Skript zur Vorlesung

Lineare Algebra

für Grund-, Mittel- und Förderschullehramt

Dr. Jan-David Hardtke

Universität Leipzig
Institut für Mathematik

Stand: Sommersemester 2018

Vorbemerkung:

Dies ist eine vorläufige Version des Vorlesungsskriptes. Der bestehende Text wird im Laufe der Zeit um weitere Kapitel und Anhänge ergänzt. Daher kann es auch passieren, dass im Text schon auf eventuell noch nicht existierende Teile des Skriptes verwiesen wird.

Falls Sie etwaige Tippfehler oder auch inhaltliche Fehler bemerken, senden Sie mir diese bitte per Email an hardtke@math.uni-leipzig.de.

Jan-David Hardtke, 20. März 2019

Inhaltsverzeichnis

I Mengen und Abbildungen

Der Mengenbegriff und der Begriff einer Abbildung (Funktion) zwischen zwei Mengen sind grundlegend nicht nur für die lineare Algebra, sondern für die gesamte Mathematik. Daher soll in diesem einleitenden Kapitel kurz das Nötigste zum Thema Mengen und Abbildungen zusammengestellt werden, wobei, im Interesse der Kürze und Einfachheit, die Diskussion an einigen Stellen bewusst etwas informal gehalten ist.

I.1 Grundlegendes über Mengen

Unter einer Menge verstehen wir hier einfach die Zusammenfassung gewisser mathematischer Objekte zu einem neuen mathematischen Objekt. Die Ausgangsobjekte bilden dabei die sogenannten Elemente der Menge. Bei diesen kann es sich z. B. um natürliche, rationale oder reelle Zahlen, aber auch um gänzlich andere Objekte handeln. So können etwa die Elemente einer Menge auch selbst wieder Mengen sein.

Um auszudrücken, dass ein Objekt x Element einer Menge A ist, schreiben wir $x \in A$, anderenfalls $x \notin A$.

Zwei Mengen A und B sind gleich ($A = B$), falls sie dieselben Elemente haben, d. h. falls jedes Element von A auch ein Element von B und umgekehrt jedes Element von B auch ein Element von A ist.

Mengen werden häufig über Eigenschaften ihrer Elemente definiert. Ist $\mathcal{E}$ eine mathematische Eigenschaft[1], so bezeichnet

$$\{x : x \text{ hat die Eigenschaft } \mathcal{E}\}$$

die Menge aller x mit der Eigenschaft $\mathcal{E}$.

Ist M eine bereits vorgegebene Menge, so schreibt man kurz

$$\{x \in M : x \text{ hat die Eigenschaft } \mathcal{E}\}$$

für die Menge

$$\{x : x \in M \text{ und } x \text{ hat die Eigenschaft } \mathcal{E}\}.$$

[1] Ich vermeide hier bewusst eine Präzisierung, in der Praxis wird man (hoffentlich) schnell verstehen, was gemeint ist.

Einige konkrete Beispiele: Bezeichnen wir wie üblich die Mengen der natürlichen, rationalen und reellen Zahlen[2] mit $\mathbb{N}$, $\mathbb{Q}$ und $\mathbb{R}$, so steht

$$\{n \in \mathbb{N} : n > 5\}$$

für die Menge aller natürlichen Zahlen größer als 5,

$$\{n \in \mathbb{N} : \text{es existiert ein } k \in \mathbb{N} \text{ mit } n = 2k\}$$

ist die Menge aller geraden Zahlen und

$$\left\{x \in \mathbb{R} : x^2 \in \mathbb{Q}\right\}$$

bezeichnet die Menge aller reellen Zahlen, deren Quadrat rational ist.

Als Nächstes kommen wir zum wichtigen Begriff der Teilmengen.

Definition I.1.1. Sind A und B zwei Mengen, so heißt A eine *Teilmenge* von B (in Zeichen: $A \subseteq B$), falls jedes Element von A auch ein Element von B ist.[3]

Das obige Gleichheitskriterium für Mengen liest sich damit kürzer wie folgt: Für alle Mengen A und B gilt[4]

$$A = B \;\Leftrightarrow\; (A \subseteq B \text{ und } B \subseteq A).$$

Wir werden dieses Kriterium zum Beispiel unten im Beweis von Lemma I.1.4 anwenden. Zuvor noch einige weitere Definitionen.

Definition I.1.2. Die *leere Menge* ist diejenige Menge, welche keine Elemente enthält. Sie wird mit $\emptyset$ bezeichnet.

Für jedes mathematische Objekt a bezeichne $\{a\}$ diejenige Menge, die a als einziges Element enthält. $\{a\}$ heißt die *Einermenge* mit Element a.

In der obigen "Eigenschaftenschreibweise" ist z. B. $\{a\} = \{x : x = a\}$.

Als kleine Übung mache man sich klar, dass die Mengen $\emptyset$, $\{\emptyset\}$ und $\{\{\emptyset\}\}$ jeweils voneinander verschieden sind.

Wir definieren als Nächstes zwei wichtige Operationen mit Mengen.

[2] Diese Zahlenbereichen werden offiziell erst später eingeführt (siehe Kapitel II), sind Ihnen aber sicherlich schon aus der Schule hinlänglich vertraut.

[3] Eine kleine Warnung hinsichtlich der Schreibweise: Manche Autoren schreiben $A \subset B$ anstelle von $A \subseteq B$, bei wieder anderen steht $A \subset B$ jedoch für eine *echte* Teilmenge, also für $A \subseteq B$ und $A \neq B$. Wir werden hier nur die Schreibweise $A \subseteq B$ verwenden und ggf. $A \neq B$ explizit dazu schreiben.

[4] Das Symbol $\Leftrightarrow$ bedeutet "genau dann, wenn", siehe Anhang A.1 zur Erklärung der Logiksymbole.

Definition I.1.3. Für zwei Mengen A und B definieren wir die *Vereinigung* von A und B durch[5]

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

und den *Durchschnitt* von A und B durch

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

A und B heißen *disjunkt*, falls $A \cap B = \emptyset$ gilt, d. h. falls A und B keine gemeinsamen Elemente haben.

Ausgehend von Einermengen können wir durch Vereinigung "größere" Mengen erzeugen. So definieren wir Paarmengen $\{a, b\}$ durch $\{a, b\} := \{a\} \cup \{b\}$, Dreiermengen durch $\{a, b, c\} := \{a, b\} \cup \{c\}$ und so fort. Dabei bezeichnen $a, b, c, \ldots$ beliebige mathematische Objekte, die nicht notwendig verschieden sein müssen. Ist z. B. $a = b$, so ist $\{a, b\} = \{a\}$. Auch die Reihenfolge der Elemente spielt keine Rolle, z. B. ist $\{a, b\} = \{b, a\}$ und $\{a, b, c\} = \{c, a, b\}$.

Hier noch ein paar konkrete Beispiele: Es ist $\{1, 2, 3\} \cup \{2, 4\} = \{1, 2, 3, 4\}$, $\{1, 2, 3\} \cap \{2, 4\} = \{2\}$ und $\{1, 3\} \cap \{2, 4\} = \emptyset$.

Als Nächstes stellen wir einige allgemeine "Rechenregeln" für Vereinigung und Durchschnitt zusammen.

Lemma I.1.4. *Für alle Mengen A, B und C gilt:*

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

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

(iii) $A \cup B = B \cup A$

(iv) $A \cap B = B \cap A$

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

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

Beweis. Wir beweisen nur exemplarisch die Aussage (v). Die übrigen Beweise sind den Leserinnen und Lesern selbst zur Übung überlassen.
Zum Beweis verwenden wir das obige Gleichheitskriterium für Mengen. Wir haben also $(A \cup B) \cap C \subseteq (A \cap C) \cup (B \cap C)$ und $(A \cap C) \cup (B \cap C) \subseteq (A \cup B) \cap C$ zu zeigen.
1) Beweis von $(A \cup B) \cap C \subseteq (A \cap C) \cup (B \cap C)$.
Sei $x \in (A \cup B) \cap C$. Dann ist $x \in A \cup B$ und $x \in C$.

[5] Hier und im Folgenden bedeutet die Schreibweise := eine Gleichheit per definitionem, d. h. das Objekt, welches links von := steht, wird durch das rechts von := stehende Objekt definiert.

Wegen $x \in A \cup B$ gilt $x \in A$ oder $x \in B$. Im ersten Fall folgt wegen $x \in C$ auch $x \in A \cap C$, im zweiten Fall folgt analog $x \in B \cap C$. Also gilt in jedem Fall $x \in (A \cap C) \cup (B \cap C)$.
2) Beweis von $(A \cap C) \cup (B \cap C) \subseteq (A \cup B) \cap C$.
Sei $x \in (A \cap C) \cup (B \cap C)$. Dann ist $x \in A \cap C$ oder $x \in B \cap C$.
Im ersten Fall ist $x \in A$ und $x \in C$, also auch $x \in A \cup B$ und $x \in C$, also $x \in (A \cup B) \cap C$.
Im zweiten Fall ist $x \in B$ und $x \in C$, folglich auch $x \in A \cup B$ und $x \in C$, also $x \in (A \cup B) \cap C$. Damit ist der Beweis abgeschlossen. □

Wir definieren nun noch die Differenz zweier Mengen.

Definition I.1.5. Sind A und B zwei Mengen, so heißt die Menge

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

die *Differenzmenge* von A und B.

Man beachte, dass bei dieser Definition nicht unbedingt $B \subseteq A$ vorausgesetzt ist. Beispielsweise ist $\{1, 2, 3\} \setminus \{1, 4\} = \{2, 3\}$.

Schließlich kommen wir noch zum Begriff der geordneten Paare. Wir hatten oben schon bemerkt, dass für Paarmengen $\{a, b\} = \{b, a\}$ gilt. Manchmal will man aber zwei Objekte auch unter Berücksichtigung der Reihenfolge zu einem neuen Objekt zusammenfassen. Dazu dient der Begriff der geordneten Paare.

Definition I.1.6. Für zwei mathematische Objekte a und b definieren wir das *geordnete Paar* (a, b) durch $(a, b) := \{\{a\}, \{a, b\}\}$.

Es gilt dann das folgende Gleichheitskriterium (das war der Sinn der Definition).

Lemma I.1.7. *Für alle mathematischen Objekte a, b, c, d gilt:*

$$(a, b) = (c, d) \;\Leftrightarrow\; a = c \text{ und } b = d.$$

Beweis. Die Schlussrichtung "$\Leftarrow$" ist klar. Wir zeigen nun "$\Rightarrow$".
Sei also $(a, b) = (c, d)$. Dann ist insbesondere $\{a\} \in \{\{c\}, \{c, d\}\}$, also $\{a\} = \{c\}$ oder $\{a\} = \{c, d\}$, woraus in jedem Fall $a = c$ folgt.
Weiter ist auch $\{a, b\} \in \{\{c\}, \{c, d\}\} = \{\{a\}, \{a, d\}\}$ (die letzte Gleichheit folgt aus der schon bewiesenen Tatsache $a = c$). Wir unterscheiden zwei Fälle.
1) Ist $a = b$, so folgt $(a, b) = (b, b) = \{\{b\}\}$. Wegen $(c, d) = (a, b)$ folgt daher $\{c, d\} = \{b\}$, also $d = b$.
2) Ist $a \neq b$, so folgt aus der oben beobachteten Tatsache $\{a, b\} \in \{\{a\}, \{a, d\}\}$, dass $\{a, b\} = \{a, d\}$ sein muss. Also ist $b \in \{a, d\}$, aber $a \neq b$, also $b = d$.
□

Für drei Objekte a, b, c definiert man das geordnete Tripel durch $(a, b, c) := ((a, b), c)$. Dann gilt offenbar $(a, b, c) = (d, e, f)$ genau dann, wenn $a = d$, $b = e$ und $c = f$ ist.

Entsprechend werden Vierertupel (Quadrupel) (a, b, c, d) erklärt durch $(a, b, c, d) := ((a, b, c), d)$ und es gilt ein analoges Gleichheitskriterium. Ebenso verfährt man für Fünfertupel, etc.

Auch geordnete Paare lassen sich natürlich wieder zu neuen Mengen zusammenfassen.

Definition I.1.8. Für zwei Mengen A und B ist ihr *kartesisches Produkt*[6] definiert durch

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

Hierzu eine kleine Bemerkung: Die obige Definition müsste eigentlich ausführlich

$$A \times B := \{x : \text{es existieren ein } a \in A \text{ und ein } b \in B \text{ mit } x = (a, b)\}$$

lauten. Allerdings verwendet man in solchen Fällen häufig abkürzende Schreibweisen wie die obige. In der Praxis sollte recht schnell klar werden, was jeweils gemeint ist.

Beispiel: $\{1, 2\} \times \{1, 2, 3\} = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3)\}$.

Natürlich kann man auch Produkte von mehr als zwei Mengen definieren. Für drei Mengen A, B, C setzt man entsprechend $A \times B \times C := \{(a, b, c) : a \in A, b \in B, c \in C\}$, usw.

I.2 Grundlegendes über Abbildungen

Wir kommen nun zum Begriff der Abbildungen (Funktionen). In den Beispielen werden wir dabei im Vorgriff schon einige elementare Funktionen (wie z. B. die Quadratfunktion und die Wurzelfunktion) verwenden, die offiziell erst später eingeführt werden, Ihnen aber sicherlich schon aus der Schule hinreichend bekannt sind, um damit zu arbeiten.

Hier nun die Definition:

Definition I.2.1. Seien A und B zwei Mengen. Eine *Abbildung* oder *Funktion* von A nach B ist ein Tripel (A, B, f), wobei f eine Zuordnungsvorschrift ist, die jedem Element $a \in A$ genau ein Element $f(a) \in B$ zuweist.
$f(a)$ heißt der Wert der Funktion an der Stelle a.
A heißt der *Definitionsbereich* und B der *Wertebereich* der Funktion.

[6] Benannt nach René Descartes (1596–1650): französischer Philosoph und Mathematiker, lieferte wichtige Beiträge zur Geometrie.

Anstelle von (A, B, f) schreibt man in der Regel $f : A \to B$ oder kurz nur f, falls Definitions- und Wertebereich implizit klar sind.[7]

Zwei Abbildungen $f : A \to B$ und $g : C \to D$ sind gleich genau dann, wenn ihre Definitions- und Wertebereiche übereinstimmen (also $A = C$ und $B = D$ gilt) und sie an jeder Stelle denselben Funktionswert haben (also $f(a) = g(a)$ für alle $a \in A = C$ gilt).

Einige Beispiele für Abbildungen:
1) $f : \{1, 2, 3\} \to \{2, 3, 4\}$ definiert durch $f(a) := a + 1$ für $a \in \{1, 2, 3\}$.
2) $f : \mathbb{N} \to \mathbb{N}$ definiert durch $f(n) := 1$ für alle $n \in \mathbb{N}$ (konstante Funktion).
3) $f : \mathbb{N} \to \mathbb{Q}$ definiert durch $f(n) := \frac{1}{n}$ für alle $n \in \mathbb{N}$.
4) $f : \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) := x$ für alle $x \in \mathbb{R}$.
5) $f : \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) := x^2$ für alle $x \in \mathbb{R}$.
6) $f : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ definiert durch $f(x) := x^2$ für alle $x \in \mathbb{R}_0^+$. Hierbei ist $\mathbb{R}_0^+ := \{x \in \mathbb{R} : x \geq 0\}$. Beachten Sie, dass diese Funktion von der aus Beispiel 5) verschieden ist (Definitions- und Wertebereich gehören ausdrücklich zu einer Funktion dazu).
7) $f : \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) := x$ für $x \geq 0$ und $f(x) := x^3$ für $x < 0$ definiert ebenfalls eine Funktion. Die Funktionswerte müssen sich nicht immer durch eine geschlossene Formel angeben lassen.

Das obige Beispiel 4) lässt sich natürlich analog auf jeder beliebigen Menge betrachten. Hierzu eine extra Definition.

Definition I.2.2. Sei A eine Menge. Die Abbildung $\mathrm{id}_A : A \to A$ definiert durch $\mathrm{id}_A(a) := a$ für alle $a \in A$ heißt die *identische Abbildung* (oder *identische Funktion*) auf A.

id_A bildet also jedes Element von A auf sich selbst ab.

Auch das obige Beispiel 2) einer konstanten Funktionen lässt sich natürlich verallgemeinern.

Definition I.2.3. Seien A und B zwei Mengen und sei $b_0 \in B$. Wir definieren eine Funktion $\underline{b_0} : A \to B$ durch $\underline{b_0}(a) := b_0$ für alle $a \in A$.
$\underline{b_0}$ heißt die *konstante Funktion* auf A mit Wert b_0.

Die Funktion $\underline{b_0}$ bildet also jedes Element aus A auf denselben Wert b_0 ab. Diese Funktion ist zu unterscheiden vom Element b_0 selbst (z. B. ist $\underline{1} : \mathbb{R} \to \mathbb{R}$, die konstante Funktion mit Wert 1 auf $\mathbb{R}$, etwas anderes als die Zahl 1). In der Praxis schreibt man dennoch häufig nur b_0 anstatt $\underline{b_0}$

[7] Zu dieser Funktionsdefinitions ist zu bemerken, dass sie eigentlich nicht mathematisch präzise ist (was genau bedeutet "Zuordnungsvorschrift"?). Die mathematisch saubere Definition lautet: f ist eine Teilmenge von $A \times B$, so dass für alle $a \in A$ genau ein $b \in B$ mit $(a, b) \in f$ existiert. Für praktische Zwecke ist die obige Definition aber gut genug und wir wollen daher den streng formalen Funktionsbegriff hier nicht weiter diskutieren.

und man muss aus dem Kontext schließen, ob b_0 selbst oder die zugehörige konstante Funktion gemeint ist.

Als Nächstes definieren wir noch Graph und Bild einer Funktion.

Definition I.2.4. Seien A und B zwei Mengen und sei $f : A \to B$ eine Funktion. Dann ist der *Graph* von f definiert durch

$$\mathrm{gr}(f) := \{(a, f(a)) : a \in A\}.$$

Das *Bild* von f ist definiert durch

$$\mathrm{Im}(f) := \{f(a) : a \in A\}.$$

Der Graph von f ist also eine Teilmenge von $A \times B$. Etwas salopp gesagt besteht er aus all jenen "Punkten" $(a, f(a))$, welche von f "getroffen" werden.

Das Bild von f ist eine Teilmenge des Wertebereichs B. Sie besteht aus denjenigen Elementen von B, welche als Funktionswerte von f auftreten. Man beachte, dass $\mathrm{Im}(f)$ deutlich kleiner sein kann als B, z. B. besteht bei einer konstanten Funktion das Bild nur aus einem einzigen Element (vergleiche auch die Definition der Surjektivität weiter unten).[8]

Nun kommen wir zur Hintereinanderausführung (Verkettung) zweier Abbildungen.

Definition I.2.5. Gegeben seien Mengen A, B, C und Abbildungen $g : A \to B$ und $f : B \to C$. Dann ist die *Verkettung* von f und g definiert durch $f \circ g : A \to C$ mit

$$(f \circ g)(a) := f(g(a)) \quad \text{für alle } a \in A.$$

Für diese Definition ist es wesentlich, dass die Funktionswerte von g im Definitionsbereich von f liegen, anderenfalls wäre $f(g(a))$ gar nicht definiert. $f \circ g$ wird übrigens gelesen als "f nach g", eben weil man erst die Abbildung g und danach die Abbildung f anwendet.

Wir betrachten wieder einige Beispiele:
1) Sei $g : \mathbb{N} \to \mathbb{Q}$ definiert durch $g(n) := 1/n$ für alle $n \in \mathbb{N}$ und $f : \mathbb{Q} \to \mathbb{Q}$ durch $f(q) := q^2$ für alle $q \in \mathbb{Q}$.

Dann ist $f \circ g$ eine Abbildung von $\mathbb{N}$ nach $\mathbb{Q}$ und es gilt $(f \circ g)(n) = f(g(n)) = f(1/n) = (1/n)^2 = 1/n^2$ für $n \in \mathbb{N}$.
2) Sei $g : \mathbb{R}_0^+ \to \mathbb{R}$ definiert durch $g(x) := \sqrt{x}$ für alle $x \in \mathbb{R}_0^+$ (zur Erinnerung $\mathbb{R}_0^+ = \{x \in \mathbb{R} : x \geq 0\}$). Weiter sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch $f(y) := y^2 + 3y + 1$ für jedes $y \in \mathbb{R}$.

Dann ist $f \circ g : \mathbb{R}_0^+ \to \mathbb{R}$ mit $(f \circ g)(x) = f(g(x)) = f(\sqrt{x}) = (\sqrt{x})^2 + 3\sqrt{x} + 1 = x + 3\sqrt{x} + 1$ für alle $x \geq 0$.

[8] Die Bezeichnung $\mathrm{Im}(f)$ für das Bild von f stammt übrigens vom englischen Wort "image". Manche Autoren schreiben stattdessen $\mathrm{ran}(f)$ für das Bild von f (von englisch "range").

3) Sei $f : \mathbb{R} \to \mathbb{R}$ erklärt durch $f(y) := \sqrt{y^2+1}$ für alle $y \in \mathbb{R}$ und sei $g : \mathbb{R} \to \mathbb{R}$ definiert durch $g(x) := x + 1$. Dann ist $f \circ g$ eine Funktion von $\mathbb{R}$ nach $\mathbb{R}$ mit $(f \circ g)(x) = f(g(x)) = f(x+1) = \sqrt{(x+1)^2+1}$, was man mittels binomischer Formel auch als $(f \circ g)(x) = \sqrt{x^2+2x+2}$ schreiben kann.

Als Nächstes wollen wir die wichtigen Begriffe der Injektivität und Surjektivität kennenlernen.

Definition I.2.6. Seien A und B zwei Mengen und sei $f : A \to B$ eine Abbildung.

(i) f heißt *injektiv*, falls für alle Elemente $a_1, a_2 \in A$ mit $a_1 \neq a_2$ auch $f(a_1) \neq f(a_2)$ gilt.

(ii) f heißt *surjektiv*, falls für alle $b \in B$ ein $a \in A$ mit $f(a) = b$ existiert.

(iii) f heißt *bijektiv*, falls f sowohl injektiv als auch surjektiv ist.

Injektivität von f bedeutet also, dass f verschiedene Elemente aus A auch auf verschiedene Elemente von B abbildet. Surjektivität bedeutet, dass jedes Element von B als Funktionswert von f auftritt. Die Formulierung "es existiert ein $a \in A$ mit $f(a) = b$" bedeutet dabei, dass mindestens ein solches a existiert, eventuell kann es mehrere (sogar unendlich viele) solche Elemente geben.

Mit Hilfe des Bildes von f lässt sich die Definition der Surjektivität kürzer fassen:

$$f \text{ ist surjektiv } \Leftrightarrow \text{ Im}(f) = B.$$

Wir betrachten wiederum einige konkrete Beispiele:
1) Für jede Menge A ist die identische Abbildung id_A bijektiv, wie sofort aus der Definition folgt.
2) Die Abbildung $f : \{1,2,3\} \to \{2,3,4\}$ mit $f(a) := a + 1$ für $a \in \{1,2,3\}$ ist bijektiv, wie man leicht sieht.
3) Die Abbildung $f : \mathbb{N} \to \mathbb{Q}$ mit $f(n) := 1/n$ für $n \in \mathbb{N}$ ist injektiv, denn aus $f(n_1) = f(n_2)$ folgt durch Kehrwertbildung $n_1 = n_2$. Hingegen ist f nicht surjektiv, da z. B. $2 \notin \text{Im}(f)$ ist.
4) Die Abbildung $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) := x^2$ für alle $x \in \mathbb{R}$ ist nicht injektiv, da z. B. $f(1) = f(-1)$ ist. Ferner ist f auch nicht surjektiv, denn es ist $f(x) \geq 0$ für alle $x \in \mathbb{R}$, das Bild $\text{Im}(f)$ enthält also keine negativen Zahlen.
5) Im Unterschied zu Beispiel 4) ist die Abbildung $f : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ mit $f(x) := x^2$ bijektiv.

Begründung: Sind $x, y \geq 0$ mit $x \neq y$, so können wir ohne Einschränkung $0 \leq x < y$ annehmen und daraus $x^2 < y^2$, also $f(x) \neq f(y)$ schließen. Das zeigt die Injektivität von f.

Für die Surjektivität nehme man ein beliebiges $y \in \mathbb{R}_0^+$ her. Dann ist $x := \sqrt{y} \in \mathbb{R}_0^+$ mit $f(x) = y$.

Als letzten Punkt in diesem Kapitel wollen wir nun noch den Begriff der Umkehrabbildung einführen: Ist $f : A \to B$ bijektiv, so existiert zu jedem $b \in B$ *genau ein* $a \in A$ mit $f(a) = b$ (wegen der Surjektivität existiert mindestens ein solches a, wegen der Injektivität kann es nicht mehr als eines geben). Das führt zu folgender Definition.

Definition I.2.7. Seien A und B zwei Mengen und sei $f : A \to B$ eine bijektive Abbildung. Die *Umkehrabbildung* (oder *Umkehrfunktion*) $f^{-1} : B \to A$ wird folgendermaßen erklärt: Für alle $b \in B$ ist $f^{-1}(b)$ dasjenige Element von A mit $f(f^{-1}(b)) = b$.

Für bijektives $f : A \to B$ ergibt sich unmittelbar aus der Definition der Umkehrabbildung:

$$f \circ f^{-1} = \mathrm{id}_B \quad \text{und} \quad f^{-1} \circ f = \mathrm{id}_A.$$

Ferner ist leicht zu sehen, dass auch f^{-1} wieder bijektiv ist und dass $(f^{-1})^{-1} = f$ gilt (die Details überlasse ich Ihnen zur Übung).

Zum Abschluss betrachten wir ein paar Beispiele, die sich an die obigen Beispiele zur Bijketivität anschließen:
1) Wir hatten oben schon festgestellt, dass für jede Menge A die identische Abbildung id_A bijektiv ist. Aus den Definitionen folgt nun unmittelbar $\mathrm{id}_A^{-1} = \mathrm{id}_A$.
2) Für die Abbildung $f : \{1, 2, 3\} \to \{2, 3, 4\}$ mit $f(a) := a + 1$ hatten wir auch schon die Bijektivität festgestellt. Die Umkehrabbildung ist gegeben durch: $f^{-1} : \{2, 3, 4\} \to \{1, 2, 3\}$ mit $f^{-1}(b) = b - 1$.
3) Ebenfalls hatten wir schon gesehen, dass die Abbildung $f : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ mit $f(x) := x^2$ bijektiv ist. Aus der obigen Rechnung folgt auch gleich $f^{-1}(y) = \sqrt{y}$ für $y \geq 0$.

II Die Zahlenbereiche von $\mathbb{N}$ bis $\mathbb{R}$

Wir wollen in diesem Kapitel das Wichtigste zu den Bereichen der natürlichen, ganzen, rationalen und reellen Zahlen zusammenstellen, wobei wir die Existenz dieser Zahlenbereiche allerdings als gegeben hinnehmen.

II.1 Natürliche, ganze, rationale und reelle Zahlen

Wir beginnen mit den natürlichen Zahlen. Zwar hatten wir diese schon bei den Beispielen in Kapitel I verwendet, wir führen sie aber noch einmal offiziell ein: Es bezeichnet

$$\mathbb{N} := \{1, 2, 3, 4, \ldots\}$$

die Menge der natürlichen Zahlen. Diese ist Ihnen sicherlich aus der Schule bestens bekannt und daher soll die Natur dieser Menge und ihrer Elemente hier auch nicht weiter hinterfragt werden. Wir setzen die natürlichen Zahlen als Grundobjekte voraus.

Manchmal will man nicht bei 1 sondern bei 0 anfangen zu zählen, daher definieren wir noch

$$\mathbb{N}_0 := \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, 4, \ldots\}.$$

(Bei der Notation ist etwas Vorsicht geboten, denn bei einigen Autoren schließt die Menge $\mathbb{N}$ die Null bereits mit ein.)

Eigentlich müsste man nun zunächst das Beweisprinzip der vollständigen Induktion und das Prinzip der rekursiven Definitionen für die Menge der natürlichen Zahlen diskutieren (ersteres werden wir noch tun, allerdings erst im nächsten Abschnitt) und müsste die üblichen arithmetischen Operationen (Addition und Multiplikation), sowie die Ordnungsstruktur der natürlichen Zahlen einführen. Anschließend müsste man aus den natürlichen Zahlen die ganzen Zahlen, aus diesen wiederum die rationalen Zahlen und schließlich aus den rationalen die reellen Zahlen konstruieren. Dieses Vorgehen ist allerdings insgesamt sehr aufwendig und wird erfahrungsgemäß nur von wenigen Studienanfängern wirklich verstanden.

Daher setzen wir hier einfach die reellen Zahlen mit ihrer üblichen Arithmetik und Ordnungsstruktur als gegeben voraus und stellen nur ihre wesentlichen Eigenschaften zusammen. Die ganzen und die rationalen Zahlen

fallen uns dann als Teilmengen in den Schoß. Im Anhang dieses Skriptes werden aber die Konstruktionen der ganzen, rationalen und reellen Zahlen zumindest kurz skizziert.

Die Menge der *reellen Zahlen* bezeichnen wir, wie schon in den Beispielen in Kapitel I, mit $\mathbb{R}$. Sie umfasst die Menge der natürlichen Zahlen inklusive der Null, also $\mathbb{N}_0 \subseteq \mathbb{R}$.

Weiter existieren auf $\mathbb{R}$ eine Addition (bezeichnet mit $+$) und eine Multiplikation (bezeichnet mit $\cdot$), die folgende Eigenschaften haben[1]:

(i) $(a+b)+c = a+(b+c)$ für alle $a, b, c \in \mathbb{R}$ (Assoziativgesetz der Addition)

(ii) $a + b = b + a$ für alle $a, b \in \mathbb{R}$ (Kommutativgesetz der Addition)

(iii) $0 + a = a$ für alle $a \in \mathbb{R}$ (Null ist neutrales Element der Addition)

(iv) Für alle $a \in \mathbb{R}$ existiert genau ein Element $-a \in \mathbb{R}$ mit $(-a) + a = 0$. (Existenz von additiven Inversen)

(v) $(ab)c = a(bc)$ für alle $a, b, c \in \mathbb{R}$ (Assoziativgesetz der Multiplikation)

(vi) $ab = ba$ für alle $a, b \in \mathbb{R}$ (Kommutativgesetz der Multiplikation)

(vii) $1a = a$ für alle $a \in \mathbb{R}$ (Eins ist neutrales Element der Multiplikation)

(viii) Für alle $a \in \mathbb{R} \setminus \{0\}$ existiert genau ein Element $a^{-1} \in \mathbb{R} \setminus \{0\}$ mit $a^{-1}a = 1$. (Existenz von multiplikativen Inversen)[2]

(ix) $a(b + c) = ab + ac$ für alle $a, b, c \in \mathbb{R}$ (Distributivgesetz)

Man beachte, dass wegen (ii) und (iii) auch $a + 0 = a$ für alle $a \in \mathbb{R}$ gilt. Ebenso ist auch $a + (-a) = 0$ und $a1 = a$ für alle $a \in \mathbb{R}$, sowie $aa^{-1} = 1$ für alle $a \in \mathbb{R} \setminus \{0\}$. Weiter folgt aus den obigen Eigenschaften (wie?): Es ist $-0 = 0$ und $-(-a) = a$, sowie $1^{-1} = 1$ und $(a^{-1})^{-1} = a$ (falls $a \neq 0$).

Auch alle weiteren bekannten Rechenregeln für die reellen Zahlen lassen sich aus den obigen Eigenschaften herleiten. Ein Beispiel:

Lemma II.1.1. *Für alle $a, b \in \mathbb{R}$ gilt:*

(a) $0a = 0 = a0$.

(b) $(-a)b = -(ab) = a(-b)$ *(insbesondere ist $(-1)b = -b = b(-1)$).*

[1] Bei $+$ und $\cdot$ handelt es sich eigentlich um Funktionen von $\mathbb{R} \times \mathbb{R}$ nach $\mathbb{R}$, wobei man die Funktionswerte an der Stelle $(a, b) \in \mathbb{R} \times \mathbb{R}$ als $a + b$ bzw. $a \cdot b$ (oder kurz ab) notiert.

[2] Das "genau ein" ist eigentlich nicht nötig. Man kann zeigen, dass die additiven und multiplikativen Inversen automatisch eindeutig bestimmt sind, falls sie existieren. Ebenso kann man beweisen, dass die neutralen Elemente 0 und 1 bereits durch ihre oben angegebene Eigenschaft eindeutig bestimmt sind, siehe dazu den Abschnitt über Gruppen in Kapitel III.

Beweis. Zu (a): Wegen der Neutralitätseigenschaft der 0 und des Distributivgesetzes ist

$$0a = (0+0)a = 0a + 0a. \tag{II.1}$$

Hier haben wir bereits das Distributivgesetz in der Form $(x+y)z = xz + yz$ benutzt. Es ergibt sich aus der ursprünglichen Form (ix) zusammen mit dem Kommutativgesetz der Multiplikation.
Nun addieren wir zu beiden Seiten der Gleichung (II.1) das Element $-(0a)$ und erhalten:

$$0 = -(0a) + 0a = -(0a) + (0a + 0a).$$

Die rechte Seite lässt sich wegen der Assoziativität der Addition weiter umformen und man erhält:

$$0 = (-(0a) + 0a) + 0a = 0 + 0a = 0a.$$

Also ist in der Tat $0a = 0$. Wegen der Kommutativität der Multiplikation ist dann auch $a0 = 0a = 0$.
Zu (b): Nach Teil (a) ist $0b = 0$ (das Element a in Teil (a) war eine beliebige reelle Zahl, also gilt die Aussage ebenso für b). Daher folgt mit dem Distributivgesetz

$$(-a)b + ab = ((-a) + a)b = 0b = 0.$$

Nun addieren wir zu beiden Seiten $-(ab)$ und erhalten:

$$((-a)b + ab) + (-(ab)) = 0 + (-(ab)) = -(ab). \tag{II.2}$$

Wegen der Assoziativität von + gilt aber

$$((-a)b + ab) + (-(ab)) = (-a)b + (ab + (-ab)) = (-a)b + 0 = (-a)b. \tag{II.3}$$

Aus (II.2) und (II.3) folgt nun $(-a)b = -(ab)$.
Da a und b beliebig waren gilt entsprechend auch $(-b)a = -(ba)$. Wegen der Kommutativität der Multiplikation folgt daraus $a(-b) = -(ab)$. □

Hier noch eine weitere Regel.

Lemma II.1.2. *Seien $a, b \in \mathbb{R}$ mit $a \neq 0$ und $b \neq 0$. Dann ist auch $ab \neq 0$ und es gilt $(ab)^{-1} = a^{-1}b^{-1}$.*

Beweis. Es ist

$$\begin{aligned}(ab)(a^{-1}b^{-1}) &= (ba)(a^{-1}b^{-1}) = ((ba)a^{-1})b^{-1} \\ &= (b(aa^{-1}))b^{-1} = (b1)b^{-1} = bb^{-1} = 1\end{aligned}$$

(machen Sie sich selbst klar, welche der obigen Regeln (i)–(ix) in jedem Rechenschritt benutzt wurden).

Wegen Lemma II.1.1 gilt $0(a^{-1}b^{-1}) = 0$, daher folgt $ab \neq 0$. Nun multiplizieren wir die obige Gleichung von links mit $(ab)^{-1}$ und erhalten

$$(ab)^{-1}((ab)(a^{-1}b^{-1})) = (ab)^{-1}.$$

Daraus folgt

$$(ab)^{-1} = ((ab)^{-1}(ab))(a^{-1}b^{-1}) = a^{-1}b^{-1}$$

(machen Sie sich wieder klar, welche der obigen Regeln (i)–(ix) hier angewendet wurden). □

Auch Differenzen und Brüche können wir nun definieren.

Definition II.1.3. Für $a, b \in \mathbb{R}$ setzen wir

$$a - b := a + (-b).$$

Ist $b \neq 0$, so setzen wir zudem

$$\frac{a}{b} := ab^{-1}.$$

Es gelten die folgenden bekannten Rechenregeln für Brüche.

Lemma II.1.4. *Für alle $a, b, c, d \in \mathbb{R}$ gilt:*

(a) $\frac{a}{1} = a$ *und* $\frac{1}{b} = b^{-1}$, *falls* $b \neq 0$.

(b) $\frac{ac}{bd} = \frac{a}{b}\frac{c}{d}$, *falls* $b \neq 0$ *und* $d \neq 0$.
Insbesondere ist $\frac{a}{b} = \frac{ac}{bc}$, *falls* $b, c \neq 0$ *(Kürzen/Erweitern).*

(c) $\frac{a+b}{c} = \frac{a}{c} + \frac{b}{c}$, *falls* $c \neq 0$.

(d) $(\frac{a}{b})^{-1} = \frac{b}{a}$, *falls* $a \neq 0$ *und* $b \neq 0$ *(Kehrwertbildung).*

(e) $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$, *falls* $b \neq 0$ *und* $d \neq 0$.

Beweis. Die Beweise für (a), (b) und (d) können Sie sich selbst zur Übung überlegen. Um (c) zu beweisen schreiben wir mit Hilfe des Distributivgesetzes:

$$\frac{a+b}{c} = (a+b)c^{-1} = ac^{-1} + bc^{-1} = \frac{a}{c} + \frac{a}{b}.$$

Zum Beweis von (e) beobachtet man zunächst, dass wegen (c)

$$\frac{ad+bc}{bd} = \frac{ad}{bd} + \frac{bc}{bd}$$

gilt. Wegen (b) folgt daraus die Behauptung. □

Als Nächstes definieren die Menge $\mathbb{Z}$ der *ganzen Zahlen*, indem wir zu $\mathbb{N}$ noch die Null und die entsprechenden additiven Inversen hinzunehmen. Wir setzen also

$$\mathbb{Z} := \mathbb{N}_0 \cup \{-n : n \in \mathbb{N}\} = \{0, 1, 2, 3, \ldots\} \cup \{-1, -2, -3, \ldots\}.$$

Schließlich definieren wir die Menge $\mathbb{Q}$ der *rationalen Zahlen* als die Menge aller Brüche zweier ganzer Zahlen, also

$$\mathbb{Q} := \left\{ \frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0 \right\}.$$

Neben den oben beschriebenen Rechenoperationen verfügen die reellen Zahlen auch über eine Ordnungsrelation $<$, mit deren Hilfe man zwei reelle Zahlen der Größe nach vergleichen kann. Wie schon zuvor bei der Addition und der Multiplikation, wollen wir nicht formal definieren, was genau "$a < b$" bedeutet. Wir nehmen die Ordnungsstruktur von $\mathbb{R}$ schlichtweg als gegeben hin und stellen nur ihre wesentlichsten Eigenschaften zusammen.

Für alle $a, b, c \in \mathbb{R}$ gilt:

(i) $a < b \Rightarrow a \neq b$ (Irreflexivität)

(ii) $a < b$ und $b < c \Rightarrow a < c$ (Transitivität)

(iii) Es gilt genau eine der drei Aussagen $a = b$, $a < b$ oder $b < a$. (Linearität)

(iv) $a < b \Rightarrow a + c < b + c$

(v) $a < b$ und $0 < c \Rightarrow ac < bc$

Die Zahl a heißt *positiv*, falls $0 < a$ gilt und *negativ*, falls $a < 0$ gilt.

Anstelle von $a < b$ schreibt man natürlich auch $b > a$. Falls $a < b$ und $b < c$ gilt, so schreibt man auch kurz $a < b < c$. Weiter schreibt man $a \leq b$ (oder $b \geq a$), falls $a < b$ oder $a = b$ gilt.[3] Für die Relation $\leq$ gelten die folgenden Regeln (für alle $a, b, c \in \mathbb{R}$):

(i') $a \leq a$ (Reflexivität)

(ii') $a \leq b$ und $b \leq c \Rightarrow a \leq c$ (Transitivität)

(iii') $a \leq b$ und $b \leq a \Rightarrow a = b$. (Antisymmetrie)

(iv') Es gilt $a \leq b$ oder $b \leq a$. (Linearität)

(v') $a \leq b \Rightarrow a + c \leq b + c$

(vi') $a \leq b$ und $0 \leq c \Rightarrow ac \leq bc$

[3] "$a \leq b$" wird gelesen als "a kleiner gleich b" (das "oder" wird verschluckt).

Diese Aussagen ergeben sich aus den obigen Regeln für die Relation $<$ (Beweis als Übung). Ferner gelten noch folgende Regeln.

Lemma II.1.5. *Für alle $a, b, c, d \in \mathbb{R}$ gilt:*

(a) $a < b$ *und* $c < d \Rightarrow a + c < b + d$

(b) $c < 0 \Leftrightarrow -c > 0$

(c) $a < b$ *und* $c < 0 \Rightarrow ac > bc$

(d) $a^2 > 0$, *falls* $a \neq 0$ *(insbesondere ist* $1 = 1^2 > 0$*)*

(e) *Ist* $a > 0$, *so ist auch* $\frac{1}{a} > 0$. *Ist* $a < 0$, *so ist* $\frac{1}{a} < 0$.

(f) $0 < a < b \Rightarrow \frac{1}{a} > \frac{1}{b}$

(g) $a < b < 0 \Rightarrow \frac{1}{a} > \frac{1}{b}$

Analoge Aussagen gelten auch für die Relation $\leq$ (soweit sinnvoll).

Beweis. (a) Angenommen es gilt $a < b$ und $c < d$. Dann ist wegen der obigen Regel (iv) auch $a + c < b + c$ und $b + c < b + d$. Die Transitivität von $<$ impliziert daher $a + c < b + d$.
(b) Ist $c < 0$, so folgt wiederum wegen der obigen Regel (iv) durch Addition von $-c$, dass $0 < -c$ gilt. Ist umgekehrt $0 < -c$, so folgt durch Addition von c analog $c < 0$.
(c) Seien $a < b$ und $c < 0$. Wegen (b) ist dann $-c > 0$ und daher folgt aus der obigen Regel (v): $-ac < -bc$. Addition von ac zu beiden Seiten liefert $0 < ac - bc$. Nun addiert man noch bc zu beiden Seiten und erhält $bc < ac$.
(d) Sei $a \neq 0$. Ist $a > 0$, so folgt aus (v), dass auch $a^2 = aa > 0$ gilt. Ist $a < 0$, so folgt aus (c) ebenfalls $a^2 = aa > 0$.
(e) Sei $a > 0$. Wäre $\frac{1}{a} < 0$, so wäre wegen (v) $1 = a\frac{1}{a} < 0$, im Widerspruch zu (d). Also muss $\frac{1}{a} > 0$ gelten. Analog sieht man: $a < 0 \Rightarrow \frac{1}{a} < 0$.
(f) Es gelte $0 < a < b$. Da nach (e) $\frac{1}{a} > 0$ gilt, folgt $1 = a\frac{1}{a} < b\frac{1}{a}$. Multiplikation mit $\frac{1}{b} > 0$ liefert $\frac{1}{b} < \frac{1}{b}b\frac{1}{a} = \frac{1}{a}$.
Aussage (g) können Sie als Übung in analoger Weise selbst beweisen. Das Formulieren und Beweisen entsprechender Regeln für $\leq$ überlasse ich Ihnen ebenfalls zur Übung. □

Zum Schluss dieses Abschnitts führen wir noch die wichtige Betragsfunktion ein.

Definition II.1.6. Für $x \in \mathbb{R}$ setzen wir

$$|x| := \begin{cases} x \text{ falls } x \geq 0 \\ -x \text{ falls } x < 0. \end{cases}$$

$|x|$ heißt der *Betrag* von x.

Die Betragsfunktion hat folgende Eigenschaften.

Lemma II.1.7. *Für alle $x, y \in \mathbb{R}$ gilt:*

(a) $|xy| = |x||y|$

(b) $|x + y| \leq |x| + |y|$ *(Dreiecksungleichung)*

(c) $||x| - |y|| \leq |x - y|$ *(umgekehrte Dreiecksungleichung)*

Beweis. (a) beweist man leicht durch Fallunterscheidung nach den Vorzeichen von x und y. Die Details überlasse ich Ihnen zur Übung.
(b) Wir bemerken zunächst, dass $a \leq |a|$ für alle $a \in \mathbb{R}$ gilt. Also ist $x \leq |x|$ und $y \leq |y|$, folglich auch

$$x + y \leq |x| + |y|.$$

Ebenso ist $-x \leq |-x| = |x|$ und $-y \leq |-y| = |y|$, also auch

$$-(x + y) = -x - y \leq |x| + |y|.$$

Da $|x + y| = x + y$ oder $|x + y| = -(x + y)$ gilt, folgt in jedem Fall $|x + y| \leq |x| + |y|$.
(c) Nach der schon bewiesenen Dreiecksungleichung gilt

$$|x| = |x - y + y| \leq |x - y| + |y|,$$

also

$$|x| - |y| \leq |x - y|.$$

Analog zeigt man auch

$$|y| - |x| \leq |y - x| = |x - y|,$$

also gilt in jedem Fall $||x| - |y|| \leq |x - y|$. □

II.2 Vollständige Induktion

In diesem Abschnitt wollen wir das wichtige Beweisprinzip der vollständigen Induktion für die Menge der natürlichen Zahlen kennenlernen.

Prinzip der vollständigen Induktion (Version 1): Sei $E \subseteq \mathbb{N}$ eine Menge von natürlichen Zahlen, welche die folgenden beiden Eigenschaften besitzt:
1) Es ist $1 \in E$.
2) Für alle $n \in E$ ist auch $n + 1 \in E$.
Dann gilt $E = \mathbb{N}$.

Dieses Prinzip kann man sich folgendermaßen veranschaulichen: Nach Eigenschaft 1) ist $1 \in E$. Wegen der Eigenschaft 2) ist dann auch $1 + 1 = 2 \in E$. Eine erneute Anwendung von 2) ergibt dann $2 + 1 = 3 \in E$, anschließend folgt $3 + 1 = 4 \in E$, $4 + 1 = 5 \in E$, $5 + 1 = 6 \in E$ etc.

Auf eine formalere Begründung dieses Prinzip wollen wir hier verzichten, wir formulieren aber noch eine leicht andere Version.

Prinzip der vollständigen Induktion (Version 2): Es sei $\mathcal{E}$ eine Eigenschaft, welche natürliche Zahlen besitzen können. Es gelte:
1) 1 hat die Eigenschaft $\mathcal{E}$.
2) Für alle natürlichen Zahlen n gilt: Hat n die Eigenschaft $\mathcal{E}$, so hat auch $n + 1$ die Eigenschaft $\mathcal{E}$.
Dann hat jede natürliche Zahl die Eigenschaft $\mathcal{E}$.

Zum Beweis wende man einfach das Prinzip der vollständigen Induktion in der Version 1 auf die Menge $E = \{n \in \mathbb{N} : n \text{ hat die Eigenschaft } \mathcal{E}\}$ an.

Will man also durch vollständige Induktion zeigen, dass jede natürliche Zahl eine bestimmte Eigenschaft $\mathcal{E}$ besitzt, so hat man zwei Schritte auszuführen: Erstens muss man nachweisen, dass 1 die fragliche Eigenschaft besitzt. Dieser erste Schritt wird auch *Induktionsanfang* genannt (er ist in der Regel einfach). Zweitens muss man zeigen, dass für jede natürliche Zahl n mit der Eigenschaft $\mathcal{E}$ auch $n + 1$ diese Eigenschaft besitzt. Das ist der sogenannte *Induktionsschritt*.

