eigenwert ist.
• $\rho := \varphi - \lambda \operatorname{Id}_V$ ist genau dann ein Isomorphismus, wenn $\det(\rho) \neq 0$, also $\chi_\varphi(\lambda) \neq 0$. □

In unendlichdimensionalen Vektorräumen kann das Spektrum aber reichhaltiger sein als die Menge der Eigenwerte:

**12.5 Beispiel (Spektrum)** Es sei $V := C([0,1], \mathbb{R})$ der $\mathbb{R}$-Vektorraum der stetigen Funktionen auf dem Intervall $[0,1]$ und $\varphi \in L(V)$ durch

$$\varphi(f)(x) := x\, f(x) \qquad \big(x \in [0,1]\big)$$

definiert. $\varphi$ multipliziert also $f \in V$ mit der Funktion $x \mapsto x$.
• Es gilt $\mathrm{Spec}(\varphi) \supset [0,1]$, denn für $\lambda \in [0,1]$ enthält das Bild $\operatorname{im}(\varphi - \lambda \operatorname{Id}_V)$ nur Funktionen $g \in V$ mit $g(\lambda) = 0$.
• Es gilt $\mathrm{Spec}(\varphi) \subset [0,1]$, denn für $\lambda \in \mathbb{R} \setminus [0,1]$ ist $\tau \in L(V)$ mit

$$\tau(g)(x) := \frac{g(x)}{x - \lambda} \qquad \big(x \in [0,1]\big)$$

Umkehrabbildung von $\varphi - \lambda \operatorname{Id}_V$.
• Wäre ein $\lambda \in [0,1]$ Eigenwert mit Eigenvektor $f \in V$, dann würde für alle $x \in [0,1]$ gelten, dass $(x - \lambda) \cdot f(x) = 0$. Dies ist für $x \in [0,1] \setminus \{\lambda\}$ nur möglich, wenn $f(x) = 0$ gilt. Dann muss wegen Stetigkeit von $f$ aber auch $f(\lambda) = 0$ gelten, sodass $f = 0$ gilt, $f$ also kein Eigenvektor ist.
Während also $\mathrm{Spec}(\varphi) = [0,1]$ ist, besitzt $\varphi$ keinen Eigenwert.

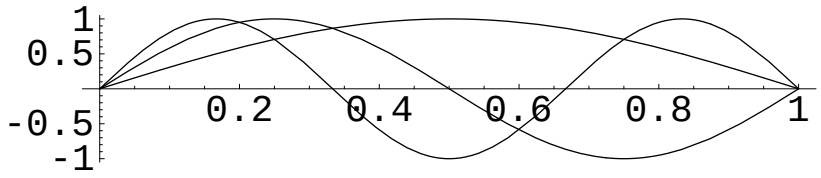
Die Berechnung des Spektrums, und insbesondere etwaiger Eigenwerte, besitzt eine große Bedeutung in der Angewandten Mathematik.

**12.6 Beispiel (Spektrum)** Es ist $V := \{f \in C^\infty(\mathbb{R}, \mathbb{R}) \mid f(x+2) = f(x)\}$ der (unendlichdimensionale) Vektorraum der glatten reellen Funktionen mit Periode 2, und

$$\varphi \in L(V) \quad , \quad \varphi(f)(x) := -\frac{d^2 f}{dx^2}(x) \qquad (x \in \mathbb{R}).$$

Dann ist $\lambda_n := (\pi n)^2$ Eigenwert zum Eigenvektor

$$f_n \in V \quad , \quad f_0(x) := 1 \quad , \quad f_n(x) := \sin(n\pi x) \qquad (n \in \mathbb{N}).$$




Nun reicht es aus, eine Funktion $f \in V$ im Intervall $[0, 2]$ zu kennen. Die angegebenen Eigenvektoren $f_n$, $n \in \mathbb{N}$, verschwinden nicht nur an den Intervallenden, sondern es ist auch $f_n(1) = 0$. Stellen wir uns die Graphen von Funktionen $f \in C^\infty([0, 1], \mathbb{R})$ mit $f(0) = f(1) = 0$ als mögliche (eindimensionale) Auslenkungen einer bei $x = 0$ und $x = 1$ eingespannten Saite vor, dann entspricht $f_1$ der Grundschwingung, $f_2$ der ersten Oberschwingung der Saite etc, und die $\sqrt{\lambda_n} = \pi n$ entsprechen den Frequenzen dieser Schwingungen.

Nicht alle Endomorphismen endlichdimensionaler Vektorräume sind diagonalisierbar:

- Die in Beispiel 12.2.1. angesprochenen Diagonalmatrizen sind offensichtlich im Sinn der obigen Definition diagonalisierbar, nämlich mit der kanonischen Basis $e_1, \ldots, e_n$ von Eigenvektoren.

- Dagegen sind die Drehmatrizen aus Beispiel 12.2.2. für einen von $0$ und $\pi$ verschiedenen Drehwinkel nicht diagonalisierbar, denn dann besitzt das charakteristische Polynom $\chi_A$ keine (reellen) Eigenwerte.

- Die Matrix $A = \left(\begin{smallmatrix} a & 1 \\ 0 & a \end{smallmatrix}\right)$ aus Beispiel 12.2.3. besitzt den Eigenwert $a$ mit Eigenvektor $e_1 = \binom{1}{0}$. Allerdings existiert kein weiterer linear unabhängiger Eigenvektor $b \in \mathbb{R}^2$. Denn wegen der linearen Unabhängigkeit könnten wir diesen in der Form $b = \binom{b_1}{1}$ mit zu bestimmendem Parameter $b_1 \in \mathbb{R}$ ansetzen. Als Eigenwert kommt nur $a$ in Frage. Damit müsste die Gleichung

$$ab_1 + 1 = ab_1$$

  erfüllt sein, was offensichtlich unmöglich ist.

Diagonalisierbare Endomorphismen sind besonders einfach, denn die Vektoren in den *Eigenräumen* $\ker(\varphi - \lambda \cdot \mathrm{Id}_V)$ zu den Eigenwerten $\lambda \in K$ werden einfach mit dem Faktor $\lambda$ multipliziert. Daher entwickeln wir im Folgenden Kriterien für die Diagonalisierbarkeit.

**12.7 Definition** $\lambda \in K$ *heißt* **Nullstelle** *des Polynoms* $p \in K[x]$, *wenn* $p(\lambda) = 0$ *gilt.* $\lambda \in K$ *heißt* $r$**–fache Nullstelle**, *wenn* $q \in K[x]$ *mit* $p(x) = (x-\lambda)^r\, q(x)$ *und* $q(\lambda) \neq 0$ *existiert.* $r$ *heißt dann die* **Vielfachheit** *oder* **Multiplizität** *von* $\lambda$.

In Beispiel 12.2.3. war $a$ zweifache Nullstelle von $\chi_A$.

**12.8 Definition** *Ist* $\lambda \in K$ *eine* $r$*–fache Nullstelle des charakteristischen Polynoms* $\chi_\varphi$, *dann heißt* $r$ *die* **algebraische Multiplizität** *und* $\operatorname{def}(\varphi - \lambda \operatorname{Id}_V)$ *die* **geometrische Multiplizität** *von* $\lambda$.

# 13 Euklidische und unitäre Vektorräume

Längen und Winkel sind geometrische Grundbegriffe. Wir wollen diese nun für eine möglichst große Klasse von Vektorräumen einführen und untersuchen.