Ein analoges Beweisprinzip gilt natürlich auch für $\mathbb{N}_0$. Dann ist der Induktionsanfang bei 0 zu wählen und im Induktionsschritt ist zu zeigen: Hat $n \in \mathbb{N}_0$ die Eigenschaft $\mathcal{E}$, so auch $n + 1$. Ebenso kann die Induktion auch bei irgendeiner natürlichen $n_0 \geq 2$ beginnen.

Wir werden sogleich ein erstes Beispiel betrachten. Zuvor führen wir aber noch folgende Schreibweise ein: Für reelle Zahlen $a_1, a_2, \ldots, a_n$ setzen wir

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

und

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

Aufgrund der Assoziativität von Addition und Multiplikation ist es gleichgültig, wo man in solch einer endlichen Summe/einem endlichen Produkt Klammern setzt. Jede Klammerung führt zu demselben Ergebnis[4], weshalb die Klammern meist von vornherein weggelassen werden.[5]

[4]Das müsste eigentlich formal bewiesen werden, ist aber ziemlich technisch. Da die Aussage intuitiv klar ist, verzichten wir hier auf einen Beweis.

[5]Die Verwendung des Buchstaben i für den Index solcher Summen oder Produkte ist

Nun kommen wir zum ersten Beispiel für einen Beweis durch vollständige Induktion.

Beispiel II.2.1. (Gaußsche Summenformel[6]) Für alle $n \in \mathbb{N}$ gilt

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

Beweis. Im Induktionsanfang haben wir die Richtigkeit der Behauptung für $n = 1$ zu überprüfen. Das ist einfach: Für $n = 1$ ergeben beide Seiten der obigen Gleichung 1.
Kommen wir nun zum Induktionsschritt: Angenommen n ist eine natürliche Zahl mit

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}. \tag{II.4}$$

Wir müssen zeigen, dass dann auch

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

gilt.
Dazu addieren wir $n + 1$ zu (II.4) und erhalten:

$$\begin{aligned}\sum_{i=1}^{n+1} i &= \sum_{i=1}^{n} i + n + 1 = \frac{n(n+1)}{2} + n + 1 \\ &= (n+1)\Big(\frac{n}{2} + 1\Big) = (n+1)\frac{n+2}{2} = \frac{(n+1)((n+1)+1)}{2}.\end{aligned}$$

Damit ist der Beweis abgeschlossen. □

Bevor wir zum nächsten Beispiel kommen, führen wir noch einmal offiziel Potenzen ein:

Definition II.2.2. Für $a \in \mathbb{R}$ und $n \in \mathbb{N}$ setzen wir

$$a^n := \prod_{i=1}^{n} a = \underbrace{a \cdot a \cdot \cdots \cdot a}_{n \text{ mal}}.$$

Außerdem setzen wir noch $a^0 := 1$.

natürlich nicht wesentlich, man kann auch jeden anderen Buchstaben zur Bezeichnung wählen (j und k sind ebenfalls sehr beliebt). Natürlich ist es auch erlaubt, dass eine Summe/ein Produkt bei irgendeinem anderen Index $m \in \mathbb{N}_0$ anstelle bei $m = 1$ beginnt.

[6] Carl Friedrich Gauß (1777–1855): deutscher Mathematiker mit zahlreichen wichtigen Beiträgen zu verschiedenen Teilgebieten der Mathematik, unter anderem zur Zahlentheorie und zur Geometrie. Der nach ihm benannte Gaußsche Integralsatz ist in der Analysis von Vektorfeldern und damit auch für Anwendungen in der Physik von großer Bedeutung.

Es gelten die bekannten Potenzgesetze

$$a^{n+m} = a^n a^m \qquad (a^m)^n = a^{nm} \qquad (ab)^n = a^n b^n,$$

wobei $a, b \in \mathbb{R}$ und $n, m \in \mathbb{N}_0$ beliebig sind. Auch diese Regeln müsste man streng genommen durch vollständige Induktion beweisen, was ich Ihnen zur Übung überlasse (führen Sie jeweils eine vollständige Induktion nach n durch, bei beliebigen, aber festen Werten a, b und m).

Hier betrachten wir stattdessen noch einige etwas interessantere Beispiele.

Beispiel II.2.3. (Geometrische Summenformel) Sei $q \in \mathbb{R}$ mit $q \neq 1$. Für alle $n \in \mathbb{N}_0$ gilt

$$\sum_{i=0}^{n} q^i = 1 + q + q^2 + \cdots + q^n = \frac{1 - q^{n+1}}{1 - q}.$$

Beweis. Der Induktionsanfang ist wieder einfach: Für $n = 0$ steht auf beiden Seiten der obigen Gleichung 1.
Induktionsschritt: Sei $n \in \mathbb{N}_0$ mit

$$\sum_{i=0}^{n} q^i = \frac{1 - q^{n+1}}{1 - q}.$$

Dann folgt:

$$\sum_{i=0}^{n+1} q^i = \sum_{i=0}^{n} q^i + q^{n+1} = \frac{1 - q^{n+1}}{1 - q} + q^{n+1} = \frac{1 - q^{n+1} + (1 - q)q^{n+1}}{1 - q}$$
$$= \frac{1 - q^{n+1} + q^{n+1} - q^{n+2}}{1 - q} = \frac{1 - q^{n+2}}{1 - q},$$

was gerade die Behauptung für $n+1$ ist, also ist der Beweis abgeschlossen. □

Vor dem nächsten Beispiel führen wir noch offiziell den Begriff der Teilbarkeit ein.

Definition II.2.4. Seien $a, b \in \mathbb{Z}$. Dann heißt a ein *Teiler* von b (in Zeichen $a \mid b$, gelesen als "a teilt b"), falls ein $k \in \mathbb{Z}$ mit $b = ka$ existiert.

Für alle ganzen Zahlen a, b, c gilt:

(i) $a \mid b$ und $a \mid c \Rightarrow a \mid b + c$

(ii) $a \mid b \Rightarrow a \mid cb$

(Beweis als Übung)

Hier nun ein Beispiel für einen Induktionsbeweis aus der Teilbarkeitstheorie.

Beispiel II.2.5. Seien $a, b, c \in \mathbb{Z}$ mit $c \mid a + b$ und $c \mid a - 1$. Dann gilt auch $c \mid a^n + b$ für alle $n \in \mathbb{N}$.

Beweis. Induktionsanfang: Für $n = 1$ ist nichts zu zeigen, da $a + b$ nach Voraussetzung teilbar durch c ist.
Induktionsschritt: Sei $n \in \mathbb{N}$ derart, dass $a^n + b$ durch c teilbar ist. Es gilt

$$a^{n+1} + b = a \cdot a^n + b = (a - 1)a^n + a^n + b \tag{II.5}$$

Da nach Voraussetzung $c \mid a - 1$ gilt, gilt auch $c \mid (a - 1)a^n$ (siehe die obige Regel (ii)). Wegen $c \mid a^n + b$ folgt daraus mit Hilfe der Regel (i) und (II.5) auch $c \mid a^{n+1} + b$. Damit ist der Beweis abgeschlossen. □

Weiter gibt es noch folgende Variante der vollständigen Induktion, die bisweilen sehr nützlich ist.

Starkes Prinzip der vollständigen Induktion: Es sei $\mathcal{E}$ eine Eigenschaft, welche natürliche Zahlen besitzen können. Es gelte:
1) 1 hat die Eigenschaft $\mathcal{E}$.
2) Für alle natürlichen Zahlen n gilt: Hat jede der Zahlen $1, \ldots, n$ die Eigenschaft $\mathcal{E}$, so hat auch $n + 1$ die Eigenschaft $\mathcal{E}$.
Dann hat jede natürliche Zahl die Eigenschaft $\mathcal{E}$.

Auch dieses Prinzip gilt natürlich entsprechend, wenn man nicht bei Eins sondern bei einer anderen Zahl $n_0 \in \mathbb{N}_0$ beginnt.

Zum Beweis des Prinzips wende man einfach das ursprüngliche vollständige Induktionsprinzip auf die Eigenschaft $\mathcal{E}'$ an, die folgendermaßen erklärt ist: Eine natürliche n hat die Eigenschaft $\mathcal{E}'$, falls jede der Zahlen $1, \ldots, n$ die Eigenschaft $\mathcal{E}$ hat.

Ein klassisches Beispiel für eine Anwendung des starken Prinzips der vollständigen Induktion ist der Beweis der Existenz der Primfaktorzerlegung. Zur Erinnerung hier noch einmal die Definition einer Primzahl.

Definition II.2.6. Eine natürliche Zahl $p \geq 2$ heißt *Primzahl*, falls 1 und p die einzigen positiven Teiler von p sind. Die Menge aller Primzahlen bezeichnen wir mit $\mathbb{P}$.

Die ersten Primzahlen sind 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31. Wesentlich längere Listen von Primzahlen finden sich in einschlägigen Tafelwerken.

Nun zur Existenz der Primfaktorzerlegung.

Satz II.2.7 (Existenz der Primfaktorzerlegung)**.** *Jede natürliche Zahl $n \geq 2$ lässt sich als Produkt von Primzahlen schreiben, d. h. es existieren Primzahlen $p_1, \ldots, p_s$ mit $n = \prod_{i=1}^{s} p_i = p_1 p_2 \ldots p_s$.*

Die Primfaktoren $p_1, \ldots, p_s$ müssen dabei natürlich nicht alle verschieden sein, z. B. ist $12 = 2 \cdot 2 \cdot 3$.

Beweis. Induktionsanfang: Da 2 selbst eine Primzahl ist, ist hier nichts weiter zu zeigen.
Induktionsschritt: Sei $n \in \mathbb{N}$ mit $n \geq 2$ derart, dass sich jede natürliche Zahl $2 \leq k \leq n$ als Produkt von Primzahlen darstellen lässt.
Ist $n+1$ selbst eine Primzahl, so muss man nichts weiter beweisen. Ist $n+1$ keine Primzahl, so existiert ein $k \in \{2, \ldots, n\}$, welches $n+1$ teilt. Also ist $n+1 = kl$ für ein $l \in \mathbb{N}$. Wegen $k \geq 2$ und $k \leq n$ ist auch $l \leq n$ und $l \geq 2$.
Laut unserer Annahme sind also k und l beide als Produkt von Primzahlen darstellbar, etwa $k = \prod_{i=1}^{s} p_i$ und $l = \prod_{i=1}^{t} q_i$.
Dann ist auch $n+1 = kl = (p_1 p_2 \ldots p_s)(q_1 q_2 \ldots q_t)$ ein Produkt von Primzahlen. □

Eine wichtige Konsequenz des obigen Satzes ist die Tatsache, dass es unendlich viele Primzahlen geben muss, der sogenannte Satz von Euklid.[7]

Satz II.2.8 (Satz von Euklid)**.** *Die Menge $\mathbb{P}$ der Primzahlen ist unendlich.*

Beweis. Wir führen den Beweis durch Widerspruch. Angenommen die Menge $\mathbb{P}$ wäre endlich. D. h. sie lässt sich in der Form $\mathbb{P} = \{p_1, p_2, \ldots, p_n\}$ aufzählen.
Nun setzen wir $a := \prod_{i=1}^{n} p_i = p_1 p_2 \ldots p_n$ und $b := a + 1$.
Nach Satz II.2.7 ist b als Produkt von Primzahlen darstellbar. Insbesondere existiert ein $p \in \mathbb{P}$ mit $p \mid b$.
Wegen $\mathbb{P} = \{p_1, p_2, \ldots, p_n\}$ ist aber $p = p_i$ für ein $i \in \{1, \ldots, n\}$.
Nach Definition von a gilt also auch $p \mid a$.
Daraus würde folgen, dass p auch ein Teiler von $b - a = 1$ ist, was natürlich unmöglich ist. Somit muss $\mathbb{P}$ unendlich sein. □

Man kann übrigens auch beweisen, dass die Primfaktorzerlegung einer Zahl eindeutig bestimmt ist (bis auf die Reihenfolge der Faktoren). Das werden wir im näcshten Anschnitt tun (siehe Satz II.3.7).

II.3 Der Euklidische Algorithmus

In diesem Abschnitt wollen wir den sogenannten Euklidischen Algorithmus, ein Verfahren zur Bestimmung des größten gemeinsamen Teilers zweier Zahlen, kennenlernen. Zunächst formulieren wir nochmal offiziell die Division mit Rest in $\mathbb{Z}$.

Satz II.3.1 (Division mit Rest)**.** *Seien $a, b \in \mathbb{Z}$ mit $b \neq 0$. Dann existieren eindeutig bestimmte Zahlen $q \in \mathbb{Z}$ und $r \in \mathbb{N}_0$ mit $a = qb + r$ und $r < |b|$.*

[7] Euklid von Alexandria: Griechischer Mathematiker, lebte vermutlich im 3.Jahrhundert v. Chr., lieferte grundlegende Beiträge zur Geometrie, Arithmetik und Zahlentheorie, sowie zur Musiktheorie (diese war im antiken Griechenland ein Teilgebiet der Mathematik, aufbauend auf der Arithmetik). Sein berühmtestes Werk "Elemente" war in vielen Ländern noch bis ins 20.Jahrhundert hinein Grundlage des Geometrieunterrichts an Schulen. Auch der im nächsten Abschnitt behandelte Euklidische Algorithmus stammt von ihm.

Im obigen Satz ist q der ganzzahlige Anteil des Quotienten a/b und r ist der verbleibende Rest. Natürlich müsste dieser Satz eigentlich formal bewiesen werden, aber da die Aussage intuitiv klar ist, verzichten wir hier darauf.

Als Nächstes definieren wir den größten gemeinsamen Teiler (ggT) zweier Zahlen.

Definition II.3.2. Seien $a, b \in \mathbb{Z}$ mit $a \neq 0$ oder $b \neq 0$. Dann bezeichnet

$$\mathrm{ggT}(a,b) := \max\{c \in \mathbb{N} : c \mid a \text{ und } c \mid b\}$$

den *größten gemeinsamen Teiler* von a und b (max steht für Maximum, also das größte Element der Menge).

Der folgende Euklidische Algorithmus[8] erlaubt es, den ggT zweier Zahlen durch wiederholte Division mit Rest zu bestimmen.

Satz II.3.3 (Euklidischer Algorithmus)**.** *Seien $a, b \in \mathbb{Z}$ mit $b \neq 0$. Wir bestimmen zunächst $q_0, r_0 \in \mathbb{Z}$ mit*

$$a = q_0 b + r_0, \quad 0 \leq r_0 < |b|.$$

Ist $r_0 = 0$, so ist $ggT(a,b) = b$.
Anderenfalls finden wir $q_1, r_1 \in \mathbb{Z}$ mit

$$b = q_1 r_0 + r_1, \quad 0 \leq r_1 < r_0.$$

Ist $r_1 = 0$, so ist $ggT(a,b) = r_0$.
Anderenfalls gibt es $q_2, r_2 \in \mathbb{Z}$ mit

$$r_0 = q_2 r_1 + r_2, \quad 0 \leq r_2 < r_1.$$

Ist $r_2 = 0$, so ist $ggT(a,b) = r_1$.
Anderenfalls bestimme $q_3, r_3 \in \mathbb{Z}$ mit

$$r_1 = q_3 r_2 + r_3, \quad 0 \leq r_3 < r_2.$$

So fortfahrend erhalten wir $q_0, \ldots, q_n, r_0 \ldots, r_n \in \mathbb{Z}$ mit

$$r_{i-1} = q_{i+1} r_i + r_{i+1}, \quad 0 \leq r_{i+1} < r_i$$

für $i = 1, \ldots, n-1$, wobei $r_0, \ldots, r_{n-1} \neq 0$ und $r_n = 0$.
Dann gilt $ggT(a,b) = r_{n-1}$, der größte gemeinsame Teiler von a und b ist also gleich dem letzten nicht verschwindenden Rest.

[8] Siehe Fußnote zum Satz von Euklid.

Beweis. Zunächst ist klar, dass das Verfahren wirklich nach endlich vielen Schritten bei $r_n = 0$ ankommt, da die Reste wegen $r_{i+1} < r_i$ in jedem Schritt um mindestens 1 kleiner werden.
Falls bereits $r_0 = 0$ ist, so gilt $b \mid a$ und somit natürlich $\mathrm{ggT}(a, b) = b$.
Ansonsten haben wir

$$r_{i-1} = q_{i+1} r_i + r_{i+1} \quad \text{für } i = 0, \ldots, n-1$$

(wobei wir $r_{-1} := b$ setzen).
Wegen $r_n = 0$ ist $r_{n-2} = q_n r_{n-1}$, also $r_{n-1} \mid r_{n-2}$.
Wegen $r_{n-3} = q_{n-1} r_{n-2} + r_{n-1}$ folgt daraus auch $r_{n-1} \mid r_{n-3}$.
Aus $r_{n-4} = q_{n-2} r_{n-3} + r_{n-2}$ folgt dann $r_{n-1} \mid r_{n-4}$.
So fortfahrend erhält man $r_{n-1} \mid r_i$ für alle $i = 0, \ldots, n-1$ und schließlich auch $r_{n-1} \mid r_{-1} = b$.
Aus $a = q_0 b + r_0$ folgt dann auch $r_{n-1} \mid a$, also ist r_{n-1} ein gemeinsamer Teiler von a und b.
Sei nun $c \in \mathbb{N}$ irgendein gemeinsamer Teiler von a und b. Wir wollen $c \leq r_{n-1}$ nachweisen. Zunächst folgt aus $r_0 = a - q_0 b$ auch $c \mid r_0$.
Wegen $r_1 = b - q_1 r_0$ folgt dann auch $c \mid r_1$.
Aus $r_2 = r_0 - q_2 r_1$ folgt dann wiederum $c \mid r_2$.
So fahren wir fort bis wir schließlich bei $c \mid r_{n-1}$ angelangt sind. Insbesondere folgt daraus $c \leq r_{n-1}$.
Also ist tatsächlich $r_{n-1} = \mathrm{ggT}(a, b)$. □

Im obigen Beweis haben wir gleich noch folgende Tatsache mitbewiesen.

Korollar II.3.4. *Seien $a, b \in \mathbb{Z}$ mit $a \neq 0$ oder $b \neq 0$. Dann gilt für alle $c \in \mathbb{N}$:*

$$c \mid a \text{ und } c \mid b \;\Rightarrow\; c \mid \mathrm{ggT}(a, b).$$

Jeder gemeinsame Teiler von a und b teilt also auch ihren größten gemeinsamen Teiler.

Es folgt ein konkretes Rechenbeispiel: Wir wollen $\mathrm{ggT}(969, 627)$ bestimmen. Der Euklidische Algorithmus liefert:

$$\begin{aligned} 969 &= 1 \cdot 627 + 342 \\ 627 &= 1 \cdot 342 + 285 \\ 342 &= 1 \cdot 285 + 57 \\ 285 &= 5 \cdot 57 + 0 \end{aligned}$$

Also ist $\mathrm{ggT}(969, 627) = 57$.

Als weiteres Beispiel bestimmen wir $\mathrm{ggT}(4828, 2624)$. Der Euklidische Algorithmus liefert in diesem Fall:

$$\begin{aligned}
4828 &= 1 \cdot 2624 + 2204 \\
2624 &= 1 \cdot 2204 + 420 \\
2204 &= 5 \cdot 420 + 104 \\
420 &= 4 \cdot 104 + 4 \\
104 &= 4 \cdot 26 + 0
\end{aligned}$$

Damit ist $\mathrm{ggT}(4828, 2624) = 4$.

Als Nächstes beweisen wir noch die folgende Eigenschaft des ggT.

Lemma II.3.5. *Seien $a, b \in \mathbb{Z}$ mit $a \neq 0$ oder $b \neq 0$. Dann gilt*

$$\mathrm{ggT}(ca, cb) = c \cdot \mathrm{ggT}(a, b)$$

für alle $c \in \mathbb{N}$.

Beweis. Ohne Einschränkung sei $b \neq 0$. Wir führen den Euklidischen Algorithmus für die Startwerte a und b aus und erhalten $q_0 \ldots, q_n \in \mathbb{Z}$ und $r_0 \ldots, r_n \in \mathbb{Z}$ mit $r_0, \ldots, r_{n-1} \neq 0$, $r_n = 0$ und

$$\begin{aligned}
&a = q_0 b + r_0, \quad 0 \leq r_0 < |b|, \\
&r_{i-1} = q_{i+1} r_i + r_{i+1}, \quad 0 \leq r_{i+1} < r_i \quad \text{für } i = 0, \ldots, n-1,
\end{aligned}$$

wobei $r_{-1} := b$.
Multipliziert man alles mit c, so erhält man

$$\begin{aligned}
&ca = q_0 cb + cr_0, \quad 0 \leq cr_0 < |cb|, \\
&cr_{i-1} = q_{i+1} cr_i + cr_{i+1}, \quad 0 \leq cr_{i+1} < cr_i \quad \text{für } i = 0, \ldots, n-1.
\end{aligned}$$

Das ist gerade der Euklidische Algorithmus für die Startwerte ca und cb. Also ist $\mathrm{ggT}(ca, cb) = cr_{n-1} = c\mathrm{ggT}(a, b)$. □

Nun können wir auch die folgende wichtige Eigenschaft von Primzahlen beweisen.

Lemma II.3.6. *Seien $a, b \in \mathbb{Z}$ und sei p eine Primzahl mit $p \mid ab$. Dann gilt $p \mid a$ oder $p \mid b$.*

Beweis. Im Fall $p \mid a$ ist nichts zu zeigen. Sei also p kein Teiler von a. Wir wollen zeigen, dass dann p ein Teiler von b sein muss, wobei wir ohne Einschränkung $b > 0$ annehmen dürfen.
Da p eine Primzahl ist, welche a nicht teilt, ist $\mathrm{ggT}(a, p) = 1$. Aus Lemma II.3.5 folgt damit $\mathrm{ggT}(ab, pb) = b \cdot \mathrm{ggT}(a, p) = b$.
Natürlich gilt $p \mid pb$ und nach Voraussetzung auch $p \mid ab$. Wegen Korollar II.3.4 folgt daraus $p \mid \mathrm{ggT}(ab, ap)$, also $p \mid b$. □

Jetzt sind wir so weit, die bereits im letzten Abschnitt erwähnte Eindeutigkeit der Primfaktorzerlegung zu beweisen.

Satz II.3.7 (Eindeutigkeit der Primfaktorzerlegung)**.** *Seien $p_1, \ldots, p_n$ und $q_1, \ldots, q_m$ Primzahlen mit $p_1 p_2 \ldots p_n = q_1 q_2 \ldots q_m$. Dann gilt $n = m$ und nach einer eventuellen Umsortierung $q'_1, \ldots, q'_n$ von $q_1, \ldots, q_n$ gilt $p_i = q'_i$ für $i = 1, \ldots, n$.*

Beweis. Wegen $p_1 p_2 \ldots p_n = q_1 q_2 \ldots q_m$ ist p_1 ein Teiler von $q_1 q_2 \ldots q_m$. Da p_1 eine Primzahl ist, folgt aus Lemma II.3.6 (mehrfach angewendet), dass p_1 einen der Faktoren $q_1, q_2, \ldots, q_m$ teilen muss. Sei also etwa $i_1 \in \{1, \ldots, m\}$ mit $p_1 \mid q_{i_1}$. Da aber q_{i_1} eine Primzahl ist, folgt daraus $p_1 = q_{i_1} =: q'_1$.
Durch Kürzen in $p_1 p_2 \ldots p_n = q_1 q_2 \ldots q_m$ folgt

$$\prod_{i=2}^{n} p_i = \prod_{i \in I} q_i,$$

wobei $I := \{1, \ldots, m\} \setminus \{i_1\}$.
Also ist p_2 ein Teiler von $\prod_{i \in I} q_i$ und dasselbe Argument wie eben liefert $p_2 = q_{i_2} =: q'_2$ für ein $i_2 \in I$.
Kürzen liefert dann

$$\prod_{i=3}^{n} p_i = \prod_{i \in J} q_i,$$

wobei $J := \{1, \ldots, m\} \setminus \{i_1, i_2\}$.
So fortfahrend erhält man $n = m$ [9] und $p_1 = q'_1, p_2 = q'_2, \ldots, p_n = q'_n$ bei geeigneter Umsortierung der q_i. □

Übrigens gibt es kein bekanntes Verfahren zur effizienten Berechnung[10] der Primfaktorzerlegung einer gegebenen (großen) Zahl. Dieser Umstand ist entscheidend für viele moderne Verschlüsselungsverfahren.

Zum Schluss dieses Abschnitts betrachten wir noch, sozusagen als Gegenstück zum ggT, das kleinste gemeinsame Vielfache (kgV).

Definition II.3.8. Seien $a, b \in \mathbb{N}$. Dann heißt

$$\mathrm{kgV}(a, b) := \min\{c \in \mathbb{N} : a \mid c \text{ und } b \mid c\}$$

das *kleinste gemeinsame Vielfache* von a und b (min steht für Minimum).

Natürlich ist ab stets ein gemeinsames Vielfaches von a und b, aber nicht unbedingt das kleinste. Den Zusammenhang zwischen kgV und ggT stellt der folgende Satz her.

[9] Anderenfalls stünde irgendwann auf einer Seite der Gleichung nur noch eine 1 und auf der anderen noch ein Produkt von Primzahlen.

[10] Wir verzichten hier auf eine Präziserung des Begriffs "effizient".

Satz II.3.9. *Für alle $a, b \in \mathbb{N}$ gilt*

$$\mathrm{kgV}(a,b) = \frac{ab}{\mathrm{ggT}(a,b)}.$$

Beweis. 1) Wir betrachten zuerst den Spezialfall $\mathrm{ggT}(a,b) = 1$. In diesem Fall wollen wir also $\mathrm{kgV}(a,b) = ab$ nachweisen. Ohne Einschränkung können wir $a \geq 2$ und $b \geq 2$ annehmen (ansonsten ist die Behauptung trivialerweise erfüllt). Klar ist auch $\mathrm{kgV}(a,b) \leq ab$.
Sei nun $c \in \mathbb{N}$ irgendein gemeinsames Vielfaches von a und b. Dann existieren $k, l \in \mathbb{N}$ mit $c = ka = lb$. Da $\mathrm{ggT}(a,b) = 1$ und $a, b \geq 2$ gilt, muss auch $k, l \geq 2$ gelten.
Wir schreiben nun die Zahlen k, l, a, b als Produkte von Primfaktoren, etwa $a = p_1 \dots p_n$, $b = q_1 \dots q_m$, $k = v_1 \dots v_s$, $l = w_1 \dots w_t$.
Dann gilt $(v_1 \dots v_s)(p_1 \dots p_n) = (w_1 \dots w_t)(q_1 \dots q_m)$. Wegen der Eindeutigkeit der Primfaktorzerlegung muss jede der Primzahlen $p_1, \dots, p_n$ auch auf der rechten Seite auftauchen. Da aber $\mathrm{ggT}(a,b) = 1$ gilt, kann keines der p_i mit einem der q_j übereinstimmen. Die Zahlen $p_1, \dots, p_n$ müssen also alle in $w_1, \dots, w_t$ auftauchen. Mit anderen Worten $a = p_1 \dots p_n$ ist ein Teiler von $w_1 \dots w_t = l$. Wir können also $l = da$ für ein geeignetes $d \in \mathbb{N}$ schreiben.
Dann folgt aber $c = lb = dab \geq ab$.
Also ist in der Tat ab das kleinste gemeinsame Vielfache von a und b.
2) Bevor wir zum allgemeinen Fall kommen, halten wir noch folgendes fest: Für alle $a, b, c \in \mathbb{N}$ gilt $\mathrm{kgV}(ca, cb) = c \cdot \mathrm{kgV}(a,b)$.
Das beweist man leicht direkt anhand der Definition des kgV (Übung).
3) Seien nun $a, b \in \mathbb{N}$ beliebig. Wir setzen $c := \mathrm{ggT}(a,b)$, sowie $a' := a/c$ und $b' := b/c$. Aus Lemma II.3.5 folgt $\mathrm{ggT}(a', b') = 1$.
Nach dem schon bewiesenen Teil 1) gilt also $\mathrm{kgV}(a', b') = a'b' = ab/c^2$.
Andererseits folgt aus 2) auch $\mathrm{kgV}(a', b') = \mathrm{kgV}(a,b)/c$.
Zusammen impliziert das $\mathrm{kgV}(a,b) = ab/c$, was zu beweisen war. □

Beispiel: Oben hatten wir schon mit Hilfe des Euklidischen Algorithmus $\mathrm{ggT}(969, 627) = 57$ nachgewiesen. Aus dem obigen Satz erhalten wir damit für das kgV:

$$\mathrm{kgV}(969, 627) = \frac{969 \cdot 627}{57} = 969 \cdot 11 = 10659$$

III Algebraische Strukturen

In diesem Kapitel wollen wir uns mit den wichtigsten algebraischen Strukturen, Gruppen und Körpern, beschäftigen. Die Grundidee dabei ist es, die beispielsweise von den ganzen oder den reellen Zahlen bekannten Rechengesetze in einen abstrakten Kontext zu übertragen.

III.1 Gruppen

Wir beginnen mit der folgenden Definition.

Definition III.1.1. Sei G eine nicht leere Menge und $* : G \times G \to G$ eine Verknüpfung (Abbildung).
1) $(G, *)$ heißt *Halbgruppe*, falls das *Assoziativgesetz*

$$(a * b) * c = a * (b * c) \quad \forall a, b, c \in G$$

gilt.
2) $(G, *)$ heißt *Monoid*, falls $(G, *)$ eine Halbgruppe ist und zusätzlich ein *neutrales Element* $e \in G$ existiert, d. h. ein Element e mit

$$e * a = a = a * e \quad \forall a \in G.$$

Bemerkung III.1.2. Ein Monoid $(G, *)$ besitzt *genau ein* neutrales Element.

Beweis. Sind $e, e' \in G$ beide neutrale Elemente, so folgt wegen der Neutralität von e' sofort $e = e' * e$. Andererseits muss wegen der Neutralität von e auch $e' * e = e'$ gelten, also ist $e = e'$. □

Nun können wir auch Gruppen definieren.

Definition III.1.3. Ein Monoid $(G, *)$ heißt *Gruppe*, falls es zu jedem $a \in G$ ein *inverses Element* gibt, d. h. ein Element $b \in G$ mit $b * a = e = a * b$, wobei e das neutrale Element von $(G, *)$ ist.

Bemerkung III.1.4. Sei $(G, *)$ eine Gruppe und sei $a \in G$. Dann besitzt a *genau ein* inverses Element, welches üblicherweise mit a^{-1} bezeichnet wird.

Beweis. Seien $b, c \in G$ beide inverse Elemente von a. Dann gilt $b = b * e = b * (a * c) = (b * a) * c = e * c = c$. □

Es gilt das folgende Lemma für Inverse Elemente.

Lemma III.1.5. *Sei $(G, *)$ eine Gruppe und seien $a, b \in G$. Dann gilt $(a * b)^{-1} = b^{-1} * a^{-1}$.*

Beweis. Es bezeichne e das neutrale Element von $(G, *)$. Dann gilt

$$\begin{aligned}(a * b) * (b^{-1} * a^{-1}) &= ((a * b) * b^{-1}) * a^{-1} = (a * (b * b^{-1})) * a^{-1} \\ &= (a * e) * a^{-1} = a * a^{-1} = e.\end{aligned}$$

Daraus folgt

$$(a * b)^{-1} * ((a * b) * (b^{-1} * a^{-1})) = (a * b)^{-1} * e = (a * b)^{-1}.$$

Andererseits ist

$$\begin{aligned}(a * b)^{-1} * ((a * b) * (b^{-1} * a^{-1})) &= ((a * b)^{-1} * (a * b)) * (b^{-1} * a^{-1}) \\ &= e * (b^{-1} * a^{-1}) = b^{-1} * a^{-1}.\end{aligned}$$

Also ist $(a * b)^{-1} = b^{-1} * a^{-1}$. □

Im obigen Lemma ist unbedingt die Reihenfolge der Elemente zu beachten, es sei denn die Gruppe ist kommutativ. Dieser Begriff ist wie folgt definiert.

Definition III.1.6. Eine Gruppe $(G, *)$ heißt *kommutative* oder *abelsche Gruppe*[1], falls das *Kommutativgesetz*

$$a * b = b * a \quad \forall a, b \in G$$

gilt.

Noch eine Bemerkung zur Schreibweise: In vielen Fällen bezeichnet man die Verknüpfung einer Gruppe mit $\cdot$ oder $+$ anstelle von $*$ und schreibt dann 1 bzw. 0 für das neutrale Element. In einer Gruppe $(G, +)$ wird das inverse Element zu a mit $-a$ bezeichnet.

Beispiele:
1) $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$ bilden jeweils eine kommutative Gruppe, $(\mathbb{N}, +)$ dagegen nur eine Halbgruppe ($+$ bezeichnet hier die übliche Addition in $\mathbb{R}$).
2) Mit der üblichen Multiplikation $\cdot$ bilden $(\mathbb{Q} \setminus \{0\}, \cdot)$ und $(\mathbb{R} \setminus \{0\}, \cdot)$ jeweils eine kommutative Gruppe, $(\mathbb{Z} \setminus \{0\}, \cdot)$ dagegen nicht. Die 0 darf man hier

[1] Benannt nach dem norwegischen Mathematiker Niels Henrik Abel (1802–1829), einem der Begründer der Gruppentheorie.

nicht mit dazu nehmen, denn sie besitzt kein multiplikatives Inverses.[2]
3) Wir erklären auf der Menge $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R} = \{(a, b) : a, b \in \mathbb{R}\}$ eine Addition durch

$$(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2) \quad \forall (a_1, b_1), (a_2, b_2) \in \mathbb{R}^2.$$

Die "Punkte" im $\mathbb{R}^2$ werden also koordinatenweise addiert, wobei rechts in den einzelnen Koordinaten die übliche Addition in $\mathbb{R}$ gemeint ist.
Dann ist $(\mathbb{R}^2, +)$ eine kommutative Gruppe, wie man leicht nachrechnet (neutrales Element ist $(0, 0)$, inverses Element zu (a, b) ist $(-a, -b)$; die Details überlasse ich Ihnen zur Übung).

III.2 Körper

Wir kommen nun zum Begriff des Körpers. Diesen hatten wir implizit schon bei der Zusammenstellung der Eigenschaften der reellen Zahlen in Kapitel II kennengelernt, führen ihn jetzt aber nochmals explizit ein.

Definition III.2.1. Sei K eine nicht leere Menge versehen mit zwei Verknüpfungen $+ : K \times K \to K$ und $\cdot : K \times K \to K$. Dann heißt $(K, +, \cdot)$ ein *Körper*, falls folgendes gilt:
1) $(K, +)$ ist eine kommutative Gruppe.
2) $(K \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe.
3) Es gilt das *Distributivgesetz*

$$a(b + c) = ab + ac \quad \forall a, b, c \in K.$$

Die in Abschnitt II.1 zusammengestellten Rechenregeln besagen also gerade, dass $(\mathbb{R}, +, \cdot)$ ein Körper ist. Ebenso bilden die rationalen Zahlen $(\mathbb{Q}, +, \cdot)$ einen Körper (Beweis?), die ganzen Zahlen dagegen nicht (hier fehlen die multiplikativen Inversen).

Der Begriff "Körper" hat hier übrigens nichts mit Körpern im geometrischen Sinne, noch etwas mit dem menschlichen Körper zu tun. Vielmehr ist der Begriff (genau wie der der Gruppe) abgeleitet aus der Soziologie (vgl. das Wort "Körperschaften").

Falls klar ist, welche Verknüpfungen auf der Menge K betrachtet werden sollen, so schreibt man meist nur K anstelle von $(K, +, \cdot)$. So wird z. B. auf $\mathbb{R}$ stets die übliche Addition und Multiplikation betrachtet, wenn nicht ausdrücklich etwas anderes gesagt ist.

Die Aussagen der Lemmata II.1.1 und II.1.2 gelten nicht nur in $\mathbb{R}$ sondern in beliebigen Körpern (mit praktisch demselben Beweis). Ebenso kann man in jedem Körper K Differenzen und Brüche definieren und die in Abschnitt II.1 zusammengestellten Rechenregeln gelten entsprechend.

[2] Denn für alle $x \in \mathbb{R}$ ist $0 \cdot x = 0 \neq 1$.

Außer $\mathbb{R}$ und $\mathbb{Q}$ kennen wir bislang keine Beispiele für Körper. Im nächsten Abschnitt führen wir aber noch den wichtigen Körper der komplexen Zahlen ein, sowie im übernächsten Abschnitt die sogenannten Restklassenkörper.

III.3 Der Körper der komplexen Zahlen

In diesem Abschnitt führen wir die sogenannten komplexen Zahlen ein. Diese sind Paare von reellen Zahlen, also Elemente von $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$. Auf der Menge $\mathbb{R}^2$ hatten wir oben schon die naheliegende Addition

$$(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2) \quad \forall (a_1, b_1), (a_2, b_2) \in \mathbb{R}^2$$

eingeführt und festgestellt, dass es sich bei $(\mathbb{R}^2, +)$ um eine kommutative Gruppe handelt.

Als Nächstes wollen wir auch eine Multiplikation auf $\mathbb{R}^2$ erklären und zwar derart, dass ein Körper entsteht. Die Definition hierzu mag zunächst etwas seltsam wirken, wir werden aber später sehen, was der Sinn dahinter ist.

Definition III.3.1. Für alle $(a_1, b_1), (a_2, b_2) \in \mathbb{R}^2$ setzen wir

$$(a_1, b_1) \cdot (a_2, b_2) := (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1).$$

Zunächst zeigen wir, dass $(\mathbb{R}^2, +, \cdot)$ tatsächlich einen Körper bildet.

Satz III.3.2. *$(\mathbb{R}^2, +, \cdot)$ ist ein Körper.*

Beweis. Wir wissen schon, dass $(\mathbb{R}^2, +)$ eine kommutative Gruppe ist. Das Assoziativ- und das Kommutativgesetz für die Multiplikation, sowie das Distributivgesetz kann man direkt nachrechnen (Übung). Ebenfalls rechnet man leicht nach, dass $(1, 0)$ neutrales Element der Multiplikation ist. Damit fehlen uns zu einem Körper nur noch die multiplikativen Inversen. Sei also $(a, b) \in \mathbb{R}^2 \setminus \{(0, 0)\}$. Dann gilt

$$(a, b) \cdot \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right) = \left(\frac{a^2}{a^2 + b^2} - \frac{-b^2}{a^2 + b^2}, \frac{-ab}{a^2 + b^2} + \frac{ab}{a^2 + b^2} \right) = (1, 0)$$

und wegen der Kommutativität der Multiplikation auch

$$\left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right) \cdot (a, b) = (1, 0).$$

Also ist

$$(a, b)^{-1} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right).$$

□

Nun können wir den Körper auch offiziell benennen.

Definition III.3.3. Der Körper $(\mathbb{R}^2, +, \cdot)$ heißt *Körper der komplexen Zahlen* und wird mit $\mathbb{C}$ bezeichnet. Die komplexe Zahl $i := (0,1)$ wird die *imaginäre Einheit* genannt.

Entscheidend ist nun die Beobachtung $i^2 = i \cdot i = (0,1) \cdot (0,1) = (-1,0)$. Ferner beachte man, dass $(a,0) + (b,0) = (a+b,0)$ und $(a,0)(b,0) = (ab,0)$ für alle $a, b \in \mathbb{R}$ gilt. Indem man also eine reelle Zahl a mit der komplexen Zahl $(a,0)$ identifiziert, kann man $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ auffassen. In diesem Sinne gilt dann also $i^2 = -1$.

Das ist die wesentliche Motivation für die Einführung der komplexen Zahlen, denn im Bereich der reellen Zahlen hat die Gleichung $x^2 = -1$ keine Lösung.

Nun können wir komplexe Zahlen auch etwas anders darstellen, es gilt nämlich $(a,b) = a + ib$, wie man leicht nachrechnet. Die Multiplikation kann man dann mit Hilfe des Distributivgesetzes und der Beziehung $i^2 = -1$ ganz einfach ausführen: Es ist

$$(a+ib)(c+id) = ac + iad + ibc + i^2bd = ac - bd + i(ad+bc).$$

Zum Beispiel ist $(1+i)(2+i) = 2 + 2i + i + i^2 = 1 + 3i$.

Als Nächstes führen wir noch die folgenden Begriffe ein.

Definition III.3.4. Sei $z = a+ib$ eine komplexe Zahl. Dann heißt $\mathrm{Re}(z) := a$ der *Realteil* und $\mathrm{Im}(z) := b$ der *Imaginärteil* von z. Ferner heißt $\bar{z} := a - ib$ die *komplex konjugierte Zahl* von z.

Stellt man sich die komplexen Zahlen geometrisch als Punkte in der Ebene vor, so ist der Realteil gerade die Koordinate auf der horizontalen Achse und der Imaginärteil die Koordinate auf der vertikalen Achse. Die Operation der komlexen Konjugation bedeutet geometrisch eine Spiegelung an der horizontalen Achse. Mit ihrer Hilfe lassen sich auch Brüche komplexer Zahlen leicht berechnen, indem man nämlich mit dem komplex Konjugierten des Nenners erweitert: Seien $z = a + ib$ und $w = c + id$ zwei komplexe Zahlen, wobei $w \neq 0$ sei. Dann gilt

$$\frac{z}{w} = \frac{z\bar{w}}{w\bar{w}} = \frac{z\bar{w}}{(c+id)(c-id)} = \frac{z\bar{w}}{c^2+d^2},$$

wobei wir im letzten Schritt die dritte binomische Formel[3], sowie die Tatsache $i^2 = -1$ ausgenutzt haben. Der Witz hierbei ist, dass nun im Nenner nur noch eine reelle Zahl steht und den Zähler $z\bar{w}$ kann man leicht ausrechnen. *Beispiel*: Es ist

$$\frac{1+i}{3+2i} = \frac{(1+i)(3-2i)}{(3+2i)(3-2i)} = \frac{3-2i+3i-2i^2}{9+4} = \frac{5}{13} + \frac{1}{13}i.$$

[3] $(x+y)(x-y) = x^2 - y^2$ Diese Formel gilt nicht nur in $\mathbb{R}$ sondern in jedem beliebigen Körper (Beweis durch direktes Nachrechnen), also insbesondere auch in $\mathbb{C}$.

Nun definieren wir noch den Betrag einer komplexen Zahl.

Definition III.3.5. Sei $z = a + ib \in \mathbb{C}$ (wobei $a, b \in \mathbb{R}$). Wir setzen $|z| := \sqrt{a^2 + b^2}$.

Aus dem Satz des Pythagoras[4] ergibt sich, dass $|z|$ gerade der Abstand des Punktes (a, b) vom Koordinatenursprung $(0, 0)$ ist.

Bezeichnet man mit φ den Winkel, der von der Verbindungsstrecke zwischen $(0, 0)$ und (a, b) und der positiven reellen Achse eingeschlossen wird, so gilt $a = |z| \cos(\varphi)$ und $b = |z| \sin(\varphi)$. Dabei steht sin für Sinus und cos für Kosinus und man verwendet die Definitionen

$$\sin(\varphi) = \text{Länge der Gegenkathete durch Länge der Hypotenuse},$$
$$\cos(\varphi) = \text{Länge der Ankathete durch Länge der Hypotenuse}.$$

Es gilt also

$$z = |z|(\cos(\varphi) + i \sin(\varphi)).$$

Diese Darstellung nennt man auch die *Polarkoordinatendarstellung* der komplexen Zahl z. Der Winkel φ ist natürlich nur bis auf Addition eines ganzzahligen Vielfachen von 2π eindeutig bestimmt.

Ohne Beweis halten wir die folgenden *Additionstheoreme* für Sinus und Kosinus fest: Für alle Winkel φ und ψ gilt

$$\sin(\varphi + \psi) = \sin(\varphi)\cos(\psi) + \cos(\varphi)\sin(\psi),$$
$$\cos(\varphi + \psi) = \cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi).$$

Sind nun zwei komplexe Zahlen z und w mit Polarkoordinatendarstellungen

$$z = |z|(\cos(\varphi) + i \sin(\varphi)) \quad \text{und} \quad w = |w|(\cos(\psi) + i \sin(\psi))$$

gegeben, so folgt durch Ausmultiplizieren mit Hilfe der Additionstheoreme leicht

$$zw = |z||w|(\cos(\varphi + \psi) + i \sin(\varphi + \psi)).$$

Damit haben wir eine anschauliche geometrische Deutung der Multiplikation in $\mathbb{C}$: Die Beträge der komplexen Zahlen werden im gewöhnlichen (reellen) Sinne multipliziert, die zugehörigen Winkel werden addiert.

Im nächsten Abschnitt wollen wir nun noch weitere Beispiele für Körper kennenlernen, die sich von den bisherigen Beispielen $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ stark unterscheiden, insofern sie nämlich nur aus endlich vielen Elementen bestehen.

[4] In einem rechtwinkligen Dreieck mit Hypotenuse der Länge c und Katheten der Längen a und b gilt $c^2 = a^2 + b^2$.

III.4 Restklassenkörper

In diesem Abschnitt wollen wir uns mit Restklassen beschäftigen. Dazu führen wir zunächst allgemein den Begriff einer Äquivalenzrelation ein.

Definition III.4.1. Es sei M eine nicht leere Menge. Unter einer *Äquivalenzrelation* auf M versteht man eine Teilmenge $R \subseteq M \times M$, die die folgenden Bedingungen erfüllt:

(i) $(a, a) \in R$ für alle $a \in M$ (Reflexivität).

(ii) Für alle $a, b \in M$ gilt: $(a, b) \in R \Rightarrow (b, a) \in R$ (Symmetrie).

(iii) Für alle $a, b, c \in M$ gilt: $(a, b) \in R$ und $(b, c) \in R \Rightarrow (a, c) \in R$ (Transitivität).

Anstelle von $(a, b) \in R$ schreibt man häufig auch aRb.

Beispiele:
1) Sei M irgendeine nicht leere Menge und $R := \{(a, a) : a \in M\}$. Mit anderen Worten es gilt $(a, b) \in R$ genau dann, wenn $a = b$ ist. R ist also gerade die Gleichheitsrelation. Diese ist offensichtlich eine Äquivalenzrelation.
2) Sei $R := \big\{(a, b) \in \mathbb{R}^2 : |a| = |b|\big\}$. Dann ist R eine Äquivalenzrelation auf $\mathbb{R}$, wie man leicht sieht.
3) Sei $R := \big\{(a, b) \in \mathbb{R}^2 : a - b \in \mathbb{Q}\big\}$. Dann ist R wiederum eine Äquivalenzrelation auf $\mathbb{R}$. Beweis:
(a) Reflexivität: Für alle $a \in \mathbb{R}$ ist $a - a = 0 \in \mathbb{Q}$, also $(a, a) \in R$.
(b) Symmetrie: Ist $(a, b) \in R$, so ist definitionsgemäß $a - b \in \mathbb{Q}$. Dann ist aber auch $b - a = -(a - b) \in \mathbb{Q}$, also $(b, a) \in R$.
(c) Transitivität: Seien $(a, b) \in R$ und $(b, c) \in R$. Dann ist $a - b \in \mathbb{Q}$ und $b - c \in \mathbb{Q}$. Folglich ist auch $a - c = a - b + (b - c) \in \mathbb{Q}$ und somit $(a, c) \in \mathbb{R}$.

Als Nächstes führen wir noch den zugehörigen Begriff der Äquivalenzklassen ein.

Definition III.4.2. Sei M eine nicht leere Menge und R eine Äquivalenzrelation auf M. Für alle $a \in M$ setzen wir

$$[a]_R := \{b \in M : (a, b) \in R\}.$$

$[a]_R$ heißt die von a erzeugte *Äquivalenzklasse* (bzgl. R).