Dazu betrachten wir zunächst die vertraute Situation des Vektorraums $\mathbb{R}^2$. Definieren wir das *Skalarprodukt* zweier Vektoren $v = \binom{v_1}{v_2}$ und $w = \binom{w_1}{w_2}$ durch

$$\langle v, w\rangle_{\mathbb{R}^2} := v_1 w_1 + v_2 w_2, \tag{13.1}$$

dann ist nach dem Satz von Pythagoras die *Länge* von $v$ gleich

$$\|v\| := \sqrt{\langle v, v\rangle_{\mathbb{R}^2}} = \sqrt{v_1^2 + v_2^2}.$$

Nur der Nullvektor besitzt also Länge $0$, und wir nehmen $v \neq 0 \neq w$ an. Dann können wir

$$v = \|v\| \cdot \left(\begin{smallmatrix}\cos\varphi\\ \sin\varphi\end{smallmatrix}\right) \quad \text{und} \quad w = \|w\| \cdot \left(\begin{smallmatrix}\cos\psi\\ \sin\psi\end{smallmatrix}\right)$$

mit eindeutig bestimmten Winkeln $\varphi, \psi \in [0, 2\pi)$ schreiben, und der *Winkel* $\rho := \psi - \varphi$ zwischen $v$ und $w$ ergibt sich wegen des Additionstheorems $\cos(\rho) = \cos(\varphi)\cos(\psi) + \sin(\varphi)\sin(\psi)$ aus

$$\cos(\rho) = \frac{\langle v, w\rangle_{\mathbb{R}^2}}{\|v\| \cdot \|w\|}.$$

Nun können wir ja den $\mathbb{R}^2$ mit $\mathbb{C}$ identifizieren, wenn wir $v$ auf $v_1 + iv_2$ abbilden. Das Skalarprodukt (13.1) von $v$ und $w \in \mathbb{C}$ ist dann gleich dem Realteil von

$$\langle v, w\rangle_{\mathbb{C}} := \bar{v}w = v_1 w_1 + v_2 w_2 + i(v_1 w_2 - v_2 w_1),$$

und $\|v\| = \sqrt{\langle v, v\rangle_{\mathbb{C}}}$. Beide Aussagen würden nicht gelten, wenn wir $\langle v, w\rangle_{\mathbb{C}}$ durch $vw$ statt durch $\bar{v}w$ definiert hätten. Es war also wichtig, die komplexe Konjugation

$$\kappa : \mathbb{C} \to \mathbb{C} \quad , \quad w \mapsto \bar{w}$$

zu benutzen, also einen Körperautomorphismus von $\mathbb{C}$.

Sowohl beim $\mathbb{R}^2$ als auch im Fall des eindimensionalen komplexen Vektorraumes $\mathbb{C}$ sind wir von einem Skalarprodukt ausgegangen, um Länge und Winkel zu definieren. Skalarprodukte sind aber Sonderfälle von so genannten Sesquilinearformen, die wir jetzt betrachten, um unsere Begriffe zu verallgemeinern.

**13.1 Definition** • *Eine bijektive Abbildung $\kappa : K \to K$ eines Körpers $K$ heißt* **(Körper-) Automorphismus**, *wenn gilt:*

$$\kappa(k_1+k_2) = \kappa(k_1)+\kappa(k_2) \quad \textit{und} \quad \kappa(k_1\cdot k_2) = \kappa(k_1)\cdot\kappa(k_2) \qquad (k_1, k_2 \in K).$$

• *Es sei $V$ ein $K$–Vektorraum und $\kappa : K \to K$ ein Körperautomorphismus. Dann heißt eine Abbildung*

$$\Phi : V \times V \to K$$

**Sesquilinearform**[12] *bezüglich $\kappa$, wenn für $v, v', w, w' \in V$ und $k \in K$ gilt:*

$$\begin{aligned}
\Phi(v + v', w) &= \Phi(v, w) + \Phi(v', w),\\
\Phi(v, w + w') &= \Phi(v, w) + \Phi(v, w'),\\
\Phi(k \cdot v, w) &= \kappa(k) \cdot \Phi(v, w) \quad \textit{und}\\
\Phi(v, k \cdot w) &= k \cdot \Phi(v, w).
\end{aligned}$$

Ist der Körperautomorphismus trivial, d.h. $\kappa = \mathrm{Id}_K$, dann wird die Sesquilinearform zu der uns schon vertrauten zweifachen Linearform oder auch *Bilinearform*.

Es reicht aus, die Werte einer Sesquilinearform auf Paaren von Basisvektoren zu kennen. Ist speziell $n := \dim(V) < \infty$, dann setzen wir für eine Basis $B = (b_1, \ldots, b_n)$ von $V$ die *darstellende Matrix* $A \in \mathrm{Mat}(n, k)$ der Sesquilinearform gleich

$$(A)_{i,k} := \Phi(b_i, b_k) \qquad (i, k = 1, \ldots, n).$$

Dann gilt für zwei beliebige Vektoren $v = \sum_{i=1}^n v_i b_i$ und $w = \sum_{k=1}^n w_k b_k$ aus $V$:

$$\Phi(v, w) = \sum_{i,k=1}^{n} \Phi(v_i b_i, w_k b_k) = \sum_{i,k=1}^{n} \kappa(v_i)(A)_{i,k} w_k. \tag{13.2}$$

[12] Sesquilinear bedeutet anderthalbfach linear.

Ist umgekehrt $A \in \mathrm{Mat}(n,k)$ eine beliebige Matrix, dann definiert die rechte Seite von (13.2) eine Sesquilinearform $\Phi$. Wir wollen Sesquilinearformen dazu benutzen, um Längen und Winkel, also reelle Zahlen zu definieren. Daher beschränken wir uns im restlichen Kapitel auf den Körper $\mathbb{R}$ und seinen Oberkörper $\mathbb{C}$, d.h.

$$K \in \{\mathbb{R}, \mathbb{C}\}.$$

Im Fall $K = \mathbb{R}$ verwenden wir $\kappa = \mathrm{Id}_{\mathbb{R}}$, d.h. untersuchen Bilinearformen, während wir für $K = \mathbb{C}$ aus den schon erwähnten Gründen die Konjugation $\kappa(w) = \bar{w}$ benutzen müssen.

**13.2 Definition** *Eine $\kappa$–Sesquilinearform $\Phi$ auf $V$ über $K \in \{\mathbb{R}, \mathbb{C}\}$ heißt*

- **hermitesch**, *wenn gilt*

$$\Phi(w,v) = \kappa(\Phi(v,w)) \qquad (v,w \in V),$$

- **antihermitesch**, *wenn gilt*

$$\Phi(w,v) = -\kappa(\Phi(v,w)) \qquad (v,w \in V).$$

- $\Phi$ *heißt* **Skalarprodukt** *auf $V$, wenn $\Phi$ hermitesch und* **positiv definit** *ist, d.h.*

$$\Phi(v,v) > 0 \qquad (v \in V \backslash \{0\}).$$

  *Statt $\Phi(v,w)$ schreiben wir dann $\langle v,w \rangle$.*

- *Ein $K$–Vektorraum $V$ mit Skalarprodukt $\langle \cdot,\cdot \rangle$ heißt für $K = \mathbb{R}$* **euklidischer Vektorraum**, *für $K = \mathbb{C}$* **unitärer Vektorraum**.

Im Fall $K = \mathbb{R}$ heißt eine (anti-)hermitesche Bilinearform auch (anti-)symmetrisch.

**13.3 Beispiele (Sesquilinearformen)** 1. Im Fall $V = \mathbb{R}^n$ ist $\Phi(v,w) := v^t A w$ genau dann hermitesch (bzw. antihermitesch), wenn $A^t = A$ (bzw. $A^t = -A$), die darstellende Matrix also (*anti-*)*symmetrisch* ist.