Die Äquivalenzklasse $[a]_R$ besteht also aus all jenen Elementen, welche bzgl. der Äquivalenzrelation R zu a äquivalent sind. Wie sehen die Äquivalenzklassen in den obigen Beispielen aus?
1) Ist $R := \{(a, a) : a \in M\}$, so ist offenbar $[a]_R = \{a\}$ für alle $a \in M$.
2) Ist $R := \big\{(a, b) \in \mathbb{R}^2 : |a| = |b|\big\}$, so gilt $[a]_R = \{b \in \mathbb{R} : |a| = |b|\} =$

$\{a, -a\}$ für alle $a \in \mathbb{R}$.
3) Ist $R := \big\{(a,b) \in \mathbb{R}^2 : a - b \in \mathbb{Q}\big\}$, so ist $[a]_R = \{b \in \mathbb{R} : a - b \in \mathbb{Q}\} = \{a + q : q \in \mathbb{Q}\}$. Das können Sie zur Übung leicht selbst nachweisen.

Hier noch ein allgemeines Lemma über Äquivalenzklassen.

Lemma III.4.3. *Sei M eine nicht leere Menge und R eine Äquivalenzrelation auf M. Weiter seien $a, b \in M$. Dann sind folgende Aussagen äquivalent:*

(i) $(a,b) \in R$

(ii) $[a]_R = [b]_R$

(iii) $[a]_R \cap [b]_R \neq \emptyset$

Beweis. (i) $\Rightarrow$ (ii): Es gelte $(a,b) \in R$. Sei $c \in [a]_R$ beliebig. Dann ist $(a,c) \in R$. Aus $(a,b) \in R$ folgt wegen der Symmetrie auch $(b,a) \in R$. Wegen der Transitivität folgt aus $(b,a) \in R$ und $(a,c) \in R$ auch $(b,c) \in R$. Also ist $c \in [b]_R$.
Damit ist $[a]_R \subseteq [b]_R$ gezeigt. Analog beweist man auch $[b]_R \subseteq [a]_R$. Somit ist $[a]_R = [b]_R$.
(ii) $\Rightarrow$ (iii): Sei $[a]_R = [b]_R$. Dann ist natürlich $[a]_R \cap [b]_R = [a]_R \neq \emptyset$, denn $a \in [a]_R$.
(iii) $\Rightarrow$ (i): Sei $[a]_R \cap [b]_R \neq \emptyset$. Dann existiert ein $c \in [a]_R \cap [b]_R$. Es folgt $(a,c) \in R$ und $(b,c) \in R$. Mit Symmetrie und Transitivität folgt daraus $(a,b) \in R$. □

Nun definieren wir die für uns entscheidenden Äquivalenzrelationen auf $\mathbb{Z}$.

Definition III.4.4. Sei $m \in \mathbb{N}$ mit $m \geq 2$ und seien $a, b \in \mathbb{Z}$. Dann heißt *a kongruent zu b modulo m* (in Zeichen: $a \equiv b \mod m$), falls $a - b$ teilbar durch m ist. Wir setzen $R_m := \big\{(a,b) \in \mathbb{Z}^2 : a \equiv b \mod m\big\}$.

Zum Beispiel ist $9 \equiv 3 \mod 2$, denn $9 - 3 = 6$ ist teilbar durch 2.

Wir weisen nun zunächst nach, dass es sich bei R_m wirklich um eine Äquivalenzrelation handelt.

Lemma III.4.5. *Sei $m \in \mathbb{N}$ mit $m \geq 2$. Dann ist R_m eine Äquivalenzrelation auf $\mathbb{Z}$.*

Beweis. 1) Reflexivität: Für alle $a \in \mathbb{Z}$ gilt $a \equiv a \mod m$, denn $a - a = 0$ ist natürlich durch m teilbar.
2) Symmetrie: Sei $a \equiv b \mod m$. Dann gilt also $m \mid a - b$ und folglich auch $m \mid -(a - b)$, also $m \mid b - a$. Daher ist auch $b \equiv a \mod m$.
3) Transitivität: Es gelte $a \equiv b \mod m$ und $b \equiv c \mod m$, also $m \mid a - b$ und $m \mid b - c$. Dann ist m auch ein Teiler von $a - b + b - c = a - c$, also gilt $a \equiv c \mod m$. □

Als Nächstes zeigen wir, dass zwei Zahlen a und b genau dann kongruent modulo m sind, wenn sie bei Division durch m denselben Rest lassen.

Lemma III.4.6. *Sei $m \in \mathbb{N}$ mit $m \geq 2$ und seien $a, b \in \mathbb{Z}$. Seien $q_1, q_2 \in \mathbb{Z}$ und $r_1, r_1 \in \mathbb{N}_0$ mit $a = q_1 m + r_1$, $b = q_2 m + r_2$ und $r_1, r_2 < m$.*
Dann gilt: $a \equiv b \mod m \Leftrightarrow r_1 = r_2$

Beweis. 1) Es gelte $r_1 = r_2$. Dann ist $a - b = m(q_1 - q_2)$, also $m \mid a - b$.
2) Es gelte $m \mid a - b$. Wegen $a - b = m(q_1 - q_2) + r_1 - r_2$ ist $r_1 - r_1 = a - b - m(q_1 - q_2)$ und es folgt $m \mid r_1 - r_2$, also $r_1 - r_2 = mk$ für ein gewisses $k \in \mathbb{Z}$.
Wäre $k \neq 0$, so wäre $|r_1 - r_2| = |k|m \geq m$. Andererseits ist wegen $0 \leq r_1, r_2 < m$ aber $|r_1 - r_2| < m$.
Also muss $k = 0$ und somit $r_1 = r_2$ gelten. □

Nun betrachten wir die Äquivalenzklassen in $\mathbb{Z}$ bzgl. der Äquivalenzrelation R_m. Anstelle von $[a]_{R_m}$ schreiben wir kurz $[a]_m$. Man nennt diese Äquivalenzklasse auch die *Restklasse von a modulo m*. Die Menge aller Restklassen modulo m bezeichnen wir mit $\mathbb{Z}_m$, also

$$\mathbb{Z}_m := \{[a]_m : a \in \mathbb{Z}\}.$$

Auch wenn es auf den ersten Blick vielleicht nicht so aussieht, ist die Menge $\mathbb{Z}_m$ tatsächlich endlich, wie das folgende Lemma zeigt.

Lemma III.4.7. *Sei $m \in \mathbb{N}$ mit $m \geq 2$. Die Menge $\mathbb{Z}_m$ ist endlich mit m Elementen, genauer gilt $\mathbb{Z}_m = \{[0]_m, [1]_m, \ldots, [m-1]_m\}$, wobei die Klassen $[0]_m, [1]_m, \ldots, [m-1]_m$ paarweise verschieden sind.*

Beweis. Ist $a \in \mathbb{Z}$, so schreiben wir per Division mit Rest $a = qm + r$, wobei $q \in \mathbb{Z}$ und $r \in \{0, \ldots, m-1\}$ ist. Somit ist $a \equiv r \mod m$ und daher nach Lemma III.4.3 $[a]_m = [r]_m$.
Damit ist $\mathbb{Z}_m = \{[0]_m, [1]_m, \ldots, [m-1]_m\}$ gezeigt.
Sind $i, j \in \{0, \ldots, m-1\}$ mit $i \neq j$, so folgt aus Lemma III.4.6 $i \not\equiv j \mod m$ und daher ist nach Lemma III.4.3 $[i]_m \neq [j]_m$. □

Wir wollen nun auf $\mathbb{Z}_m$ eine Addition und eine Multiplikation einführen. Dazu liegt folgende Definition nahe:

$$[a]_m + [b]_m := [a + b]_m \quad \text{und} \quad [a]_m \cdot [b]_m := [ab]_m \quad \text{für } a, b \in \mathbb{Z}.$$

Allerdings gibt es hier noch ein kleines Problem: Was wenn $[a]_m = [c]_m$ und $[b]_m = [d]_m$ ist? Ist dann auch $[a + b]_m = [c + d]_m$ und $[ab]_m = [cd]_m$? Anderenfalls wäre die Definition nicht sinnvoll.

Sei also $[a]_m = [c]_m$ und $[b]_m = [d]_m$. Dann ist $m \mid a - c$ und $m \mid b - d$. Wegen $a+b-(c+d) = a-c+b-d$ ist dann auch $a+b-(c+d)$ durch m teilbar, also $a + b \equiv c + d \mod m$ und Lemma III.4.3 impliziert $[a + b]_m = [c + d]_m$.

Weiter gilt $ab - cd = a(b - d) + d(a - c)$, also ist m auch ein Teiler von $ab - cd$ und es folgt wie eben $[ab]_m = [cd]_m$.

Addition und Multiplikation in $\mathbb{Z}_m$ sind also wohldefiniert.

Nun ist es nicht schwierig, folgendes zu beweisen.

Lemma III.4.8. *Sei $m \in \mathbb{N}$ mit $m \geq 2$. Dann ist $(\mathbb{Z}_m, +)$ eine kommutative Gruppe und $(\mathbb{Z}_m, \cdot)$ ist ein kommutatives Monoid. Ferner gilt das Distributivgesetz in $(\mathbb{Z}_m, +, \cdot)$.*

Beweis. Das können Sie zur Übung leicht selbst nachrechnen (neutrales Element in $(\mathbb{Z}_m, +)$ ist natürlich $[0]_m$, das additive Inverse von $[a]_m$ ist $[-a]_m$, neutrales Element in $(\mathbb{Z}_m, \cdot)$ ist $[1]_m$). □

Den Inhalt des obigen Lemmas fasst man in der Sprache der Algebra auch folgendermaßen zusammen: $(\mathbb{Z}_m, +, \cdot)$ ist ein kommutativer Ring mit Einselement.

Das einzige, was $\mathbb{Z}_m$ zu einem Körper noch fehlt, sind die multiplikativen Inversen. Es stellt sich jedoch heraus, dass diese nicht immer existieren. Der folgende Satz gibt Auskunft darüber, wann $\mathbb{Z}_m$ ein Körper ist.

Satz III.4.9. *Sei $m \in \mathbb{N}$ mit $m \geq 2$. Dann gilt: $(\mathbb{Z}_m, +, \cdot)$ ist ein Körper genau dann, wenn m eine Primzahl ist.*

Beweis. 1) Sei m eine Primzahl. Wir zeigen zunächst folgendes: Sind $[a]_m, [b]_m \in \mathbb{Z}_m$ mit $[a]_m[b]_m = [0]_m$, so ist $[a]_m = [0]_m$ oder $[b]_m = [0]_m$.
Ist nämlich $[0]_m = [a]_m[b]_m = [ab]_m$, so gilt $m \mid ab$ und da m eine Primzahl ist, folgt aus Lemma II.3.6 $m \mid a$ oder $m \mid b$, also ist $[a]_m = [0]_m$ oder $[b]_m = [0]_m$.
Nun setzen wir zur Abkürzung $\mathbb{Z}_m^* := \mathbb{Z}_m \setminus \{[0]_m\}$, nehmen uns ein beliebiges Element $[a]_m \in \mathbb{Z}_m^*$ her und betrachten die Abbildung $f : \mathbb{Z}_m^* \to \mathbb{Z}_m^*$, die durch $f([b]_m) := [a]_m[b]_m$ definiert ist. Aufgrund unserer Vorüberlegung bildet f wirklich nach $\mathbb{Z}_m^*$ ab.
Weiter ist die Abbildung f auch injektiv, denn ist $f([b]_m) = f([c]_m)$, so folgt $[a]_m[b - c]_m = [0]_m$ und wegen $[a]_m \neq [0]_m$ folgt $[b - c]_m = [0]_m$, also $[b]_m = [c]_m$.
f bildet also verschiedene Elemente von $\mathbb{Z}_m^*$ wieder auf verschiedene Elemente von $\mathbb{Z}_m^*$ ab. Da die Menge $\mathbb{Z}_m^*$ aber endlich ist (sie hat $m - 1$ Elemente), ist das nur möglich, wenn auch jedes Element von $\mathbb{Z}_m^*$ von der Abbildung f getroffen wird. Mit anderen Worten: f ist auch surjektiv.
Insbesondere muss es ein $[b]_m \in \mathbb{Z}_m^*$ mit $f([b]_m) = [1]_m$ geben. Es folgt $[a]_m[b]_m = [1]_m$, also ist $[b]_m$ inverses Element zu $[a]_m$.
Damit bewiesen, dass $(\mathbb{Z}_m, +, \cdot)$ einen Körper bildet.
2) Sei nun m keine Primzahl. Dann existieren natürliche Zahlen k und l mit $1 < k, l < m$ und $kl = m$. Es folgt $[k]_m[l]_m = [kl]_m = [m]_m = [0]_m$.
Wegen $1 < k < m$ ist $[k]_m \neq [0]_m$. Wäre $(\mathbb{Z}_m, +, \cdot)$ ein Körper, so gäbe es also ein $[a]_m \in \mathbb{Z}_m$ mit $[a]_m[k]_m = [1]_m$.

Es folgt $[l]_m = [a]_m([k]_m[l]_m) = [a]_m[0]_m = [0]_m$, was aber wegen $1 < l < m$ unmöglich ist.
Also ist $(\mathbb{Z}_m, +, \cdot)$ in diesem Fall kein Körper. □

Für eine Primzahl p bezeichnet man $\mathbb{Z}_p$ als den zugehörigen *Restklassenkörper*. Dieser besteht aus p Elementen.

Als eine Anwendung der Restklassen beweisen wir nun den *Kleinen Satz von Fermat*.[5]

Satz III.4.10 (Kleiner Satz von Fermat)**.** *Sei p eine Primzahl und a eine natürliche Zahl, die nicht durch p teilbar ist. Dann gilt $a^{p-1} \equiv 1 \mod p$.*

Beweis. Da p eine Primzahl ist, ist wie oben bewiesen $\mathbb{Z}_p$ ein Körper, d. h. jedes Element aus $\mathbb{Z}_p^* := \mathbb{Z}_p \setminus \{[0]_p\} = \{[1]_p, \ldots, [p-1]_p\}$ besitzt ein multiplikatives Inverses. Da a nicht durch p teilbar ist, ist $[a]_p \in \mathbb{Z}_p^*$.
Wir definieren nun eine Abbildung $\sigma : \mathbb{Z}_p^* \to \mathbb{Z}_p^*$ durch $\sigma([b]_p) := [a]_p[b]_p$.
Dann ist σ bijektiv. Beweis dazu: Ist $\sigma([b]_p) = \sigma([c]_p)$, so folgt durch Multiplikation mit $[a]_p^{-1}$ auch $[b]_p = [c]_p$, also ist σ injektiv. Ist ferner $[c]_p \in \mathbb{Z}_p^*$, so ist auch $[a]_p^{-1}[c]_p \in \mathbb{Z}_p^*$ mit $\sigma([a]_p^{-1}[c]_p) = [c]_p$, also ist σ auch surjektiv.
Wegen der Bijektivität von σ ist also $\sigma([1]_p), \ldots, \sigma([p-1]_p)$ eine Umordnung von $[1]_p, \ldots, [p-1]_p$ und wegen der Kommutativität der Multiplikation folgt

$$\begin{aligned}[1]_p[2]_p \ldots [p-1]_p &= \sigma([1]_p)\sigma([2]_p) \ldots \sigma([p-1]_p) \\ &= [a]_p[1]_p[a]_p[2]_p \ldots [a]_p[p-1]_p = [a^{p-1}]_p[1]_p[2]_p \ldots [p-1]_p\end{aligned}$$

Multipliziert man nun mit dem Inversen von $[1]_p[2]_p \ldots [p-1]_p$, so folgt $[1]_p = [a^{p-1}]_p$, also $a^{p-1} \equiv 1 \mod p$. □

Dieser Satz liefert das folgende Korollar.

Korollar III.4.11. *Für alle Primzahlen p und alle $a \in \mathbb{N}$ gilt $a^p \equiv a \mod p$.*

Beweis. Ist p ein Teiler von a, so gilt natürlich auch $p \mid a^p - a$, also $a^p \equiv a \mod p$.
Anderenfalls gilt nach dem Kleinen Satz von Fermat $p \mid a^{p-1} - 1$, also auch $p \mid a(a^{p-1} - 1)$, also $a^p \equiv a \mod p$. □

Eine Verallgemeinerung des Kleinen Satzes von Fermat ist der *Satz von Euler*[6]. Für $n \in \mathbb{N}$ bezeichne $\varphi(n)$ die Anzahl aller $a \in \{1, \ldots, n\}$ mit $\mathrm{ggT}(a, n) = 1$ (Eulersche φ-Funktion). Zum Beispiel gilt für Primzahlen p stets $\varphi(p) = p - 1$.

[5] Benannt nach Pierre de Fermat (1607–1665): französischer Mathematiker und Jurist. Nach ihm ist auch das Fermatsche Prinzip benannt (ein Variationsprinzip der theoretischen Physik, das die Ausbreitung von Licht in einem Medium beschreibt). Der Große Satz von Fermat besagt, dass für natürliche Zahlen $n \geq 3$ keine natürlichen Zahlen a, b, c mit $a^n + b^n = c^n$ existieren. Bewiesen wurde dieser Satz allerdings erst 1994 von Andrew Wiles.

[6] Leonhard Euler (1707–1783): Schweizer Mathematiker und Physiker mit diversen wichtigen Beiträgen u. a. zur Analysis und zur Zahlentheorie.

Satz III.4.12 (Satz von Euler)**.** *Für alle natürlichen Zahlen $n \geq 2$ und alle $a \in \mathbb{N}$ mit* $\mathrm{ggT}(a, n) = 1$ *gilt* $a^{\varphi(n)} \equiv 1 \mod n$.

Auf einen Beweis dieses Satzes wollen wir hier verzichten.

Als Letztes befassen wir uns noch mit dem *Satz von Wilson*[7]. Dazu zunächst noch eine Definition.

Definition III.4.13. Für $n \in \mathbb{N}$ sei

$$n! := \prod_{i=1}^{n} i = 1 \cdot 2 \ldots n.$$

$n!$ wird gelesen als n *Fakultät*. Zusätzlich definiert man noch $0! := 1$.

Zum Beispiel ist $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$.

Der Satz von Wilson lautet nun wie folgt.

Satz III.4.14 (Satz von Wilson)**.** *Sei $n \in \mathbb{N}$ mit $n \geq 2$. Dann gilt: n ist eine Primzahl genau dann, wenn $(n-1)! + 1$ durch n teilbar ist.*

Beweis. 1) Wir nehmen zunächst an, dass n keine Primzahl ist. Dann existiert ein $k \in \{2, \ldots, n-1\}$ mit $k \mid n$. Wegen $(n-1)! = 1 \cdot 2 \ldots k \ldots n-1$ gilt $k \mid (n-1)!$.
Wäre k auch ein Teiler von $(n-1)! + 1$, so wäre k ein Teiler von $(n-1)! + 1 - (n-1)! = 1$, was wegen $k \geq 2$ unmöglich ist. Also ist k kein Teiler von $(n-1)! + 1$. Wegen $k \mid n$ ist daher erst recht n kein Teiler von $(n-1)! + 1$.
D. h. umgekehrt: Ist n ein Teiler von $(n-1)! + 1$, so ist n eine Primzahl.
2) Sei nun n eine Primzahl. Dann ist $\mathbb{Z}_n$ ein Körper (Satz III.4.9), also besitzt jedes Element in $\mathbb{Z}_n^* := \mathbb{Z}_n \setminus \{[0]_n\} = \{[1]_n, \ldots, [n-1]_n\}$ ein multiplikatives Inverses. Es sei $A := \left\{v \in \mathbb{Z}_n^* : v^{-1} = v\right\}$
Natürlich gilt $[1]_n \in A$ und $[n-1]_n = [-1]_n \in A$. Ist umgekehrt $v \in A$, so folgt $v^2 = vv = vv^{-1} = [1]_n$ und somit $(v - [1]_n)(v + [1]_n) = v^2 - [1]_n = [0]_n$. Ist $v \neq [1]_n$, so ist $v - [1]_n \in \mathbb{Z}_n^*$ und Multiplikation mit $(v - [1]_n)^{-1}$ liefert $v + [1]_n = [0]_n$, also $v = [-1]_n = [n-1]_n$.
Es gilt also $A = \{[1]_n, [n-1]_n\}$.
Nun betrachten wir das Produkt $[(n-1)!]_n = [1]_n[2]_n \ldots [n-1]_n$. Hier lassen sich alle Elemente von $\mathbb{Z}_n^* \setminus A$ zu Paaren vv^{-1} zusammenfassen. Diese heben sich also insgesamt auf und es bleibt nur $[(n-1)!]_n = [1]_n[n-1]_n = [n-1]_n = [-1]_n$.
Also gilt $(n-1)! \equiv -1 \mod n$, d. h. $(n-1)! + 1$ ist durch n teilbar. □

Für die Praxis ist der obige Primzahltest allerdings wenig hilfreich, da die Zahlen $(n-1)!$ mit wachsendem n rasch so groß werden, dass eine Berechnung faktisch unmöglich ist (auch nicht mit modernsten Computern).

[7] Benannt nach John Wilson (1741–1793): britischer Mathematiker und Jurist.

IV Vektorräume

Mit diesem Kapitel beginnt nun die eigentliche lineare Algebra. Ihre zentralen Objekte sind die sogenannten Vektorräume, die wir im Folgenden diskutieren wollen.

IV.1 Vektorräume: Definition und Beispiele

Wir geben zunächst einfach die formale Definition an.

Definition IV.1.1. Sei K ein Körper und V eine Menge, die mit einer Addition $+ : V \times V \to V$ und einer weiteren Verknüpfung $\cdot : K \times V \to V$ versehen ist. Es gelte:
1) $(V, +)$ ist eine kommutative Gruppe.
2) $\lambda(\mu v) = (\lambda\mu)v$ für alle $\lambda, \mu \in K$ und alle $v \in V$.
3) $(\lambda + \mu)v = \lambda v + \mu v$ für alle $\lambda, \mu \in K$ und alle $v \in V$.
4) $\lambda(v + w) = \lambda v + \lambda w$ für alle $\lambda \in K$ und alle $v, w \in V$.
5) $1v = v$ für alle $v \in V$.
Dann heißt $(V, +, \cdot)$ ein *Vektorraum* über dem Körper K oder kurz ein K-Vektorraum. Die Elemente von V nennt man *Vektoren*[1]. Die Elemente von K werden in diesem Zusammenhang auch als *Skalare* bezeichnet. Die Eigenschaften 1)–5) nennt man *Vektorraumaxiome*.

Es ist zu beachten, dass die Symbole $+$ und $\cdot$ hier in zwei unterschiedlichen Bedeutungen auftreten, einmal als die Addition und Multiplikation im Körper K, einmal als die Addition und die Multiplikation mit Skalaren im Vektorraum V. Beispielsweise bezeichnet das $+$ auf der linken Seite von 3) die Addition in K, das $+$ auf der rechten Seite von 3) ist die Addition in V.

Der für uns wichtigste Fall ist $K = \mathbb{R}$, also der Fall reeller Vektorräume. Als Erstes betrachten wir nun einige Beispiele.

Beispiele:
1) Das wichtigste Beispiel zuerst: Es sei n eine natürliche Zahl und wir betrachten die Menge $\mathbb{R}^n$ aller Spalten der Länge n mit reellen Einträgen,

[1] Das Wort Vektor leitet sich ab vom lateinischen Wort "vehere", was "tragen" oder "transportieren" bedeutet.

also

$$\mathbb{R}^n := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \ldots, x_n \in \mathbb{R} \right\}.$$

Das ist im Grunde nur das n-fache kartesische Produkt $\mathbb{R} \times \cdots \times \mathbb{R}$, allerdings werden die Elemente jetzt nicht als n-Tupel (also zeilenweise), sondern als Spalten aufgeschrieben (der Grund dafür wird später bei der Matrizenrechnung klar werden).

Ein Element $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ des $\mathbb{R}^2$ kann man sich als Pfeil in der Ebene veranschaulichen, der vom Koordinatenurpsrung $(0, 0)$ zum Punkt (x_1, x_2) weist (entsprechend für Elemente des $\mathbb{R}^3$ mit Pfeilen im Raum).

Auf der Menge $\mathbb{R}^n$ definieren wir nun in naheliegender Weise eine Addition durch

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

und eine Multiplikation mit $\lambda \in \mathbb{R}$ durch

$$\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

Die Addition und die Multiplikation mit λ erfolgen also koordinatenweise.

Es ist leicht nachzuweisen, dass $(\mathbb{R}^n, +, \cdot)$ tatsächlich ein Vektorraum über dem Körper $\mathbb{R}$ ist. Die Rechenregeln beweist man, indem man sie koordinatenweise auf die entsprechenden Rechenregeln in $\mathbb{R}$ zurückführt (Übung). Das neutrale Element der Addition (der Nullvektor) ist

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

(diesen bezeichnet man üblicherweise wieder mit 0). Ferner ist natürlich

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

Wie schon gesagt, ist $\mathbb{R}^n$ das für uns wichtigste Beispiel eines Vektorraumes. Man kann aber auch zu einem beliebigen Körper K in völlig analoger Weise einen Vektorraum K^n definieren. Zumindest der komplexe Vektorraum $\mathbb{C}^n$ wird uns später auch noch begegnen.

2) Als weiteres Beispiel betrachten wir einen Vektorraum von Abbildungen. Sei M eine beliebige nicht leere Menge und sei V_M die Menge aller Abbildungen von M nach $\mathbb{R}$.
Für zwei Abbildungen $f, g : M \to \mathbb{R}$ definieren wir $f + g : M \to \mathbb{R}$ durch $(f + g)(x) := f(x) + g(x)$ für alle $x \in M$. Ist ferner $\lambda \in \mathbb{R}$, so definieren wir $\lambda f : M \to \mathbb{R}$ durch $(\lambda f)(x) := \lambda f(x)$ für alle $x \in M$.
Wiederum kann man leicht nachweisen, dass V_M auf diese Weise zu einem Vektorraum über $\mathbb{R}$ wird. In diesem Fall sind die Vektoren also Abbildungen.
3) Hier noch ein etwas exotischeres Beispiel: Man kann die Menge $\mathbb{R}$ der reellen Zahlen als einen Vektorraum über dem Körper $\mathbb{Q}$ der rationalen Zahlen auffassen. Die Addition ist einfach die übliche Addition reeller Zahlen und die Multiplikation von $x \in \mathbb{R}$ mit einem Skalar $q \in \mathbb{Q}$ ist durch das gewöhnliche Produkt qx erklärt. Aus den Körpereigenschaften von $\mathbb{R}$ folgt unmittelbar, dass auf diese Weise ein Vektorraum über $\mathbb{Q}$ entsteht.

Im nächsten Abschnitt werden wir den Begriff der Unterräume einführen, mit dem sich leicht diverse weitere Beispiele finden lassen. Zunächst halten wir aber noch ein paar einfache Aussagen fest, die in jedem beliebigen Vektorraum gelten.

Lemma IV.1.2. *Sei K ein Körper und V ein Vektorraum über K. Dann gilt für alle $v \in V$ und alle $\lambda \in K$:*

(i) $\lambda v = 0 \Leftrightarrow \lambda = 0$ *oder* $v = 0$

(ii) $(-\lambda)v = -(\lambda v) = \lambda(-v)$ *(insbesondere ist* $(-1)v = -v$*)*

Beweis. (i) $0v = 0$ und $\lambda 0 = 0$ beweist man ganz ähnlich wie Teil (a) von Lemma II.1.1 (Übung). Ist umgekehrt $\lambda v = 0$, aber $\lambda \neq 0$, so folgt $0 = \lambda^{-1}0 = \lambda^{-1}(\lambda v) = (\lambda^{-1}\lambda)v = 1v = v$.
(ii) Das beweist man ähnlich wie Teil (b) von Lemma II.1.1 (Übung). □

IV.2 Unterräume

In diesem Abschnitt geht es um Untervektorräume, also solche Vektorräume, die in einem gegebenen Vektorraum enthalten sind. Die genaue Definition lautet wie folgt.

Definition IV.2.1. Sei K ein Körper und sei V ein Vektorraum über K. Ferner sei U eine Teilmenge von V. Dann heißt U ein *Untervektorraum* oder kurz *Unterraum* von V, falls U versehen mit der Addition und Skalarmultiplikation von V selbst wieder einen Vektorraum über K bildet.

Die obige Bedingung beinhaltet natürlich insbesondere, dass U bezüglich der Addition und Skalarmultiplikation von V abgeschlossen ist, d. h. mit u und v sind auch $u + v$ und λu ($\lambda \in K$) wieder Elemente von U. Tatsächlich

ist diese Bedingung auch schon hinreichend (abgesehen von der trivialen Zusatzbedingung $U \neq \emptyset$), wie das folgende Lemma zeigt.

Lemma IV.2.2. *Seien K ein Körper, V ein Vektorraum über K und $U \subseteq V$. Dann sind folgende Aussagen äquivalent:*
(a) U ist ein Unterraum von V.
(b) $U \neq \emptyset$ und für alle $u, v \in U$ und alle $\lambda \in K$ gilt auch $u + v \in U$ und $\lambda u \in U$.

Beweis. (a) $\Rightarrow$ (b) ist klar. Gelte nun umgekehrt (b). Wir müssen die Vektorraumaxiome für U nachweisen. Da aber die Vektorraumaxiome 2)–5) ohnehin für alle Elemente von V gelten, gelten sie natürlich erst recht für alle Elemente von U, hier ist also nichts zu zeigen. Dasselbe gilt für das Assoziativ- und das Kommutativgesetz der Addition.
Weiter existiert wegen $U \neq \emptyset$ ein $u_0 \in U$ und nach Voraussetzung ist dann auch $0 = 0u_0 \in U$. Das zeigt die Existenz des neutralen Elements in U. Ebenso gilt für alle $u \in U$ auch $-u = (-1)u \in U$, also existieren auch die additiven Inversen in U. □

Beispiele:
1) Sei V ein beliebiger Vektorraum über irgendeinem Körper K. Natürlich ist V selbst stets ein Unterraum von V. Ebenso ist $\{0\}$ stets ein Unterraum, der sogenannte Nullraum.
2) Sei

$$U := \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}.$$

Dann ist U ein Unterraum des $\mathbb{R}^2$, wie man anhand des obigen Lemmas leicht einsieht. U ist gerade die x-Achse in der Ebene $\mathbb{R}^2$.
3) Verallgemeinerung von 2): Sei $n \in \mathbb{N}$ und $v \in \mathbb{R}^n \setminus \{0\}$. Dann ist

$$U := \{tv : t \in \mathbb{R}\}$$

ein Unterraum des $\mathbb{R}^n$ (Beweis wieder mit Hilfe des obigen Kriteriums). Im Falle $n = 2$ oder $n = 3$ ist U eine Gerade durch den Koordinatenursprung in Richtung des Vektors v. Dieses Beispiel funktioniert natürlich analog auch in jedem beliebigen Vektorraum.
4) Für alle $a, b \in \mathbb{R}$ sei $f_{a,b} : \mathbb{R} \to \mathbb{R}$ definiert durch $f_{a,b}(x) := ax + b$ für $x \in \mathbb{R}$. Dann ist

$$U := \{f_{a,b} : a, b \in \mathbb{R}\}$$

ein Unterraum des oben eingeführten Vektorraumes $V_{\mathbb{R}}$ aller Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ (Beweis als Übung, wieder mit dem obigen Lemma). U besteht gerade aus allen Polynomen vom Grad≤ 1.

Als Nächstes halten wir noch folgende einfache aber wichtige Beobachtung fest.

Lemma IV.2.3. *Sei K ein Körper und sei V ein Vektorraum über K. Seien $U_1, U_2 \subseteq V$ Unterräume von V. Dann ist auch $U_1 \cap U_2$ wieder ein Unterraum von V.*

Beweis. Es ist $0 \in U_1$ und $0 \in U_2$, also auch $0 \in U_1 \cap U_2$.
Seien nun $u, v \in U_1 \cap U_2$. Dann gilt $u, v \in U_1$ und $u, v \in U_2$. Da U_1 und U_2 Unterräume sind, folgt $u + v \in U_1$ und $u + v \in U_2$, sowie $\lambda u \in U_1$ und $\lambda u \in U_2$ für alle $\lambda \in K$. Somit ist auch $u + v \in U_1 \cap U_2$ und $\lambda u \in U_1 \cap U_2$ für alle $\lambda \in K$. □

Bei Vereinigungen von Unterräumen sieht die Sache dagegen deutlich anders aus: Sind U_1 und U_2 Unterräume von V, so gilt

$$U_1 \cup U_2 \text{ ist ein Unterraum von } V \;\Leftrightarrow\; U_1 \subseteq U_2 \text{ oder } U_2 \subseteq U_1.$$

Die Vereinigung von zwei Unterräumen ist also nur dann wieder ein Unterraum, wenn der eine bereits im anderen enthalten ist. Den Beweis überlasse ich Ihnen zur Übung.

Nun definieren wir noch den Begriff des von einer Teilmenge aufgespannten Unterraumes.

Definition IV.2.4. Seien K ein Körper, V ein Vektorraum über K und $A \subseteq V$ eine Teilmenge von V mit $A \neq \emptyset$. Dann heißt

$$\operatorname{span}(A) := \left\{ \sum_{i=1}^{n} \lambda_i x_i : n \in \mathbb{N}, \lambda_1, \ldots, \lambda_n \in K, x_1, \ldots, x_n \in A \right\}$$

der von A *aufgspannte (oder erzeugte) Unterraum von V* oder auch die *lineare Hülle von A*. Der Vollständigkeit halber setzt man noch $\operatorname{span}(\emptyset) := \{0\}$.

$\operatorname{span}(A)$ besteht also aus allen Vektoren der Form

$$\sum_{i=1}^{n} \lambda_i x_i = \lambda_1 x_1 + \cdots + \lambda_n x_n$$

mit $x_1, \ldots, x_n \in A$ und $\lambda_1, \ldots, \lambda_n \in K$ (das Summenzeichen in V ist analog zum Summenzeichen in $\mathbb{R}$ definiert). Einen Ausdruck der Form $\sum_{i=1}^{n} \lambda_i x_i$ nennt man auch eine *Linearkombination* der Vektoren $x_1, \ldots, x_n$. Die Länge n der Linearkombinationen in der Definition von $\operatorname{span}(A)$ ist dabei nicht fest vorgegeben sondern darf über ganz $\mathbb{N}$ variieren.

Um die Bezeichnung zu rechtfertigen, muss natürlich noch gezeigt werden, dass es sich bei $\operatorname{span}(A)$ wirklich um einen Unterraum handelt. Das tun wir im folgenden Lemma. Tatsächlich zeigen wir sogar mehr, nämlich, dass $\operatorname{span}(A)$ der kleinste Unterraum von V ist, der A enthält.

Lemma IV.2.5. *Sei K ein Körper und sei V ein Vektorraum über K. Ferner sei $A \subseteq V$. Dann gilt:*
1) $\operatorname{span}(A)$ *ist ein Unterraum von V mit $A \subseteq \operatorname{span}(A)$.*
2) Für alle Unterräume U von V mit $A \subseteq U$ gilt $\operatorname{span}(A) \subseteq U$.

Beweis. Im Fall $A = \emptyset$ sind die beiden Aussagen trivialerweise erfüllt. Sei also $A \neq \emptyset$.
1) Klar ist $0 \in \operatorname{span}(A)$. Seien nun $x, y \in \operatorname{span}(A)$ beliebig. Dann existieren $n, m \in \mathbb{N}$, $x_1, \ldots, x_n \in A$, $y_1, \ldots, y_m \in A$ und $\lambda_1, \ldots, \lambda_n \in K$, $\mu_1, \ldots, \mu_m \in K$ mit

$$x = \sum_{i=1}^{n} \lambda_i x_i \quad \text{und} \quad y = \sum_{i=1}^{m} \mu_i y_i.$$

Dann ist $x+y = \lambda_1 x_1 + \ldots \lambda_n x_n + \mu_1 y_1 + \cdots + \mu_m y_m$ eine Linearkombination von $n+m$ Elementen aus A, also gilt auch $x + y \in \operatorname{span}(A)$.
Weiter gilt für $\lambda \in K$ auch $\lambda x = \sum_{i=1}^{n} (\lambda\lambda_i) x_i \in \operatorname{span}(A)$. Also ist $\operatorname{span}(A)$ ein Unterraum von V.
Ferner gilt für alle $x \in A$ auch $x = 1x \in \operatorname{span}(A)$, also ist $A \subseteq \operatorname{span}(A)$.
2) Sei U irgendein Unterraum von V mit $A \subseteq U$. Sei $x \in \operatorname{span}(A)$ beliebig. Wieder schreiben wir x in der Form $x = \sum_{i=1}^{n} \lambda_i x_i$ mit $x_1, \ldots, x_n \in A$ und $\lambda_1, \ldots, \lambda_n \in K$. Wegen $A \subseteq U$ folgt $x_1, \ldots, x_n \in U$. Da U ein Unterraum ist, folgt daraus auch $\lambda_1 x_1, \ldots, \lambda_n x_n \in U$. Wiederum wegen der Unterraumeigenschaft von U folgt daraus $\lambda_1 x_1 + \lambda_2 x_2 \in U$, dann $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 \in U$, usw. bis man schließlich $x = \lambda_1 x_1 + \cdots + \lambda_n x_n \in U$ erhält. Also gilt $\operatorname{span}(A) \subseteq U$. □

Ist V ein Vektorraum über irgendeinem Körper K und $v \in V$, so ist $\operatorname{span}(\{v\}) = \{\lambda v : \lambda \in K\}$. Allgemeiner gilt für alle $v_1, \ldots, v_n \in V$

$$\operatorname{span}(\{v_1, \ldots, v_n\}) = \left\{ \sum_{i=1}^{n} \lambda_i v_i : \lambda_1, \ldots, \lambda_n \in K \right\}$$

(das ergibt sich leicht aus der Definition, indem man in einer beliebigen Linearkombination $\sum_{i=1}^{m} \alpha_i x_i$ mit $x_1, \ldots, x_m \in \{v_1, \ldots, v_n\}$ einfach die Terme nach den v_i sortiert und zusammenfasst).

Beispiele:
1) Wir definieren im $\mathbb{R}^n$ Vektoren $e_1, \ldots, e_n$ wie folgt:

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; e_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Der Vektor e_i hat also an der i-ten Stelle eine 1 und alle anderen Einträge sind 0.
Nun gilt für alle Vektoren $x \in \mathbb{R}^n$

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n.$$

Also gilt $\operatorname{span}(\{e_1, \ldots, e_n\}) = \mathbb{R}^n$.
2) Betrachtet man zum Beispiel nur die beiden Vektoren $e_1, e_2 \in \mathbb{R}^3$, so gilt

$$\operatorname{span}\{e_1, e_2\} = \{xe_1 + ye_2 : x, y \in \mathbb{R}\} = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\}.$$

Der aufgespannte Raum ist also gerade die x-y-Ebene.
3) Es ist

$$\operatorname{span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} \lambda - \mu \\ 2\lambda \\ 2(\lambda + \mu) \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\}.$$

Das ist eine schief im Raum liegende, durch den Ursprung verlaufende Ebene.
4) Zum Schluß noch ein Beispiel im Vektorraum $V_{\mathbb{R}}$ aller Funktionen von $\mathbb{R}$ nach $\mathbb{R}$. Für alle $k \in \mathbb{N}_0$ sei $p_k(x) := x^k$ für $x \in \mathbb{R}$. p_k ist also die k-te Potenzfunktion. Dann gilt für alle $n \in \mathbb{N}_0$

$$\operatorname{span}(\{p_0, \ldots, p_n\}) = \left\{ \sum_{i=0}^{n} a_i p_i : a_0, \ldots, a_n \in \mathbb{R} \right\}.$$

$\operatorname{span}(\{p_0, \ldots, p_n\})$ besteht also aus all jenen Funktionen $f : \mathbb{R} \to \mathbb{R}$, die sich in der Form $f(x) = \sum_{i=0}^{n} a_i x^i$ darstellen lassen, also aus allen Polynomen vom Grad$\leq n$.

Als Nächstes führen wir noch den wichtigen Begriff eines Erzeugendensystems ein.

Definition IV.2.6. Sei K ein Körper und V ein Vektorraum über K. Eine Teilmenge $A \subseteq V$ heißt *Erzeugendensystem* von V, falls $\operatorname{span}(A) = V$ gilt.

Beispiele:
1) Wir hatten oben schon gesehen, dass $\{e_1, \ldots, e_n\}$ ein Erzeugendensystem des $\mathbb{R}^n$ ist.
2) Selbstverständlich kann ein Vektorraum nicht nur auf eine Weise erzeugt werden. So ist z. B. auch

$$E := \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

ein Erzeugendensystem des $\mathbb{R}^3$, denn für alle

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$$

gilt

$$x = \frac{x_1 + x_2 - x_3}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{x_2 + x_3 - x_1}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \frac{x_1 + x_3 - x_2}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

wie man leicht nachrechnet. Also ist $\text{span}(E) = \mathbb{R}^3$.
3) Wie wir oben schon gesehen hatten, bilden die Potenzfunktionen $\{p_0, \ldots, p_n\}$ ein Erzeugendensystem für den Vektorraum aller Polynomfunktionen vom Grad$\leq n$.

Als letzten Punkt in diesem Abschnitt wollen wir noch Summen von Unterräumen betrachten. Dazu definieren wir zunächst allgemein Summen von Mengen.

Definition IV.2.7. Sei V ein Vektorraum über einem Körper K. Seien $A, B \subseteq V$. Wir setzen

$$A + B := \{a + b : a \in A, b \in B\}.$$

$A + B$ heißt die *Summe* von A und B.

Nun zeigen wir, dass die Summe von zwei Unterräumen wieder ein Unterraum ist.

Lemma IV.2.8. *Sei V ein Vektorraum über einem Körper K und seien U_1 und U_2 Unterräume von V. Dann gilt $U_1 + U_2 = \text{span}(U_1 \cup U_2)$. Insbesondere ist $U_1 + U_2$ wieder ein Unterraum von V.*

Beweis. 1) Sei $x \in U_1 + U_2$. Dann ist $x = u + v$ für gewisse $u \in U_1$, $v \in U_2$. Es folgt $u, v \in U_1 \cup U_2 \subseteq \text{span}(U_1 \cup U_2)$ und da $\text{span}(U_1 \cup U_2)$ ein Unterraum von V, muss auch $x = u + v \in \text{span}(U_1 \cup U_2)$ gelten. Also ist $U_1 + U_2 \subseteq \text{span}(U_1 \cup U_2)$.
2) Anhand des üblichen Unterraumkriteriums (Lemma IV.2.2) weist man leicht nach, dass $U_1 + U_2$ ein Unterraum von V ist (Übung).
Ferner ist $u = u + 0 \in U_1 + U_2$ für alle $u \in U_1$, also $U_1 \subseteq U_1 + U_2$. Ebenso sieht man $U_2 \subseteq U_1 + U_2$ und somit ist auch $U_1 \cup U_2 \subseteq U_1 + U_2$. Aus Lemma IV.2.5 folgt nun $\text{span}(U_1 \cup U_2) \subseteq U_1 + U_2$.
Insgesamt gilt also $\text{span}(U_1 \cup U_2) = U_1 + U_2$. □

Besonders wichtig ist der Fall einer Zerlegung des Vektorraumes V in eine sogenannte direkte Summe. Dies ist wie folgt erklärt.

Definition IV.2.9. Sei V ein Vektorraum über einem Körper K und seien U_1 und U_2 Unterräume von V. Dann nennt man V die *direkte Summe von U_1 und U_2* (in Zeichen $V = U_1 \oplus U_2$), falls $U_1 + U_2 = V$ und $U_1 \cap U_2 = \{0\}$ gilt.

Ein Beispiel: Sei $U_1 := \text{span}(\{e_1\})$ und $U_2 := \text{span}(\{e_2\})$ im $\mathbb{R}^2$. Dann ist $\mathbb{R}^2 = U_1 \oplus U_2$, wie man leicht sieht.

Direkte Summen lassen sich wie folgt charakterisieren.

Lemma IV.2.10. *Sei V ein Vektorraum über einem Körper K und seien U_1 und U_2 Unterräume von V. Dann sind folgende Aussagen äquivalent:*
1) $V = U_1 \oplus U_2$
2) Für alle $v \in V$ existiert genau ein *Paar $(u_1, u_2) \in U_1 \times U_2$ mit $u_1 + u_2 = v$.*

Beweis. 1) $\Rightarrow$ 2): Sei $V = U_1 \oplus U_2$. Dann gilt insbesondere $V = U_1 + U_2$ und somit ist die Existenzaussage in 2) klar. Seien nun $u_1, u_1' \in U_1$ und $u_2, u_2' \in U_2$ mit $u_1 + u_2 = u_1' + u_2'$. Dann ist $u_1 - u_1' \in U_1$, denn U_1 ist ein Unterraum. Ebenso ist $u_2' - u_2 \in U_2$. Es gilt aber $u_1 - u_1' = u_2' - u_2$.
Also ist $u_1 - u_1' \in U_1 \cap U_2 = \{0\}$ und es folgt $u_1 = u_1'$ und $u_2 = u_2'$. Damit ist auch die Eindeutigkeitsaussage in 2) gezeigt.
2) $\Rightarrow$ 1): Es gelte 2). Dann ist natürlich $V = U_1 + U_2$. Sei nun $v \in U_1 \cap U_2$. Nach Voraussetzung läßt sich 0 *eindeutig* als $0 = u_1 + u_2$ mit $u_1 \in U_1$ und $u_2 \in U_2$ darstellen. Eine solche Darstellung ist natürlich $0 = 0 + 0$. Eine weitere ist aber $0 = v - v$, denn $v \in U_1$ und $-v \in U_2$. Folglich muss $v = 0$ sein. Also ist $U_1 \cap U_2 = \{0\}$. □

IV.3 Lineare Unabhängigkeit, Basen und Dimension

In diesem Abschnitt beschäftigen wir uns mit dem zentralen Begriff der linearen Unabhängigkeit von Vektoren, sowie (darauf aufbauend) dem Konzept einer Basis eines Vektorraums und dem Dimensionsbegriff.

Hier nun zunächst die Definition linearer Unabhängigkeit.

Definition IV.3.1. Sei V ein Vektorraum über einem Körper K und sei $(v_1, \ldots, v_n)$ ein n-Tupel von Vektoren aus V. Dann heißt $(v_1, \ldots, v_n)$ *linear unabhängig*, falls folgendes gilt: Sind $\lambda_1, \ldots, \lambda_n \in K$, so gilt

$$\sum_{i=1}^{n} \lambda_i v_i = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0.$$

Anderenfalls heißt $(v_1, \ldots, v_n)$ *linear abhängig.*

Man kann diese Definition natürlich auch umgekehrt lesen: Lineare Unabhängigkeit von $(v_1, \ldots, v_n)$ bedeutet: Ist wenigstens einer der Skalare $\lambda_1, \ldots, \lambda_n$ von Null verschieden, so ist auch $\sum_{i=1}^{n} \lambda_i v_i \neq 0$.

Zuerst einige Bemerkungen: Ist eines der v_i gleich 0, so ist $(v_1, \ldots, v_n)$ offenbar linear abhängig, denn dann gilt $0v_1 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_n = 0$.

Taucht in $(v_1, \ldots, v_n)$ ein Vektor mehrfach auf (etwa $v_k = v_l$ mit $k \neq l$), so ist $(v_1, \ldots, v_n)$ ebenfalls linear abhängig (setze $\lambda_i := 0$ für $k \neq i \neq l$ und $\lambda_k := 1$, $\lambda_l := -1$).

Die Reihenfolge der Vektoren $v_1, \ldots, v_n$ spielt dagegen für die Frage der linearen Unabhängigkeit keine Rolle (da die Addition in V kommutativ ist).