2. Im Fall $V = \mathbb{C}^n$ ist $\Phi(v,w) := \bar{v}^t A w$ (mit $\bar{v} := \begin{pmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_n \end{pmatrix}$) genau dann hermitesch (bzw. antihermitesch), wenn die Matrix $A \in \mathrm{Mat}(n,\mathbb{C})$ selbst *hermitesch* (bzw. *antihermitesch*) ist, d.h. $A^* = A$ (bzw. $A^* = -A$) mit $A^* := \bar{A}^t$ und $(\overline{A})_{i,k} := \overline{(A)_{i,k}}$ gilt.

3. Beispiele für Skalarprodukte auf dem $K^n$ sind die *kanonischen Skalarprodukte*

$$\langle v, w\rangle := \sum_{i=1}^{n} v_i w_i \qquad (v, w \in \mathbb{R}^n)$$

und

$$\langle v, w\rangle := \sum_{i=1}^{n} \bar{v}_i w_i \qquad (v, w \in \mathbb{C}^n).$$

Diese verallgemeinern die zu Beginn des Kapitels diskutierten Skalarprodukte auf $\mathbb{R}^2$ bzw. $\mathbb{C}$.

4. Wir betrachten den $\mathbb{R}$–Vektorraum $V := C([a, b], \mathbb{R})$ der auf dem Intervall $[a, b]$ (mit $a < b$) stetigen Funktionen.

$$\langle v, w\rangle := \int_a^b v(x)w(x)\,dx \qquad (v, w \in V) \tag{13.3}$$

definiert eine symmetrische Bilinearform. Diese ist positiv definit, denn für $v \neq 0$ gibt es ein $x_0 \in [a, b]$ mit $\varepsilon := |v(x_0)| > 0$ und wegen Stetigkeit von $v$ ein $\delta > 0$ mit

$$|v(x) - v(x_0)| < \frac{\varepsilon}{2} \quad \text{für} \quad |x - x_0| < \delta.$$

Wählen wir nun ein $\delta$, das kleiner als die Intervallbreite $b - a$ ist, dann gilt

$$\langle v, v\rangle = \int_a^b |v(x)|^2\,dx \geq \delta \cdot (\varepsilon/2)^2 > 0.$$

Damit definiert (13.3) ein Skalarprodukt auf $V$.

Ähnlich können wir auf dem $\mathbb{C}$–Vektorraum

$$V := C([a, b], \mathbb{C})$$

durch

$$\langle v, w\rangle := \int_a^b \bar{v}(x)w(x)\,dx \qquad (v, w \in V)$$

ein Skalarprodukt einführen.

**13.4 Definition** *Es sei* $\|v\| := \sqrt{\langle v, v\rangle}$.

**13.5 Satz (Cauchy–Schwarz–Ungleichung)** *Für einen euklidischen oder unitären Vektorraum $(V, \langle\cdot,\cdot\rangle)$ gilt*

$$|\langle v, w\rangle| \leq \|v\| \cdot \|w\| \qquad (v, w \in V).$$

**Beweis:**

- Falls einer der Vektoren der Nullvektor ist, sind beide Seiten $0$, und die Behauptung ist wahr.
- Sonst setzen wir $k := \frac{\langle w,v\rangle}{\langle w,w\rangle} \in K$. Für $K = \mathbb{C}$ bezeichnet $\overline{k}$ die komplexkonjugierte Zahl, für $K = \mathbb{R}$ ist $\overline{k} = k$. Es gilt

  $$0 \leq \langle v - k \cdot w, v - k \cdot w\rangle = \langle v, v\rangle - \overline{k}\langle w, v\rangle - k\langle v, w\rangle + k\overline{k}\langle w, w\rangle .$$

  Multiplizieren wir diese Ungleichung mit $\langle w, w\rangle > 0$, dann ergibt sich

  $$0 \leq \langle v, v\rangle\langle w, w\rangle - \overline{\langle w, v\rangle}\langle w, v\rangle - \langle w, v\rangle\langle v, w\rangle + \langle w, v\rangle\langle v, w\rangle ,$$

  also

  $$|\langle v, w\rangle|^2 \leq \langle v, v\rangle\langle w, w\rangle . \qquad \square$$