Beispiele:
1) Ein einzelner Vektor $v \in V$ ist offenbar genau dann linear unabhängig, wenn $v \neq 0$ gilt.
2) $(e_1, \ldots, e_n)$ ist linear unabhängig im $\mathbb{R}^n$, denn sind $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit $\sum_{i=1}^n \lambda_i e_i = 0$, so folgt

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

also $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.
3) Die Vektoren

$$v_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad v_2 := \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

sind linear unabhängig im $\mathbb{R}^2$, denn sind $\lambda_1, \lambda_2 \in \mathbb{R}$ mit $\lambda_1 v_1 + \lambda_2 v_2 = 0$, so folgt $\lambda_1 + \lambda_2 = 0$ und $\lambda_1 + 2\lambda_2 = 0$. Das impliziert aber $\lambda_1 = -2\lambda_2 = 2\lambda_1$ und daraus folgt $\lambda_1 = 0 = \lambda_2$.
4) Im Gegensatz zu 2) sind die Vektoren

$$w_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad w_2 := \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

natürlich linear abhängig, denn $2w_1 - w_2 = 0$.
Allgemein sind zwei Vektoren genau dann linear abhängig, wenn der eine ein Vielfaches des anderen ist.
5) Die Vektoren

$$v_1 := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, v_2 := \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}, v_3 := \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

sind linear abhängig, da z. B. $3v_1 - v_2 - v_3 = 0$ ist.
6) Die Vektoren

$$w_1 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, w_2 := \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, w_3 := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

sind linear unabhängig im $\mathbb{R}^3$.
Beweis: Seien $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ mit $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0$. Dann folgt

$$\begin{aligned} \lambda_1 + \lambda_3 &= 0 \\ \lambda_1 + \lambda_2 &= 0 \\ \lambda_2 + \lambda_3 &= 0 \end{aligned}$$

Das impliziert $\lambda_1 = -\lambda_3 = \lambda_2 = -\lambda_1$. Daraus folgt $\lambda_1 = 0 = \lambda_2 = \lambda_3$.
7) Die Funktionen $f, g : \mathbb{R} \to \mathbb{R}$ seien definiert durch $f(x) := x^2$ und $g(x) := (x+1)^2$ für $x \in \mathbb{R}$. Dann sind f und g linear unabhängig im Vektorraum $V_{\mathbb{R}}$ aller Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$.
Beweis: Seien $\lambda_1, \lambda_2 \in \mathbb{R}$ mit $\lambda_1 f + \lambda_2 g = 0$. Es folgt $\lambda_1 x^2 + \lambda_2 (x+1)^2 = 0$ für alle $x \in \mathbb{R}$. Das lässt sich mittels binomischer Formel umformen zu

$$(\lambda_1 + \lambda_2) x^2 + 2\lambda_2 x + \lambda_2 = 0 \quad \forall x \in \mathbb{R}.$$

Insbesondere folgt für $x = 0$, dass $\lambda_2 = 0$ sein muss. Dann folgt aber $\lambda_1 x^2 = 0$ für alle $x \in \mathbb{R}$ und somit auch $\lambda_1 = 0$.

Nun definieren wir noch den Begriff der linearen Unabhängigkeit für Mengen.

Definition IV.3.2. Sei V ein Vektorraum über einem Körper K und sei $A \subseteq V$. Dann heißt A *linear unabhängig*, falls für alle paarweise verschiedenen[2] $a_1, \ldots, a_n \in A$ gilt: $(a_1, \ldots, a_n)$ ist linear unabhängig.
Anderenfalls heißt A *linear abhängig*.

Man beachte, dass gemäß dieser Definition die leere Menge $\emptyset$ linear unabhängig ist. Sind ferner $a_1, \ldots, a_n \in V$ paarweise verschieden, so ist offenbar $\{a_1, \ldots, a_n\}$ linear unabhängig genau dann, wenn $(a_1, \ldots, a_n)$ linear unabhängig ist. In manchen Vektorräumen gibt es aber auch unendliche, linear unabhängige Mengen. Hierzu ein Beispiel: Wir betrachten den Vektorraum $V_{\mathbb{R}}$ aller Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$ und darin die Menge $A := \{p_n : n \in \mathbb{N}_0\}$ der Potenzfunktionen. Wir wollen zeigen, dass A linear unabhängig ist. Dazu genügt es zu zeigen, dass $(p_0, \ldots, p_n)$ linear unabhängig ist für alle $n \in \mathbb{N}$ (wieso?). Sei also $n \in \mathbb{N}$ beliebig und seien $\lambda_0, \ldots, \lambda_n \in \mathbb{R}$ mit $\sum_{i=0}^n \lambda_i p_i = 0$, d. h. $\sum_{i=0}^n \lambda_i x^i = 0$ für alle $x \in \mathbb{R}$. Aus dem Prinzip des Koeffizientenvergleichs (siehe den Anhang über Polynome) folgt dann aber $\lambda_i = 0$ für alle $i = 0, \ldots, n$.

Nun kommen wir zur Definition einer Basis eines Vektorraums.

Definition IV.3.3. Sei V ein Vektorraum über einem Körper K und sei $B \subseteq V$. Dann heißt B eine *Basis* von V, falls B sowohl linear unabhängig als auch ein Erzeugendensystem von V ist.

[2] Das heißt $a_i \neq a_j$ für $i \neq j$.

Aus unseren bisherigen Beispielen zu Erzeugendensystemen und linearer Unabhängigkeit ergeben sich nun unmittelbar die folgenden Beispiele für Basen:
1) Die leere $\emptyset$ ist eine Basis des Nullraumes $\{0\}$.
2) Die Menge $\{e_1, \ldots, e_n\}$ ist eine Basis des $\mathbb{R}^n$. Diese wird auch die *kanonische Basis* des $\mathbb{R}^n$ genannt.
3) Die Menge

$$E := \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

ist eine Basis des $\mathbb{R}^3$.
4) Für alle $n \in \mathbb{N}_0$ ist $\{p_0, \ldots, p_n\}$ eine Basis des Vektorraums aller Polynomfunktionen vom Grad$\leq n$.
5) Die Menge $\{p_n : n \in \mathbb{N}_0\}$ ist eine Basis des Vektorraums aller Polynomfunktionen.

Basen von Vektorräumen lassen sich folgendermaßen charakterisieren.

Satz IV.3.4. *Sei V ein Vektorraum über einem Körper K und sei $B \subseteq V$. Dann sind folgende Aussagen äquivalent:*
1) B ist eine Basis von V.
2) B ist maximal linear unabhängig, *d. h. B ist linear unabhängig und ist $A \subseteq V$ eine linear unabhängige Menge mit $B \subseteq A$, so folgt $B = A$.*
3) B ist ein minimales Erzeugendensystem *von V, d. h. B ist ein Erzeugendensystem von V und ist A ein Erzeugendensystem von V mit $A \subseteq B$, so folgt $A = B$.*

Beweis. 1) $\Rightarrow$ 2): Sei B eine Basis von V. Dann ist B definitionsgemäß linear unabhängig. Sei nun $A \subseteq V$ linear unabhängig mit $B \subseteq A$ und sei $v \in A$ beliebig. Wir wollen zeigen, dass bereits $v \in B$ gelten muss.
Da B eine Basis von V ist, gilt $V = \text{span}(B)$ und folglich existieren paarweise verschiedene Vektoren $b_1, \ldots, b_n \in B$ und Skalare $\lambda_1, \ldots, \lambda_n \in K$ mit $v = \sum_{i=1}^n \lambda_i b_i$.
Nun gilt aber auch $b_1, \ldots, b_n \in A$, da $B \subseteq A$. Wäre also $v \neq b_i$ für alle $i = 1, \ldots, n$, so wären $b_1, \ldots, b_n, v \in A$ paarweise verschieden und wegen $\sum_{i=1}^n \lambda_i b_i - v = 0$ wäre $(b_1, \ldots, b_n, v)$ linear abhängig, im Widerspruch zur linearen Unabhängigkeit von A.
Also muss $v = b_i$ für ein $i \in \{1, \ldots, n\}$ gelten und somit ist $v \in B$.
2) $\Rightarrow$ 1): Sei B maximal linear unabhängig. Wir wollen zeigen, dass B auch ein Erzeugendensystem und somit eine Basis von V ist. Sei dazu $v \in V$ beliebig. Ist $v \in B$, so ist natürlich auch $v \in \text{span}(B)$, wir können also ohne Einschränkung $v \notin B$ annehmen.
Nach Voraussetzung ist dann aber die Menge $B \cup \{v\}$ linear abhängig. Daher existieren paarweise verschiedene $b_1, \ldots, b_n \in B \cup \{v\}$, so dass $(b_1, \ldots, b_n)$

linear abhängig ist. Da aber B linear unabhängig ist, muss eines der b_i gleich v sein. Ohne Einschränkung können wir $b_n = v$ annehmen.
Wegen der linearen Abhängigkeit von $(b_1, \ldots, b_n)$ existieren $\lambda_1, \ldots, \lambda_n \in K$ mit $\sum_{i=1}^{n} \lambda_i b_i = 0$, wobei nicht alle λ_i gleich 0 sind.
Wäre aber $\lambda_n = 0$, so wäre $\sum_{i=1}^{n-1} \lambda_i b_i = 0$ und wegen der linearen Unabhängigkeit von B müsste dann auch $\lambda_1 = \cdots = \lambda_{n-1} = 0$ gelten.
Also ist $\lambda_n \neq 0$ und aus $\sum_{i=1}^{n} \lambda_i b_i = 0$ folgt daher

$$v = b_n = -\frac{1}{\lambda_n} \sum_{i=1}^{n-1} \lambda_i b_i.$$

Somit ist $v \in \operatorname{span}(B)$, was die Argumentation abschließt.
1) $\Rightarrow$ 3): Sei wieder B eine Basis von V. Dann ist B auch ein Erzeugendensystem von V. Sei $A \subseteq B$ ebenfalls ein Erzeugendensystem von V und sei $b \in B$ beliebig. Wir haben $b \in A$ zu zeigen.
Wegen $\operatorname{span}(A) = V$ existieren paarweise verschiedene $a_1, \ldots, a_n \in A$ und Skalare $\lambda_1, \ldots, \lambda_n \in K$ mit $b = \sum_{i=1}^{n} \lambda_i a_i$.
Wegen $A \subseteq B$ gilt auch $a_1, \ldots, a_n \in B$. Wäre also $b \notin \{a_1, \ldots, a_n\}$, so wären $a_1, \ldots, a_n, b \in B$ paarweise verschieden und wegen $b - \sum_{i=1}^{n} \lambda_i a_i = 0$ wäre $(a_1, \ldots, a_n, b)$ linear abhängig, was der linearen Unabhängigleit von B widerspricht. Also ist $b = a_i$ für ein $i \in \{1, \ldots, n\}$ und somit $b \in A$.
3) $\Rightarrow$ 1): Sei B ein minimales Erzeugendensystem von V. Wir wollen die lineare Unabhängigkeit von B nachweisen. Seien dazu $b_1, \ldots, b_n \in B$ paarweise verschieden und seien $\lambda_1, \ldots, \lambda_n \in K$ mit $\sum_{i=1}^{n} \lambda_i b_i = 0$.
Angenommen es existiert ein $i_0 \in \{1, \ldots, n\}$ mit $\lambda_{i_0} \neq 0$. Dann folgt

$$b_{i_0} = -\frac{1}{\lambda_{i_0}} \sum_{i \in I} \lambda_i b_i, \tag{IV.1}$$

wobei $I := \{1, \ldots, n\} \setminus \{i_0\}$.
Setze $A := B \setminus \{b_{i_0}\}$. Dann folgt aus (IV.1) $b_{i_0} \in \operatorname{span}(A)$.
Ferner ist natürlich auch $A \subseteq \operatorname{span}(A)$, also $B = A \cup \{b_{i_0}\} \subseteq \operatorname{span}(A)$.
Aus Lemma IV.2.5 folgt dann aber $V = \operatorname{span}(B) \subseteq \operatorname{span}(A)$. Also ist $\operatorname{span}(A) = V$, d. h. A ist ein Erzeugendensystem von V.
Es ist aber $A \subseteq B$ und $A \neq B$ (denn $b_{i_0} \in B \setminus A$), im Widerspruch zur Voraussetzung 3).
Also muss $\lambda_i = 0$ für alle $i \in \{1, \ldots, n\}$ gelten, was zu beweisen war. □

Als Nächstes wollen wir den sogenannten Austauschsatz von Steinitz[3] beweisen. Als Vorbereitung zeigen wir zunächst ein entsprechendes Austauschlemma.

[3] Ernst Steinitz (1871–1928): deutscher Mathematiker, lieferte wichtige Beiträge vor allem zur Algebra (insbesondere zur Körpertheorie), aber zum Beispiel auch zur Graphentheorie.

Lemma IV.3.5 (Austauschlemma von Steinitz)**.** *Sei V ein Vektorraum über einem Körper K und sei B eine Basis von V. Sei $v \in V$ und seien $\lambda_1, \ldots, \lambda_n \in K$ und $b_1, \ldots, b_n \in B$ paarweise verschieden mit $v = \sum_{i=1}^{n} \lambda_i b_i$.*

Ist $j \in \{1, \ldots, n\}$ mit $\lambda_j \neq 0$, so ist $(B \setminus \{b_j\}) \cup \{v\}$ wieder eine Basis von V.

Beweis. Zur Abkürzung setzen wir $C := (B \setminus \{b_j\}) \cup \{v\}$.
1) Wir wollen zuerst zeigen, dass C ein Erzeugendensystem von V ist. Aus $v = \sum_{i=1}^{n} \lambda_i b_i$ und $\lambda_j \neq 0$ folgt durch umstellen

$$b_j = \frac{1}{\lambda_j} v - \sum_{i \in I} \frac{\lambda_i}{\lambda_j} b_i,$$

wobei $I := \{1, \ldots, n\} \setminus \{j\}$. Das zeigt $b_j \in \operatorname{span}(C)$.
Ferner ist natürlich auch $B \setminus \{b_j\} \subseteq \operatorname{span}(C)$, also $B \subseteq \operatorname{span}(C)$.
Dann muss aber auch $\operatorname{span}(B) \subseteq \operatorname{span}(C)$ gelten. Da B eine Basis von V ist, ist aber $\operatorname{span}(B) = V$. Also ist auch $\operatorname{span}(C) = V$.
2) Nun zeigen wir noch die lineare Unabhängigkeit von C. Seien dazu $c_1, \ldots, c_m \in C$ paarweise verschieden und $\alpha_1, \ldots, \alpha_m \in K$ mit $\sum_{i=1}^{m} \alpha_i c_i = 0$.
1. Fall: $c_i \neq v$ für alle $i = 1, \ldots, m$. Dann gilt $c_1, \ldots, c_m \in B \setminus \{b_j\}$ und wegen der linearen Unabhängigkeit von B folgt $\alpha_1 = \cdots = \alpha_m = 0$.
2. Fall: Es existiert ein $i_0 \in \{1, \ldots, m\}$ mit $c_{i_0} = v$. Sei $J := \{1, \ldots, m\} \setminus \{i_0\}$. Es folgt

$$0 = \sum_{i=1}^{m} \alpha_i c_i = \alpha_{i_0} v + \sum_{i \in J} \alpha_i c_i = \sum_{i=1}^{n} \alpha_{i_0} \lambda_i b_i + \sum_{i \in J} \alpha_i c_i.$$

Dies lässt sich wiederum zusammenfassen zu einer Linearkombination von paarweise verschiedenen Elementen aus der Menge B. Der Koeffizient von b_j ist dabei einfach $\alpha_{i_0} \lambda_j$, denn in der Summe $\sum_{i \in J} \alpha_i c_i$ kommt kein Term mit b_j vor. Wegen der linearen Unabhängigkeit von B folgt also $\alpha_{i_0} \lambda_j = 0$. Wegen $\lambda_j \neq 0$ muss somit $\alpha_{i_0} = 0$ gelten.
Dann folgt aber $\sum_{i \in J} \alpha_i c_i = 0$ und die lineare Unabhängigleit von B impliziert dann auch $\alpha_i = 0$ für alle $i \in J = \{1, \ldots, m\} \setminus \{i_0\}$.
Also ist auch C linear unabhängig. □

Nun kommen wir zum Austauschsatz.

Satz IV.3.6 (Austauschsatz von Steinitz)**.** *Sei V ein Vektorraum über einem Körper K. Sei B eine Basis von V und sei A eine endliche, linear unabhängige Teilmenge von V. Dann gibt es eine endliche Teilmenge $C \subseteq B$, die genauso viele Elemente wie A hat, so dass $(B \setminus C) \cup A$ wieder eine Basis von V ist.*

Der Satz besagt also, dass man die linear unabhängige Menge A in die Basis B "hineintauschen"kann.

Beweis. Der Beweis erfolgt durch vollständige Induktion nach der Anzahl der Elemente von A.
Induktionsanfang: A hat keine Elemente, also $A = \emptyset$. Dann kann (bzw. muss) man natürlich $C = \emptyset$ wählen.
Induktionsschritt: Angenommen nun die Behauptung stimmt für linear unabhängige Mengen mit n Elementen.
Sei $A \subseteq V$ linear unabhängig mit $n+1$ Elementen. Sei $v \in A$. Dann ist $A' := A \setminus \{v\}$ eine n-elementige linear unabhängige Teilmenge von V. Nach Voraussetzung existiert also eine n-elementige Teilmenge $C \subseteq B$, so dass $(B \setminus C) \cup A'$ wieder eine Basis von V ist.
Folglich können wir auch v schreiben als

$$v = \sum_{i=1}^{n} \lambda_i b_i + \sum_{i=1}^{m} \mu_i a_i,$$

wobei $\lambda_1, \ldots, \lambda_n, \mu_1 \ldots, \mu_m \in K$ und $b_1, \ldots, b_n \in B \setminus (C \cup A')$ und $a_1, \ldots, a_m \in A'$ jeweils paarweise verschieden sind.
Wäre $\lambda_1 = \cdots = \lambda_n = 0$, so wäre $0 = -v + \sum_{i=1}^{m} \mu_i a_i$, was der linearen Unabhängigkeit von $A = A' \cup \{v\}$ widerspricht.
Also existiert ein $j \in \{1, \ldots, n\}$ mit $\lambda_j \neq 0$. Aus dem Austauschlemma folgt dann, dass $(((B \setminus C) \cup A') \setminus \{b_j\}) \cup \{v\} = (B \setminus (C \cup \{b_j\})) \cup A$ wieder eine Basis von V ist. Also ist $C' := C \cup \{b_j\}$ die gesuchte $(n+1)$-elementige Teilmenge von B. □

Bislang wissen wir noch nicht, ob überhaupt jeder Vektorraum eine Basis besitzt. Es gelten jedoch die beiden folgenden Sätze.

Satz IV.3.7 (Basisergänzungssatz)**.** *Sei K ein Körper und V ein Vektorraum über K. Sei $A \subseteq V$ linear unabhängig. Dann existiert eine Basis B von V mit $A \subseteq B$.*

Satz IV.3.8 (Basisauswahlsatz)**.** *Sei K ein Körper und V ein Vektorraum über K. Sei $E \subseteq V$ ein Erzeugendensystem von V. Dann existiert eine Basis B von V mit $B \subseteq E$.*

Daraus ergibt sich insbesondere, dass jeder Vektorraum wenigstens eine Basis besitzt.

Satz IV.3.9 (Basisexistenzsatz)**.** *Sei K ein Körper und V ein Vektorraum über K. Dann besitzt V eine Basis.*

Beweis. Wende den Basisergänzungssatz auf die linear unabhängige Menge $\emptyset$ oder den Basisauswahlsatz auf das Erzeugendensystem V an. □

Die Beweise für Basisergänzungs- und Basisauswahlsatz sind jedoch in voller Allgemeinheit relativ schwierig (sie beruhen auf dem sogenannten

Zornschen Lemma, das selbst erst mit Hilfe des sogenannten Auswahlaxioms bewiesen werden muss).[4] Wir wollen sie daher hier nicht führen, sondern beschränken uns im wesentlich auf den für uns wichtigsten Fall, dass der Vektorraum ein *endliches* Erzeugendensystem besitzt, was die Sache einfacher macht. Solche Vektorräume lassen sich wie folgt charakterisieren.

Satz IV.3.10. *Sei V ein Vektorraum über einem Körper K. Dann sind folgende Aussagen äquivalent:*

(a) *V hat ein endliches Erzeugendensystem.*

(b) *Für alle Erzeugendensysteme E von V gilt: Es existiert eine endliche Teilmenge $E' \subseteq E$, so dass E' immer noch ein Erzeugendensystem von V ist.*

(c) *Für alle Erzeugendensysteme E von V gilt: Es existiert eine endliche Basis B von V mit $B \subseteq E$.*

Beweis. $(a) \Rightarrow (b)$: Es existiere ein endliches Erzeugendensystem von V, etwa $A = \{a_1, \ldots, a_n\}$. Sei nun E irgendein Erzeugendensystem von V, also $\operatorname{span}(E) = V$. Dann lässt sich insbesondere jedes a_i als Linearkombination von endlich vielen Elementen aus E schreiben, d. h. es existieren endliche Mengen $E_1, \ldots, E_n \subseteq E$ mit $a_i \in \operatorname{span}(E_i)$ für $i = 1, \ldots, n$.
Die Menge $E' := E_1 \cup E_2 \cup \cdots \cup E_n$ ist als Vereinigung endlich vieler endlicher Teilmengen von E wieder eine endliche Teilmenge von E. Ferner gilt $a_i \in \operatorname{span}(E_i) \subseteq \operatorname{span}(E')$ für alle $i = 1, \ldots, n$, also $A \subseteq \operatorname{span}(E')$ und daher auch $\operatorname{span}(A) \subseteq \operatorname{span}(E')$. Wegen $\operatorname{span}(A) = V$ folgt $\operatorname{span}(E') = V$, also ist E' ein Erzeugendensystem von V.
$(b) \Rightarrow (c)$: Es gelte (b). Wir zeigen zunächst, dass jedes endliche Erzeugendensystem C von V eine Basis von V enthält. Dazu führen wir eine vollständige Induktion nach der Anzahl der Elemente von C durch.
Induktionsanfang: Ist die Anzahl der Elemente von C gleich 0, so folgt $C = \emptyset$ und somit $V = \operatorname{span}(C) = \{0\}$. Dann ist aber $C = \emptyset$ auch eine Basis von V.
Induktionsschritt: Die Behauptung gelte für alle Erzeugendensysteme von V mit höchstens n Elementen und es sei C ein Erzeugendensystem von V mit $n + 1$ Elementen. Wir unterscheiden zwei Fälle:
1.Fall: Für alle echten Teilmengen von C, also alle $D \subseteq C$ mit $D \neq C$, gilt: D ist kein Erzeugendensytem von V. Mit anderen Worten C ist ein minimales Erzeugendensystem und daher nach Satz IV.3.4 eine Basis von V.
2.Fall: Es existiert ein $D \subseteq C$ mit $D \neq C$ und $\operatorname{span}(D) = V$. Dann ist D endlich und die Anzahl der Elemente von D ist echt kleiner die Anzahl der

[4] Es ist ferner zu bemerken, dass es sich bei diesen Sätzen um reine Existenzsätze handelt. Im Allgemeinen gibt es keine Methode, eine Basis eines Vektorraums explizit zu konstruieren. Beispielsweise hat niemand jemals eine Basis des Vektorraums $V_{\mathbb{R}}$ tatsächlich "gesehen".

Elemente von C, also höchstens n. Nach Voraussetzung existiert dann eine Basis B von V mit $B \subseteq D \subseteq C$.
Sei nun E ein beliebiges Erzeugendensystem von V. Wegen (b) existiert eine endliche Teilmenge $E' \subseteq E$, die immer noch ein Erzeugendensystem von V ist und nach unserer obigen Überlegung muss E' eine Basis von V enthalten.
$(c) \Rightarrow (a)$: Es gelte (c). Da V selbst natürlich ein Erzeugendensystem von V ist, muss also eine endliche Basis $B \subseteq V$ existieren, insbesondere gilt (a). □

Dieser Satz beinhaltet insbesondere den Basisauswahlsatz für Vektorräume mit einem endlichen Erzeugendensystem. In Kombination mit dem Austauschsatz von Steinitz erhält man nun folgenden wichtigen Anzahlsatz.

Satz IV.3.11. *Sei V ein Vektorraum über einem Körper K, sei $A \subseteq V$ linear unabhängig und sei $E \subseteq V$ ein Erzeugendensystem von V. Ist E endlich, so ist auch A endlich und hat höchstens so viele Elemente wie E.*

Beweis. Sei E endlich, sagen wir mit m Elementen. Nach Satz IV.3.10 existiert dann eine endliche Basis B von V mit $B \subseteq E$. Die Anzahl der Elemente von B sei $n \leq m$.
Angenommen nun A wäre unendlich oder aber endlich mit mehr als n Elementen. Dann gäbe es eine $(n+1)$-elementige Teilmenge $A' \subseteq A$. Da A linear unabhängig ist, trifft dies natürlich erst recht auf A' zu. Nach dem Austauschsatz von Steinitz gäbe es dann aber eine endliche Teilmenge $C \subseteq B$ mit $n+1$ Elementen, so dass $(B \setminus C) \cup A'$ wieder eine Basis von V ist. Das ist aber ein Widerspruch, denn B hat nur n Elemente.
Also ist A endlich mit höchsten $n \leq m$ Elementen. □

Damit erhalten wir nun das folgende wichtige Korollar.

Korollar IV.3.12. *Sei V ein Vektorraum über einem Körper K, welcher ein endliches Erzeugendensystem besitzt. Dann gilt:*
1) V besitzt eine endliche Basis.
2) Jede Basis von V ist endlich und je zwei Basen von V haben die gleiche Anzahl von Elementen.

Beweis. 1) Das folgt bereits aus Satz IV.3.10.
2) Sei B irgendeine Basis von V. Nach 1) existiert eine endliche Basis B_0 von V, etwa mit n Elementen. Insbesondere ist B_0 ein Erzeugendensystem und B linear unabhängig, also folgt aus Satz IV.3.11, dass B endlich ist und aus höchstens n Elementen besteht.
Andererseits ist auch B ein Erzeugendensystem und B_0 linear unabhängig, also folgt aus Satz IV.3.11 auch, dass die Anzahl der Elemente von B mindestens n sein muss. Also hat B genau n Elemente. □

Dieses Korollar gestattet es nun, den Begriff der Dimension eines Vektorraumes einzuführen.

Definition IV.3.13. Sei V ein Vektorraum über einem Körper K. Besitzt V ein endliches Erzeugendensystem, so bezeichne mit $\dim(V)$ die Anzahl der Elemente in einer Basis von V (wegen Korollar IV.3.12 ist das wohldefiniert). Anderenfalls setze $\dim(V) := \infty$.

$\dim(V)$ heißt die *Dimension* von V.

Beispiele:
1) Wir hatten oben schon gesehen, dass $\{e_1, \dots, e_n\}$ eine Basis des $\mathbb{R}^n$ bildet, also ist $\dim(\mathbb{R}^n) = n$ (wie man es erwarten würde).
2) Ebenfalls hatten iwr oben schon gesehen, dass die Potenzfunktionen $\{p_0, \dots, p_n\}$ eine Basis des Vektorraumes aller Polynomfunktionen vom Grad$\leq n$ bilden. Dieser Vektorraum hat also die Dimension $n + 1$.
3) Der Vektorraum aller Polynomfunktionen ist dagegen unendlich-dimensional, denn er hat z. B. $\{p_n : n \in \mathbb{N}_0\}$ als Basis (siehe oben).

Jetzt beweisen wir noch folgende Charakterisierung.

Lemma IV.3.14. *Sei V ein Vektorraum über einem Körper K. Dann sind folgende Aussagen äquivalent:*
1) $\dim(V) = \infty$
2) Es gibt eine unendliche Folge $v_1, v_2, \dots$ von Vektoren in V, deren Anfangsstücke $(v_1, \dots, v_n)$ jeweils linear unabhängig sind (für alle $n \in \mathbb{N}$).

Beweis. Sei zunächst $\dim(V) = \infty$. Dann besitzt V also kein endliches Erzeugendensystem. Insbesondere ist $V \neq \{0\}$ und somit existiert ein $v_1 \in V$ mit $v_1 \neq 0$. Dieses ist dann natürlich linear unabhängig.
Angenommen nun es sind bereits linear unabhängig Vektoren $(v_1, \dots, v_n)$ in V konstruiert. Da V kein endliches Erzeugendensystem besitzt, existiert ein $v_{n+1} \in V$ mit $v_{n+1} \notin \operatorname{span}\{v_1, \dots, v_n\}$. Wir wollen zeigen, dass auch $(v_1, \dots, v_{n+1})$ noch linear unabhängig ist. Seien also $\lambda_1, \dots, \lambda_{n+1} \in K$ mit $\sum_{i=1}^{n+1} \lambda_i v_i = 0$. Wäre $\lambda_{n+1} \neq 0$, so könnte man dies umstellen zu $v_{n+1} = -\sum_{i=1}^{n} \lambda_i \lambda_{n+1}^{-1} v_i$, was im Widerspruch zu $v_{n+1} \notin \operatorname{span}\{v_1, \dots, v_n\}$ steht.
Also ist $\lambda_{n+1} = 0$ und somit $\sum_{i=1}^{n} \lambda_i v_i = 0$. Wegen der linearen Unabhängigkeit von $(v_1, \dots, v_n)$ folgt daraus auch $\lambda_1 = \dots = \lambda_n = 0$ und wir sind fertig.
Ist V endlich-dimensional, etwa $\dim(V) = n$, so besitzt V eine Basis aus n Elementen und aus Satz IV.3.11 folgt daher, dass es keine $n + 1$ linear unabhängigen Vektoren in V geben kann. □

Als Folgerung erhält man, dass Unterräume von endlich-dimensionalen Vektorräumen wieder endlich-dimensional sind.

Korollar IV.3.15. *Sei V ein Vektorraum über einem Körper K mit* $\dim(V) < \infty$ *und sei $U \subseteq V$ ein Unterraum. Dann ist auch* $\dim(U) < \infty$.

Beweis. Wäre $\dim(U) = \infty$, so gäbe es nach dem vorigen Satz eine unendliche Folge $u_1, u_2, \ldots$ in U, so dass $(u_1, \ldots, u_n)$ linear unabhängig in U (und folglich auch in V) ist für alle $n \in \mathbb{N}$, was aber der Tatsache $\dim(V) < \infty$ widerspricht. □

Als Nächstes zeigen wir noch den Basisergänzungssatz für endlich-dimensionale Vektorräume.

Satz IV.3.16. *Sei V ein Vektorraum über einem Körper K mit $\dim(V) < \infty$ und sei $A \subseteq V$ linear unabhängig. Dann existiert eine Basis B von V mit $A \subseteq B$.*

Beweis. Da V endlich-dimensional ist, muss wegen Satz IV.3.11 die Menge A endlich sein. Ist B_0 eine Basis von V, so existiert also nach dem Austauschsatz von Steinitz eine Teilmenge $C \subseteq B_0$, so dass $B := (B_0 \setminus C) \cup A$ eine Basis von V ist. □

Schließlich erhalten wir noch folgende Aussage über die Dimension von Unterräumen.

Korollar IV.3.17. *Sei V ein Vektorraum über einem Körper K und sei $U \subseteq V$ ein Unterraum. Dann gilt $\dim(U) \leq \dim(V)$.*

Ist ferner $\dim(V) < \infty$, so gilt $U = V \Leftrightarrow \dim(U) = \dim(V)$.

Beweis. Die Aussage ist klar für $\dim(V) = \infty$. Sei also $\dim(V) < \infty$. Nach Korollar IV.3.15 ist dann auch $\dim(U) < \infty$ und folglich existiert eine endliche Basis A von U. Diese ist insbesondere in U und daher auch in V linear unabhängig, also existiert nach Satz IV.3.16 eine Basis B von V mit $A \subseteq B$. Folglich ist die Anzahl der Elemente von A kleiner oder gleich der Anzahl der Elemente von B, also $\dim(U) \leq \dim(V)$.
Gilt sogar $\dim(U) = \dim(V)$, so kann A keine echte Teilmenge von B sein. Also $A = B$ und es folgt $U = \operatorname{span}(A) = \operatorname{span}(B) = V$. □

Als letzten Punkt in diesem Kapitel betrachten wir noch die Dimension von direkten Summen.

Lemma IV.3.18. *Sei V ein endlichdimensionaler Vektorraum über einem Körper K und seien $U_1, U_2 \subseteq V$ Unterräume mit $V = U_1 \oplus U_2$. Dann gilt $\dim(V) = \dim(U_1) + \dim(U_2)$.*

Beweis. Seien $n_1 := \dim(U_1)$ und $n_2 := \dim(U_2)$. Sei B_1 eine Basis von U_1 und B_2 eine Basis von U_2. Dann hat B_1 genau n_1 Elemente und B_2 genau n_2 Elemente. Setze $B := B_1 \cup B_2$.
Es gilt $U_i = \operatorname{span}(B_i) \subseteq \operatorname{span}(B)$ für $i = 1, 2$, also auch $U_1 + U_2 \subseteq \operatorname{span}(B)$ (denn $\operatorname{span}(B)$ ist ein Unterraum von V). Wegen $V = U_1 \oplus U_2$ folgt daher $\operatorname{span}(B) = V$, d. h. B ist ein Erzeugendensystem von V.

Ferner gilt $B_1 \cap B_2 \subseteq U_1 \cap U_2 = \{0\}$ und da B_1 und B_2 linear unabhängig sind (folglich die 0 nicht enthalten), folgt $B_1 \cap B_2 = \emptyset$.
Schreibe $B_1 = \{v_1, \ldots, v_{n_1}\}$ und $B_2 = \{w_1, \ldots, w_{n_2}\}$. Wegen $B_1 \cap B_2 = \emptyset$ hat $B = \{v_1, \ldots, v_{n_1}, w_1, \ldots, w_{n_2}\}$ also $n_1 + n_2$ Elemente.
Außerdem ist B linear unabhängig.
Beweis dazu: Seien $\lambda_1, \ldots, \lambda_{n_1}, \mu_1, \ldots, \mu_{n_2} \in K$ mit $\sum_{i=1}^{n_1} \lambda_i v_i + \sum_{i=1}^{n_2} \mu_i w_i = 0$. Die erste Summe gehört zu U_1, die zweite zu U_2. Wegen $V = U_1 \oplus U_2$ folgt also $\sum_{i=1}^{n_1} \lambda_i v_i = 0$ und $\sum_{i=1}^{n_2} \mu_i w_i = 0$. Da B_1 und B_2 jeweils linear unabhängig sind, folgt $\lambda_i = 0$ für alle $i = 1, \ldots, n_1$ und $\mu_i = 0$ für alle $i = 1, \ldots, n_2$.
Also ist B eine Basis von V und daher $\dim(V) = n_1 + n_2$. □

V Lineare Abbildungen und Matrizen

In diesem Kapitel befassen wir uns mit sogenannten linearen Abbildungen zwischen Vektorräumen und der eng damit verwandten Matrizenrechnung.

V.1 Lineare Abbildungen

Unter einer linearen Abbildung von einem Vektorraum in einen anderen versteht man eine Abbildung, die die Vektorraumstruktur (also die Addition und Skalarmultiplikation) erhält. Die genaue Definition lautet wie folgt.

Definition V.1.1. Seien V und W zwei Vektorräume über demselben Körper K und sei $F : V \to W$ eine Abbildung. F heißt *linear*, falls folgendes gilt:

(a) $F(v + w) = F(v) + F(w)$ für alle $v, w \in V$.

(b) $F(\lambda v) = \lambda F(v)$ für alle $v \in V$ und alle $\lambda \in K$.

Die Addition und die Skalarmultiplikation in V und in W werden hier jeweils mit dem gleichen Symbol bezeichnet, was erfahrungsgemäß nicht zu Verwechslungen führt. Ebenso verfährt man hinsichtlich der neutralen und inversen Elemente von V und W.

Anstelle von einer linearen Abbildung spricht man auch von einem *Vektorraumhomomorphismus* oder kurz *Homomorphismus*. Die Menge aller linearen Abbildungen von V nach W bezeichnet man daher auch mit $\mathrm{Hom}(V, W)$. Im Fall $V = W$ spricht man auch von *Endomorphismen* und schreibt $\mathrm{End}(V)$ anstelle von $\mathrm{Hom}(V, V)$

Bevor wir zu den Beispielen kommen zunächst noch eine allgemeine Bemerkung.

Bemerkung V.1.2. Seien V und W zwei Vektorräume über demselben Körper K und sei $F : V \to W$ eine lineare Abbildung. Dann gilt $F(0) = 0$ und $F(-v) = -F(v)$ für alle $v \in V$.

Beweis. Sei v_0 irgendein Element von V. Aus der Linearität von F folgt $F(0) = F(0v_0) = 0F(v_0) = 0$.
Weiter gilt für alle $v \in V$ auch $F(-v) = F((-1)v) = (-1)F(v) = -F(v)$. □

Beispiele:
1) Für je zwei Vektorräume V und W ist die konstante Abbildung von V nach W mit Wert 0 offensichtlich linear.
2) Die Abbildung $F : \mathbb{R}^2 \to \mathbb{R}^2$ definiert durch

$$F\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) := \begin{pmatrix} x \\ 0 \end{pmatrix}$$

ist linear, denn es ist

$$F\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 + x_2 \\ 0 \end{pmatrix} = F\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right) + F\left(\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right)$$

und

$$F\left(\lambda \begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} \lambda x \\ 0 \end{pmatrix} = \lambda F\left(\begin{pmatrix} x \\ y \end{pmatrix}\right)$$

für $\lambda \in \mathbb{R}$.
3) Für alle Vektorräume V über einem Körper K und alle $\mu \in K$ gilt: Die Abbildung $F : V \to V$ definiert durch $F(v) := \mu v$ für alle $v \in V$ ist linear, wie man leicht nachrechnet.
Ferner ist für alle $v_0 \in V$ die Abbildung $G : K \to V$ definiert durch $G(a) := av_0$ für $a \in K$ linear, wie man ebenfalls leicht nachrechnet.
4) Die Abbildung $F : \mathbb{R}^3 \to \mathbb{R}^2$ definiert durch

$$F\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) := \begin{pmatrix} 2x + y \\ 3x + z \end{pmatrix}$$

ist linear (Nachrechnen zur Übung).
5) Es bezeichne wieder $V_{\mathbb{R}}$ den Vektorraum aller Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ und es sei $x_0 \in \mathbb{R}$. Die Abbildung $F : V_{\mathbb{R}} \to \mathbb{R}$ sei definiert durch $F(f) := f(x_0)$ für alle $f \in V_{\mathbb{R}}$ (Auswertung an der Stelle x_0). Diese Abbildung ist linear, wie man leicht sieht.

Wir wollen nun einige einfache Eigenschaften linearer Abbildungen zusammenstellen. Dazu zuerst noch folgende allgemeine Definition.

Definition V.1.3. Seien A und B beliebige Mengen und sei $f : A \to B$ eine beliebige Abbildung. Für Teilmengen $C \subseteq A$ heißt

$$f[C] := \{f(x) : x \in C\}$$

das *Bild* von C unter f.
Für Teilmengen $D \subseteq B$ heißt

$$f^{-1}[D] := \{x \in A : f(x) \in D\}$$

das *Urbild* von D unter f.

Es ist insbesondere $f[A] = \mathrm{Im}(f)$ das Bild von f.

Nun wollen wir zeigen, dass Bilder und Urbilder von Unterräumen unter linearen Abbildungen wieder Unterräume sind.

Lemma V.1.4. *Seien V und W Vektorräume über dem Körper K und sei $F : V \to W$ eine lineare Abbildung. Sei $U \subseteq V$ ein Unterraum von V und sei $U' \subseteq W$ ein Unterraum von W. Dann ist auch $F[U] \subseteq W$ ein Unterraum von W und $F^{-1}[U'] \subseteq V$ ein Unterraum von V.*

Beweis. Es ist $0 = F(0) \in F[U]$, also $F[U] \neq \emptyset$. Sind ferner $w_1, w_2 \in F[U]$, so existieren definitionsgemäß $u_1, u_2 \in U$ mit $w_1 = F(u_1)$, $w_2 = F(u_2)$. Wegen der Linearität von F folgt $w_1 + w_2 = F(u_1) + F(u_2) = F(u_1 + u_2)$, sowie $\lambda w_1 = \lambda F(u_1) = F(\lambda u_1)$ für $\lambda \in K$. Da U ein Unterraum von V ist, gilt auch $u_1 + u_2 \in U$ und $\lambda u_1 \in U$. Daher folgt $w_1 + w_2 \in F[U]$ und $\lambda w_1 \in F[U]$. Also ist $F[U]$ ein Unterraum.
Den Beweis für das Urbild können Sie zur Übung selbst durchführen. □

Als Nächstes definieren wir noch den Kern einer linearen Abbildung.

Definition V.1.5. Seien K ein Körper und V und W Vektorräume über K. Sei $F : V \to W$ eine lineare Abbildung. Dann heißt die Menge

$$\ker(F) := \{v \in V : F(v) = 0\}$$

der *Kern* von F.

Es ist also gerade $\ker(F) = F^{-1}[\{0\}]$ und daher ist $\ker(F)$ nach dem vorigen Lemma ein Unterraum von V. Mit Hilfe des Kerns lässt sich die Injektivität einer linearen Abbildung leicht wie folgt charakterisieren.

Lemma V.1.6. *Seien K ein Körper und V und W Vektorräume über K. Sei $F : V \to W$ eine lineare Abbildung. Dann gilt: F ist injektiv genau dann, wenn $\ker(F) = \{0\}$.*

Beweis. 1) Sei F injektiv und sei $v \in \ker(F)$. Dann ist $F(v) = 0 = F(0)$ und wegen der Injektivität folgt $v = 0$.
2) Sei $\ker(F) = \{0\}$ und seien $v, w \in V$ mit $F(v) = F(w)$. Wegen der Linearität von F folgt $F(v - w) = F(v) - F(w) = 0$, also ist $v - w \in \ker(F) = \{0\}$. Es folgt $v - w = 0$, also $v = w$. Das zeigt die Injektivität von F. □

Sei wieder K ein Körper und seien V und W Vektorräume über K. Für zwei lineare Abbildungen $F, G : V \to W$ und $\lambda \in K$ werden $F + G : V \to W$ und $\lambda F : V \to W$ definiert durch $(F + G)(v) := F(v) + G(v)$ und $(\lambda F)(v) := \lambda F(v)$ für alle $v \in V$.

Es ist nicht schwer nachzuweisen, dass auch $F + G$ und λF wieder lineare Abbildungen sind und dass $\mathrm{Hom}(V, W)$ auf diese Weise zu einem Vektorraum über K wird (Übung).

Das nächste Lemma zeigt, dass auch die Verkettung linearer Abbildungen wieder linear ist.

Lemma V.1.7. *Sei K ein Körper und seien U, V und W Vektorräume über K. Seien $G : V \to U$ und $F : U \to W$ lineare Abbildungen. Dann ist auch $F \circ G : V \to W$ linear.*

Beweis. Seien $v, w \in V$ und sei $\lambda \in K$. Wegen der Linearität von G gilt $G(v + w) = G(v) + G(w)$ und $G(\lambda v) = \lambda G(v)$. Wegen der Linearität von F folgt daraus $(F \circ G)(v + w) = F(G(v + w)) = F(G(v) + G(w)) = F(G(v)) + F(G(w)) = (F \circ G)(v) + (F \circ G)(w)$ und $(F \circ G)(\lambda v) = F(G(\lambda v)) = F(\lambda G(v)) = \lambda F(G(v)) = \lambda(F \circ G)(v)$. □

Als Nächstes betrachten wir den folgenden Satz, dessen erster Teil besagt, dass eine lineare Abbildung bereits eindeutig bestimmt ist, sobald man weiß, welche Werte sie auf einem Erzeugendensystem annimmt. Der zweite Teil besagt, dass man jede Abbildung auf einer Basis eines Vektorraumes zu einer linearen Abbildung auf dem ganzen Raum fortsetzen kann.

Satz V.1.8. *Sei K ein Körper und seien V und W Vektorräume über K. Dann gilt:*
1) Ist E ein Erzeugendensystem von V und sind $F, G : V \to W$ lineare Abbildungen mit $F(v) = G(v)$ für alle $v \in E$, so gilt $F = G$.
2) Ist B eine Basis von V und $f : B \to W$ eine Abbildung, so existiert genau eine lineare Abbildung $F : V \to W$ mit $F(b) = f(b)$ für alle $b \in B$.

Beweis. 1) Seien E, F und G wie oben und sei $v \in V$ beliebig. Wegen $\operatorname{span}(E) = V$ existieren $v_1, \ldots, v_n \in E$ und $\lambda_1, \ldots, \lambda_n \in K$ mit $v = \sum_{i=1}^{n} \lambda_i v_i$. Da F und G linear sind, gilt

$$F(v) = F(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 F(v_1) + \cdots + \lambda_n F(v_n),$$
$$G(v) = G(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 G(v_1) + \cdots + \lambda_n G(v_n).$$

Wegen $v_i \in E$ gilt nach Voraussetzung aber $F(v_i) = G(v_i)$ für $i = 1, \ldots, n$, daher folgt $F(v) = G(v)$.
2) Sei B eine Basis von V und sei $f : B \to W$ irgendeine Abbildung. Ich führe den Beweis der Einfachheit halber hier nur für den Fall $n := \dim(V) < \infty$, die Aussage gilt aber allgemein.
Im endlichdimensionalen Fall können wir B in der Form $B = \{b_1, \ldots, b_n\}$ schreiben. Jedes $v \in V$ lässt sich dann als Linearkombination $v = \sum_{i=1}^{n} \lambda_i b_i$ mit $\lambda_1, \ldots, \lambda_n \in K$ schreiben und diese Darstellung ist auch eindeutig bestimmt, denn sind auch $\mu_1, \ldots, \mu_n \in K$ mit $\sum_{i=1}^{n} \mu_i b_i = v = \sum_{i=1}^{n} \lambda_i b_i$, so folgt $\sum_{i=1}^{n} (\lambda_i - \mu_i) b_i = \sum_{i=1}^{n} \lambda_i b_i - \sum_{i=1}^{n} \mu_i b_i = 0$ und die lineare Unabhängigkeit von B impliziert dann $\lambda_i = \mu_i$ für $i = 1, \ldots, n$.
Daher ist es wohldefinert, $F(v) := \sum_{i=1}^{n} \lambda_i f(b_i) \in W$ zu setzen.

Um die Linearität von F nachzuweisen, nehmen wir uns zwei Vektoren $v = \sum_{i=1}^{n} \lambda_i b_i$ und $v' = \sum_{i=1}^{n} \lambda_i' b_i$ her. Dann gilt $v + v' = \sum_{i=1}^{n} (\lambda_i + \lambda_i') b_i$ und daher

$$F(v + v') = \sum_{i=1}^{n} (\lambda_i + \lambda_i') f(b_i) = \sum_{i=1}^{n} \lambda_i f(b_i) + \sum_{i=1}^{n} \lambda_i' f(b_i) = F(v) + F(v').$$

Ähnlich zeigt man auch $F(\lambda v) = \lambda F(v)$ für alle $\lambda \in K$.
Ferner hat natürlich b_j die Darstellung $b_j = \sum_{i=1}^{n} \alpha_i b_i$, wobei $\alpha_i = 0$ für $i \neq j$ und $\alpha_j = 1$ ist, also gilt $F(b_j) = \sum_{i=1}^{n} \alpha_i f(b_i) = f(b_j)$. Das zeigt $F(b) = f(b)$ für alle $b \in B$.
Nun noch zur Eindeutigkeit: Ist G eine weitere lineare Abbildung von V nach W mit $G(b) = f(b) = F(b)$ für alle $b \in B$, so folgt aus 1) sofort $F = G$. □

Nun führen wir noch das wichtige Konzept der Isomorphie von Vektorräumen ein.

Definition V.1.9. Seien K ein Körper und V und W Vektorräume über K. Eine bijektive lineare Abbildung $F : V \to W$ nennt man einen *Isomorphismus* von V nach W.

V und W heißen *isomorph* (in Zeichen: $V \cong W$), falls es einen Isomorphismus von V nach W gibt.

Die Idee dabei ist, dass zwei isomorphe Vektorräume sich nicht wesentlich voneinander unterscheiden (es unterscheiden sich sozusagen nur die Namen ihrer Elemente und Verknüpfungen).

Als Erstes wollen wir nun folgende Beobachtung festhalten.

Lemma V.1.10. *Seien K ein Körper und V und W Vektorräume über K. Sei $F : V \to W$ ein Isomorphismus. Dann ist auch $F^{-1} : W \to V$ wieder ein Isomorphismus.*

Beweis. Klar ist F^{-1} bijektiv, wir müssen also nur die Linearität von F^{-1} nachweisen. Seien dazu $w_1, w_2 \in W$ beliebig. Dann sind $F^{-1}(w_1), F^{-1}(w_2) \in V$ und da F linear ist, gilt

$$F(F^{-1}(w_1) + F^{-1}(w_2)) = F(F^{-1}(w_1)) + F(F^{-1}(w_2)) = w_1 + w_2.$$

Wendet man hier nun auf beiden Seiten F^{-1} an, so folgt

$$F^{-1}(w_1) + F^{-1}(w_2) = F^{-1}(F(F^{-1}(w_1) + F^{-1}(w_2))) = F^{-1}(w_1 + w_2).$$

Analog zeigt man auch $F^{-1}(\lambda w) = \lambda F^{-1}(w)$ für $\lambda \in K$ und $w \in W$. □

Für Vektorräume U, V und W über demselben Körper K gelten folgende Regeln bzgl. der Isomorphie:

(i) $V \cong V$

(ii) $V \cong W \Rightarrow W \cong V$

(iii) $V \cong W$ und $W \cong U \Rightarrow V \cong U$

(i) ist klar, denn die identische Abbildung $\mathrm{id}_V : V \to V$ ist ein Isomorphismus. (ii) folgt direkt aus Lemma V.1.10 und für (iii) beachte man, dass die Verkettung zweier Isomorphismen wieder ein Isomorphismus ist (Übung).

Nun zeigen wir, dass ein Isomorphismus stets Basen wieder in Basen überführt.

Lemma V.1.11. *Sei K ein Körper und seien V und W Vektorräume über K. Sei $F : V \to W$ ein Isomorphismus und sei $B \subseteq V$ eine Basis von V. Dann ist $F[B]$ eine Basis von W.*

Insbesondere gilt $\dim(V) = \dim(W)$*, falls* $V \cong W$*.*

Beweis. 1) Sei $w \in W$ beliebig. Dann ist $v := F^{-1}(w) \in V$ und da B eine Basis, also insbesondere ein Erzeugendensystem, von V ist, existieren $b_1, \ldots, b_n \in B$ und $\lambda_1, \ldots, \lambda_n \in K$ mit $v = \sum_{i=1}^{n} \lambda_i b_i$. Wegen der Linearität von F folgt

$$w = F(v) = F(\lambda_1 b_1 + \cdots + \lambda_n b_n) = \lambda_1 F(b_1) + \cdots + \lambda_n F(b_n).$$

Wegen $F(b_1), \ldots, F(b_n) \in F[B]$ folgt daraus $w \in \mathrm{span}(F[B])$. Also ist $F[B]$ ein Erzeugendensystem von W.
2) Seien nun $w_1, \ldots, w_n \in F[B]$ paarweise verschieden und seien $\lambda_1, \ldots, \lambda_n \in K$ mit $\sum_{i=1}^{n} \lambda_i w_i = 0$. Wegen $w_1, \ldots, w_n \in F[B]$ gilt $b_i := F^{-1}(w_i) \in B$ für alle $i = 1, \ldots, n$. Nach Lemma V.1.10 ist auch F^{-1} linear, daher folgt

$$0 = F^{-1}(\lambda_1 w_1 + \cdots + \lambda_n w_n) = \lambda_1 F^{-1}(w_1) + \cdots + \lambda_n F^{-1}(w_n) = \lambda_1 b_1 + \cdots + \lambda_n b_n.$$

Wegen der Injektivität von F^{-1} sind auch die Elemente $b_1, \ldots, b_n$ paarweise verschieden. Die lineare Unabhängigkeit von B impliziert daher $\lambda_1 = \cdots = \lambda_n = 0$.
Also ist $F[B]$ auch linear unabhängig und damit eine Basis von W.
Der Zusatz ergibt sich so: Ist $\dim(V) = n \in \mathbb{N}$, so besteht B aus n Elementen. Wegen der Bijektivität von F ist dann auch die Anzahl der Elemente von $F[B]$ gleich n, also $\dim(W) = n$.
Ist dagegen V unendlichdimensional, so muss auch W unendlichdimensional sein, denn anderenfalls gäbe es eine endliche Basis C für W und nach derselben Überlegung wie oben wäre dann $F^{-1}[C]$ eine endliche Basis für V, die es aber nicht geben kann. □

Als Nächstes wollen wir begründen, dass tatsächlich jeder n-dimensionale Vektorraum über K isomorph zum K^n ist.[1] Dazu führen wir zunächst den Begriff einer geordneten Basis ein.

[1] Zur Erinnerung: Der Vektorraum K^n für einen beliebigen Körper K ist in völliger Analogie zum Vektorraum $\mathbb{R}^n$ definiert.

Definition V.1.12. Sei K ein Körper und V ein Vektorraum über K mit $\dim(V) = n \in \mathbb{N}$. Ein n-Tupel $\mathcal{A} = (a_1, \ldots, a_n)$ von Vektoren in V heißt eine *geordnete Basis* von V, falls $\{a_1, \ldots, a_n\}$ eine Basis von V ist.

Natürlich $\mathcal{A} = (a_1, \ldots, a_n)$ genau dann eine geordnete Basis von V, wenn $(a_1, \ldots, a_n)$ linear unabhängig ist, denn aus der linearen Unabhängigkeit folgt wegen des Basisergänzungssatzes die Existenz einer Basis B mit $\{a_1, \ldots, a_n\} \subseteq B$, aber B muss n Elemente haben, also $B = \{a_1, \ldots, a_n\}$. Ebenso folgt aus dem Basisauswahlsatz, dass $\mathcal{A}$ genau dann eine geordnete Basis von V ist, wenn $\operatorname{span}\{a_1, \ldots, a_n\} = V$ gilt.

Ist nun $\mathcal{A} = (a_1, \ldots, a_n)$ eine geordnete Basis von V, so definieren wir eine Abbildung $\Phi_{\mathcal{A}} : K^n \to V$ durch

$$\Phi_{\mathcal{A}}\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) := \sum_{i=1}^{n} x_i a_i.$$

Es ist leicht nachzurechnen, dass $\Phi_{\mathcal{A}}$ linear ist. Ferner ist die Abbildung wegen $\operatorname{span}\{a_1, \ldots, a_n\} = V$ auch surjektiv. Ist zudem $x \in K^n$ mit $\Phi_{\mathcal{A}}(x) = 0$, so impliziert die lineare Unabhängigkeit von $(a_1, \ldots, a_n)$, dass $x = 0$ sein muss. Also ist $\ker(\Phi_{\mathcal{A}}) = \{0\}$ und daher ist $\Phi_{\mathcal{A}}$ nach Lemma V.1.6 auch injektiv. Also ist $\Phi_{\mathcal{A}}$ ein Isomorphismus.

Es gilt also $K^n \cong V$, falls V die Dimension n hat. Dieses Resultat lässt sich noch wie folgt ausbauen.

Satz V.1.13. *Sei K ein Körper und seien V und W Vektorräume über K mit $n := \dim(V) = \dim(W) < \infty$. Dann gilt $V \cong W$.*

Beweis. Nach unserer obigen Überlegung gilt $K^n \cong V$ und $K^n \cong W$. Aus den oben erwähnten Regeln für Isomorphie folgt dann aber auch $V \cong W$. □

Die Umkehrabbildung von $\Phi_{\mathcal{A}}$ ist nach Lemma V.1.10 ebenfalls ein Isomorphismus. Sie wird auch mit $\mathcal{K}_{\mathcal{A}}$ bezeichnet. Für $v \in V$ heißt $\mathcal{K}_{\mathcal{A}}(v) \in K^n$ der Koordinatenvektor von v bzgl. $\mathcal{A}$.

Als Nächstes wollen wir noch die wichtige Dimensionsformel für lineare Abbildungen herleiten. Dazu zeigen wir zunächst folgenden Satz.

Satz V.1.14. *Sei K ein Körper und seien V und W Vektorräume über K. Sei $F : V \to W$ eine lineare Abbildung. Sei A eine Basis für $\ker(F)$ und B eine Basis für $\operatorname{Im}(F)$. Ferner sei $C \subseteq F^{-1}[B]$ mit folgender Eiegnschaft: Für alle $b \in B$ existiert genau ein $c \in C$ mit $F(c) = b$.*

Dann gilt $A \cap C = \emptyset$ und $A \cup C$ ist eine Basis von V.

Beweis. 1) Seien $a \in A$ und $c \in C$ beliebig. Wäre $a = c$, so wäre $F(a) = F(c) \in B$, da $c \in F^{-1}[B]$. Da aber $a \in \ker(F)$ ist, ist $F(a) = 0$. Also wäre $0 \in B$, was der linearen Unabhängigkeit von B widerspricht. Also ist $a \neq c$. Damit ist $A \cap C = \emptyset$ gezeigt.

2) Seien nun $a_1, \ldots, a_n \in A$ paarweise verschieden und $c_1, \ldots, c_m \in C$ paarweise verschieden und seien $\lambda_1, \ldots, \lambda_n, \mu_1 \ldots, \mu_m \in K$ mit $\sum_{i=1}^n \lambda_i a_i + \sum_{i=1}^m \mu_i c_i = 0$. Wegen $A \subseteq \ker(F)$ ist $F(a_i) = 0$ für alle $i = 1, \ldots, n$ und wegen der Linearität von F folgt

$$0 = F(0) = F\left(\sum_{i=1}^n \lambda_i a_i + \sum_{i=1}^m \mu_i c_i\right) = \sum_{i=1}^n \lambda_i F(a_i) + \sum_{i=1}^m \mu_i F(c_i) = \sum_{i=1}^m \mu_i F(c_i).$$

Nach Voraussetzung ist $b_i := F(c_i) \in B$ für alle $i = 1, \ldots, n$ und die Vektoren $b_1, \ldots, b_n$ sind paarweise verschieden. Wegen $\sum_{i=1}^m \mu_i b_i = 0$ und der linearen Unabhängigkeit von B folgt $\mu_1 = \cdots = \mu_m = 0$.
Dann folgt aber $\sum_{i=1}^n \lambda_i a_i = 0$ und die lineare Unabhängigkeit von A impliziert $\lambda_1 = \cdots = \lambda_n = 0$.
Damit ist die lineare Unabhängigkeit von $A \cup C$ bewiesen.
3) Es sei nun $v \in V$ beliebig. Dann ist $F(v) \in \mathrm{Im}(F)$ und folglich existieren $b_1, \ldots, b_n \in B$ und $\alpha_1, \ldots, \alpha_n \in K$ mit $F(v) = \sum_{i=1}^n \alpha_i b_i$. Nach Voraussetzung existieren $c_1, \ldots, c_n \in C$ mit $F(c_i) = b_i$ für $i = 1, \ldots, n$.
Sei $\tilde{v} := \sum_{i=1}^n \alpha_i c_i \in V$. Dann ist $F(\tilde{v}) = \sum_{i=1}^n \alpha_i F(c_i) = F(v)$, also $F(v - \tilde{v}) = 0$, d. h. $v - \tilde{v} \in \ker(F) = \mathrm{span}(A)$.
Nach Definition ist auch $\tilde{v} \in \mathrm{span}(C)$, also folgt $v = \tilde{v} + v - \tilde{v} \in \mathrm{span}(A \cup C)$.
Also ist $A \cup C$ auch ein Erzeugendensystem von V. □

Nun kommen wir zur angekündigten *Dimensionsformel für lineare Abbildungen*.

Satz V.1.15. *Sei K ein Körper und sei V ein Vektorraum über K mit $\dim(V) < \infty$. Ferner sei W ein beliebiger Vektorraum über K und es sei $F : V \to W$ eine lineare Abbildung. Dann sind auch $\ker(F)$ und $\mathrm{Im}(F)$ endlich-dimensional und es gilt*

$$\dim(V) = \dim(\ker(F)) + \dim(\mathrm{Im}(F)).$$

Beweis. Da V endlich-dimensional ist, ist natürlich auch der Unterraum $\ker(F)$ endlich-dimensional. Ist ferner E ein endliches Erzeugendensystem von V, so ist auch $F[E] \subseteq \mathrm{Im}(F)$ endlich und es gilt $\mathrm{Im}(F) = \mathrm{span}(F[E])$ (Übung), also ist auch $\mathrm{Im}(F)$ endlich-dimensional.
Nun wählen wir eine Basis A von $\ker(F)$ und eine Basis B von $\mathrm{Im}(F)$. Die Anzahl der Elemente von A sei k, die Anzahl der Elemente von B sei m. Wir schreiben B in der Form $B = \{b_1, \ldots, b_m\}$ und wählen zu jedem $i \in \{1, \ldots, m\}$ ein $c_i \in V$ mit $F(c_i) = b_i$. Setze $C := \{c_1, \ldots, c_m\}$.
Nach Satz V.1.14 gilt $A \cap C = \emptyset$ und $A \cup C$ ist eine Basis von V. Daher ist die Anzahl der Elemente von $A \cup C$ einerseits gleich $k + m$ und andererseits gleich $\dim(V)$, also ist

$$\dim(V) = k + m = \dim(\ker(F)) + \dim(\mathrm{Im}(F)).$$

□

Die Dimensionsformel liefert folgendes wichtiges Korollar.

Korollar V.1.16. *Sei K ein Körper und seien V und W Vektorräume über K mit $n := \dim(V) = \dim(W) < \infty$. Sei $F : V \to W$ eine lineare Abbildung. Dann gilt: F ist injektiv $\Leftrightarrow$ F ist surjektiv.*

Beweis. Laut Dimensionsformel gilt $\dim(\ker(F)) + \dim(\mathrm{Im}(F)) = n$. Injektivität von F ist äquivalent zu $\ker(F) = \{0\}$ (Lemma V.1.6), also zu $\dim(\ker(F)) = 0$ und damit zu $\dim(\mathrm{Im}(F)) = n = \dim(W)$. Wegen Korollar IV.3.17 ist das aber äquivalent zu $\mathrm{Im}(F) = W$, also zur Surjektivität von F. □

V.2 Matrizen

In diesem Abschnitt beginnen wir mit der Matrizenrechnung, die für die lineare Algebra von größter Bedeutung ist, da sich mit ihrer Hilfe sämtliche linearen Abbildungen zwischen endlich-dimensionalen Vektorräumen beschreiben lassen. Zudem ist sie wichtig für die Behandlung linearer Gleichungssysteme (siehe den folgenden Abschnitt). Nun zunächst zur grundlegenden Definition.

Definition V.2.1. Sei K ein Körper und seien $m, n \in \mathbb{N}$. Eine *$m \times n$-Matrix* A mit Einträgen aus K ist nichts anderes als rechteckiges Schema (eine Tabelle[2]) mit Einträgen aus K, die aus m Zeilen und n Spalten besteht:

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

Man schreibt dafür kurz $A = (a_{ij})_{i,j=1}^{m,n}$ (der erste Index gibt die Zeile i an, in der das Element in der Matrix A auftaucht, der zweite Index j gibt die entsprechende Spalte an).

Die Menge aller $m \times n$-Matrizen mit Einträgen aus K bezeichnen wir mit $M(m \times n, K)$.

Im Falle $m = n$ spricht man von *quadratischen Matrizen*. Auch die Fälle $m = 1$ (einzeilige Matrizen) oder $n = 1$ (einspaltige Matrizen) sind zugelassen. Insbesondere ist $M(m \times 1, K) = K^m$. Ganz konkrete Beispiele für Matrizen sind etwa

$$\begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix} \quad \begin{pmatrix} 1 & 4 & -2 \\ 1 & 0 & 3 \\ -1 & 7 & 9 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 2 & -2 \\ 0 & 4 & 4 & -2 \end{pmatrix}$$

[2]Ganz formal handelt es sich um eine Abbildung von $\{1, \dots, m\} \times \{1, \dots, n\}$ nach K.

oder

$$\begin{pmatrix} 1/2 \\ 1/3 \\ 1/4 \end{pmatrix} \quad \begin{pmatrix} 2 & 3 & 4 \end{pmatrix} \quad \begin{pmatrix} 1 & 3 \\ -1 & 0 \\ 3 & 0 \end{pmatrix}.$$

Auf der Menge $M(m\times n, K)$ können wir in naheliegender Weise eine Addition und eine Multiplikation mit Skalaren definieren: Für $A = (a_{ij})_{i,j=1}^{m,n}$, $B = (b_{ij})_{i,j=1}^{m,n}$ und $\lambda \in K$ setzen wir $A+B := (a_{ij}+b_{ij})_{i,j=1}^{m,n}$ und $\lambda A := (\lambda a_{ij})_{i,j=1}^{m,n}$ oder schematisch

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

und

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

Es ist leicht nachzurechnen, dass $M(m \times n, K)$ auf diese Weise zu einem Vektorraum über K wird (Übung). Dessen Dimension ist mn, wie der folgende Satz zeigt.

Satz V.2.2. *Seien K ein Körper und $m, n \in \mathbb{N}$. Es gilt* $\dim(M(m\times n, K)) = mn$.

Beweis. Für alle Paare $(i, j) \in \{1, \dots, m\}\times\{1, \dots, n\}$ bezeichne E_{ij} diejenige $m \times n$-Matrix, deren Eintrag in der i-ten Zeile und j-ten Spalte gleich 1 ist, während aller anderen Einträge gleich 0 sind. Dann gilt offenbar für alle Matrizen $A = (a_{ij})_{i,j=1}^{m,n}$

$$A = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}E_{ij}.$$

Also ist $B := \{E_{ij} : (i, j) \in \{1, \dots, m\} \times \{1, \dots, n\}\}$ ein Erzeugendensystem von $M(m \times n, K)$. Dasselbe Argument rückwärts gelesen zeigt auch, dass aus

$$\sum_{i=1}^{m}\sum_{j=1}^{n} \lambda_{ij}E_{ij} = 0$$

$\lambda_{ij} = 0$ für alle i und j folgt. Also ist B auch linear unabhängig und somit eine Basis für $M(m \times n, K)$. Da B genau mn Elemente besitzt, folgt die Behauptung. □

Als Nächstes definieren wir noch die Multiplikation von Matrizen.

Definition V.2.3. Seien K ein Körper und seien $m, n, k \in \mathbb{N}$. Sei $A = (a_{ij})_{i,j=1}^{m,k}$ eine $m \times k$-Matrix mit Einträgen aus K und sei $B = (b_{ij})_{i,j=1}^{k,n}$ eine $k \times n$-Matrix mit Einträgen aus K. Dann ist das Produkt AB eine $m \times n$-Matrix, deren Eintrag $(AB)_{ij}$ an der Stelle $(i, j) \in \{1, \ldots, m\} \times \{1, \ldots, n\}$ gegeben ist durch

$$(AB)_{ij} := \sum_{l=1}^{k} a_{il} b_{lj}.$$

Der Eintrag $(AB)_{ij}$ entsteht also dadurch, dass man jedes Element aus der i-ten Zeile von A mit dem entsprechenden Element aus der j-ten Spalte von B multipliziert und diese Produkte anschließend aufsummiert. Dazu ist es erforderlich, dass die Anzahl der Spalten von A mit der Anzahl der Zeilen von B übereinstimmt. Anderenfalls ist das Produkt AB nicht definiert. Hier ein paar konkrete Beispiele:

1) Es ist

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 9 & 1 \end{pmatrix}.$$

Dagegen ist

$$\begin{pmatrix} 3 & -1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 4 \\ 6 & 4 \end{pmatrix}.$$

Die Matrixmultiplikation ist also im Allgemeinen *nicht kommutativ*.
2) Es gilt

$$\begin{pmatrix} 1 & -2 & 1/2 \\ 2 & -3 & 2 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1/2 & 0 & 1 \\ 2 & 3 & -1 \\ 5 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -11/2 & 7/2 \\ 5 & -7 & 7 \\ -9/2 & -1 & 0 \end{pmatrix}.$$

3) Es ist

$$\begin{pmatrix} 1 & 2 & 3 \\ -1 & 4 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} 11 & 16 \\ 11 & 19 \end{pmatrix}.$$

4) Es gilt

$$\begin{pmatrix} 1 & 2 & 3 \\ -1 & 4 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 1 \\ 1 & 4 & -1 \end{pmatrix} = \begin{pmatrix} 8 & 20 & -1 \\ 11 & 26 & 0 \end{pmatrix}.$$

Hier noch ein wichtiger Spezialfall der Matrizenmultiplikation: Elemente des K^n sind, wie gesagt, nichts anderes als Matrizen mit nur einer Spalte und n Zeilen. Ist also $A = (a_{ij})_{i,j=1}^{m,n}$ eine $m \times n$-Matrix mit Einträgen aus K und $x \in K^n$ mit den Koordinaten $x_1, \ldots, x_n$, so ist das Produkt Ax definiert. Ax

eine Matrix mit m Zeilen und einer Spalte, also $Ax \in K^m$. Für $i \in \{1, \ldots, m\}$ ist die i-te Koordinate von Ax gegeben durch $\sum_{j=1}^{n} a_{ij}x_j$. Zum Beispiel ist

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 7 \end{pmatrix}$$

und

$$\begin{pmatrix} 1 & 4 & 3 \\ 3 & -2 & 0 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix}.$$

Nun stellen wir ein paar allgemeine Rechenregeln für die Multiplikation von Matrizen zusammen.

Lemma V.2.4. *Sei K ein Körper und seien $m, n, k, l \in \mathbb{N}$. Ferner seien $A, A' \in M(m \times k, K)$, $B, B' \in M(k \times n, K)$ und $C \in M(n \times l, K)$. Dann gilt:*
1) $(\lambda A)B = \lambda(AB) = A(\lambda B)$ für alle $\lambda \in K$.
2) $A(B + B') = AB + AB'$
3) $(A + A')B = AB + A'B$
4) $A(BC) = (AB)C$

Beweis. 1), 2) und 3) können Sie zur Übung selbst beweisen.
4) Man beachte zunächst, dass beide Produkte $A(BC)$ und $(AB)C$ wohldefiniert sind und jeweils eine $m \times l$-Matrix ergeben. Die Einträge von A, B und C bezeichnen wir wie üblich mit a_{is}, b_{st}, c_{tj}.
Für $i \in \{1, \ldots, m\}$ und $j \in \{1, \ldots, l\}$ ist der Eintrag von $A(BC)$ an der Stelle (i, j) gegeben durch

$$(A(BC))_{ij} = \sum_{s=1}^{k} a_{is}(BC)_{sj} = \sum_{s=1}^{k} a_{is} \sum_{t=1}^{n} b_{st}c_{tj} = \sum_{s=1}^{k} \sum_{t=1}^{n} a_{is}b_{st}c_{tj}.$$

In dieser Doppelsumme kann man nun die Reihenfolge der Summation vertauschen (Kommutativität der Addition in K). So erhält man

$$(A(BC))_{ij} = \sum_{t=1}^{n} \sum_{s=1}^{k} a_{is}b_{st}c_{tj} = \sum_{t=1}^{n} c_{tj} \sum_{s=1}^{k} a_{is}b_{st} = \sum_{t=1}^{n} c_{tj}(AB)_{it}.$$

Das ist aber gerade $((AB)C)_{ij}$, der Eintrag von $(AB)C$ an der Stelle (i, j). Also ist $((AB)C)_{ij} = (A(BC))_{ij}$ und der Beweis ist abgeschlossen. □

Nach dem obigen Lemma ist die Matrixmultiplikation also assoziativ und (von beiden Seiten) distributiv. Hingegen hatten wir oben schon gesehen, dass das Kommutativgesetz im Allgemeinen nicht gilt.

Bei einer quadratischen Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

bezeichnet man die von links oben nach rechts unten verlaufende Diagonale mit den Einträgen $a_{11}, a_{22}, \dots, a_{nn}$ als die *Hauptdiagonale* von A. Sind alle Einträge von A außerhalb der Hauptdiagonalen gleich Null ($a_{ij} = 0$ für $i \neq j$), so nennt man A eine *Diagonalmatrix*. Sind alle Einträge unterhalb der Hauptdiagonalen gleich Null ($a_{ij} = 0$ für $j < i$), so heißt A eine *obere Dreiecksmatrix*. Sind alle Einträge oberhalb der Hauptdiagonalen gleich Null ($a_{ij} = 0$ für $j > i$), so heißt A eine *untere Dreiecksmatrix*. Eine spezielle Diagonalmatrix ist die sogenannte Einheitsmatrix.

Definition V.2.5. Für $n \in \mathbb{N}$ bezeichne mit E_n diejenige $n \times n$-Matrix, deren Einträge auf der Hauptdiagonalen alle gleich 1 sind, während die Einträge außerhalb der Hauptdiagonalen alle gleich 0 sind. Es ist also

$$E_n := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

E_n nennt man die $n \times n$-*Einheitsmatrix*.

Den Eintrag von E_n an der Stelle $(i, j) \in \{1, \dots, n\} \times \{1, \dots, n\}$ bezeichnet man auch mit δ_{ij} (das sogenannte *Kronecker-Symbol*[3]). Es gilt also $\delta_{ij} = 0$ für $i \neq j$ und $\delta_{ii} = 1$.

Für alle $m \times n$-Matrizen A rechnet man leicht folgendes nach:

$$E_m A = A \quad \text{und} \quad A E_n = A.$$

Insbesondere fungiert die Einheitsmatrix E_n also als neutrales Element in $M(n \times n, K)$. Das wirft dann auch sofort die Frage nach der Existenz von inversen Matrizen auf.

Definition V.2.6. Sei K ein Körper und sei $n \in \mathbb{N}$. Eine Matrix $A \in M(n \times n, K)$ heißt *invertierbar*, falls es eine Matrix $B \in M(n \times n, K)$ mit $AB = E_n = BA$ gibt.

Bemerkung V.2.7. Ist $A \in M(n \times n, K)$ invertierbar, so existiert *genau ein* $B \in M(n \times n, K)$ mit $AB = E_n = BA$. Diese Matrix nennt man dann die *inverse Matrix* von A und bezeichnet sie mit A^{-1}.

[3]Benannt nach dem deutschen Mathematiker Leopold Kronecker (1823–1891).

Beweis. Genauso wie Bemerkung III.1.4. □

Beispiele:
1) Die Einheitsmatrix E_n ist natürlich invertierbar mit $E_n^{-1} = E_n$. Dagegen ist die Nullmatrix 0 (alle Einträge gleich Null) nicht invertierbar, denn $0B = 0 \neq E_n$ für alle $B \in M(n \times n, K)$.
2) Die Matrix

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

ist invertierbar mit

$$A^{-1} = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix},$$

denn

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

3) Neben der Nullmatrix gibt es noch diverse weitere Matrizen, die nicht invertierbar sind, z. B. ist die Matrix

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

nicht invertierbar. Anderenfalls gäbe es nämlich $a, b, c, d \in \mathbb{R}$ mit

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

und dann müsste einerseits $a + 2c = 1$ und andererseits $a + 2c = 0$ sein.

Wie man allgemein feststellen kann, ob eine gegebene Matrix invertierbar ist (und ggf. auch die inverse Matrix explizit bestimmt), werden wir erst im nächsten Abschnitt sehen. Hier halten wir noch folgendes fest.

Lemma V.2.8. *Sei K ein Körper, sei $n \in \mathbb{N}$ und seien $A, B \in M(n \times n, K)$ invertierbar. Dann ist auch AB invertierbar und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.*

Beweis. Es gilt $(AB)(B^{-1}A^{-1}) = ((AB)B^{-1})A^{-1} = (A(BB^{-1}))A^{-1} = (AE_n)A^{-1} = AA^{-1} = E_n$ und analog zeigt man auch $(B^{-1}A^{-1})(AB) = E_n$. Daraus folgt die Behauptung. □

Wir setzen

$$\mathrm{GL}(n, K) := \{A \in M(n \times n, K) : A \text{ ist invertierbar}\}.$$

Nach dem obigen Lemma gilt für $A, B \in \mathrm{GL}(n, K)$ auch $AB \in \mathrm{GL}(n, K)$. Es ist leicht nachzuweisen, dass $\mathrm{GL}(n, K)$ bezüglich der Matrixmultiplikation eine Gruppe bildet (GL steht für "general linear group").

Eine weitere wichtige Operation ist die Transposition von Matrizen.

Definition V.2.9. Sei $A = (a_{ij})_{i,j=1}^{m,n}$ eine $m \times n$-Matrix mit Einträgen aus einem Körper K. Die *transponierte Matrix* A^T von A ist eine $n \times m$-Matrix, deren Eintrag an der Stelle $(j,i) \in \{1,\dots,n\} \times \{1,\dots,m\}$ gerade a_{ij} ist.

Nach der obigen Vorschrift entsteht die i-te Spalte von A^T also dadurch, dass man die i-te Zeile von A "senkrecht aufstellt." Zum Beispiel ist

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 0 & 4 \\ 1 & 2 & -2 \\ 3 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 1 & 3 \\ 0 & 2 & 5 \\ 4 & -2 & 6 \end{pmatrix}.$$

Für quadratische Matrizen ist die Transposition nichts anderes als eine Spiegelung an der Hauptdiagonalen. Aber auch nicht-quadratische Matrizen lassen sich transponieren, z. B. ist

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 3 & 4 \end{pmatrix}^T = \begin{pmatrix} 1 & 0 \\ 1 & 3 \\ 2 & 4 \end{pmatrix}.$$

Für die Transposition von Matrizen gelten die folgenden Rechenregeln.

Lemma V.2.10. *Sei K ein Körper und seien $m,n,k \in \mathbb{N}$. Für alle $A,B \in M(m \times n, K)$ und alle $C \in M(n \times k, K)$ gilt:*
1) $(\lambda A)^T = \lambda A^T$ für alle $\lambda \in K$.
2) $(A+B)^T = A^T + B^T$
3) $(AC)^T = C^T A^T$.

Den Beweis dieses Lemmas überlasse ich Ihnen zur Übung.
Nun kommen wir zum entscheidenden Zusammenhang zwischen Matrizen und linearen Abbildungen. Ist $A \in M(m \times n, K)$, so ist die Abbildung $F_A : K^n \to K^m$ definiert durch $F_A(x) := Ax$ für alle $x \in K^n$ linear, wie sofort aus den Rechenregeln in Lemma V.2.4 folgt. Tatsächlich ist jede lineare Abbildung von K^n nach K^m von dieser Form, wie der folgende Satz zeigt.

Satz V.2.11. *Sei K ein Körper und seien $m,n \in \mathbb{N}$. Sei ferner $F : K^n \to K^m$ linear. Dann existiert genau eine Matrix $A \in M(m \times n, K)$ mit $F = F_A$. Die j-te Spalte von A ist gegeben durch den Vektor $F(e_j)$ für alle $j \in \{1,\dots,n\}$.*

Beweis. Sei $A = (a_{ij})_{i,j=1}^{m,n}$ diejenige $m \times n$-Matrix, deren j-te Spalte gerade gleich $F(e_j)$ ist für alle $j \in \{1,\dots,n\}$. Für $i \in \{1,\dots,m\}$ bezeichne $F(e_j)_i$ die i-te Koordinate von $F(e_j)$. Mit anderen Worten ist $F(e_j)_i = a_{ij}$. Sei

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$$

beliebig. Dann gilt wegen der Linearität von F aber $F(x) = F(\sum_{j=1}^n x_j e_j) = \sum_{j=1}^n x_j F(e_j)$. Für alle $i \in \{1, \ldots, m\}$ ist die i-te Koordinate von $F(x)$ also gerade

$$\sum_{j=1}^n x_j F(e_j)_i = \sum_{j=1}^n a_{ij} x_j,$$

was genau die i-te Koordinate von Ax ist. Also ist $F(x) = Ax = F_A(x)$.
Zur Eindeutigkeit: Ist $B = (b_{ij})_{i,j=1}^{m,n}$ eine weitere Matrix mit $F_B = F$, so folgt $F(e_j) = Be_j$ für alle $j \in \{1, \ldots, n\}$. Der i-te Eintrag von Be_j ist aber gerade

$$\sum_{k=1}^n b_{ik}\delta_{kj} = b_{ij}$$

für alle $i \in \{1, \ldots, m\}$. Also ist Be_j gerade der j-te Spaltenvektor von B. Dieser stimmt überein mit $F(e_j)$, dem j-ten Spaltenvektor von A (für alle $j \in \{1, \ldots, n\}$), also ist $A = B$. □

Beispiel: Sei $F : \mathbb{R}^3 \to \mathbb{R}^3$ definiert durch

$$F\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) := \begin{pmatrix} 2x - 4y \\ 2z \\ 3x - z \end{pmatrix}.$$

Dann ist F linear, wie man leicht nachrechnet. Es gilt

$$F(e_1) = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}, \quad F(e_2) = \begin{pmatrix} -4 \\ 0 \\ 0 \end{pmatrix}, \quad F(e_3) = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}.$$

Für

$$A := \begin{pmatrix} 2 & -4 & 0 \\ 0 & 0 & 2 \\ 3 & 0 & -1 \end{pmatrix}$$

gilt also nach dem obigen Satz $F_A = F$.

Jetzt zeigen wir noch das folgende Ergebnis, dass die Suche nach inversen Matrizen erleichtert, weil man nur noch eine Seite prüfen muss.

Satz V.2.12. *Seien K ein Körper, $n \in \mathbb{N}$ und $A, B \in M(n \times n, K)$. Dann sind folgende Aussagen äquivalent:*
1) A ist invertierbar mit $A^{-1} = B$.
2) $AB = E_n$
3) $BA = E_n$

Beweis. Aus 1) folgt natürlich 2) und 3).
Gelte nun $AB = E_n$. Ist $x \in K^n$ mit $F_B(x) = Bx = 0$, so folgt $x = E_n x = (AB)x = A(Bx) = 0$. Also ist $\ker(F_B) = \{0\}$, d. h. $F_B : K^n \to K^n$ ist injektiv. Nach Korollar V.1.16 ist $F := F_B$ dann auch surjektiv, also ein Isomorphismus. Wir wissen, dass auch F^{-1} wieder linear ist, also existiert nach Satz V.2.11 eine Matrix $C \in M(n \times n, K)$ mit $F^{-1} = F_C$. Für alle $x \in K^n$ gilt dann $(BC)x = (F_B \circ F_C)(x) = (F \circ F^{-1})(x) = x$.
Insbesondere ist $(BC)e_i = e_i$ für alle $i \in \{1, \ldots, n\}$ und $(BC)e_i$ ist gerade der i-te Spaltenvektor von BC. Also gilt $BC = E_n$.
Es folgt $A = AE_n = A(BC) = (AB)C = E_n C = C$ und somit $BA = BC = E_n$.
Gilt umgekehrt $BA = E_n$, so zeigt dasselbe Argument wie eben (nur mit vertauschten Rollen von A und B), dass auch $AB = E_n$ gilt. □

Als Nächstes wollen wir das Prinzip der Darstellung linearer Abbildungen durch Matrizen auch auf abstrakte endlich-dimensionale Vektorräume verallgemeinern. Sei also wieder K ein beliebiger Körper und seien V und W Vektorräume über K mit $\dim(V) = n \in \mathbb{N}$ und $\dim(W) = m \in \mathbb{N}$. Wir wählen eine geordnete Basis $\mathcal{A} = (a_1, \ldots, a_n)$ von V und eine geordnete Basis $\mathcal{B} = (b_1, \ldots, b_m)$ von W und betrachten die im letzten Abschnitt definierten Isomorphismen $\Phi_\mathcal{A} : K^n \to V$ und $\Phi_\mathcal{B} : K^m \to W$, sowie ihre Umkehrabbildungen $\mathcal{K}_\mathcal{A} := \Phi_\mathcal{A}^{-1}$ und $\mathcal{K}_\mathcal{B} := \Phi_\mathcal{B}^{-1}$.

Ist nun $F : V \to W$ eine lineare Abbildung, so ist auch $\mathcal{K}_\mathcal{B} \circ (F \circ \Phi_\mathcal{A}) : K^n \to K^m$ linear. Also existiert nach Satz V.2.11 genau eine $m \times n$-Matrix A mit $F_A = \mathcal{K}_\mathcal{B} \circ (F \circ \Phi_\mathcal{A})$. Diese Matrix A nennen wir die *darstellende Matrix* von F bezüglich $\mathcal{A}$ unf $\mathcal{B}$ und bezeichnen sie mit $\mathcal{M}_\mathcal{B}^\mathcal{A}(F)$.

Nach Satz V.2.11 ist die j-te Spalte von $\mathcal{M}_\mathcal{B}^\mathcal{A}(F)$ gerade der Vektor $\mathcal{K}_\mathcal{B}(F(\Phi_\mathcal{A}(e_j))) = \mathcal{K}_\mathcal{B}(F(a_j))$ (für alle $j \in \{1, \ldots, n\}$).

Außerdem gilt für alle $v \in V$:

$$\mathcal{K}_\mathcal{B}(F(v)) = (\mathcal{K}_\mathcal{B} \circ (F \circ \Phi_\mathcal{A}))(\mathcal{K}_\mathcal{A}(v)) = F_A(\mathcal{K}_\mathcal{A}(v)) = \mathcal{M}_\mathcal{B}^\mathcal{A}(F)\mathcal{K}_\mathcal{A}(v).$$

Ist $V = W$ und $\mathcal{A} = \mathcal{B}$, so schreibt man auch kurz $\mathcal{M}_\mathcal{B}(F)$ anstelle von $\mathcal{M}_\mathcal{B}^\mathcal{B}(F)$.

Ist $V = K^n$ und $W = K^m$ und verwendet man für $\mathcal{A}$ und $\mathcal{B}$ jeweils die kanonische Basis, so ist $\mathcal{M}_\mathcal{B}^\mathcal{A}(F)$ natürlich nichts anderes als die Matrix mit den Spalten $F(e_1), \ldots, F(e_n)$ aus Satz V.2.11. Manchmal will man allerdings auch hier eine andere Basis als die kanonische verwenden.

Beispiel:
Sei $F : \mathbb{R}^3 \to \mathbb{R}^2$ definiert durch

$$F\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) := \begin{pmatrix} x + y \\ 2x - z \end{pmatrix}.$$

Dann ist F linear, wie man leicht nachrechnet. Es sei $\mathcal{A} := (e_1, e_2, e_3)$ die kanonische Basis des $\mathbb{R}^3$ und es sei $\mathcal{B} := (b_1, b_2)$, wobei

$$b_1 := \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{und} \quad b_2 := \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

($\mathcal{B}$ ist eine geordnete Basis des $\mathbb{R}^2$, wie man leicht sieht).
Wir wollen die darstellende Matrix $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)$ bestimmen. Zunächst ist

$$\begin{aligned} F(e_1) &= \begin{pmatrix} 1 \\ 2 \end{pmatrix} = b_2 = \Phi_{\mathcal{B}}\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right), \\ F(e_2) &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{3}(2b_1 + b_2) = \Phi_{\mathcal{B}}\left(\begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}\right), \\ F(e_3) &= \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \frac{1}{3}(b_1 - b_2) = \Phi_{\mathcal{B}}\left(\begin{pmatrix} 1/3 \\ -1/3 \end{pmatrix}\right). \end{aligned}$$

Die i-te Spalte von $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)$ ist gerade $\mathcal{K}_{\mathcal{B}}(F(e_i)) = \Phi_{\mathcal{B}}^{-1}(F(e_i))$ für alle $i = 1, 2, 3$. Also gilt

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 1 & 1/3 & -1/3 \end{pmatrix}.$$

Als Nächstes wollen wir noch die darstellende Matrix der Komposition zweier linearer Abbildungen bestimmen.

Lemma V.2.13. *Sei K ein Körper und seien U, V, W Vektorräume über K mit den Dimensionen $\dim(U) = k \in \mathbb{N}$, $\dim(V) = n \in \mathbb{N}$ und $\dim(W) = m \in \mathbb{N}$. Seien $\mathcal{A}$ eine geordnete Basis von V, $\mathcal{B}$ eine geordnete Basis von W und $\mathcal{C}$ eine geordnete Basis von U. Ferner seien $F : V \to W$ und $G : U \to V$ lineare Abbildungen. Dann gilt:*

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(F \circ G) = \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)\mathcal{M}_{\mathcal{A}}^{\mathcal{C}}(G)$$

Beweis. Wir schreiben $\mathcal{C} = (c_1, \ldots, c_k)$. Nach unseren obigen Überlegungen ist für $j \in \{1, \ldots, k\}$ die j-te Spalte von $\mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(F \circ G)$ gerade $\mathcal{K}_{\mathcal{B}}(F(G(c_j))) = \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)\mathcal{K}_{\mathcal{A}}(G(c_j))$ und $\mathcal{K}_{\mathcal{A}}(G(c_j)) = \mathcal{M}_{\mathcal{A}}^{\mathcal{C}}(G)\mathcal{K}_{\mathcal{C}}(c_j) = \mathcal{M}_{\mathcal{A}}^{\mathcal{C}}(G)e_j$.
Also gilt $\mathcal{K}_{\mathcal{B}}(F(G(c_j))) = \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)\mathcal{M}_{\mathcal{A}}^{\mathcal{C}}(G)e_j$ und das ist genau die j-te Spalte von $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)\mathcal{M}_{\mathcal{A}}^{\mathcal{C}}(G)$. □

Mit Hilfe darstellender Matrizen können wir auch das folgende wichtige Ergebnis zeigen.

Satz V.2.14. *Seien V und W Vektorräume über demselben Körper K mit $\dim(V) = n \in \mathbb{N}$ und $\dim(W) = m \in \mathbb{N}$. Dann gilt $\mathrm{Hom}(V, W) \cong M(m \times n, K)$. Insbesondere ist $\dim(\mathrm{Hom}(V, W)) = mn$.*

Beweis. Wir wählen eine geordnete Basis $\mathcal{A}$ von V und eine geordnete Basis $\mathcal{B}$ von W. Es ist leicht nachzurechnen, dass $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F+G) = \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F) + \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(G)$ und $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(\lambda F) = \lambda \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)$ für alle $F, G \in \mathrm{Hom}(V,W)$ und alle $\lambda \in K$ gilt (Übung). Mit anderen Worten die Abbildung $T : \mathrm{Hom}(V,W) \to M(m \times n, K)$ definiert durch $T(F) := \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)$ für $F \in \mathrm{Hom}(V,W)$ ist linear.
Ist $F : V \to W$ linear mit $T(F) = \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F) = 0$, so folgt $\mathcal{K}_{\mathcal{B}}(F(v)) = \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)\mathcal{K}_{\mathcal{A}}(v) = 0$ und folglich $F(v) = 0$ für alle $v \in V$. Also ist $\ker(T) = \{0\}$ und somit ist T injektiv.
Sei nun $A = (a_{ij})_{i,j=1}^{m,n}$ eine beliebige $m \times n$-Matrix mit Einträgen aus K. Sei $\mathcal{A} = (a_1, \ldots, a_n)$. Aus Satz V.1.8 folgt: Es existiert genau eine lineare Abbildung $F : V \to W$ mit

$$F(a_j) = \Phi_{\mathcal{B}}\left(\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}\right) \quad \forall j = 1, \ldots, n.$$

Dann ist der j-te Spaltenvektor von $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)$ gerade gleich $\mathcal{K}_{\mathcal{B}}(F(a_j))$ und somit gleich dem j-ten Spaltenvektor von A. Also gilt $T(F) = \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F) = A$. Damit ist auch die Surjektivität von T gezeigt. T ist also ein Isomorphismus.
Zum Zusatz: Wegen $\mathrm{Hom}(V,W) \cong M(m \times n, K)$ gilt nach Lemma V.1.11 $\dim(\mathrm{Hom}(V,W)) = \dim(M(m \times n, K))$.
Aus Satz V.2.2 folgt damit $\dim(\mathrm{Hom}(V,W)) = mn$. □

Das nächste Lemma zeigt den Zusammenhang zwischen Isomorphismen und der Invertierbarkeit von Matrizen.

Lemma V.2.15. *Seien V und W Vektorräume über demselben Körper K mit $\dim(V) = \dim(W) = n \in \mathbb{N}$. Sei $\mathcal{A}$ eine geordnete Basis von V und sei $\mathcal{B}$ eine geordnete Basis von W. Ferner sei $F : V \to W$ linear. Dann gilt: F ist ein Isomorphismus genau dann, wenn $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)$ invertierbar ist. Ggf. gilt dann $(\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F))^{-1} = \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(F^{-1})$.*

Beweis. 1) Sei $C := \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)$ invertierbar. Ist $v \in V$ mit $F(v) = 0$, so folgt $C\mathcal{K}_{\mathcal{A}}(v) = \mathcal{K}_{\mathcal{B}}(F(v)) = 0$ und somit $0 = C^{-1}(C\mathcal{K}_{\mathcal{A}}(v)) = \mathcal{K}_{\mathcal{A}}(v)$. Daraus folgt $v = 0$. Es ist also $\ker(F) = \{0\}$ und daher ist F injektiv.
Nun sei $w \in W$ beliebig. Wir setzen $v := \Phi_{\mathcal{A}}(C^{-1}\mathcal{K}_{\mathcal{B}}(w)) \in V$. Dann gilt $\mathcal{K}_{\mathcal{B}}(F(v)) = C\mathcal{K}_{\mathcal{A}}(v) = C(C^{-1}\mathcal{K}_{\mathcal{B}}(w)) = \mathcal{K}_{\mathcal{B}}(w)$ und deshalb auch $F(v) = w$. Also ist F auch surjektiv.
2) Sei nun F ein Isomorphismus. Aus Lemma V.2.13 folgt $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(F^{-1}) = \mathcal{M}_{\mathcal{B}}(F \circ F^{-1}) = \mathcal{M}_{\mathcal{B}}(\mathrm{id}_W) = E_n$. Nach Satz V.2.12 ist also $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F)$ invertierbar mit $(\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F))^{-1} = \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(F^{-1})$. □

Schließlich wollen wir noch untersuchen, wie sich die darstellende Matrix einer linearen Abbildung verändert, wenn man die verwendeten Basen verändert. Dazu führen wir zuerst folgenden Begriff ein.

Definition V.2.16. Sei K ein Körper und sei V ein Vektorraum über K der Dimension $n \in \mathbb{N}$. Seien $\mathcal{B} = (b_1, \ldots, b_n)$ und $\mathcal{B}' = (b'_1, \ldots, b'_n)$ zwei geordnete Basen von V. Dann heißt $T^{\mathcal{B}}_{\mathcal{B}'} := \mathcal{M}^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_V)$ die *Basiswechselmatrix* oder *Transformationsmatrix* von $\mathcal{B}$ zu $\mathcal{B}'$.

$T^{\mathcal{B}}_{\mathcal{B}'}$ ist also eine $n \times n$-Matrix, deren j-te Spalte der Koordinatenvektor $\mathcal{K}_{\mathcal{B}'}(b_j)$ ist. Es gilt $\mathcal{K}_{\mathcal{B}'}(v) = T^{\mathcal{B}}_{\mathcal{B}'}\mathcal{K}_{\mathcal{B}}(v)$ für alle $v \in V$ (daher der Name Basiswechselmatrix). Zudem folgt aus Lemma V.2.15: $T^{\mathcal{B}}_{\mathcal{B}'}$ ist invertierbar mit $(T^{\mathcal{B}}_{\mathcal{B}'})^{-1} = T^{\mathcal{B}'}_{\mathcal{B}}$.