**13.6 Definition** • *Eine Abbildung $\|\cdot\| : V \to \mathbb{R}$ eines $k$–Vektorraumes $V$ heißt* **Länge** *oder* **Norm**, *wenn für $v, w \in V$ und $k \in K$ gilt*

1. $\|v\| \geq 0$, $\|v\| = 0 \Longleftrightarrow v = 0$
2. $\|kv\| = |k|\|v\|$
3. $\|v + w\| \leq \|v\| + \|w\|$

*Die letzte Ungleichung heißt* **Dreiecksungleichung**.
• $\|v\|$ *heißt die* **Norm** *(auch* **Länge** *oder* **Betrag***)* **von** $v$.
• *Ist $\|\cdot\| : V \to \mathbb{R}$ eine Norm auf $V$, dann heißt $(V, \|\cdot\|)$* **normierter Vektorraum**.

**13.7 Satz** $\|v\| = \sqrt{\langle v, v\rangle}$ *ist eine Norm.*

**Beweis:**

1. Das Skalarprodukt ist positiv definit.

2. $\|kv\| = \sqrt{\langle kv, kv\rangle} = \sqrt{\overline{k}k\,\langle v, v\rangle} = \sqrt{\overline{k}k}\sqrt{\langle v, v\rangle} = |k|\|v\|$.

3. $\|v + w\|^2 = \langle v + w, v + w\rangle = \langle v, v\rangle + \langle w, w\rangle + \langle w, v\rangle + \langle v, w\rangle$
$\leq \|v\|^2 + \|w\|^2 + 2\|v\| \cdot \|w\| = (\|v\| + \|w\|)^2$. □

Allerdings kommt nicht jede Norm von einem Skalarprodukt:

**13.8 Beispiel (Norm)** $V = \mathbb{R}^2$, $\|\binom{v_1}{v_2}\| := \max(|v_1|, |v_2|)$.

**13.9 Definition** *Es sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder unitärer Vektorraum.*

- *Zwei Vektoren $v, w \in V$ heißen* **orthogonal** *$(v \perp w)$, wenn $\langle v, w\rangle = 0$ gilt.*
- *Zwei Teilmengen $M, N \subset V$ heißen* **orthogonal** *$(M \perp N)$, wenn alle $v$ und $w$ orthogonal sind, falls $v \in M$ und $w \in N$.*
- *Das* **orthogonale Komplement** *einer Teilmenge $M \subset V$ ist*

$$M^\perp := \{v \in V \mid \forall w \in M : v \perp w\}.$$

Die leere Menge und der Nullraum stehen damit senkrecht (orthogonal) zu allen Vektoren.

**13.10 Satz** • *Für alle $M \subset V$ ist $M^\perp$ ein Unterraum, und $(M^\perp)^\perp \supset M$.*

**Beweis:**

- Ist $v \in M^\perp$ und $k \in K$, dann ist für $w \in M$

$$\langle w, kv\rangle = k\,\langle w, v\rangle = 0.$$

- Sind $v_1, v_2 \in M^\perp$, dann ist $\langle w, v_1 + v_2\rangle = \langle w, v_1\rangle + \langle w, v\rangle = 0$.
- Weiterhin ist $0 \in M^\perp$.

Damit ist $M^\perp$ ein Unterraum. Da $M \perp M^\perp$ gilt, ist $M \subset (M^\perp)^\perp$. □

# Literatur


[1] Blatter, Ch.: Analysis 1. Springer

[2] Brieskorn, E.: Lineare Algebra und analytische Geometrie. Vieweg 1982

[3] Ebbinghaus H.D. et al: Zahlen. Springer 1992

[4] Fischer, G.: Lineare Algebra. Vieweg 1995

[5] Kostrikin, A.I.; Manin, Y.I.: Linear algebra and geometry. Gordon & Breach, 1997

[6] Kowalsky, H.-J.: Einführung in die Lineare Algebra. de Gruyter 1971

# Index