Es gilt nun die folgende Formel.

Lemma V.2.17. *Seien V und W zwei endlich-dimensionale Vektorräume über demselben Körper K und sei $F : V \to W$ linear. Seien $\mathcal{A}$ und $\mathcal{A}'$ geordnete Basen von V und $\mathcal{B}$ und $\mathcal{B}'$ geordnete Basen von W. Dann gilt:*

$$\mathcal{M}^{\mathcal{A}'}_{\mathcal{B}'}(F) = T^{\mathcal{B}}_{\mathcal{B}'}\mathcal{M}^{\mathcal{A}}_{\mathcal{B}}(F)T^{\mathcal{A}'}_{\mathcal{A}}$$

Beweis. Aus Lemma V.2.13 folgt

$$T^{\mathcal{B}}_{\mathcal{B}'}\mathcal{M}^{\mathcal{A}}_{\mathcal{B}}(F) = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_W)\mathcal{M}^{\mathcal{A}}_{\mathcal{B}}(F) = \mathcal{M}^{\mathcal{A}}_{\mathcal{B}'}(\mathrm{id}_W \circ F) = \mathcal{M}^{\mathcal{A}}_{\mathcal{B}'}(F).$$

Eine weitere Anwendung dieses Lemmas liefert

$$\begin{aligned} T^{\mathcal{B}}_{\mathcal{B}'}\mathcal{M}^{\mathcal{A}}_{\mathcal{B}}(F)T^{\mathcal{A}'}_{\mathcal{A}} &= \mathcal{M}^{\mathcal{A}}_{\mathcal{B}'}(F)T^{\mathcal{A}'}_{\mathcal{A}} = \mathcal{M}^{\mathcal{A}}_{\mathcal{B}'}(F)\mathcal{M}^{\mathcal{A}'}_{\mathcal{A}}(\mathrm{id}_V) \\ &= \mathcal{M}^{\mathcal{A}'}_{\mathcal{B}'}(F \circ \mathrm{id}_V) = \mathcal{M}^{\mathcal{A}'}_{\mathcal{B}'}(F). \end{aligned}$$

□

Als letzten Punkt in diesem Abschnitt führen wir die wichtigen Begriffe des Zeilen- und des Spaltenrangs einer Matrix ein.

Definition V.2.18. Seien K ein Körper, $m, n \in \mathbb{N}$ und $A \in M(m \times n, K)$. Wir bezeichnen die Spaltenvektoren von A mit $a_1, \ldots, a_n \in K^m$ und die Zeilenvektoren von A mit $z_1, \ldots, z_m \in M(1 \times n, K)$ und setzen

$$S(A) := \mathrm{span}\{a_1, \ldots, a_n\}$$

und

$$Z(A) := \mathrm{span}\{z_1, \ldots, z_m\}.$$

$S(A)$ heißt der *Spaltenraum* und $Z(A)$ der *Zeilenraum* von A. Ferner heißt S-Rang$(A) := \dim(S(A))$ der *Spaltenrang* und Z-Rang$(A) := \dim(Z(A))$ der *Zeilenrang* von A.

Wir werden später zeigen, dass Zeilen- und Spaltenrang einer Matrix stets übereinstimmen. Einstweilen müssen wir die Unterscheidung aber noch vornehmen.

Für alle Vektoren $x \in K^n$ (mit Koordinaten $x_1, \ldots, x_n$) gilt $Ax = \sum_{j=1}^n x_j a_j$, wie man leicht nachrechnet. Daher gilt also

$$S(A) = \operatorname{span}\{a_1, \ldots, a_n\} = \{Ax : x \in K^n\} = \operatorname{Im}(F_A).$$

Die Invertierbarkeit einer quadratischen Matrix lässt sich nun wie folgt charakterisieren.

Lemma V.2.19. *Seien K ein Körper, $n \in \mathbb{N}$ und $A \in M(n \times n, K)$. A ist invertierbar genau dann, wenn S-Rang$(A) = n$ gilt.*

Beweis. Wegen $S(A) = \operatorname{Im}(F_A)$ gilt S-Rang$(A) = n$ genau dann, wenn $\operatorname{Im}(F_A) = K^n$ ist. Diese Bedingung, die Surjektivität von F_A, ist aber wegen Korollar V.1.16 äquivalent dazu, dass F_A ein Isomorphismus ist. Das ist wegen Lemma V.2.15 aber äquivalent zur Invertierbarkeit von A. □

V.3 Der Gaußsche Algorithmus

In diesem Abschnitt beschäftigen wir uns mit *linearen Gleichungssystemen*, d. h. mit Systemen der Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\quad \vdots \qquad\qquad & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

bestehend aus m Gleichungen für die n Variablen $x_1, \ldots, x_n \in K$, wobei die *Koeffizienten* $a_{ij} \in K$ und die $b_1, \ldots, b_m \in K$ vorgegeben sind (K kann ein beliebiger Körper sein, der wichtigste Fall ist aber wieder $K = \mathbb{R}$). Solche Systeme treten in verschiedensten Anwendungsproblemen auf, vom Ausbalancieren chemischer Reaktionsgleichungen, über Wirtschaftsmodelle bis hin zur Kryptographie.

Setzt man $A := (a_{ij})_{i,j=1}^{m,n}$ und

$$x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad b := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

so ist das obige lineare Gleichungssystem offenbar äquivalent zu $Ax = b$.

Im Falle $b = 0$ heißt das System *homogen*, anderenfalls *inhomogen*. Die Matrix A heißt die *Koeffizientenmatrix*. Fügt man an A zusätzlich noch den Vektor b als $(n+1)$-te Spalte an (symbolisch $(A|b)$), so erhält man die sogenannte *erweiterte Koeffizientenmatrix*. Weiter setzen wir

$$L(A, b) := \{x \in K^n : Ax = b\}.$$

$L(A,b)$ ist die *Lösungsmenge* des linearen Gleichungssystems $Ax = b$. Sie stimmt offenbar gerade mit dem Urbild $F_A^{-1}(\{b\})$ überein. Für die Lösungsmenge des entsprechenden homogenen Systems schreiben wir kurz $L(A) := L(A,0)$. Mit anderen Worten ist $L(A) = \ker(F_A)$.

Die Mengen $L(A)$ und $L(A,b)$ hängen wie folgt zusammen.

Lemma V.3.1. *Seien K ein Körper und $m,n \in \mathbb{N}$, sowie $A \in M(m \times n, K)$ und $b \in K^m$. Ist $x_0 \in L(A,b)$, so gilt*

$$L(A,b) = \{x_0 + x : x \in L(A)\}.$$

Beweis. Sei $S := \{x_0 + x : x \in L(A)\}$. Für alle $x \in L(A)$ gilt $Ax = 0$ und somit $A(x_0 + x) = Ax_0 + Ax = Ax_0 = b$ (letzteres wegen $x_0 \in L(A,b)$). Also ist $x_0 + x \in L(A,b)$. Das zeigt $S \subseteq L(A,b)$.
Sei nun umgekehrt $y \in L(A,b)$. Setze $x := y - x_0$. Dann gilt $Ax = A(y - x_0) = Ay - Ax_0 = b - b = 0$, also $x \in L(A)$ und somit ist $y = x_0 + x \in S$. Also gilt auch $L(A,b) \subseteq S$. □

Um die allgemeine Lösung eines inhomogenen linearen Gleichungssystem $Ax = b$ zu bestimmen, braucht man also nur eine spezielle Lösung x_0 desselben zu finden und anschließend die allgemeine Lösungsmenge $L(A)$ des zugehörigen homogenen Systems zu bestimmen. Sämtliche Lösungen des inhomogenen Systems sind dann von der Form $x_0 + x$ mit $x \in L(A)$.

Es kann aber auch vorkommen, dass das inhomogene System $Ax = b$ überhaupt keine Lösung besitzt, dass also $L(A,b) = \emptyset$ gilt (siehe die Beispiele unten).

Um ein lineares Gleichungssystem zu lösen, geht man in der Regel so vor, dass man schrittweise Variablen zu eliminieren versucht, bis in einer Gleichung nur noch eine Variable auftaucht. Dabei darf man zu einer Gleichung ein Vielfaches einer anderen Gleichung addieren, man darf eine Gleichung mit einem von Null verschiedenen Faktor multiplizieren und man darf zwei Gleichungen vertauschen. All diese Operationen ändern nichts an der Lösungsmenge des Gleichungssystems. Für eine Matrix A sehen die entsprechenden Umformungen wie folgt aus:

Typ (I) Umformungen: Vertauschen zweier Zeilen von A

Typ (II) Umformungen: Addition des λ-fachen einer Zeile von A zu einer anderen Zeile von A ($\lambda \in K$)

Typ (III) Umformungen: Multiplizieren einer Zeile von A mit einem Faktor $\lambda \in K \setminus \{0\}$

Solche Umformungen nennt man *elementare Zeilenumformungen.*

Sind $A, A' \in M(m \times n, K)$ und $b, b' \in K^m$ derart, dass $(A'|b')$ durch elementare Zeilenumformungen aus $(A|b)$ hervorgeht, so gilt $L(A, b) = L(A', b')$.

Diesen Umstand wollen wir ausnutzen, indem wir $(A|b)$ mittels elementarer Zeilenumformungen auf eine besonders einfache Form bringen, in der sich die Lösungsmenge des Gleichungssystems leicht ablesen lässt.

Sei also $A = (a_{ij})_{i,j=1}^{m,n}$.

Zunächst können wir nach einer eventuellen Zeilenvertauschung (Typ (I) Umformung) $a_{11} \neq 0$ annehmen (falls tatsächlich alle Einträge in der ersten Spalte gleich Null sind, gehen wir im folgenden Schema gleich zur nächsten Spalte über).

Nun ziehen wir für alle $i = 2, \ldots, m$ von der i-ten Zeile jeweils das a_{i1}/a_{11}-fache der ersten Zeile ab (Typ (II) Umformungen). Dadurch werden alle Einträge in der ersten Spalte unterhalb von a_{11} zu Null. Die neue Matrix nennen wir provisorisch $A^{(1)}$ und ihre Einträge $a_{ij}^{(1)}$.

Nun betrachten wir die zweite Spalte von $A^{(1)}$. Ist $a_{i2}^{(1)} \neq 0$ für wenigstens ein $i \in \{2, \ldots, m\}$, so kann man mittels Zeilenvertauschung ohne Einschränkung auch $a_{22}^{(1)} \neq 0$ annehmen. Dann ziehen wir für alle $i = 3, \ldots, m$ das a_{i2}/a_{22}-fache der zweiten Zeile von der i-ten Zeiel ab (Typ (II) Umformungen). Auf diese Weise erhalten wir eine Matrix $A^{(2)}$, in der alle Einträge in der ersten Spalte unterhalb des ersten Eintrags gleich Null sind und auch alle Einträge in der zweiten Spalte unterhalb des zweiten Eintrags gleich Null sind. Anschließend verfährt man genauso mit der dritten Spalte etc. Ist dagegen $a_{i2}^{(1)} = 0$ für alle $i = 2, \ldots, m$, so geht man einfach gleich zur dritten Spalte über.

Letztendlich erhält man eine Matrix $A' = (a'_{ij})_{i,j=1}^{m,n}$ der folgenden Form: Es existieren ein $r \in \{1, \ldots, m\}$ und Indizes $1 \leq j_1 < j_2 < \cdots < j_r \leq n$, so dass folgendes gilt:

1) Für alle $r < i \leq m$ enthält die i-te Zeile von A' nur Nullen als Einträge.
2) Für alle $i = 1, \ldots, r$ ist $a'_{ij_i} \neq 0$ und $a'_{ij} = 0$ für $j < j_i$.

Man sagt A' habe *Zeilenstufenform*. Das oben beschriebene Verfahren zum Erreichen dieser Zeilenstufenform nennt man den *Gaußschen Algorithmus*.[4] Man beachte, dass dieser Algorithmus nur Typ (I) und Typ (II) Umformungen (keine vom Typ (III)) verwendet. Natürlich dürfen aber auch Typ (III) Umformungen verwendet werden, um die Rechnung ggf. zu vereinfachen.

Die obigen Elemente a_{ij_i} heißen die *Pivotelemente* und die zugehörigen Spalten die *Pivotspalten*.

Der Witz ist nun, dass sich bei einer Koeffizientenmatrix in Zeilenstufenform die Lösungsmenge des zugehörigen linearen Gleichungssystems leicht

[4] Nach C. F. Gauß, siehe Fußnote zur Gaußschen Summenformel.

ablesen lässt. Es folgen dazu einige konkrete Beispiele.

Beispiele:
1) Wir betrachten das folgende homogene lineare Gleichungssystem:

$$\begin{aligned} 2x + y + 4z &= 0 \\ 2x + 3y - z &= 0 \\ 4x + 4y - z &= 0 \end{aligned}$$

Die Koeffizientenmatrix ist

$$A = \begin{pmatrix} 2 & 1 & 4 \\ 2 & 3 & -1 \\ 4 & 4 & -1 \end{pmatrix}.$$

Diese wollen wir mit Hilfe des Gaußschen Algorithmus in Zeilenstufenform überführen. Wir ziehen dazu zunächst von der zweiten Zeile die erste Zeile und von der dritten Zeile das 2-fache der ersten Zeile ab und erhalten die Matrix:

$$\begin{pmatrix} 2 & 1 & 4 \\ 0 & 2 & -5 \\ 0 & 2 & -9 \end{pmatrix}$$

Als Nächstes ziehen wir von der dritten Zeile die zweite Zeile ab und erhalten:

$$\begin{pmatrix} 2 & 1 & 4 \\ 0 & 2 & -5 \\ 0 & 0 & -4 \end{pmatrix}$$

Das entsprechende homogene lineare Gleichungssystem lautet:

$$\begin{aligned} 2x + y + 4z &= 0 \\ 2y - 5z &= 0 \\ -4z &= 0 \end{aligned}$$

Aus der letzten Gleichung folgt sofort $z = 0$. Aus der zweiten Gleichung folgt dann auch $y = 0$ und aus der ersten Gleichung schließlich auch $x = 0$. Das Gleichungssystem hat also nur die triviale Lösung $x = y = z = 0$, d. h. es ist $L(A) = \{0\}$.
2) Betrachten wir nun das folgende inhomogene lineare Gleichungssystem:

$$\begin{aligned} x + y + 2z &= 1 \\ 2x + 3y - z &= 4 \\ 3x + y + 3z &= 12 \end{aligned}$$

Die erweiterte Koeffizientenmatrix ist

$$(A|b) = \left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 2 & 3 & -1 & 4 \\ 3 & 1 & 3 & 12 \end{array}\right).$$

Wir wollen nun A mit dem Gaußschen Algorithmus in Zeilenstufenform überführen und wenden dabei jeweils auch dieselben Zeilenumformungen auf die zusätzliche letzte Spalte an. Zunächst ziehen wir von der 2.Zeile das 2-fache der 1.Zeile und von der 3.Zeile das 3-fache der 1.Zeile ab und erhalten:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & 1 & -5 & 2 \\ 0 & -2 & -3 & 9 \end{array}\right).$$

Nun addieren wir zur 3.Zeile das 2-fache der 2.Zeile. Das liefert

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & 1 & -5 & 2 \\ 0 & 0 & -13 & 13 \end{array}\right).$$

Das zugehörige lineare Gleichungssystem lautet:

$$\begin{aligned} x + y + 2z &= 1 \\ y - 5z &= 2 \\ -13z &= 13 \end{aligned}$$

Es folgt sofort $z = -1$, $y = 2 + 5z = -3$ und $x = 1 - y - 2z = 6$. Also ist unser Gleichungssystem eindeutig lösbar und es gilt

$$L(A, b) = \left\{ \begin{pmatrix} 6 \\ -3 \\ -1 \end{pmatrix} \right\}.$$

3) Nun betrachten wir das homogene lineare Gleichungssystem

$$\begin{aligned} -y + z &= 0 \\ 3x + y + 4z &= 0 \\ 3x + 2y + 3z &= 0. \end{aligned}$$

Die Koeffizientenmatrix ist

$$A = \begin{pmatrix} 0 & -1 & 1 \\ 3 & 1 & 4 \\ 3 & 2 & 3 \end{pmatrix}.$$

Wir vertauschen zuerst die erste und die zweite Zeile und erhalten

$$\begin{pmatrix} 3 & 1 & 4 \\ 0 & -1 & 1 \\ 3 & 2 & 3 \end{pmatrix}.$$

Ziehen wir von der dritten Zeile die erste Zeile ab, so erhalten wir

$$\begin{pmatrix} 3 & 1 & 4 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

Nun addieren wir noch zur dritten Zeile die zweite Zeile, was

$$\begin{pmatrix} 3 & 1 & 4 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

liefert. Das entsprechende homogene lineare Gleichungssystem ist also:

$$\begin{aligned} 3x + y + 4z &= 0 \\ -y + z &= 0 \end{aligned}$$

Es folgt $y = z$ und $x = (1/3)(-y - 4z) = -5z/3$, wobei $z \in \mathbb{R}$ beliebig sein kann. Also gilt

$$L(A) = \left\{ z \begin{pmatrix} -5/3 \\ 1 \\ 1 \end{pmatrix} : z \in \mathbb{R} \right\} = \operatorname{span}\left\{ \begin{pmatrix} -5/3 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

$L(A)$ ist also eindimensional.
4) Nun betrachten wir das folgende inhomogene System, dass dieselbe Koeffizientenmatrix wie in 3) hat:

$$\begin{aligned} -y + z &= 1 \\ 3x + y + 4z &= 3 \\ 3x + 2y + 3z &= 3 \end{aligned}$$

Die erweiterte Koeffizientenmatrix ist

$$(A|b) = \left(\begin{array}{ccc|c} 0 & -1 & 1 & 1 \\ 3 & 1 & 4 & 3 \\ 3 & 2 & 3 & 3 \end{array}\right).$$

Dieselben Zeilenumformungen wie in 3) führen zu

$$\left(\begin{array}{ccc|c} 3 & 1 & 4 & 3 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

Die dritte Gleichung im zugehörigen System führt aber auf den Widerspruch $0 = 0x + 0y + 0z = 1$. Also kann dieses Gleichungssystem (und daher auch

das ursprüngliche) keine Lösung besitzen, d. h. es gilt $L(A,b) = \emptyset$.
5) Wie betrachten das inhomogene System

$$\begin{aligned} -y + z &= 1 \\ 3x + y + 4z &= 1 \\ 3x + 2y + 3z &= 0 \end{aligned}$$

mit der erweiterten Koeffizientenmatrix

$$(A|b) = \left(\begin{array}{ccc|c} 0 & -1 & 1 & 1 \\ 3 & 1 & 4 & 1 \\ 3 & 2 & 3 & 0 \end{array}\right).$$

Wieder führne wir dieselben Zeilenumformungen wie bei 3) durch und erhalten

$$\left(\begin{array}{ccc|c} 3 & 1 & 4 & 1 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Das entsprechende lineare Gleichungssystem ist

$$\begin{aligned} 3x + y + 4z &= 1 \\ -y + z &= 1. \end{aligned}$$

Dieses hat speziell die Lösung $z = 0$, $y = -1$, $x = (1/3)(1 - y) = 2/3$. Nach Lemma V.3.1 gilt dann

$$L(A,b) = \left\{ \begin{pmatrix} 2/3 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} x \\ y \\ z \end{pmatrix} : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in L(A) \right\}.$$

Die Lösungsmenge $L(A)$ des homogenen Systems hatten wir schon in 3) bestimmt. Es folgt

$$L(A,b) = \left\{ \begin{pmatrix} 2/3 \\ -1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -5/3 \\ 1 \\ 1 \end{pmatrix} : z \in \mathbb{R} \right\}.$$

6) Betrachten wir nun das homogene lineare Gleichungssytem

$$\begin{aligned} 2x - 3y &= 0 \\ x + y &= 0 \\ 4x - y &= 0. \end{aligned}$$

Die Koeffizientenmatrix ist

$$A = \begin{pmatrix} 2 & -3 \\ 1 & 1 \\ 4 & -1 \end{pmatrix}.$$

Zieht man von der zweiten Zeile das 1/2-fache der ersten Zeile und von der dritten Zeile das 2-fache der ersten Zeile ab, so erhält man die Matrix

$$\begin{pmatrix} 2 & -3 \\ 0 & 5/2 \\ 0 & 5 \end{pmatrix}.$$

Zieht man nun von der dritten Zeile noch das 2-fache der zweiten Zeile ab, so erhält man

$$\begin{pmatrix} 2 & -3 \\ 0 & 5/2 \\ 0 & 0 \end{pmatrix}$$

mit dem zugehörigen Gleichungssystem

$$\begin{aligned} 2x - 3y &= 0 \\ \frac{5}{2}y &= 0. \end{aligned}$$

Es folgt $y = 0 = x$, d. h. $L(A) = \{0\}$.
7) Zum Schluß betrachten wir noch das inhomogene lineare Gleichungssystem

$$\begin{aligned} x_1 + x_2 + x_3 - x_4 &= 1 \\ 2x_1 + 2x_2 - x_3 + 4x_4 &= 2 \\ 5x_1 + 5x_2 + 8x_3 - 11x_4 &= 5. \end{aligned}$$

Die erweiterte Koeffizientenmatrix ist

$$(A|b) = \left(\begin{array}{cccc|c} 1 & 1 & 1 & -1 & 1 \\ 2 & 2 & -1 & 4 & 2 \\ 5 & 5 & 8 & -11 & 5 \end{array}\right).$$

Wir ziehen von der zweiten Zeile das 2-fache der ersten Zeile und von der dritten Zeile das 5-fache der ersten Zeile ab und erhalten

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & -1 & 1 \\ 0 & 0 & -3 & 6 & 0 \\ 0 & 0 & 3 & -6 & 0 \end{array}\right).$$

Nun addieren wir noch zur dritten Zeile die zweite Zeile und kommen auf die Matrix

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & -1 & 1 \\ 0 & 0 & -3 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

mit dem zugehörigen Gleichungssystem

$$\begin{aligned} x_1 + x_2 + x_3 - x_4 &= 1 \\ -3x_3 + 6x_4 &= 0. \end{aligned}$$

Eine spezielle Lösung ist offensichtlich gegeben durch $x_3 = x_4 = 0 = x_2$ und $x_1 = 1$. Das entsprechende homogene System lautet

$$\begin{aligned} x_1 + x_2 + x_3 - x_4 &= 0 \\ -3x_3 + 6x_4 &= 0. \end{aligned}$$

Setzt man hier $\lambda := x_2$ und $\mu := x_4$, so folgt $x_3 = 2\mu$ und $x_1 = -\lambda - \mu$, also ist

$$L(A) = \left\{ \begin{pmatrix} -\lambda - \mu \\ \lambda \\ 2\mu \\ \mu \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\} = \operatorname{span} \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 2 \\ 1 \end{pmatrix} \right\}.$$

Man sieht leicht, dass diese beiden Vektoren auch linear unabhängig sind, also ist $L(A)$ zweidimensional.
Nach Lemma V.3.1 gilt für die allgemeine Lösung des inhomogenen Systems dann

$$L(A, b) = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -1 \\ 0 \\ 2 \\ 1 \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\}.$$

Hat man eine Matrix $A = (a_{ij})_{i,j=1}^{m,n}$ in Zeilenstufenform vorliegen (mit r von Null verschiedenen Zeilen), so kann man noch alle Pivotelemente a_{ij_i}, $i = 1, \ldots, r$, zu 1 machen, indem man die i-te Zeile jeweils mit $1/a_{ij_i}$ multipliziert (Typ (III) Umformungen). Die so erhaltene Matrix sei $A' = (a'_{ij})_{i,j=1}^{m,n}$. Nun kann man in A' noch alle Einträge oberhalb des r-ten Pivotelements durch Typ (II) Umformungen zu Null machen, indem man für $i = 1, \ldots, r-1$ jeweils das a'_{ij_r}-fache der r-ten Zeile von der i-ten Zeile abzieht. Danach macht man entsprechend alle Einträge oberhalb des $(r-1)$-ten Pivotelements zu Null usw., bis oberhalb aller Pivotelemente nur noch Nullen stehen. Diese Form nennt man dann auch *reduzierte Zeilenstufenform*. Bei einer Koeffizientenmatrix in reduzierter Zeilenstufenform ist die Lösungsmenge des linearen Gleichungssystems noch leichter abzulesen. Hierzu ein Beispiel:

Wir betrachten das lineare Gleichungssystem

$$\begin{aligned} 4x_1 + 2x_2 - x_3 - x_4 &= 1 \\ 2x_1 + 2x_2 - x_3 + 5x_4 &= 1 \\ 2x_1 + x_2 + x_3 - 6x_4 &= 1 \\ 4x_1 + x_2 + x_3 - 12x_4 &= 1, \end{aligned}$$

das die erweiterte Koeffizientenmatrix

$$(A|b) = \left(\begin{array}{cccc|c} 4 & 2 & -1 & -1 & 1 \\ 2 & 2 & -1 & 5 & 1 \\ 2 & 1 & 1 & -6 & 1 \\ 4 & 1 & 1 & -12 & 1 \end{array}\right)$$

hat. Der Gaußsche Algorithmus führt zunächst auf

$$\left(\begin{array}{cccc|c} 4 & 2 & -1 & -1 & 1 \\ 0 & 1 & -1/2 & 11/2 & 1/2 \\ 0 & 0 & 3/2 & -11/2 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

(nachrechnen). Nun multiplizieren wir noch die dritte Zeile mit $2/3$ und die erste Zeile mit $1/4$ und erhalten

$$\left(\begin{array}{cccc|c} 1 & 1/2 & -1/4 & -1/4 & 1/4 \\ 0 & 1 & -1/2 & 11/2 & 1/2 \\ 0 & 0 & 1 & -11/3 & 1/3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Jetzt addieren wir zur zweiten Zeile das $1/2$-fache der dritten Zeile und zur ersten Zeile das $1/4$-fache der dritten Zeile. Das führt auf

$$\left(\begin{array}{cccc|c} 1 & 1/2 & 0 & -7/6 & 1/3 \\ 0 & 1 & 0 & 11/3 & 2/3 \\ 0 & 0 & 1 & -11/3 & 1/3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Schließlich ziehen wir noch von der ersten Zeile das $1/2$-fache der zweiten Zeile ab und erhalten so die reduzierte Zeilenstufenform

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & -3 & 0 \\ 0 & 1 & 0 & 11/3 & 2/3 \\ 0 & 0 & 1 & -11/3 & 1/3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Setzt man $\lambda := x_4$, so erhält man aus dem zugehörigen linearen Gleichungssystem sofort $x_3 = 1/3 + (11/3)\lambda$, $x_2 = 2/3 - (11/3)\lambda$ und $x_1 = 3\lambda$. Es ist also

$$L(A,b) = \left\{ \begin{pmatrix} 0 \\ 2/3 \\ 1/3 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ -11/3 \\ 11/3 \\ 1 \end{pmatrix} : \lambda \in \mathbb{R} \right\}.$$

Nun betrachten wir noch den Zusammenhang zwischen der Lösbarkeit quadratischer linearer Gleichungssysteme und der Invertierbarkeit der Koeffizientenmatrix.

Lemma V.3.2. *Sei K ein Körper und sei $n \in \mathbb{N}$. Ferner sei $A \in M(n \times n, K)$. Dann sind folgende Aussagen äquivalent:*
1) A ist invertierbar.
2) Für alle $b \in K^n$ ist das lineare Gleichungssystem $Ax = b$ eindeutig lösbar.

Beweis. Sei zunächst A invertierbar und sei $b \in K^n$ beliebig. Dann gilt wegen $A^{-1}A = E_n$ aber

$$Ax = b \;\Leftrightarrow\; A^{-1}(Ax) = A^{-1}b \;\Leftrightarrow x = A^{-1}b.$$

Also hat $Ax = b$ genau eine Lösung, nämlich $x = A^{-1}b$.
Nehmen wir nun umgekehrt an es gelte 2). Dann ist insbesondere F_A surjektiv, also S-Rang$(A) = n$ und somit folgt aus Lemma V.2.19 die Invertierbarkeit von A. □

Wir kommen nun nochmal auf Zeilenrang und Spaltenrang einer Matrix zu sprechen. Wir wollen darauf hinaus, dass diese beiden Zahlen stets übereinstimmen. Dies zeigen wir zunächst für Matrizen in Zeilenstufenform.

Lemma V.3.3. *Sei K ein Körper, seien $m, n \in \mathbb{N}$ und sei $A \in M(m \times n, K)$ eine Matrix in Zeilenstufenform. Dann gilt Z-Rang(A) =S-Rang$(A) = r$, wobei r die Anzahl der von Null verschiedenen Zeilen von A ist.*

Beweis. Die Zeilen von A seien $z_1, \ldots, z_m$, die Spalten seien $a_1, \ldots, a_n$, die Pivotspalten seien $a_{j_1}, \ldots, a_{j_r}$, wobei $j_1 < j_2 < \cdots < j_r$. Es gilt also $z_i = 0$ für $i > r$ und somit ist der Zeilenraum $Z(A) = \text{span}\{z_1, \ldots, z_r\}$. Wegen der Stufenform von A sind die Zeilen $z_1, \ldots, z_r$ offensichtlich auch linear unabhängig, also gilt Z-Rang$(A) = r$. Ebenso sind aufgrund der Stufenform die Pivotspalten $a_{j_1}, \ldots, a_{j_r}$ linear unabhängig. Da alle Zeilen unterhalb der r-ten Zeile gleich Null sind, kann man die Spalten auch als Vektoren im K^r auffassen. Als r linear unabhängige Vektoren bilden $a_{j_1}, \ldots, a_{j_r}$ dann bereits eine Basis des K^r, also lässt sich jeder Spaltenvektor aus ihnen linear kombinieren. D. h. es gilt $S(A) = \text{span}\{a_{j_1}, \ldots, a_{j_r}\}$ und somit S-Rang$(A) = r$. □

Als Nächstes halten wir fest, dass sich der Zeilenraum (und folglich auch der Zeilenrang) bei elementaren Zeilenumformungen nicht ändert.

Lemma V.3.4. *Sei K ein Körper, seien $m, n \in \mathbb{N}$ und seien $A, A' \in M(m \times n, K)$. Falls A' durch elementare Zeilenumformungen aus A hervorgeht, so gilt $Z(A) = Z(A')$.*

Beweis. Es genügt, die Behauptung für den Fall einer einzigen elementaren Zeilenumformung zu zeigen. Für den Fall einer Typ (I) Umformung (Zeilenvertauschung) ist das klar. Betrachten wir nun den Fall einer Typ (II) Umformung: Es seien $z_1, \ldots, z_m$ die Zeilen von A und A' gehe aus A durch Addition

von λz_i zu z_j hervor (wobei $i \neq j$). Dann ist also $Z(A) = \text{span}\{z_1, \ldots, z_m\}$ und

$$Z(A') = \text{span}\{z_1, \ldots, z_{j-1}, z_j + \lambda z_i, z_{j+1}, \ldots, z_m\}.$$

Es ist aber $z_j + \lambda z_i \in Z(A)$ und folglich $Z(A') \subseteq Z(A)$. Umgekehrt ist auch $z_j = (z_j + \lambda z_i) - \lambda z_i \in Z(A')$ und somit auch $Z(A) \subseteq Z(A')$.
Den Fall einer Typ (III) Umformung können Sie zur Übung selbst beweisen. □

Das nächste Lemma besagt, dass sich der Spaltenrang einer Matrix bei Multiplikation von links mit einer invertierbaren Matrix nicht verändert.

Lemma V.3.5. *Sei K ein Körper, seien $m, n \in \mathbb{N}$ und sei $A \in M(m \times n, K)$. Ferner sei $B \in M(m \times m, K)$ invertierbar. Dann gilt S-Rang$(BA) = S$-Rang(A).*

Beweis. Wir wissen $S(A) = \text{Im}(F_A)$ und $S(BA) = \text{Im}(F_{BA})$. Daher ist die Abbildung $G : S(A) \to S(BA)$ mit $G(y) := By$ für $y \in S(A)$ wohldefiniert. Ferner ist G natürlich linear und aus $G(y) = By = 0$ folgt durch Multiplikation mit B^{-1} sofort $y = 0$. Also ist $\ker(G) = \{0\}$ und daher ist G injektiv.
Ist nun $z \in S(BA)$, so existiert ein $x \in K^n$ mit $z = (BA)x$. Dann ist $y := Ax \in S(A)$ mit $G(y) = By = z$. Also ist G auch surjektiv und somit ein Isomorphimus. Es ist also $S(A) \cong S(BA)$. Daraus folgt S-Rang$(A) = \dim(S(A)) = \dim(S(BA)) =$ S-Rang(BA). □

Als Nächstes zeigen wir, dass elementare Zeilenumformungen sich durch Multiplikation von links mit einer geeigneten invertierbaren Matrix ausdrücken lassen.

Lemma V.3.6. *Sei K ein Körper, seien $m, n \in \mathbb{N}$ und seien $A, A' \in M(m \times n, K)$ derart, dass A' durch elementare Zeilenumformungen aus A hervorgeht. Dann existiert eine invertierbare Matrix $B \in M(m \times m, K)$ mit $A' = BA$.*

Beweis. 1) Wir betrachten zunächst den Fall, dass A' durch eine einzige elementare Zeilenumformung aus A hervorgeht.
Typ (II) Umformung: A' gehe aus A durch Addition des λ-fachen der i-ten Zeile zur j-ten Zeile hervor (wobei $i \neq j$ und $\lambda \in K$).
Es sei $A = (a_{kl})_{k,l=1}^{m,n}$ und wir definieren die $m \times m$-Matrix $B = (b_{kl})_{k,l=1}^{m,m}$ durch

$$b_{kl} := \begin{cases} \delta_{kl} & \text{für } k \neq j, \\ \delta_{jl} + \lambda \delta_{il} & \text{für } k = j. \end{cases}$$

B ist gerade diejenige Matrix, die entsteht, wenn man in E_m das λ-fache der i-ten Zeile zur j-ten Zeile addiert.

Nun gilt für alle $k \in \{1, \ldots, m\}$ mit $k \neq j$ und alle $l \in \{1, \ldots, n\}$:

$$(BA)_{kl} = \sum_{s=1}^{m} b_{ks} a_{sl} = \sum_{s=1}^{m} \delta_{ks} a_{sl} = a_{kl}$$

und außerdem

$$(BA)_{jl} = \sum_{s=1}^{m} b_{js} a_{sl} = \sum_{s=1}^{m} (\delta_{js} + \lambda \delta_{is}) a_{sl} = \sum_{s=1}^{m} \delta_{js} a_{sl} + \lambda \sum_{s=1}^{m} \delta_{is} a_{sl} = a_{jl} + \lambda a_{il}.$$

Die k-te Zeile von BA stimmt also mit der k-ten Zeile von A' überein für alle $k = 1, \ldots, m$, also ist $A' = BA$.
Definiert man ferner $C \in M(m \times m, K)$ analog zu B nur mit $-\lambda$ anstelle von λ, so kann man leicht $BC = E_m = CB$ nachrechnen (Übung). Also sit B invertierbar mit $B^{-1} = C$.
Typ (I) Umformungen: A' gehe aus A durch Vertauschung der i-ten und der j-ten Zeile hervor.
Es sei B diejenige $m \times m$-Matrix, die aus der Einheitsmatrix E_m durch Vertauschung der i-ten und der j-ten Zeile entsteht. Dann rechnet man leicht $A' = BA$ und $B^2 = E_m$ nach (Übung). Letzteres impliziert die Invertierbarkeit von B mit $B^{-1} = B$.
Typ (III) Umformungen: A' entstehe aus A durch Multiplikation der i-ten Zeile mit $\lambda \in K \setminus \{0\}$. Sei B diejenige Matrix, die aus E_m mittels Multiplikation der i-ten Zeile mit λ hervorgeht (d. h. B ist eine Diagonalmatrix, der i-te Diagonaleintrag ist λ, alle anderen Diagonaleinträge sind 1). Durch direktes Nachrechnen sieht man $A' = BA$ (Übung). Ferner zeigt man leicht $BC = CB = E_m$, wobei C diejenige Matrix ist, die aus E_m mittels Multiplikation der i-ten Zeile mit $1/\lambda$ entsteht. Also ist B invertierbar mit $B^{-1} = C$.
2) Angenommen nun A' entsteht durch ν elementare Zeilenumformungen aus A. Nach 1) existieren dann invertierbare $m \times m$-Matrizen $B_1, \ldots, B_\nu$ mit $A' = (B_\nu B_{\nu-1} \ldots B_1)A$.
Wir setzen $B := B_\nu B_{\nu-1} \ldots B_1$. Dann ist $A' = BA$ und aus Lemma V.2.8 (mehrfach angewendet) folgt: B ist invertierbar mit $B^{-1} = B_1^{-1} B_2^{-1} \ldots B_\nu^{-1}$. □

Nun können wir endlich zeigen, dass Zeilenrang und Spaltenrang einer Matrix stets übereinstimmen.

Satz V.3.7. *Sei K ein Körper, seien $m, n \in \mathbb{N}$ und sei $A \in M(m \times n, K)$. Dann gilt S-*Rang(A) *= Z-*Rang(A).

Beweis. Mit Hilfe des Gaußschen Algorithmus kann man A mittels elementarer Zeilenumformungen in Zeilenstufenform überführen. Diese Matrix in Zeilenstufenform heiße A'. Nach Lemma V.3.4 gilt Z-Rang(A) = Z-Rang(A').

Da A' Zeilenstufenform hat, gilt nach Lemma V.3.3 aber Z-Rang(A') = S-Rang(A').
Außerdem existiert nach Lemma V.3.6 eine invertierbare Matrix $B \in M(m \times m, K)$ mit $A' = BA$. Wegen Lemma V.3.5 folgt daraus S-Rang(A') = S-Rang(A).
Insgesamt folgt also S-Rang(A) = Z-Rang(A). □

Da wir nun wissen, dass stets S-Rang(A) = Z-Rang(A) gilt, können wir diese Zahl einfach mit Rang(A) bezeichnen und nennen sie den *Rang* von A. Für diesen gelten folgende Aussagen.

Lemma V.3.8. *Sei K ein Körper, seien $m, n \in \mathbb{N}$ und sei $A \in M(m \times n, K)$. Dann gilt:*
1) Rang$(A) \leq \min\{m, n\}$[5]
2) Rang(A) = Rang(A^T)
3) Rang(BA) = Rang(A) *für invertierbares* $B \in M(m \times m, K)$
4) Rang(AC) = Rang(A) *für invertierbares* $C \in M(n \times n, K)$
5) Ist $m = n$, so ist A invertierbar genau dann, wenn Rang$(A) = n$ *gilt.*

Beweis. 1) Da der Zeilenraum $Z(A)$ ein Unterraum des n-dimensionalen Vektorraumes $M(1 \times n, K)$ ist, ist Rang(A) = dim$(Z(A)) \leq n$. Andererseits ist der Spaltenraum $S(A)$ ein Unterraum des m-dimensionalen Vektorraumes K^m und daher gilt auch Rang(A) = dim$(S(A)) \leq m$. Daraus folgt Rang$(A) \leq \min\{m, n\}$.
2) Es seien $a_1, \ldots, a_n$ die Spaltenvektoren von A. Dann sind $a_1^T, \ldots, a_n^T$ gerade die Zeilenvektoren von A^T. Daraus folgt leicht, dass die Abbildung $G : S(A) \to Z(A^T)$ mit $G(y) := y^T$ für $y \in S(A)$ ein Isomorphismus ist (Details als Übung). Somit gilt Rang(A) = dim$(S(A))$ = dim$(Z(A^T))$ = Rang(A^T).
3) folgt aus Lemma V.3.5.
4) Sei $C \in M(n \times n, K)$ invertierbar. Es gilt $C^T(C^{-1})^T = (C^{-1}C)^T = E_n^T = E_n$ und analog auch $(C^{-1})^T C^T = E_n$. Also ist auch C^T invertierbar mit $(C^T)^{-1} = (C^{-1})^T$. Mit Hilfe von 2) und 3) folgt nun Rang(AC) = Rang$((AC)^T)$ = Rang$(C^T A^T)$ = Rang(A^T) = Rang(A).
5) folgt aus Lemma V.2.19. □

Das folgende Lemma charakterisiert die Lösbarkeit eines linearen Gleichungssystems anhand des Ranges der Koeffizientenmatrix und der erweiterten Koeffizientenmatrix.

Lemma V.3.9. *Sei K ein Körper, seien $m, n \in \mathbb{N}$ und sei $A \in M(m \times n, K)$. Ferner sei $b \in K^m$. Dann gilt* Rang$(A) \leq$ Rang$(A|b)$. *Das lineare Gleichungssystem $Ax = b$ ist lösbar genau dann, wenn* Rang(A) = Rang$(A|b)$ *gilt.*

Weiterhin gilt für die Dimension des Lösungsraumes $L(A)$ des homogenen Systems: dim$(L(A)) = n -$ Rang(A).

[5] $\min\{m, n\}$ steht für das Minimum von m und n, also die kleinere der beiden Zahlen.

Beweis. Die Spaltenvektoren von A seien $a_1, \ldots, a_n$. Dann gilt für die Spaltenräume $S(A) = \text{span}\{a_1, \ldots, a_n\}$ und $S(A|b) = \text{span}\{a_1, \ldots, a_n, b\}$, also $S(A) \subseteq S(A|b)$ und folglich $\text{Rang}(A) \leq \text{Rang}(A|b)$.
Angenommen nun es gibt ein $x \in K^n$ mit $Ax = b$. Die Koordinaten von x seien $x_1, \ldots, x_n$. Dann gilt $b = Ax = \sum_{j=1}^{n} x_j a_j$ und daher ist $b \in S(A)$. Also ist $\{a_1, \ldots, a_n, b\} \subseteq S(A)$ und es folgt $S(A|b) = S(A)$, also auch $\text{Rang}(A) = \text{Rang}(A|b)$.
Sei nun umgekehrt $\text{Rang}(A) = \text{Rang}(A|b)$, also $\dim(S(A)) = \dim(S(A|b))$. Wegen $S(A) \subseteq S(A|b)$ folgt daraus $S(A) = S(A|b)$. Insbesondere ist $b \in S(A)$. D. h. es existieren $x_1, \ldots, x_n \in K$ mit $b = \sum_{j=1}^{n} x_j a_j$. Bezeichnet wieder $x \in K^n$ den Vektor mit den Koordinaten $x_1, \ldots, x_n$, so heißt das gerade $Ax = b$.
Wegen $L(A) = \ker(F_A)$ und $\text{Rang}(A) = \dim(S(A)) = \dim(\text{Im}(F_A))$ folgt aus der Dimensionsformel für lineare Abbildungen $\dim(L(A)) + \text{Rang}(A) = n$. □

Als letzten Punkt wollen wir noch das Problem der konkreten Berechnung der Inversen einer Matrix angehen. Auch dieses kann mit Hilfe des Gaußschen Algorithmus gelöst werden. Es sei also A eine $n \times n$-Matrix mit Einträgen aus einem Körper K. Wir bilden zunächst die $n \times 2n$-Matrix $(A|E_n)$. Diese formen wir nun mittels elementarer Zeilentransformationen so weit um, dass die linke $n \times n$-Matrix Zeilenstufenform hat. Das Ergebnis sei etwa $(A'|B)$. Falls A' mindestens eine Nullzeile enthält, so folgt aus unseren bisherigen Überlegungen $\text{Rang}(A) = \text{Rang}(A') < n$ und somit ist A nicht invertierbar.

Enthält A' dagegen keine Nullzeilen, so ist $\text{Rang}(A) = \text{Rang}(A') = n$, was die Invertierbarkeit von A impliziert. Um A^{-1} zu bestimmen, formen wir $(A'|B)$ nun noch weiter um und überführen A' mittels elementarer Zeilenumformungen in reduzierte Zeilenstufenform, welche bei einer quadratischen Matrix ohne Nullzeilen aber die Einheitsmatrix E_n sein muss. Als Ergebnis erhalten wir also eine $n \times 2n$-Matrix der Form $(E_n|C)$.

Wir wollen begründen, dass $A^{-1} = C$ gilt. Die Spalten von C seien $c_1, \ldots, c_n$. Dann geht die Matrix $(E_n|c_i)$ also jeweils durch elementare Zeilenumformungen aus $(A|e_i)$ hervor (denn wir hatten mit der Matrix $(A|E_n)$ begonnen und die i-te Spalte von E_n ist gerade e_i). Es folgt $L(A, e_i) = L(E_n, c_i)$ für alle $i = 1, \ldots, n$. Aber natürlich ist $L(E_n, c_i) = \{c_i\}$. Es folgt $Ac_i = e_i$ für alle $i = 1, \ldots, n$. Ac_i ist aber gerade die i-te Spalte von AC. Also gilt $AC = E_n$. Multiplikation mit von links mit A^{-1} liefert nun $C = A^{-1}$.

Zur Illustration des oben beschriebenen Verfahrens betrachten wir zwei Beispiele.

Beispiele:
1) Sei

$$A = \begin{pmatrix} 2 & -1 & -1 \\ 3 & 2 & 4 \\ 2 & 6 & 10 \end{pmatrix}.$$

Wir bilden die Matrix

$$(A|E_3) = \left(\begin{array}{ccc|ccc} 2 & -1 & -1 & 1 & 0 & 0 \\ 3 & 2 & 4 & 0 & 1 & 0 \\ 2 & 6 & 10 & 0 & 0 & 1 \end{array}\right).$$

Hier ziehen wir zunächst das 3/2-fache der ersten Zeile von der zweiten Zeile und von der dritten Zeile die erste Zeile ab. Das liefert

$$\left(\begin{array}{ccc|ccc} 2 & -1 & -1 & 1 & 0 & 0 \\ 0 & 7/2 & 11/2 & -3/2 & 1 & 0 \\ 0 & 7 & 11 & -1 & 0 & 1 \end{array}\right).$$

Nun ziehen wir von der dritten Zeile das 2-fache der zweiten Zeile ab und erhalten

$$\left(\begin{array}{ccc|ccc} 2 & -1 & -1 & 1 & 0 & 0 \\ 0 & 7/2 & 11/2 & -3/2 & 1 & 0 \\ 0 & 0 & 0 & 2 & -2 & 1 \end{array}\right).$$

Da links eine Nullzeile entstanden ist, ist A nicht invertierbar.
2) Sei

$$A = \left(\begin{array}{ccc} 1 & 3 & 4 \\ 2 & 4 & 1 \\ -1 & 3 & 2 \end{array}\right).$$

Wir bilden

$$(A|E_3) = \left(\begin{array}{ccc|ccc} 1 & 3 & 4 & 1 & 0 & 0 \\ 2 & 4 & 1 & 0 & 1 & 0 \\ -1 & 3 & 2 & 0 & 0 & 1 \end{array}\right).$$

Von der zweiten Zeile ziehen wir das 2-fache der ersten Zeile ab und zur dritten Zeile addieren wir die erste Zeile. Das liefert

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 4 & 1 & 0 & 0 \\ 0 & -2 & -7 & -2 & 1 & 0 \\ 0 & 6 & 6 & 1 & 0 & 1 \end{array}\right).$$

Nun addieren wir zur dritten Zeile das 3-fache der zweiten Zeile. Damit erhalten wir

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 4 & 1 & 0 & 0 \\ 0 & -2 & -7 & -2 & 1 & 0 \\ 0 & 0 & -15 & -5 & 3 & 1 \end{array}\right).$$

Jetzt multiplizieren wir die dritte Zeile mit $-1/15$ und die zweite Zeile mit $-1/2$ und erhalten dadurch

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 4 & 1 & 0 & 0 \\ 0 & 1 & 7/2 & 1 & -1/2 & 0 \\ 0 & 0 & 1 & 1/3 & -1/5 & -1/15 \end{array}\right).$$

Von der zweiten Zeile ziehen wir nun das 7/2-fache der dritten Zeile ab und von der ersten Zeile subtrahieren wir das 4-fache der dritten Zeile. Das führt auf

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 0 & -1/3 & 4/5 & 4/15 \\ 0 & 1 & 0 & -1/6 & 1/5 & 7/30 \\ 0 & 0 & 1 & 1/3 & -1/5 & -1/15 \end{array}\right).$$

Schließlich ziehen wir noch von der ersten Zeile das 3-fache der zweiten Zeile ab und erhalten

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/6 & 1/5 & -13/30 \\ 0 & 1 & 0 & -1/6 & 1/5 & 7/30 \\ 0 & 0 & 1 & 1/3 & -1/5 & -1/15 \end{array}\right).$$

A ist also invertierbar und es gilt

$$A^{-1} = \left(\begin{array}{ccc} 1/6 & 1/5 & -13/30 \\ -1/6 & 1/5 & 7/30 \\ 1/3 & -1/5 & -1/15 \end{array}\right).$$

VI Determinanten

In diesem Kapitel wollen wir eine wichtige Kennzahl einer Matrix, ihre sogenannte Determinante, einführen. Als Vorbereitung dazu betrachten wir Permutationen, welche auch für sich genommen ein interessantes Thema darstellen.

VI.1 Vorbereitung: Permutationen

Unter einer Permutation von n Elementen (sagen wir $1, \ldots, n$) versteht man eine Umordnung, in der jedes Element genau einmal vorkommt. Formal ist eine Permutation von $1, \ldots, n$ also nichts anderes als eine bijektive Abbildung $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. Die Menge all dieser Permutationen σ bezeichnen wir mit S_n.

Ein Element $\sigma \in S_n$ gibt man häufig auch als eine $2 \times n$-Matrix an:

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

Es ist nicht schwierig zu zeigen, dass S_n versehen mit der Verkettung $\circ$ als Verknüpfung eine Gruppe bildet, die sogenannte *symmetrische Gruppe* der Ordnung der n (neutrales Element ist die identische Abbildung id, Inverses zu $\sigma \in S_n$ ist die Umkehrabbildung σ^{-1}). Die Gruppe S_n ist allerdings für $n \geq 3$ nicht kommutativ, wie wir gleich zeigen werden. Zuvor definieren wir erst noch spezielle Permutationen, nämlich solche, die zwei Elemente vertauschen, aber die anderen unverändert lassen: Seien $i, j \in \{1, \ldots, n\}$ mit $i \neq j$. Wir definieren $\tau_{ij} : \{1, \ldots, n\} \to \{1, \ldots, n\}$ durch

$$\tau_{ij}(k) := \begin{cases} i & \text{für } k = j, \\ j & \text{für } k = i, \\ k & \text{für } k \in \{1, \ldots, n\} \setminus \{i, j\}. \end{cases}$$

Offensichtlich ist τ_{ij} bijektiv, also $\tau_{ij} \in S_n$. Es gilt $\tau_{ij} \circ \tau_{ij} = \mathrm{id}$, also $\tau_{ij}^{-1} = \tau_{ij}$. Permutationen der Form τ_{ij} nennt man auch *Transpositionen.*

Offenbar gilt $S_2 = \{\mathrm{id}, \tau_{12}\}$ und folglich ist S_2 kommutativ. Ist aber $n \geq 3$, so gilt z. B. $(\tau_{12} \circ \tau_{13})(1) = 3$ und $(\tau_{13} \circ \tau_{12})(1) = 2$, also $\tau_{12} \circ \tau_{13} \neq \tau_{13} \circ \tau_{12}$.

Wie viele Permutationen von $\{1,\dots,n\}$ gibt es? Zur Konstruktion einer Permutation $\sigma : \{1,\dots,n\} \to \{1,\dots,n\}$ hat man zunächst n Wahlmöglichkeiten zur Definition von $\sigma(1)$, anschließend noch $n-1$ Wahlmöglichkeiten für $\sigma(2)$, dann noch $n-2$ Wahlmöglichkeiten für $\sigma(3)$ usw., bis man schließlich nur noch 2 Möglichkeiten für $\sigma(n-1)$ und nur eine einzige Möglichkeit für $\sigma(n)$ hat. Das ergibt insgesamt $n(n-1)(n-2)\dots 2\cdot 1$ Möglichkeiten. Diese Zahl bezeichnet man üblicherweise mit $n!$ (gelesen als "n Fakultät"). Die Menge S_n hat also $n!$ Elemente.

Als Nächstes zeigen wir das wichtige Ergebnis, dass sich jede Permutation als Produkt (Verkettung) von Transpositionen schreiben lässt, d. h. man kann jede Umordnung der Elemente $1,\dots,n$ durch sukzessives Vertauschen von je zwei Elementen erreichen.

Satz VI.1.1. *Für alle $n \geq 2$ und alle $\sigma \in S_n$ gilt: Es existieren Transpositionen $\sigma_1,\dots,\sigma_m \in S_n$ mit $\sigma = \sigma_1 \circ \dots \circ \sigma_m$.*

Beweis. Ist $\sigma \in S_n$ mit $\sigma \neq \mathrm{id}$, so gibt es ein $k \in \{1,\dots,n\}$ mit $\sigma(k) \neq k$. Wir setzen

$$k_0(\sigma) := \min\{k \in \{1,\dots,n\} : \sigma(k) \neq k\}$$

(min steht für das Minimum, also das kleinste Element der Menge). Weiter setzen wir noch $k_0(\mathrm{id}) := n$.
Wir zeigen nun induktiv folgendes: Für alle $i \in \{0.\dots,n-1\}$ lässt sich jedes $\sigma \in S_n$ mit $k_0(\sigma) \geq n-i$ als Produkt von Transpositionen schreiben. Damit ist dann unsere Behauptung bewiesen (denn für jedes $\sigma \in S_n$ gilt $k_0(\sigma) \in \{1,\dots,n\}$, also $k_0(\sigma) = n-i$ für ein $i \in \{0,\dots,n-1\}$).
Induktionsanfang: $i = 0$ Es sei also $\sigma \in S_n$ mit $k_0(\sigma) = n$. Wäre $\sigma \neq \mathrm{id}$, so hieße das nach Definition von k_0 aber $\sigma(k) = k$ für $k = 1,\dots,n-1$ und $\sigma(n) \neq n$, was aber wegen der Bijektivität von σ nicht sein kann. Also muss $\sigma = \mathrm{id}$ gelten und es ist z. B. $\mathrm{id} = \tau_{12} \circ \tau_{12}$.
Induktionsschritt: Die Behauptung gelte für ein $i \in \{0,\dots,n-2\}$. Es sei $\sigma \in S_n$ mit $k_0(\sigma) \geq n-(i+1) = n-i-1$. Ist sogar $k_0(\sigma) \geq n-i$, so wissen wir nach Induktionsvoraussetzung schon, dass σ ein Produkt von Transpositionen ist. Sei also $k_0(\sigma) = n-i-1$.
Die Definition von k_0 impliziert dann $\sigma(k) = k$ für $k = 1,\dots,n-i-2$ und $\sigma(n-i-1) \neq n-i-1$. Es folgt $\sigma(n-i-1) > n-i-1$.
Es sei $\tau \in S_n$ die Vertauschung von $n-i-1$ und $\sigma(n-i-1)$ und $\rho := \tau \circ \sigma$. Dann gilt $\rho(k) = k$ für alle $k = 1,\dots,n-i-1$, wie man leicht durch einsetzen bestätigt. Daher gilt $k_0(\rho) \geq n-i$.
Nach Induktionsvoraussetzung lässt sich also ρ als Produkt von Transpositionen schreiben und folglich ist auch $\sigma = \tau \circ \rho$ wieder ein Produkt von Transpositionen. □

Als Nächstes wollen wir das Vorzeichen (oder Signum) einer Permutation definieren.

Definition VI.1.2. Seien $n \in \mathbb{N}$ und $\sigma \in S_n$. Wir setzen

$$F_\sigma := \Big\{(i,j) \in \{1,\dots,n\}^2 : i < j \text{ und } \sigma(i) > \sigma(j)\Big\}$$

und bezeichnen mit $f(\sigma)$ die Anzahl der Elemente von F_σ. $f(\sigma)$ heißt die Anzahl der *Fehlstände* von σ. Schließlich setzen wir $\operatorname{sign}(\sigma) := (-1)^{f(\sigma)}$. $\operatorname{sign}(\sigma)$ heißt das *Vorzeichen* oder *Signum* von σ.

Zum Beispiel gilt $\operatorname{sign}(\mathrm{id}) = 1$ und $\operatorname{sign}(\tau) = -1$ für alle Transpositionen $\tau \in S_n$ (Übung).

Wir zeigen nun folgende nützliche Darstellung des Vorzeichens einer Permutation.

Lemma VI.1.3. *Seien $n \in \mathbb{N}$ und $\sigma \in S_n$. Dann gilt*

$$\operatorname{sign}(\sigma) = \prod_{(i,j)\in I_n} \frac{\sigma(i) - \sigma(j)}{i - j},$$

wobei $I_n := \Big\{(i,j) \in \{1,\dots,n\}^2 : i < j\Big\}$.

Beweis. Wir definieren $\varphi : I_n \to I_n$ wie folgt:

$$\varphi(i,j) := \begin{cases} (\sigma(i),\sigma(j)), & \text{falls } \sigma(i) < \sigma(j), \\ (\sigma(j),\sigma(i)), & \text{falls } \sigma(i) > \sigma(j). \end{cases}$$

Es ist nicht schwierig zu zeigen, dass φ bijektiv ist (Übung).
Wir setzen weiter

$$A := \{(i,j) \in I_n : \sigma(i) > \sigma(j)\} \quad \text{und} \quad B := I_n \setminus A.$$

Ferner sei $G(i,j) := i - j$ für alle $(i,j) \in I_n$. Da $\varphi : I_n \to I_n$ bijektiv ist, gilt

$$\prod_{(i,j)\in I_n} G(i,j) = \prod_{(i,j)\in I_n} G(\varphi(i,j)),$$

denn links wie rechts stehen dieselben Faktoren, nur in einer anderen Reihenfolge. Nun ist aber nach Definition von φ

$$G(\varphi(i,j)) = \begin{cases} \sigma(i) - \sigma(j) & \text{für } (i,j) \in B, \\ -(\sigma(i) - \sigma(j)) & \text{für } (i,j) \in A. \end{cases}$$

Es folgt:

$$\begin{aligned} \prod_{(i,j)\in I_n} (i-j) &= \prod_{(i,j)\in I_n} G(i,j) = \prod_{(i,j)\in I_n} G(\varphi(i,j)) \\ &= \prod_{(i,j)\in B} (\sigma(i) - \sigma(j)) \prod_{(i,j)\in A} (-(\sigma(i) - \sigma(j))) \\ &= (-1)^{f(\sigma)} \prod_{(i,j)\in B} (\sigma(i) - \sigma(j)) \prod_{(i,j)\in A} (\sigma(i) - \sigma(j)) \\ &= \operatorname{sign}(\sigma) \prod_{(i,j)\in I_n} (\sigma(i) - \sigma(j)), \end{aligned}$$

denn die Anzahl der Elemente von A ist gerade $f(\sigma)$.
Nun folgt aber

$$\prod_{(i,j)\in I_n} \frac{\sigma(i)-\sigma(j)}{i-j} = \frac{\prod_{(i,j)\in I_n}(\sigma(i)-\sigma(j))}{\prod_{(i,j)\in I_n}(i-j)} = \frac{1}{\operatorname{sign}(\sigma)} = \operatorname{sign}(\sigma).$$

□

Damit können wir nun den folgenden entscheidenden Satz zeigen.

Satz VI.1.4. *Sei $n \in \mathbb{N}$ und seien $\sigma_1, \sigma_2 \in S_n$. Dann gilt*

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

Beweis. Nach dem obigen Lemma gilt

$$\begin{aligned}\operatorname{sign}(\sigma_1 \circ \sigma_2) &= \prod_{(i,j)\in I_n} \frac{\sigma_1(\sigma_2(i)) - \sigma_1(\sigma_2(j))}{i-j} \\ &= \prod_{(i,j)\in I_n} \frac{\sigma_2(i)-\sigma_2(j)}{i-j} \prod_{(i,j)\in I_n} \frac{\sigma_1(\sigma_2(i)) - \sigma_1(\sigma_2(j))}{\sigma_2(i)-\sigma_2(j)} \\ &= \operatorname{sign}(\sigma_2) \prod_{(i,j)\in I_n} \frac{\sigma_1(\sigma_2(i)) - \sigma_1(\sigma_2(j))}{\sigma_2(i)-\sigma_2(j)}.\end{aligned}$$

Wie im vorigen Beweis definieren wir eine bijektive Abbildung $\varphi : I_n \to I_n$ wie folgt:

$$\varphi(i,j) := \begin{cases} (\sigma_2(i), \sigma_2(j)), & \text{falls } \sigma_2(i) < \sigma_2(j), \\ (\sigma_2(j), \sigma_2(i)), & \text{falls } \sigma_2(i) > \sigma_2(j). \end{cases}$$

Ferner sei

$$H(k,l) := \frac{\sigma_1(k) - \sigma_1(l)}{k-l} \quad \forall (k,l) \in I_n.$$

Dann gilt $H(\varphi(i,j)) = (\sigma_1(\sigma_2(i)) - \sigma_1(\sigma_2(j))/(\sigma_2(i) - \sigma_2(j))$ für alle $(i,j) \in I_n$, wie man sofort nachprüft. Es folgt

$$\operatorname{sign}(\sigma_1 \circ \sigma_2) = \operatorname{sign}(\sigma_2) \prod_{(i,j)\in I_n} H(\varphi(i,j)) = \operatorname{sign}(\sigma_2) \prod_{(i,j)\in I_n} H(i,j).$$

Nach dem obigen Lemma ist aber $\prod_{(i,j)\in I_n} H(i,j) = \operatorname{sign}(\sigma_1)$. □

Aus dem obigen Satz ergibt sich sofort folgendes Korollar.

Korollar VI.1.5. *Sei $n \in \mathbb{N}$ und sei $\sigma \in S_n$. Sind $\sigma_1, \ldots, \sigma_m \in S_n$ Transpositionen mit $\sigma = \sigma_1 \circ \cdots \circ \sigma_m$, so gilt* $\operatorname{sign}(\sigma) = (-1)^m$.

Beweis. Aus Satz VI.1.4 folgt $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma_1 \circ \cdots \circ \sigma_m) = \prod_{i=1}^{m} \operatorname{sign}(\sigma_i) = (-1)^m$, da Transpositionen stets das Signum -1 haben. □

Dieses Korollar impliziert insbesondere, dass man eine Permutation σ nicht einerseits als Produkt von einer geraden Anzahl von Transpositionen und andererseits als ein Produkt von einer ungeraden Anzahl Transpositionen darstellen kann.

VI.2 Die Determinante einer Matrix

Nach der Vorbereitung im letzten Abschnitt können wir nun direkt die Determinante definieren. Diese ist nur für quadratische Matrizen erklärt.

Definition VI.2.1. Sei K ein Körper und sei $A = (a_{ij})_{i,j=1}^{n,n}$ eine $n \times n$-Matrix mit Einträgen in K. Die *Determinante* von A ist definiert durch

$$\det(A) := \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$$

(diese Formel heißt *Leibniz-Formel*[1]). Anstelle von $\det(A)$ schreibt man auch $|A|$.

Die obige Summe erstreckt sich über alle Elemente σ von S_n, sie besteht also aus $n!$ Summanden. In jedem Summanden bildet man das Produkt $a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$ und versieht dieses mit dem Vorzeichen der Permutation σ.

Wir wollen nun zunächst die Determinante einer 2×2-Matrix bestimmen. Da S_2 nur die beiden Elemente id und τ_{12} hat (mit Vorzeichen 1 bzw. -1), folgt

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

(bei der Schreibweise mit den senkrechten Strichen lässt man die runden Klammern um die Matrix in der Regel weg). Die Determinante einer 2×2-Matrix berechnet sich also, indem man die Einträge "über Kreuz multipliziert" und anschließend die Differenz bildet. Zum Beispiel gilt

$$\begin{vmatrix} 2 & 3 \\ 4 & 2 \end{vmatrix} = 2 \cdot 2 - 3 \cdot 4 = -8.$$

[1] Gottfried Wilhelm Leibniz (1646–1716): deutscher Universalgelehrter (unter anderem Mathematiker, Philosoph, Historiker), Erfinder der Differentialrechnung (unabhängig von Isaac Newton).

Für 3×3-Matrizen wird es bereits deutlich komplizierter. Die Menge S_3 hat $3! = 6$ Elemente, nämlich

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Von diesen haben die ersten drei das Vorzeichen 1, die anderen das Vorzeichen -1, wie man leicht nachprüft.

Damit erhält man folgenden Ausdruck für die Determinante einer 3×3-Matrix:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$
$$- a_{13}a_{22}a_{31} = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$
$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Tatsächlich ist dies nur ein Spezialfall eines allgemeinen Entwicklungssatzes (siehe Satz VI.2.6).

Beispiel: Es gilt

$$\begin{vmatrix} 1 & 2 & 0 \\ 2 & -1 & 3 \\ 3 & 4 & -2 \end{vmatrix} = 1 \cdot (2 - 12) - 2 \cdot (-4 - 9) + 0 \cdot (8 + 3) = -10 + 26 = 16.$$

Als Nächstes stellen wir einige wichtige Eigenschaften von Determinanten zusammen.

Lemma VI.2.2. *Es sei K ein Körper und es seien $A, A' \in M(n \times n, K)$. Ferner sei $b \in M(1 \times n, K)$ ein Zeilenvektor der Länge n und $A_{i,b}$ sei diejenige Matrix, in aus A entsteht, indem man die i-te Zeile durch b ersetzt. Dann gilt:*
1) Entsteht A' aus A durch Multiplikation einer Zeile mit $\lambda \in K$, so gilt $\det(A') = \lambda \det(A)$.
2) Es gilt $\det(\lambda A) = \lambda^n \det(A)$ *für alle $\lambda \in K$.*
3) Ensteht aus A' aus A durch Addition des Zeilenvektors b zur i-ten Zeile, so gilt $\det(A') = \det(A) + \det(A_{i,b})$.
4) Sind zwei Zeilen in A identisch, so ist $\det(A) = 0$.
5) Entsteht A' aus A durch Vertauschung zweier Zeilen, so ist $\det(A') = -\det(A)$.
6) Ensteht aus A' aus A durch eine Zeilenumformung vom Typ (II), so gilt $\det(A') = \det(A)$.
7) Es ist $\det(E_n) = 1$.

Beweis. 1) A' entstehe aus A durch Multiplikation der i-ten Zeile mit λ. Dann gilt:

$$\begin{aligned}\det(A') &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots (\lambda a_{i\sigma(i)}) \dots a_{n\sigma(n)} \\ &= \lambda \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)} = \lambda \det(A).\end{aligned}$$

2) folgt durch n-fache Anwendung von 1).
3) Sei $b = (b_1 \ \dots \ b_n)$. Es gilt

$$\begin{aligned}&\det(A') = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots (a_{i\sigma(i)} + b_{\sigma(i)}) \dots a_{n\sigma(n)} \\ &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)(a_{1\sigma(1)} \dots a_{n\sigma(n)} + a_{1\sigma(1)} \dots b_{\sigma(i)} \dots a_{n\sigma(n)}) \\ &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)} + \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \dots b_{\sigma(i)} \dots a_{n\sigma(n)} \\ &= \det(A) + \det(A_{i,b}).\end{aligned}$$

4) Seien $i, j \in \{1, \dots, n\}$ mit $i \neq j$ und $a_{ik} = a_{jk}$ für alle $k = 1, \dots, n$. Es sei τ die Transposition τ_{ij}. Wir setzen

$$\begin{aligned}A_n &:= \{\sigma \in S_n : \operatorname{sign}(\sigma) = 1\}, \\ B_n &:= \{\sigma \in S_n : \operatorname{sign}(\sigma) = -1\}.\end{aligned}$$

Ferner sei $\varphi : A_n \to B_n$ definiert durch $\varphi(\sigma) := \sigma \circ \tau$ für alle $\sigma \in A_n$. Dann ist φ wohldefiniert (d. h. φ bildet tatsächlich nach B_n ab) und bijektiv (Übung). Es gilt

$$\begin{aligned}&\det(A) = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \prod_{s=1}^{n} a_{s\sigma(s)} = \sum_{\sigma\in A_n} \prod_{s=1}^{n} a_{s\sigma(s)} - \sum_{\sigma\in B_n} \prod_{s=1}^{n} a_{s\sigma(s)} \\ &= \sum_{\sigma\in A_n} \prod_{s=1}^{n} a_{s\sigma(s)} - \sum_{\sigma\in A_n} \prod_{s=1}^{n} a_{s\sigma(\tau(s))}\end{aligned}$$

(Letzteres wegen der Bijektivität von φ). Wegen der Bijektivität von τ und $\tau \circ \tau = \mathrm{id}$ ist aber

$$\prod_{s=1}^{n} a_{s\sigma(\tau(s))} = \prod_{s=1}^{n} a_{\tau(s)\sigma(\tau(\tau(s)))} = \prod_{s=1}^{n} a_{\tau(s)\sigma(s)} = \prod_{s=1}^{n} a_{s\sigma(s)}$$

für alle $\sigma \in A_n$, wobei wir im letzten Schritt die Voraussetzung "i-te Zeile von A" = "j-te Zeile von A" ausgenutzt haben.
Insgesamt folgt also $\det(A) = 0$.

5) Sind $v_1, \ldots, v_n \in M(1 \times n, K)$, so schreiben wir $\det(v_1, \ldots, v_n)$ für die Determinante der Matrix mit den Zeilenvektoren $v_1, \ldots, v_n$.
Es seien nun $z_1, \ldots, z_n$ die Zeilenvektoren von A und A' entstehe aus A durch Vertauschung von Zeile i und Zeile j (wobei $i \neq j$). Dann gilt wegen 4) $\det(z_1, \ldots, z_i + z_j, \ldots, z_i + z_j, \ldots, z_n) = 0$, wobei hier an der i-ten und der j-ten Stelle jeweils $z_i + z_j$ stehen soll, an allen Stellen $k \in \{1, \ldots, n\} \setminus \{i, j\}$ steht z_k.
Andererseits folgt aus 3)

$$\begin{aligned}
&0 = \det(z_1, \ldots, z_i + z_j, \ldots, z_i + z_j, \ldots, z_n) = \\
&\det(z_1, \ldots, z_i, \ldots, z_i, \ldots, z_n) + \det(z_1, \ldots, z_i, \ldots, z_j, \ldots, z_n) \\
&+ \det(z_1, \ldots, z_j, \ldots, z_i, \ldots, z_n) + \det(z_1, \ldots, z_j, \ldots, z_j, \ldots, z_n) \\
&= \det(A) + \det(A'),
\end{aligned}$$

denn wiederum wegen 4) gilt

$$\det(z_1, \ldots, z_i, \ldots, z_i, \ldots, z_n) = 0 = \det(z_1, \ldots, z_j, \ldots, z_j, \ldots, z_n).$$

Es folgt $\det(A') = -\det(A)$.
6) A' entstehe aus A durch Addition des λ-fachen der j-ten Zeile zur i-ten Zeile (wobei $i \neq j$). Die Zeilenvektoren von A seien wieder $z_1, \ldots, z_n$. Dann gilt wegen 3) und 1)

$$\begin{aligned}
&\det(A') = \det(z_1 \ldots, z_i + \lambda z_j, \ldots, z_n) \\
&= \det(z_1, \ldots, z_i, \ldots, z_n) + \lambda \det(z_1 \ldots, z_j, \ldots, z_n) \\
&= \det(A) + \lambda \det(z_1 \ldots, z_j, \ldots, z_n).
\end{aligned}$$

In $\det(z_1 \ldots, z_j, \ldots, z_n)$ steht sowohl in der i-ten als auch in der j-ten Zeile z_j, also ist wegen 4) $\det(z_1 \ldots, z_j, \ldots, z_n) = 0$ und es folgt $\det(A') = \det(A)$.
7) Für alle $\sigma \in S_n$ mit $\sigma \neq \mathrm{id}$ existiert ein $k \in \{1, \ldots, n\}$ mit $\sigma(k) \neq k$, also $\delta_{k\sigma(k)} = 0$, also $\prod_{s=1}^{n} \delta_{s\sigma(s)} = 0$. Es folgt

$$\det(E_n) = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{s=1}^{n} \delta_{s\sigma(s)} = \mathrm{sign}(\mathrm{id}) \prod_{s=1}^{n} \delta_{ss} = 1.$$

□

Die Eigenschaften 1) und 3) besagen gerade, dass die Determinante in jeder Zeile linear ist. Als Abbildung vom Vektorraum $M(n \times n, K)$ nach K ist det dagegen nicht linear. Das folgt allein schon aus Punkt 2) des obigen Lemmas, aber ferner ist im Allgemeinen auch $\det(A+B) \neq \det(A) + \det(B)$, wie man sich leicht anhand von Beispielen klar macht.

Als Nächstes zeigen wir noch, dass die Determinante einer Matrix stets mit der Determinante ihrer Transponierten übereinstimmt.

Satz VI.2.3. *Sei K ein Körper und sei $A \in M(n \times n, K)$. Dann gilt* $\det(A^T) = \det(A)$.

Beweis. Wie man leicht sieht, ist die Abbildung $\psi : S_n \to S_n$ mit $\psi(\sigma) := \sigma^{-1}$ für alle $\sigma \in S_n$ bijektiv. Daher gilt

$$\det(A^T) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma^{-1}) \prod_{i=1}^{n} a_{\sigma^{-1}(i)i}$$

(denn man hat nur die Summanden mittels ψ umgeordnet).
Nun gilt aber

$$1 = \operatorname{sign}(\mathrm{id}) = \operatorname{sign}(\sigma^{-1} \circ \sigma) = \operatorname{sign}(\sigma^{-1})\operatorname{sign}(\sigma),$$

also

$$\operatorname{sign}(\sigma^{-1}) = 1/\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma)$$

und ferner

$$\prod_{i=1}^{n} a_{\sigma^{-1}(i)i} = \prod_{i=1}^{n} a_{i\sigma(i)}$$

(wegen der Bijektivität von σ).
Es folgt

$$\det(A^T) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = \det(A).$$

□

Aus diesem Satz folgt sofort, dass alles was wir in Lemma VI.2.2 hinsichtlich des Verhaltens der Determinante bei Veränderung der Zeilen der Matrix festgestellt hatten, sinngemäß auch für die Spalten der Matrix gilt.

Das folgende Invertierbarkeitskriterium ist eine Hauptanwendung der Determinanten.

Satz VI.2.4. *Sei K ein Körper und sei $A \in M(n \times n, K)$. Dann gilt:*
1) A ist invertierbar genau dann, wenn $\det(A) \neq 0$ *gilt.*
2) Es existiert ein $x \in K^n \setminus \{0\}$ mit $Ax = 0$ genau dann, wenn $\det(A) = 0$ *gilt.*

Beweis. 1) Seien $z_1, \ldots, z_n$ die Zeilenvektoren von A. Wir wissen schon, dass A genau dann invertierbar ist, wenn $\operatorname{Rang}(A) = n$ gilt. Das wiederum ist äquivalent zur linearen Unabhängigkeit von $z_1, \ldots, z_n$.
a) Seien $z_1, \ldots, z_n$ linear unabhängig. Dann ist $(z_1, \ldots, z_n)$ eine geordnete Basis von $M(1 \times n, K)$ und folglich existieren $\alpha_{ij} \in K$ $(i, j = 1, \ldots, n)$ mit $e_i^T = \sum_{j=1}^{n} \alpha_{ij} z_j$ für alle $i = 1, \ldots, n$.

Da die Determinante in jeder Zeile linear ist, ergibt sich

$$\begin{aligned}
1 &= \det(E_n) = \det(e_1^T, \dots, e_n^T) \\
&= \det\left(\sum_{j_1=1}^{n} \alpha_{1j_1} z_{j_1}, \sum_{j_2=1}^{n} \alpha_{2j_2} z_{j_2}, \dots, \sum_{j_n=1}^{n} \alpha_{nj_n} z_{j_n}\right) \\
&= \sum_{j_1=1}^{n} \alpha_{1j_1} \det\left(z_{j_1}, \sum_{j_2=1}^{n} \alpha_{2j_2} z_{j_2}, \dots, \sum_{j_n=1}^{n} \alpha_{nj_n} z_{j_n}\right) \\
&= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \alpha_{1j_1}\alpha_{2j_2} \det\left(z_{j_1}, z_{j_2}, \sum_{j_3=1}^{n} \alpha_{3j_3} z_{j_3}, \dots, \sum_{j_n=1}^{n} \alpha_{nj_n} z_{j_n}\right) = \dots = \\
&= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_n=1}^{n} \alpha_{1j_1}\alpha_{2j_2} \dots \alpha_{nj_n} \det(z_{j_1}, \dots, z_{j_n}).
\end{aligned}$$

Nun ist aber $\det(z_{j_1}, \dots, z_{j_n}) = 0$, falls zwei der Indizes $j_1, \dots, j_n$ gleich sind. Wir müssen also nur über alle n-Tupel $(j_1, \dots, j_n)$ von paarweise verschiedenen Indizes aus $\{1, \dots, n\}$, sprich über alle Permutationen von $\{1, \dots, n\}$ summieren. Es folgt

$$1 = \sum_{\sigma \in S_n} \alpha_{1\sigma(1)}\alpha_{2\sigma(2)} \dots \alpha_{n\sigma(n)} \det(z_{\sigma(1)}, \dots, z_{\sigma(n)}).$$

Stellt man eine Permutation $\sigma \in S_n$ als Produkt von Transpositionen $\sigma = \sigma \circ \dots \circ \sigma_m$ dar, so ist $\det(z_{\sigma(1)}, \dots, z_{\sigma(n)}) = (-1)^m \det(z_1, \dots, z_n) = \operatorname{sign}(\sigma)\det(A)$, denn jede der m Zeilenvertauschungen ändert das Vorzeichen um -1. Es folgt

$$1 = \det(A) \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \alpha_{1\sigma(1)}\alpha_{2\sigma(2)} \dots \alpha_{n\sigma(n)}$$

und folglich muss $\det(A) \neq 0$ gelten.
b) Seien nun $z_1, \dots, z_n$ linear abhängig. Dann existieren $\lambda_1, \dots, \lambda_n \in K$ mit $\sum_{i=1}^{n} \lambda_i z_i = 0$, wobei nicht alle λ_i gleich Null sind. Sei $j \in \{1, \dots, n\}$ mit $\lambda_j \neq 0$. Dann folgt $z_j = -\lambda_j^{-1} \sum_{i \in I} \lambda_i z_i$, wobei $I := \{1, \dots, n\} \setminus \{j\}$. Damit folgt

$$\det(A) = \det(z_1, \dots, z_j, \dots, z_n) = -\frac{1}{\lambda_j} \sum_{i \in I} \lambda_i \det(z_1, \dots, z_i, \dots, z_n) = 0,$$

denn in $\det(z_1, \dots, z_i, \dots, z_n)$ (z_i steht an der j-ten Stelle) stimmt jeweils die j-te Zeile mit der i-ten überein, also ist diese Determinante gleich Null für alle $i \in I$.

2) Seien $a_1, \ldots, a_n$ die Spaltenvektoren von A. Dann gilt:

$$\begin{aligned}
&\exists x \in K^n \setminus \{0\}\ Ax = 0 \\
&\Leftrightarrow\ \exists (x_1, \ldots, x_n)^T \in K^n \setminus \{0\}\ \sum_{j=1}^{n} x_j a_j = 0 \\
&\Leftrightarrow\ a_1, \ldots, a_n \text{ sind linear abhängig} \\
&\Leftrightarrow\ \mathrm{Rang}(A) < n \\
&\Leftrightarrow\ A \text{ ist nicht invertierbar.}
\end{aligned}$$

Letzteres ist nach 1) äquivalent zu $\det(A) = 0$. □

Ohne Beweis geben wir noch den sogenannten *Determinantenmultiplikationssatz* an.

Satz VI.2.5. *Sei K ein Körper. Für alle $A, B \in M(n \times n, K)$ gilt* $\det(AB) = \det(A)\det(B)$.

Ist $A \in M(n \times n, K)$ und sind $i, j \in \{1, \ldots, n\}$, so bezeichnen wir mit A_{ij} diejenige $(n-1) \times (n-1)$-Matrix, die aus A durch Streichen der i-ten Zeile und der j-ten Spalte hervorgeht. Mit dieser Notation gilt dann der folgende Satz.

Satz VI.2.6 (Entwicklungssatz für Determinanten)**.** *Sei K ein Körper und sei $A = (a_{ij})_{i,j=1}^{n,n} \in M(n \times n, K)$. Dann gilt:*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \quad \forall i = 1, \ldots, n$$

(*Entwicklung nach der i-ten Zeile*),

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \quad \forall j = 1, \ldots, n$$

(*Entwicklung nach der j-ten Spalte*).

Dieser Satz erlaubt es also, die Berechnung einer $n \times n$-Determinante auf die Berechnung von (höchstens) n Determinanten des Formats $(n-1) \times (n-1)$ zurückzuführen. Die oben gefundene Formel für 3×3-Determinanten ist nichts anderes als die Entwicklung derselben nach der ersten Zeile. Auf einen Beweis des allgemeinen Entwicklungssatzes wollen wir hier aus Gründen der Zeit und Einfachheit verzichten.

Um den Rechenaufwand möglichst gering zu halten, empfiehlt es sich (wenn möglich) nach einer Zeile oder Spalte zu entwickeln, die besonders viele Nullen enthält.

Beispiel: Durch Entwicklung nach der zweiten Spalte erhält man

$$\begin{vmatrix} 2 & 0 & 3 & 1 \\ -1 & 1 & 4 & 1 \\ 2 & 0 & -3 & 2 \\ 1 & 0 & 4 & -1 \end{vmatrix} = -0 + \begin{vmatrix} 2 & 3 & 1 \\ 2 & -3 & 2 \\ 1 & 4 & -1 \end{vmatrix} - 0 + 0$$
$$= 2(3-8) - 3(-2-2) + 8 + 3 = -10 + 12 + 11 = 13.$$

Eine wichtige Folgerung aus dem Entwicklungssatz ist die folgende Formel zur Berechnung der Determinante einer Dreiecksmatrix: Diese ist gleich dem Produkt der Einträge auf der Hauptdiagonalen.

Korollar VI.2.7. *Sei K ein Körper. Für alle $n \in \mathbb{N}$ und alle oberen/unteren Dreiecksmatrizen $A = (a_{ij})_{i,j=1}^{n,n} \in M(n \times n, K)$ gilt*

$$\det(A) = \prod_{i=1}^{n} a_{ii}.$$

Beweis. 1) Wir betrachten zunächst den Fall oberer Dreiecksmatrizen und argumentieren mittels vollständiger Induktion nach n. Für $n = 1$ ist nichts weiter zeigen. Angenommen nun die Behauptung gilt für ein $n \in \mathbb{N}$ und es sei $A = (a_{ij})_{i,j=1}^{n+1,n+1}$ eine obere Dreiecksmatrix mit $n+1$ Zeilen und Spalten. Dann ist $a_{i1} = 0$ für alle $i = 2, \ldots, n+1$ und daher liefert eine Entwicklung nach der ersten Spalte $\det(A) = a_{11}\det(A_{11})$.
Aber A_{11} ist wieder eine obere Dreiecksmatrix vom Format $n \times n$ mit den Hauptdiagonaleinträgen $a_{22}, \ldots, a_{n+1n+1}$. Nach Induktionsvoraussetzung gilt also $\det(A_{11}) = a_{22} \ldots a_{n+1n+1}$ und es folgt $\det(A) = a_{11}a_{22} \ldots a_{n+1n+1}$.
2) Ist nun $A = (a_{ij})_{i,j=1}^{n,n}$ eine untere Dreiecksmatrix, so ist A^T eine obere Dreiecksmatrix. Somit folgt aus 1) $\det(A^T) = a_{11} \ldots a_{nn}$. Nach Satz VI.2.3 ist aber $\det(A) = \det(A^T)$, also $\det(A) = a_{11} \ldots a_{nn}$. □

Beispiel: Es ist

$$\begin{vmatrix} 3 & 2 & 6 \\ 0 & -1 & 8 \\ 0 & 0 & 5 \end{vmatrix} = 3 \cdot (-1) \cdot 5 = -15.$$

Mit Hilfe des Gaußschen Algorithmus kann man jede quadratische Matrix mittels elementarer Zeilenumformungen in eine obere Dreiecksmatrix überführen, für welche sich die Determinante leicht ausrechnen lässt. Allerdings ist dabei etwas Vorsicht geboten: Wir hatten oben schon festgestellt, dass Typ (II) Umformungen die Determinante unverändert lassen. Dagegen

ändert sich bei Typ (I) Umformungen (Zeilenvertauschungen) das Vorzeichen der Determinante und bei Typ (III) Umformungen muss man den entsprechenden Faktor aus der Determinante herausziehen.

Beispiel: Es gilt

$$\begin{vmatrix} 0 & 3 & -1 & 2 \\ 1 & 2 & 4 & 1 \\ 3 & 1 & 2 & 1 \\ 2 & 1 & 4 & 1 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 4 & 1 \\ 0 & 3 & -1 & 2 \\ 3 & 1 & 2 & 1 \\ 2 & 1 & 4 & 1 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 4 & 1 \\ 0 & 3 & -1 & 2 \\ 0 & -5 & -10 & -2 \\ 0 & -3 & -4 & -1 \end{vmatrix}$$

$$= - \begin{vmatrix} 1 & 2 & 4 & 1 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & -35/3 & 4/3 \\ 0 & 0 & -5 & 1 \end{vmatrix} = -\frac{1}{3} \begin{vmatrix} 1 & 2 & 4 & 1 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & -35 & 4 \\ 0 & 0 & -5 & 1 \end{vmatrix} = -\frac{1}{3} \begin{vmatrix} 1 & 2 & 4 & 1 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & -35 & 4 \\ 0 & 0 & 0 & 3/7 \end{vmatrix}$$

$$= -\frac{1}{3} \cdot 1 \cdot 3 \cdot (-35) \cdot \frac{3}{7} = 15$$

(machen Sie sich klar, welche Umformungen in den einzelnen Schritten verwendet wurden).

Ohne Beweis geben wir noch den folgenden Satz an.

Satz VI.2.8. *Sei K ein Körper und seien $A \in M(n \times n, K)$, $B \in M(m \times m, K)$, $C \in M(n \times m, K)$. Ferner bezeichne N die $m \times n$-Nullmatrix. Es sei M die durch*

$$M := \begin{pmatrix} A & C \\ N & B \end{pmatrix}$$

definierte $(n+m) \times (n+m)$-Matrix.
Dann gilt $\det(M) = \det(A)\det(B)$.

Oben hatten wir schon gezeigt, dass eine quadratische Matrix A genau dann invertierbar ist, wenn $\det(A) \neq 0$ gilt. Es ist auch möglich, die Inverse von A mit Hilfe von Determinanten zu berechnen, nämlich wie folgt.

Satz VI.2.9. *Sei K ein Körper und sei $A \in M(n \times n, K)$ mit* $\det(A) \neq 0$. *Dann ist A invertierbar und es gilt*

$$A^{-1} = \left(\frac{(-1)^{i+j} \det(A_{ji})}{\det(A)} \right)_{i,j=1}^{n,n}$$

(beachten Sie die Reihenfolge der Indizes).

Für die konkrete Berechnung einer inversen Matrix ist diese Formel allerdings nicht zu empfehlen, da die Determinantenberechnung sehr aufwendig ist. Die Methode der Inversenbestimmung mit dem Gaußschen Algorithmus

ist sehr viel effizienter. Wir verzichten daher auch auf einen Beweis dieses Satzes.

Zum Schluß geben wir noch, ebenfalls ohne Beweis, die sogenannte *Cramersche Regel*[2] an.

Satz VI.2.10 (Cramersche Regel)**.** *Sei K ein Körper und sei $A \in M(n \times n, K)$ mit $\det(A) \neq 0$. Sei $b \in K^n$ und für alle $i = 1, \ldots, n$ sei $A^{i,b}$ diejenige Matrix, die aus A ensteht, indem man die i-te Spalte durch b ersetzt. Sei*

$$x_i := \frac{\det(A^{i,b})}{\det(A)} \quad \forall i = 1, \ldots, n$$

und

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

Dann gilt $Ax = b$.

Auch diese Formel ist jedoch für die konkrete Berechnung der Lösung eines linearen Gleichungssystems eher ungeeignet (wegen des hohen Aufwands zur Determinantenberechnung). Auch hier ist der Gauß-Algorithmus wesentlich effizienter.

[2]Benannt nach Gabriel Cramer (1704–1752), einem Mathematiker aus Genf.

VII Skalarprodukte

In diesem Kapitel betrachten wir Vektorräume mit einer zusätzlichen Verknüpfung, einem sogenannten Skalarprodukt. Auch den dadurch induzierten Abstandsbegriff und das Konzept der Orthogonalität wollen wir kurz besprechen.

VII.1 Skalarprodukte und ihre Eigenschaften

Wir beginnen mit der Definition der Bilinearität.

Definition VII.1.1. Es sei K ein Körper und V ein Vektorraum über K. Eine Abbildung $\varphi : V \times V \to K$ heißt *Bilinearform*, falls folgendes gilt:

(i) $\varphi(\lambda v, w) = \lambda\varphi(v, w) = \varphi(v, \lambda w)$ für alle $v, w \in V$ und alle $\lambda \in K$.

(ii) $\varphi(v_1 + v_2, w) = \varphi(v_1, w) + \varphi(v_2, w)$ für alle $v_1, v_2, w \in V$.

(iii) $\varphi(v, w_1 + w_2) = \varphi(v, w_1) + \varphi(v, w_2)$ für alle $w_1, w_2, v \in V$.

Bilinearität von φ bedeutet also nichts anderes, als dass φ in jeder der beiden Variablen linear ist, wenn man die jeweils andere Variable festhält.

Beispiel: Die Abbildung $\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ mit

$$\varphi\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) := \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1y_2 - x_2y_1$$

ist eine Bilinearform.

Skalarprodukte sind nur auf reellen oder komplexen Vektorräumen definiert. Wir beginnen mit dem reellen Fall.

Definition VII.1.2. Sei V ein Vektorraum über $\mathbb{R}$. Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ heißt *Skalarprodukt* auf V, falls folgendes gilt:

(a) $\langle \cdot, \cdot \rangle$ ist eine Bilinearform.

(b) $\langle \cdot, \cdot \rangle$ ist *symmetrisch*, d. h. $\langle v, w \rangle = \langle w, v \rangle$ für alle $v, w \in V$.

(c) $\langle \cdot, \cdot \rangle$ ist *positiv definit*, d. h. $\langle v, v \rangle > 0$ für alle $v \in V \setminus \{0\}$.

Beispiele:
1) Für $x = (x_1 \ \dots \ x_n)^T, y = (y_1 \ \dots \ y_n)^T \in \mathbb{R}^n$ setzen wir

$$\langle x, y\rangle := \sum_{i=1}^{n} x_i y_i.$$

Dann ist $\langle \cdot, \cdot \rangle$ eine symmetrische Bilinearform auf $\mathbb{R}^n$, wie man leicht nachrechnet (Übung). Ist ferner $x \neq 0$, so ist $x_j \neq 0$ und somit $x_j^2 > 0$ für mindestens ein $j \in \{1, \dots, n\}$, während $x_i^2 \geq 0$ für alle $i \in \{1, \dots, n\} \setminus \{j\}$. Das impliziert $\langle x, x\rangle = \sum_{i=1}^{n} x_i^2 > 0$.
Also ist $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf $\mathbb{R}^n$, das sogenannte *euklidische* (oder *kanonische*) Skalarprodukt.
2) Es sei $[0, 1]$ das Intervall $\{x \in \mathbb{R} : 0 \leq x \leq 1\}$ und V der Vektorraum aller Polynomfunktionen $f : [0, 1] \to \mathbb{R}$ vom Grad≤ 2 mit $f(0) = 0$ (V ist ein Unterraum des Raumes aller Abbildungen von $[0, 1]$ nach $\mathbb{R}$, wie man leicht nachweist). Für alle $f, g \in V$ existieren eindeutig bestimmte $a, b, c, d \in \mathbb{R}$ mit $f(x) = ax^2 + bx$ und $g(x) = cx^2 + dx$ für alle $x \in [0, 1]$. Wir setzen

$$\langle f, g\rangle := \frac{1}{5}ac + \frac{1}{4}(ad + bc) + \frac{1}{3}bd.$$

Das mag zunächst etwas seltsam aussehen, wer aber bereits ein wenig mit der Integralrechnung vertraut ist, der kann leicht nachrechnen, dass

$$\langle f, g\rangle = \int_0^1 f(x)g(x)\,\mathrm{d}x$$

gilt. In dieser Version wirkt die Definition sicherlich etwas natürlicher.
Nun kann man entweder direkt anhand der Definition oder anhand der Integraldarstellung nachweisen, dass $\langle \cdot, \cdot \rangle$ eine symmetrische Bilinearform auf V ist (Übung). Ferner gilt

$$\begin{aligned}\langle f, f\rangle &= \frac{1}{5}a^2 + \frac{1}{2}ab + \frac{1}{3}b^2 = \frac{1}{5}(a + 5b/4)^2 + \frac{1}{3}b^2 - \frac{5}{16}b^2 \\ &= \frac{1}{5}(a + 5b/4)^2 + \frac{1}{48}b^2.\end{aligned}$$

Ist $f \neq 0$, so folgt $b \neq 0$ oder $a \neq 0$. Im ersten Fall ist $\langle f, f\rangle \geq b^2/48 > 0$. Ist dagegen $b = 0$, so ist $\langle f, f\rangle = a^2/5 > 0$. Also ist $\langle \cdot, \cdot \rangle$ auch positiv definit und damit ein Skalarprodukt auf V.

Bevor wie auch Skalarprodukte auf komplexen Vektorräumen definieren können, müssen wir noch den Begriff einer Sesquilinearform einführen.

Definition VII.1.3. Es sei V ein Vektorraum über $\mathbb{C}$. Eine Abbildung $\varphi : V \times V \to \mathbb{C}$ heißt *Sesquilinearform*, falls folgendes gilt:

(i) $\varphi(\lambda v, w) = \lambda\varphi(v, w)$ und $\varphi(v, \lambda w) = \bar{\lambda}\varphi(v, w)$ für alle $v, w \in V$ und alle $\lambda \in \mathbb{C}$.

(ii) $\varphi(v_1 + v_2, w) = \varphi(v_1, w) + \varphi(v_2, w)$ für alle $v_1, v_2, w \in V$.

(iii) $\varphi(v, w_1 + w_2) = \varphi(v, w_1) + \varphi(v, w_2)$ für alle $w_1, w_2, v \in V$.

Beispiel: Die Abbildung $\varphi : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ mit

$$\varphi\left(\begin{pmatrix} z_1 \\ u_1 \end{pmatrix}, \begin{pmatrix} z_2 \\ u_2 \end{pmatrix}\right) := \begin{vmatrix} z_1 & \bar{z_2} \\ u_1 & \bar{u_2} \end{vmatrix} = z_1\bar{u}_2 - \bar{z_2}u_1$$

ist eine Sesquilinearform.

Nun können wir auch Skalarprodukte auf Vektorräumen über $\mathbb{C}$ definieren.

Definition VII.1.4. Sei V ein Vektorraum über $\mathbb{C}$. Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ heißt *Skalarprodukt* auf V, falls folgendes gilt:

(a) $\langle \cdot, \cdot \rangle$ ist eine Sesquilinearform.

(b) $\langle \cdot, \cdot \rangle$ ist *hermitesch*[1], d. h. $\langle v, w \rangle = \overline{\langle w, v \rangle}$ für alle $v, w \in V$.

(c) $\langle \cdot, \cdot \rangle$ ist *positiv definit*, d. h. $\langle v, v \rangle > 0$ für alle $v \in V \setminus \{0\}$.

Beispiel: Für $z = (z_1 \ \dots \ z_n)^T, u = (u_1 \ \dots \ u_n)^T \in \mathbb{C}^n$ setzen wir

$$\langle z, u \rangle := \sum_{i=1}^{n} z_i\bar{u}_i.$$

Dann ist $\langle \cdot, \cdot \rangle$ eine hermitesche Sesquilinearform (Übung) und für alle $z \in \mathbb{C}^n \setminus \{0\}$ gilt $\langle z, z \rangle = \sum_{i=1}^{n} z_i\bar{z_i} = \sum_{i=1}^{n} |z_i|^2 > 0$. Also ist $\langle \cdot; \cdot \rangle$ ein Skalarprodukt, das *kanonische Skalarprodukt* auf $\mathbb{C}^n$.

Als Nächstes führen wir noch den wichtigen Begriff einer Norm ein.

Definition VII.1.5. Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $\|\cdot\| : V \to \mathbb{R}_0^+$ eine Abbildung ($\mathbb{R}_0^+ := \{x \in \mathbb{R} : x \geq 0\}$). Für alle $v, w \in V$ und alle $\lambda \in \mathbb{K}$ gelte:

(i) $\|\lambda v\| = |\lambda|\|v\|$

(ii) $\|v\| = 0 \quad \Leftrightarrow \quad v = 0$

[1] Nach dem französischen Mathematiker Charles Hermite (1822–1901).

(iii) $\|v+w\| \leq \|v\| + \|w\|$ (Dreiecksungleichung)

Dann heißt $\|\cdot\|$ eine *Norm* auf V.

Die Anschauung dabei ist, dass die Zahl $\|v\|$ die "Länge" des Vektors v beschreibt.

Beispiel: Für $x = (x_1 \ \dots \ x_n)^T \in \mathbb{R}^n$ sei

$$\|x\|_1 := \sum_{i=1}^{n} |x_i|.$$

Dann ist $\|\cdot\|_1$ eine Norm auf $\mathbb{R}^n$, wie Sie zur Übung leicht selbst nachweisen können.

Wir wollen nun zeigen, dass jedes Skalarprodukt auch eine Norm induziert, nämlich wie folgt.

Definition VII.1.6. Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V. Setze $\|v\| := \sqrt{\langle v, v \rangle}$ für alle $v \in V$.

Wegen $\langle v, v \rangle \geq 0$ ist $\|\cdot\|$ in jedem Fall wohldefiniert. Wir müssen aber natürlich noch nachweisen, dass es sich tatsächlich um eine Norm handelt. Zunächst gilt wegen der positiven Definitheit von $\langle \cdot, \cdot \rangle$

$$\|v\| = 0 \quad \Leftrightarrow \quad \langle v, v \rangle = 0 \quad \Leftrightarrow \quad v = 0.$$

Ferner gilt für alle Skalare λ und alle $v \in V$

$$\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{\lambda \bar{\lambda} \langle v, v \rangle} = \sqrt{|\lambda|^2 \langle v, v \rangle} = |\lambda| \sqrt{\langle v, v \rangle} = |\lambda| \|v\|$$

(im reellen Fall muss man sich die komplexe Konjugation einfach wegdenken).

Damit bleibt nur noch die Dreiecksungleichung zu zeigen, was allerdings nicht ganz einfach ist. Wir zeigen dazu zunächst eine andere Ungleichung, nämlich die sogenannte *Cauchy-Schwarz-Ungleichung*.[2]

Satz VII.1.7 (Cauchy-Schwarz-Ungleichung)**.** *Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V. Sei $\|v\| := \sqrt{\langle v, v \rangle}$ für alle $v \in V$. Dann gilt*

$$|\langle v, w \rangle| \leq \|v\| \|w\| \quad \forall v, w \in V.$$

[2]Benannt ist sie nach dem französischen Mathematiker Augustin-Louis Cauchy (1789–1857) und dem deutschen Mathematiker Hermann Amandus Schwarz (1843–1921).

Beweis. Für $w = 0$ ist die Aussage klar. Sei nun also $w \neq 0$. Wir setzen $\lambda := -\langle v, w\rangle / \|w\|^2$. Dann gilt

$$\begin{aligned}
&0 \leq \|v + \lambda w\|^2 = \langle v + \lambda w, v + \lambda w\rangle = \langle v, v\rangle + \langle v, \lambda w\rangle + \langle \lambda w, v\rangle + \langle \lambda w, \lambda w\rangle \\
&= \|v\|^2 + \bar{\lambda}\langle v, w\rangle + \lambda\overline{\langle v, w\rangle} + \lambda\bar{\lambda}\|w\|^2 = \|v\|^2 + \bar{\lambda}\langle v, w\rangle + \lambda\overline{\langle v, w\rangle} + |\lambda|^2\|w\|^2 \\
&= \|v\|^2 - 2\frac{\overline{\langle v, w\rangle}\langle v, w\rangle}{\|w\|^2} + \frac{|\langle v, w\rangle|^2}{\|w\|^4}\|w\|^2 = \|v\|^2 - \frac{|\langle v, w\rangle|^2}{\|w\|^2}.
\end{aligned}$$

Es folgt $|\langle v, w\rangle|^2 \leq \|v\|^2\|w\|^2$ und somit $|\langle v, w\rangle| \leq \|v\|\|w\|$ (im reellen Fall kann man sich wieder einfach die komplexe Konjugation wegdenken). □

Nun können wir auch beweisen, dass die Vorschrift $\|v\| := \sqrt{\langle v, v\rangle}$ wirklich eine Norm definiert.

Satz VII.1.8. *Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V. Sei $\|v\| := \sqrt{\langle v, v\rangle}$ für alle $v \in V$. Dann ist $\|\cdot\|$ eine Norm auf V.*

Beweis. Nach unseren obigen Überlegungen müssen wir nur noch die Dreiecksungleichung beweisen. Wir führen das wieder nur für den Fall $\mathbb{K} = \mathbb{C}$ aus, der Fall $\mathbb{K} = \mathbb{R}$ ist völlig analog.
Seien also $v, w \in V$ beliebig. Es gilt

$$\begin{aligned}
&\|v + w\|^2 = \langle v + w, v + w\rangle = \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\
&= \|v\|^2 + \langle v, w\rangle + \overline{\langle v, w\rangle} + \|w\|^2 = \|v\|^2 + 2\mathrm{Re}(\langle v, w\rangle) + \|w\|^2 \\
&\leq \|v\|^2 + 2|\langle v, w\rangle| + \|w\|^2,
\end{aligned}$$

wobei wir die Tatsachen $z + \bar{z} = 2\mathrm{Re}(z)$ und $\mathrm{Re}(z) \leq |z|$ für alle $z \in \mathbb{C}$ verwendet haben, die Sie zur Übung leicht selbst beweisen können.
Nach der Cauchy-Schwarz-Ungleichung ist $|\langle v, w\rangle| \leq \|v\|\|w\|$ und somit folgt

$$\|v + w\|^2 \leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2,$$

also $\|v + w\| \leq \|v\| + \|w\|$. □

Jedes Skalarprodukt induziert also eine Norm. Jedoch stammt umgekehrt nicht jede Norm von einem Skalarprodukt. Zum Beispiel kann man zeigen, dass die obige Norm $\|\cdot\|_1$ auf $\mathbb{R}^n$ nicht durch ein Skalarprodukt induziert wird.

Wichtigstes Beispiel ist wieder das euklidische Skalarprodukt auf dem $\mathbb{R}^n$. Hier ist die induzierte Norm die *euklidische Norm*

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

(wobei $x_1, \ldots, x_n$ natürlich die Koordinaten des Vektors x sind).

Stellt man Vektoren im $\mathbb{R}^2$ oder $\mathbb{R}^3$ wieder als Pfeile dar, so ergibt sich aus dem Satz des Pythagoras, dass $\|x\|_2$ gerade gleich der Länge des Pfeils ist (im Sinne der üblichen, euklidischen Geometrie).

Jede Norm $\|\cdot\|$ auf einem (reellen oder komplexen) Vektorraum V induziert auch einen Abstandsbegriff, nämlich wie folgt: Für alle $v, w \in V$ setzen wir $d(v, w) := \|v - w\|$ und nennen dies den *Abstand* von v zu w. Es gilt dann

(a) $d(v, w) \geq 0$ für alle $v, w \in V$,

(b) $d(v, w) = 0 \quad \Leftrightarrow \quad v = w$,

(c) $d(v, w) = d(w, v)$ für alle $v, w \in V$,

(d) $d(v, w) \leq d(v, u) + d(u, w)$ für alle $v, w, u \in V$.

Das können Sie zur Übung selbst beweisen. Eigenschaft (d) nennt man wieder Dreiecksungleichung. Insgesamnt fasst man die Aussagen (a)–(d) auch folgendermaßen zusammen: d ist eine *Metrik* auf V. Im Falle des $\mathbb{R}^2$ oder $\mathbb{R}^3$ mit der euklidischen Norm erhält man auf diese Weise gerade den üblichen, euklidischen Abstand.

Weiterhin eröffnen Skalarprodukte auch die Möglcihkeit, Winkel zwischen zwei Vektoren zu definieren.

Definition VII.1.9. Sei V ein *reeller* Vektorraum mit einem Skalarprodukt $\langle \cdot, \cdot \rangle$. Seien $v, w \in V \setminus \{0\}$. Wir setzen

$$\angle(v, w) := \arccos\left(\frac{\langle v, w \rangle}{\|v\|\|w\|}\right)$$

und nennen dies den *Winkel* zwischen v und w. Hierbei bezeichnet arccos die Arcus-Kosinusfunktion, also die Umkehrfunktion des Kosinus auf der Menge $\{\varphi \in \mathbb{R} : 0 \leq \varphi \leq \pi\}$.

Mit der obigen Definition gilt dann also

$$\langle v, w \rangle = \|v\|\|w\| \cos(\angle(v, w)).$$

Zum Schluß dieses Abschnitts zeigen wir noch das folgende wichtige Lemma.

Lemma VII.1.10. *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $A = (a_{ij})_{i,j=1}^{m,n} \in M(m \times n, \mathbb{K})$. Wir bezeichnen das kanonische Skalarprodukt sowohl auf dem $\mathbb{K}^n$ als auch auf dem $\mathbb{K}^m$ mit $\langle \cdot, \cdot \rangle$.*
1) Im Fall $\mathbb{K} = \mathbb{R}$ gilt: $\langle Ax, y \rangle = \langle x, A^T y \rangle$.
2) Im Fall $\mathbb{K} = \mathbb{C}$ gilt: $\langle Ax, y \rangle = \langle x, \bar{A}^T y \rangle$ (hierbei bezeichnet $\bar{A}$ die Matrix $(\overline{a_{ij}})_{i,j=1}^{m,n}$).

Beweis. 1) Sei $x = (x_1 \ \dots \ x_n)^T$ und $y = (y_1 \ \dots \ y_m)^T$. Dann gilt

$$\langle Ax, y\rangle = \sum_{i=1}^{m} (Ax)_i y_i = \sum_{i=1}^{m} \left(y_i \sum_{j=1}^{n} a_{ij} x_j \right) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} y_i x_j$$
$$= \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} y_i x_j = \sum_{j=1}^{n} \left(x_j \sum_{i=1}^{m} a_{ij} y_i \right) = \sum_{j=1}^{n} x_j (A^T y)_j = \langle x, A^T y\rangle.$$

2) beweist man analog. □

VII.2 Orthogonalität

In diesem Abschnitt führen wir das wichtige Konzept der Orthogonalität ein. Die Definition lautet wie folgt.

Definition VII.2.1. Sei V ein reeller oder komplexer Vektorraum und sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V. Seien $v, w \in V$. Man sagt, dass v *senkrecht* auf w steht (oder auch: v ist *orthogonal* zu w), falls $\langle v, w\rangle = 0$ gilt. In Zeichen: $v \perp w$.

Für einen reellen Vektorraum mit Skalarprodukt gilt mit unserer Definition des Zwischenwinkels also $v \perp w \Leftrightarrow \angle(v, w) = \pi/2$ (was genau 90° entspricht).

Mit diesem abstrakten Begriff von Orthogonalität erhalten wir auch eine abstrakte Version des *Satzes von Pythagoras.*

Satz VII.2.2 (Satz des Pythagoras)**.** *Sei V ein reeller oder komplexer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Seien $v, w \in V$ mit $v \perp w$. Dann gilt $\|v + w\|^2 = \|v\|^2 + \|w\|^2$.*

Beweis. Es gilt

$$\|v + w\|^2 = \|v\|^2 + 2\mathrm{Re}(\langle v, w\rangle) + \|w\|^2$$

(siehe den Beweis von Satz VII.1.8). Nach Voraussetzung ist aber $\langle v, w\rangle = 0$, also folgt die Behauptung. □

Als Nächstes definieren wir gleich noch den Begriff eines Orthonormalsystems.

Definition VII.2.3. Sei V ein reeller oder komplexer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Seien $v_1, \dots, v_n \in V$. Wir sagen, dass $(v_1, \dots, v_n)$ ein *Orthonormalsystem* (kurz ONS) bildet, falls $v_i \perp v_j$ für alle $i, j \in \{1, \dots, n\}$ mit $i \neq j$ und $\|v_i\| = 1$ für alle $i = 1, \dots, n$ gilt.

$(v_1, \ldots, v_n)$ ist also ein ONS, falls die Vektoren paarweise senkrecht aufeinander stehen und alle die Norm 1 haben. Das kann man auch äquivalent als $\langle v_i, v_j \rangle = \delta_{ij}$ für alle $i, j = 1, \ldots, n$ ausdrücken.

Beispiele:
1) $(e_1, \ldots, e_n)$ bildet ein ONS im $\mathbb{R}^n$ bzgl. des euklidischen Skalarprodukts.
2) Die Vektoren

$$v_1 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \; v_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \; v_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

bilden ein ONS im $\mathbb{R}^3$ bzgl. des euklidischen Skalarprodukts.

Eine einfache aber wichtige Beobachtung besteht nun darin, dass jedes ONS automatisch linear unabhängig ist.

Lemma VII.2.4. *Sei V ein reeller oder komplexer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und sei $(v_1, \ldots, v_n)$ ein ONS in V. Dann ist $(v_1, \ldots, v_n)$ linear unabhängig.*

Beweis. Seien $\lambda_1, \ldots, \lambda_n$ Skalare mit $\sum_{i=1}^n \lambda_i v_i = 0$. Dann gilt für alle $j = 1, \ldots, n$:

$$0 = \left\langle \sum_{i=1}^n \lambda_i v_i, v_j \right\rangle = \sum_{i=1}^n \lambda_i \langle v_i, v_j \rangle = \sum_{i=1}^n \lambda_i \delta_{ij} = \lambda_j.$$

□

Daraus ergibt sich sofort folgendes Korollar.

Korollar VII.2.5. *Sei V ein reeller oder komplexer Vektorraum mit Skalarprodukt und sei $\dim(V) = n$. Ist $(v_1, \ldots, v_n)$ ein ONS in V, so ist $(v_1, \ldots, v_n)$ eine geordnete Basis von V.*

In diesem Fall spricht man auch von einer *Orthonormalbasis* (kurz ONB). Zum Beispiel ist $(e_1, \ldots, e_n)$ eine ONB des $\mathbb{R}^n$ bzgl. des euklidischen Skalarprodukts. Bezüglich einer ONB lassen sich die Koordinaten eines Vektors leicht berechnen, nämlich wie folgt.

Lemma VII.2.6. *Sei V ein reeller oder komplexer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine ONB von V. Dann ist*

$$\mathcal{K}_{\mathcal{B}}(v) = \begin{pmatrix} \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_n \rangle \end{pmatrix}$$

für alle $v \in V$.

Beweis. Sei $\mathcal{K}_{\mathcal{B}}(v) = (x_1 \ \dots \ x_n)^T$. Dann gilt $v = \sum_{i=1}^n x_i v_i$. Es folgt für alle $j \in \{1, \dots, n\}$:

$$\langle v, v_j \rangle = \left\langle \sum_{i=1}^n x_i v_i, v_j \right\rangle = \sum_{i=1}^n x_i \langle v_i, v_j \rangle = \sum_{i=1}^n x_i \delta_{ij} = x_j.$$

□

Noch wissen wir allerdings nicht, ob jeder endlich-dimensionale Vektorraum mit Skalarprodukt wirklich eine ONB besitzt. Dies folgt jedoch aus dem folgenden Satz.

Satz VII.2.7. *Sei V ein reeller oder komplexer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und seien $v_1, \dots, v_n \in V$ linear unabhängig. Dann existieren Vektoren $w_1, \dots, w_n \in V$ mit den folgenden beiden Eigenschaften:*
1) $(w_1, \dots, w_n)$ ist ein ONS.
2) Für alle $k \in \{1, \dots, n\}$ gilt $\operatorname{span}\{v_1, \dots, v_k\} = \operatorname{span}\{w_1, \dots, w_k\}$.

Beweis. Wir konstruieren die Vektoren $w_1, \dots, w_n$ induktiv. Zunächst setzen wir $w_1 := v_1/\|v_1\|$ (das ist möglich, da wegen der linearen Unabhängigkeit $v_1 \neq 0$ gilt). Dann gilt natürlich $\|w_1\| = 1$ und $\operatorname{span}\{w_1\} = \operatorname{span}\{v_1\}$.
Nun nehmen wir an, dass fúr ein $m \in \{1, \dots, n-1\}$ die Vektoren $w_1, \dots, w_m$ bereits wie gewünscht konstruiert sind, d. h. $(w_1, \dots, w_m)$ ist ein ONS und $\operatorname{span}\{w_1, \dots, w_k\} = \operatorname{span}\{v_1, \dots, v_k\}$ für alle $k = 1, \dots, m$.
Wäre $v_{m+1} = \sum_{k=1}^m \langle v_{m+1}, w_k \rangle w_k$, so wäre $v_{m+1} \in \operatorname{span}\{w_1, \dots, w_m\} = \operatorname{span}\{v_1, \dots, v_m\}$, was der linearen Unabhängigkeit von $(v_1, \dots, v_n)$ widerspricht. Also muss $v_{m+1} \neq \sum_{k=1}^m \langle v_{m+1}, w_k \rangle w_k$ gelten und somit können wir

$$w_{m+1} := \frac{v_{m+1} - \sum_{k=1}^m \langle v_{m+1}, w_k \rangle w_k}{\|v_{m+1} - \sum_{k=1}^m \langle v_{m+1}, w_k \rangle w_k\|} \tag{VII.1}$$

setzen. Dann ist natürlich $\|w_{m+1}\| = 1$ und bezeichnen wir den Nenner in (VII.1) kurz mit λ, so gilt für alle $i = 1, \dots, m$

$$\begin{aligned}
\langle w_{m+1}, w_i \rangle &= \frac{1}{\lambda} \langle v_{m+1}, w_i \rangle - \frac{1}{\lambda} \left\langle \sum_{k=1}^m \langle v_{m+1}, w_k \rangle w_k, w_i \right\rangle \\
&= \frac{1}{\lambda} \langle v_{m+1}, w_i \rangle - \frac{1}{\lambda} \sum_{k=1}^m \langle v_{m+1}, w_k \rangle \langle w_k, w_i \rangle \\
&= \frac{1}{\lambda} \langle v_{m+1}, w_i \rangle - \frac{1}{\lambda} \sum_{k=1}^m \langle v_{m+1}, w_k \rangle \delta_{ik} \\
&= \frac{1}{\lambda} \langle v_{m+1}, w_i \rangle - \frac{1}{\lambda} \langle v_{m+1}, w_i \rangle = 0.
\end{aligned}$$

Also ist $(w_1, \dots, w_{m+1})$ ein ONS.

Ferner gilt $\sum_{k=1}^{m}\langle v_{m+1}, w_k\rangle w_k \in \operatorname{span}\{w_1,\dots,w_m\} = \operatorname{span}\{v_1,\dots,v_m\}$ und folglich $w_{m+1} \in \operatorname{span}\{v_1,\dots,v_m,v_{m+1}\}$. Es folgt $\operatorname{span}\{w_1,\dots,w_{m+1}\} \subseteq \operatorname{span}\{v_1,\dots,v_{m+1}\}$.
Umgekehrt ist $v_{m+1} = \lambda w_{m+1} + \sum_{k=1}^{m}\langle v_{m+1}, w_k\rangle w_k \in \operatorname{span}\{w_1,\dots,w_{m+1}\}$.
Damit folgt auch $\operatorname{span}\{v_1,\dots,v_{m+1}\} \subseteq \operatorname{span}\{w_1,\dots,w_{m+1}\}$. □

Das im obigen Beweis angewendete Verfahren zur Konstruktion der w_k nennt man auch das *Gram-Schmidt-Orthonormalisierungsverfahren.*[3]

Als Nächstes führen wir noch orthogonale Komplemente ein.

Definition VII.2.8. Sei V ein reeller oder komplexer Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$ und sei $A \subseteq V$. Dann heißt

$$A^{\perp} := \{v \in V : \langle v, a\rangle = 0 \ \forall a \in A\}$$

das *orthogonale Komplement* von A.

$A^{\perp}$ besteht also aus all jenen Vektoren, die auf der gesamten Menge A senkrecht stehen. Als Übung können Sie folgende Aussagen beweisen:
1) $A^{\perp}$ ist ein Unterraum von V.
2) Es gilt $A^{\perp} = (\operatorname{span}(A))^{\perp}$.

Wichtig ist der folgende Zerlegungssatz.

Satz VII.2.9. *Sei V ein reeller oder komplexer Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$ und sei $U \subseteq V$ ein Unterraum. Es gelte $\dim(V) < \infty$. Dann ist $V = U \oplus U^{\perp}$. Insbesondere ist $\dim(V) = \dim(U) + \dim(U^{\perp})$.*

Beweis. Die Aussage ist klar für $U = \{0\}$ (denn dann ist $U^{\perp} = V$). Sei also $U \neq \{0\}$ und sei $(u_1,\dots,u_k)$ eine geordnete ONB von U.
Ist $v \in V$, so setzen wir $u := \sum_{i=1}^{k}\langle v, u_i\rangle u_i$. Dann ist $u \in U$ und für alle $j = 1,\dots,k$ gilt

$$\begin{aligned}\langle v-u, u_j\rangle &= \langle v, u_j\rangle - \sum_{i=1}^{k}\langle v, u_i\rangle\langle u_i, u_j\rangle = \langle v, u_j\rangle - \sum_{i=1}^{k}\langle v, u_i\rangle\delta_{ij}\\ &= \langle v, u_j\rangle - \langle v, u_j\rangle = 0.\end{aligned}$$

Also ist $v - u \in \{u_1,\dots,u_k\}^{\perp} = (\operatorname{span}\{u_1,\dots,u_k\})^{\perp} = U^{\perp}$ und es folgt $v = u + v - u \in U + U^{\perp}$.
Also ist $V = U + U^{\perp}$.
Ist ferner $v \in U \cap U^{\perp}$, so folgt $\langle v, v\rangle = 0$ und somit $v = 0$, also ist $U \cap U^{\perp} = \{0\}$ und es gilt $V = U \oplus U^{\perp}$.
Wegen Lemma IV.3.18 folgt daraus $\dim(V) = \dim(U) + \dim(U^{\perp})$. □

[3] Benannt nach dem dänischen Mathematiker Jørgen Pedersen Gram (1850–1916) und dem deutschen Mathematiker Erhard Schmidt (1876–1959).

VIII Eigenwerttheorie

In diesem Kapitel befassen wir uns mit sogenannten Eigenwertproblemen. Diese haben vielfältige Anwendungen z. B. in der Physik (etwa bei der Bestimmung der Eigenfrequenzen eines schwingungsfähigen mechanischen Systems), aber auch in anderen Bereichen (beispielsweise beim PageRank-Algorithmus von Google, siehe Abschnitt VIII.3).

VIII.1 Eigenwerte und Eigenvektoren

Die Eigenwerte einer Matrix sind wie folgt definiert.

Definition VIII.1.1. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $A \in M(n \times n, \mathbb{K})$. Eine Zahl $\lambda \in \mathbb{K}$ heißt ein *Eigenwert* von A, falls es ein $x \in \mathbb{K}^n \setminus \{0\}$ mit $Ax = \lambda x$ gibt. In diesem Fall heißt x ein *Eigenvektor* von A zum Eigenwert λ.

Es geht bei Eigenwertproblemen also darum, einen von Null verschiedenen Vektor zu finden, der von der Matrix A wieder auf ein Vielfaches von sich selbst abgebildet wird.

Beispiel:
Seien

$$A := \begin{pmatrix} 2 & 1 \\ 4 & -1 \end{pmatrix} \quad \text{und} \quad x := \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Dann gilt

$$Ax = \begin{pmatrix} 2 & 1 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3x.$$

Also ist 3 ein Eigenwert von A und x ist ein zugehöriger Eigenvektor.

Zur systematischen Bestimmung sämtlicher Eigenwerte einer Matrix verwendet man in der Regel das sogenannte charakteristische Polynom.

Definition VIII.1.2. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $A \in M(n \times n, \mathbb{K})$. Setze

$$\chi_A(\lambda) := \det(A - \lambda E_n) \quad \forall \lambda \in \mathbb{K}.$$

Die Funktion χ_A nennt man das *charakteristische Polynom* von A.

Um die Bezeichnung "charakteristisches Polynom" zu rechtfertigen, muss man natürlich noch zeigen, dass es sich bei χ_A wirklich um ein Polynom handelt. Das besagt Teil 1) des folgenden Lemmas. Der zweite, entscheidende Teil besagt, dass die Nullstellen von χ_A genau die Eigenwerte von A sind.

Lemma VIII.1.3. *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $A = (a_{ij})_{i,j=1}^{n,n} \in M(n \times n, \mathbb{K})$. Dann gilt:*
1) χ_A ist ein Polynom n-ten Grades.
2) Eine Zahl $\lambda \in \mathbb{K}$ ist ein Eigenwert von A genau dann, wenn $\chi_A(\lambda) = 0$ gilt.

Beweis. 1) Für alle $\sigma \in S_n$ und alle $\lambda \in \mathbb{K}$ sei

$$p_\sigma(\lambda) := \operatorname{sign}(\sigma) \prod_{i=1}^{n} (a_{i\sigma(i)} - \lambda \delta_{i\sigma(i)}).$$

Dann gilt

$$\chi_A(\lambda) = \det(A - \lambda E_n) = \sum_{\sigma \in S_n} p_\sigma(\lambda).$$

p_{id} ist als Produkt von n Polynomen ersten Grades ein Polynom vom Grad n. Für alle $\sigma \in S_n \setminus \{\mathrm{id}\}$ existiert mindestens ein $i \in \{1, \ldots, n\}$ mit $\delta_{i\sigma(i)} = 0$, so dass p_σ ein Polynom vom Grad$\leq n-1$ ist.
Insgesamt ist also χ_A als Summe der p_σ ein Polynom vom Grad n.
2) Sei $\lambda \in \mathbb{K}$. Die Gleichung $Ax = \lambda x$ ist äquivalent zu $(A - \lambda E_n)x = 0$. Nach Satz VI.2.4 besitzt diese eine nichttriviale Lösung (d. h. λ ist ein Eigenwert von A) genau dann, wenn $\chi_A(\lambda) = \det(A - \lambda E_n) = 0$ gilt. □

Aus diesem Lemma folgt insbesondere, dass eine $n \times n$-Matrix A höchstens n verschiedene Eigenwerte besitzen kann, denn ein Polynom n-ten Grades hat höchstens n paarweise verschiedene Nullstellen.

Beispiel: Sei wieder

$$A := \begin{pmatrix} 2 & 1 \\ 4 & -1 \end{pmatrix}.$$

Wir hatten oben schon gesehen, dass 3 ein Eigenwert von A ist. Das charakteristische Polynom von A ist

$$\chi_A(\lambda) = \begin{vmatrix} 2-\lambda & 1 \\ 4 & -1-\lambda \end{vmatrix} = (2-\lambda)(-1-\lambda) - 4$$
$$= -2 - 2\lambda + \lambda + \lambda^2 - 4 = \lambda^2 - \lambda - 6.$$

Dieses quadratische Polynom hat die beiden Nullstellen

$$\lambda_{1/2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 6} = \frac{1}{2} \pm \sqrt{\frac{25}{4}} = \frac{1}{2} \pm \frac{5}{2},$$

also $\lambda_1 = 3$ und $\lambda_2 = -2$. Dies sind die Eigenwerte von A.

Als Nächstes definieren wir noch den zu einem Eigenwert gehörigen Eigenraum.

Definition VIII.1.4. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $A \in M(n \times n, \mathbb{K})$. Ist $\lambda \in \mathbb{K}$ ein Eigenwert von A, so setzen wir

$$E_A(\lambda) := \{x \in \mathbb{K}^n : Ax = \lambda x\} = \{x \in \mathbb{K}^n : (A - \lambda E_n)x = 0\} = L(A - \lambda E_n).$$

$E_A(\lambda)$ heißt der *Eigenraum* von A zum Eigenwert λ. Weiter heißt $g_A(\lambda) := \dim(E_A(\lambda))$ die *geometrische Vielfachheit* des Eigenwertes λ von A.

$E_A(\lambda)$ besteht also aus allen Eigenvektoren von A zum Eigenwert λ zuzüglich des Nullvektors. Hat man einen Eigenwert λ von A bestimmt, so kann man den zugehörigen Eigenraum als Lösungsmenge des homogenen linearen Gleichungssystems $(A - \lambda E_n)x = 0$ wie gewohnt mit Hilfe des Gaußschen Algorithmus bestimmen.

Bevor wir fortfahren können, benötigen wir noch einen kurzen Einschub über Polynome (einige Beweise finden sich im Anhang): Die Menge aller Polynome $p : \mathbb{K} \to \mathbb{K}$ bezeichnen wir mit $\mathbb{K}[x]$ (wobei wieder $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ ist). Sei nun $p \in \mathbb{K}[x]$ ein nicht konstantes Polynom. Eine Zahl $a \in \mathbb{K}$ ist eine Nullstelle von p genau dann, wenn es ein Polynom $q \in \mathbb{K}[x]$ mit $p(x) = q(x)(x - a)$ für alle $x \in \mathbb{K}$ gibt.

Ist a eine Nullstelle von p, so heißt

$$k(a, p) := \max\Big\{k \in \mathbb{N} : \exists q \in \mathbb{K}[x]\ \forall x \in \mathbb{K}\ p(x) = q(x)(x - a)^k\Big\}$$

die *Vielfachheit* von a als Nullstelle von p.

Man sagt, p zerfalle über $\mathbb{K}$ in *Linearfaktoren*, falls es paarweise verschiedene Zahlen $a_1, \ldots, a_m \in \mathbb{K}$, sowie $k_1, \ldots, k_m \in \mathbb{N}$ und $a \in \mathbb{K}$ gibt, so dass

$$p(x) = a(x - a_1)^{k_1} \ldots (x - a_m)^{k_m} \quad \forall x \in \mathbb{K}$$

gilt. In diesem Fall sind $a_1, \ldots, a_m$ natürlich genau die Nullstellen von p und es gilt $k(a_i, p) = k_i$ für alle $i = 1, \ldots, m$.

Nicht jedes reelle Polynom zerfällt über $\mathbb{R}$ in Linearfaktoren, z. B. hat $p(x) := x^2 + 1$ keine reelle Nullstelle. Hingegen besagt der *Fundamentalsatz der Algebra*, dass jedes nicht konstante Polynom $p \in \mathbb{C}[x]$ über $\mathbb{C}$ in Linearfaktoren zerfällt (z. B. ist $z^2 + 1 = (z + i)(z - i)$ für alle $z \in \mathbb{C}$). Dieser Satz ist allerdings ziemlich tiefliegend und soll in dieser Vorlesung nicht bewiesen werden.

Nun kehren wir zur Thematik der Eigenwerte zurück und geben zunächst folgende Definition an.

Definition VIII.1.5. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $A \in M(n \times n, \mathbb{K})$. Ist $\lambda \in \mathbb{K}$ ein Eigenwert von A, so heißt $k_A(\lambda) := k(\lambda, \chi_A)$ die *algebraische Vielfachheit* von λ als Eigenwert von A.

Die algebraische Vielfachheit eines Eigenwertes ist also seine Vielfachheit als Nullstelle des charakteristischen Polynoms, seine geometrische Vielfachheit ist die Dimension des zugehörigen Eigenraumes. Das folgende Lemma stellt einen Zusammenhang zwischen beiden Vielfachheiten her.

Lemma VIII.1.6. *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$, sei $A \in M(n \times n, \mathbb{K})$ und sei $\lambda \in \mathbb{K}$ ein Eigenwert von A. Dann gilt $g_A(\lambda) \leq k_A(\lambda)$.*

Beweis. Sei $l := g_A(\lambda) = \dim(E_A(\lambda))$. Sei $(x_1, \ldots, x_l)$ eine geordnete Basis von $E_A(\lambda)$. Nach dem Basisergänzungssatz gibt es Vektoren $x_{l+1}, \ldots, x_n \in \mathbb{K}^n$, so dass $\mathcal{B} := (x_1, \ldots, x_l, x_{l+1}, \ldots, x_n)$ eine geordnete Basis des $\mathbb{K}^n$ bildet. Es sei $\mathcal{A} := (e_1, \ldots, e_n)$ die kanonische Basis des $\mathbb{K}^n$ und $T := T_{\mathcal{B}}^{\mathcal{A}}$ die zugehörige Basiswechselmatrix. Dann ist T invertierbar mit $T^{-1} = T_{\mathcal{A}}^{\mathcal{B}}$. Setze $B := TAT^{-1}$.
Dann gilt $B = T_{\mathcal{B}}^{\mathcal{A}} \mathcal{M}_{\mathcal{A}}(F_A) T_{\mathcal{A}}^{\mathcal{B}} = \mathcal{M}_{\mathcal{B}}(F_A)$.
Die i-te Spalte von $\mathcal{M}_{\mathcal{B}}(F_A)$ ist $\mathcal{K}_{\mathcal{B}}(F_A(x_i)) = \mathcal{K}_{\mathcal{B}}(Ax_i) = \mathcal{K}_{\mathcal{B}}(\lambda x_i) = \lambda e_i$ für alle $i = 1, \ldots, l$, denn x_i ist jeweils ein Eigenvektor von A zum Eigenwert λ. Damit hat also B die Form

$$B = \begin{pmatrix} \lambda E_l & B_1 \\ 0 & B_2 \end{pmatrix}$$

für gewisse Matrizen $B_1 \in M(l \times (n-l), \mathbb{K})$ und $B_2 \in M((n-l) \times (n-l), \mathbb{K})$. Mit Hilfe von Satz VI.2.8 folgt

$$\begin{aligned} \chi_B(t) &= \det \begin{pmatrix} (\lambda - t)E_l & B_1 \\ 0 & B_2 - tE_{n-l} \end{pmatrix} = \det((\lambda - t)E_l)\det(B_2 - tE_{n-l}) \\ &= (\lambda - t)^l \chi_{B_2}(t) = (-1)^l (t - \lambda)^l \chi_{B_2}(t) \end{aligned}$$

für alle $t \in \mathbb{K}$.
Andererseits gilt aber wegen des Determinantenmultiplikationssatzes

$$\begin{aligned} \chi_B(t) &= \det(B - tE_n) = \det(TAT^{-1} - tE_n) = \det(T(A - tE_n)T^{-1}) \\ &= \det(T)\det(A - tE_n)\det(T^{-1}) = \det(TT^{-1})\chi_A(t) \\ &= \det(E_n)\chi_A(t) = \chi_A(t) \end{aligned}$$

für alle $t \in \mathbb{K}$.
Also ist $\chi_A(t) = (-1)^l (t - \lambda)^l \chi_{B_2}(t)$ für alle $t \in \mathbb{K}$ und daher gilt für die algebraische Vielfachheit $k_A(\lambda) \geq l = g_A(\lambda)$. □

Als Nächstes zeigen wir noch, dass Eigenvektoren zu paarweise verschiedenen Eigenwerten stets linear unabhängig sind.

Lemma VIII.1.7. *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $A \in M(n \times n, \mathbb{K})$. Sind $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$ paarweise verschiedene Eigenwerte von A und ist $x_i \in \mathbb{K}^n$ ein Eigenvektor von A zum Eigenwert λ_i für alle $i = 1, \ldots, m$, so ist $(x_1, \ldots, x_m)$ linear unabhängig.*

Beweis. Wir argumentieren mittels vollständiger Induktion nach m. Für $m = 1$ ist nur zu bemerken, dass x_1 als Eigenvektor von Null verschieden und somit linear unabhängig ist.
Angenommen nun die Behauptung gilt für m paarweise verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_m$ und es sei λ_{m+1} ein weiterer Eigenwert mit $\lambda_{m+1} \neq \lambda_i$ für alle $i = 1, \ldots, m$. Sei x_i ein Eigenvektor von A zum Eigenwert λ_i für alle $i = 1, \ldots, m+1$.
Seien $\alpha_1, \ldots, \alpha_{m+1} \in \mathbb{K}$ mit $\sum_{i=1}^{m+1} \alpha_i x_i = 0$. Es folgt

$$\begin{aligned}
&\sum_{i=1}^{m} \alpha_i(\lambda_{m+1} - \lambda_i)x_i = \lambda_{m+1} \sum_{i=1}^{m} \alpha_i x_i - \sum_{i=1}^{m} \alpha_i \lambda_i x_i \\
&= -\lambda_{m+1}\alpha_{m+1}x_{m+1} - \sum_{i=1}^{m} \alpha_i A x_i = -\lambda_{m+1}\alpha_{m+1}x_{m+1} - A\left(\sum_{i=1}^{m} \alpha_i x_i\right) \\
&= -\lambda_{m+1}\alpha_{m+1}x_{m+1} - A(-\alpha_{m+1}x_{m+1}) \\
&= -\lambda_{m+1}\alpha_{m+1}x_{m+1} + \lambda_{m+1}\alpha_{m+1}x_{m+1} = 0.
\end{aligned}$$

Da nach Induktionsvoraussetzung $(x_1, \ldots, x_m)$ linear unabhängig ist, folgt $\alpha_i(\lambda_{m+1} - \lambda_i) = 0$ für alle $i = 1, \ldots, m$. Wegen $\lambda_{m+1} \neq \lambda_i$ impliziert das $\alpha_i = 0$ für alle $i = 1, \ldots, m$.
Aus $\sum_{i=1}^{m+1} \alpha_i x_i = 0$ folgt damit $\alpha_{m+1}x_{m+1} = 0$ und wegen $x_{m+1} \neq 0$ muss also auch $\alpha_{m+1} = 0$ gelten. □

VIII.2 Diagonalisierbarkeit

In diesem Abschnitt geht es um Diagonalisierbarkeit von Matrizen, d. h. man möchte eine gegebene Matrix auf eine möglichst einfache Form, nämlich auf Diagonalform, bringen. Dazu führen wir zunächst den Begriff der Ähnlichkeit zweier Matrizen ein.

Definition VIII.2.1. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und seien $A, B \in M(n \times n, \mathbb{K})$. A heißt *ähnlich* zu B (in Zeichen: $A \sim B$), falls es eine invertierbare Matrix $T \in M(n \times n, \mathbb{K})$ mit $TAT^{-1} = B$ gibt.

Als Übung können Sie beweisen, dass es sich bei der Äihnlichkeit um eine Äquivalenzrelation auf der Menge $M(n \times n, \mathbb{K})$ handelt.

Diagonalisierbarkeit wird nun einfach definiert als Ähnlichkeit zu einer Diagonalmatrix.

Definition VIII.2.2. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$. Eine Matrix $A \in M(n \times n, \mathbb{K})$ heißt *diagonalisierbar*, falls es eine Diagonalmatrix D mit $A \sim D$ gibt.

Zuerst zeigen wir nun das folgende Lemma.

Lemma VIII.2.3. *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $A \in M(n \times n, \mathbb{K})$. Dann gilt: A ist diagonalisierbar genau dann, wenn es eine Basis des $\mathbb{K}^n$ gibt, die aus Eigenvektoren von A besteht.*

Beweis. 1) Sei zunächst A diagonalisierbar. Dann existieren eine invertierbare Matrix T und $\mu_1, \ldots, \mu_n \in \mathbb{K}$ mit
Es folgt $\mu_i e_i = De_i = (TAT^{-1})e_i$ und somit $\mu_i T^{-1} e_i = A(T^{-1}e_i)$, d. h. μ_i ist ein Eigenwert von A und $T^{-1}e_i$ ein zugehöriger Eigenvektor (für alle $i = 1, \ldots, n$).
Da T^{-1} invertierbar ist, gilt $\text{Rang}(T^{-1}) = n$ und somit sind die Spalten $T^{-1}e_1, \ldots, T^{-1}e_n$ von T^{-1} linear unabhängig, d. h. sie bilden eine Basis des $\mathbb{K}^n$, die aus Eigenvektoren von A besteht.
2) Nehmen wir nun umgekehrt an es gibt eine Basis $\mathcal{B} = (b_1, \ldots, b_n)$ des $\mathbb{K}^n$ aus Eigenvektoren von A. Für alle $i = 1, \ldots, n$ sei μ_i der zu b_i gehörige Eigenwert von A. Ferner sei $\mathcal{A} = (e_1, \ldots, e_n)$ die kanonische Basis des $\mathbb{K}^n$ und $T := T_{\mathcal{B}}^{\mathcal{A}}$ die Basiswechselmatrix.
T ist invertierbar mit $T^{-1} = T_{\mathcal{A}}^{\mathcal{B}}$ und es gilt $TAT^{-1} = T_{\mathcal{B}}^{\mathcal{A}} \mathcal{M}_{\mathcal{A}}(F_A) T_{\mathcal{A}}^{\mathcal{B}} = \mathcal{M}_{\mathcal{B}}(F_A)$.
Die i-te Spalte von $\mathcal{M}_{\mathcal{B}}(F_A)$ ist $\mathcal{K}_{\mathcal{B}}(F_A(b_i)) = \mathcal{K}_{\mathcal{B}}(Ab_i) = \mathcal{K}_{\mathcal{B}}(\mu_i b_i) = \mu_i e_i$ (für alle $i = 1, \ldots, n$), also gilt

$$TAT^{-1} = \mathcal{M}_{\mathcal{B}}(F_A) = \begin{pmatrix} \mu_1 & 0 & \ldots & 0 \\ 0 & \mu_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \mu_n \end{pmatrix}.$$

□

Nun kommen wir zum Hauptkriterium für Diagonalisierbarkeit.

Satz VIII.2.4. *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $A \in M(n \times n, \mathbb{K})$. Dann gilt: A ist diagonalisierbar genau dann, wenn das charakteristische Polynom χ_A über $\mathbb{K}$ in Linearfaktoren zerfällt und $g_A(\lambda) = k_A(\lambda)$ für alle Eigenwerte λ von A gilt.*

Beweis. 1) Nehmen wir zunächst an A sei diagonalisierbar. Dann existieren eine invertierbare Matrix T und $\mu_1, \ldots, \mu_n \in \mathbb{K}$ mit

$$TAT^{-1} = \begin{pmatrix} \mu_1 & 0 & \ldots & 0 \\ 0 & \mu_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \mu_n \end{pmatrix} =: D$$

Es sei m die Anzahl der Elemente von $\{\mu_1, \ldots, \mu_n\}$ und $\{\lambda_1, \ldots, \lambda_m\} = \{\mu_1, \ldots, \mu_n\}$. Für alle $i \in \{1, \ldots, m\}$ sei $J_i := \{j \in \{1, \ldots, n\} : \mu_j = \lambda_i\}$ und k_i die Anzahl der Elemente von J_i. D. h. in der Folge $\mu_1, \ldots, \mu_n$ kommt der Wert λ_i genau k_i mal vor.

Eine analoge Rechnung wie im Beweis von Lemma VIII.1.6 zeigt $\chi_A = \chi_D$. Es folgt

$$\chi_A(t) = \begin{vmatrix} \mu_1 - t & 0 & \dots & 0 \\ 0 & \mu_2 - t & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mu_n - t \end{vmatrix} = (\mu_1 - t)(\mu_2 - t)\dots(\mu_n - t)$$
$$= (-1)^n(t - \mu_1)(t - \mu_2)\dots(t - \mu_n) = (-1)^n(t - \lambda_1)^{k_1}\dots(t - \lambda_m)^{k_m}$$

für alle $t \in \mathbb{K}$. Das charakteristische Polynom von A zerfällt also in Linearfaktoren, die Eigenwerte von A sind genau $\lambda_1, \dots, \lambda_m$ und für die algebraischen Vielfachheiten gilt $k_A(\lambda_i) = k_i$ für $i = 1, \dots, m$.
Sei nun $i \in \{1, \dots, m\}$ beliebig und $x \in \mathbb{K}^n$ beliebig. Dann gilt

$$Dx = \lambda_i x \quad \Leftrightarrow \quad T^{-1}Dx = \lambda_i T^{-1}x \quad \Leftrightarrow \quad AT^{-1}x = \lambda_i T^{-1}x.$$

Es folgt $U_i := L(D - \lambda_i E_n) = \{x \in \mathbb{K}^n : T^{-1}x \in E_A(\lambda_i)\}$. Daher ist die Abbildung $\psi_i : E_A(\lambda_i) \to U_i$ mit $\psi_i(y) := Ty$ für $y \in E_A(\lambda_i)$ ein Isomorphismus. Folglich gilt $g_A(\lambda_i) = \dim(E_A(\lambda_i)) = \dim(U_i)$.
Ferner gilt für alle $x = (x_1 \ \dots \ x_n)^T \in \mathbb{K}^n$

$$x \in U_i \quad \Leftrightarrow \quad Dx = \lambda_i x \quad \Leftrightarrow \quad \begin{pmatrix} \mu_1 x_1 \\ \vdots \\ \mu_n x_n \end{pmatrix} = \begin{pmatrix} \lambda_i x_1 \\ \vdots \\ \lambda_i x_n \end{pmatrix}$$
$$\Leftrightarrow \quad \mu_j x_j = \lambda_i x_j \ \forall j = 1, \dots, n \quad \Leftrightarrow \quad x_j = 0 \ \forall j \in \{1, \dots, n\} \setminus J_i$$
$$\Leftrightarrow \quad x \in \operatorname{span}\{e_j : j \in J_i\}.$$

Es gilt also $U_i = \operatorname{span}\{e_j : j \in J_i\}$ und daher $g_A(\lambda_i) = \dim(U_i) = k_i = k_A(\lambda_i)$.
2) Nehmen wir nun umgekehrt an, dass das charakteristische Polynom in Linearfaktoren zerfällt, etwa

$$\chi_A(t) = (-1)^n(t - \lambda_1)^{k_1}\dots(t - \lambda_m)^{k_m},$$

wobei $k_1, \dots, k_m \in \mathbb{N}$ und $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ paarweise verschieden sein sollen. Weiterhin nehmen wir $g_A(\lambda_i) = k_A(\lambda_i) = k_i$ für alle $i = 1, \dots, m$ an.
Für alle $i = 1, \dots, m$ sei $(b_1^{(i)}, \dots, b_{k_i}^{(i)})$ eine geordnete Basis von $E_A(\lambda_i)$.
Da χ_A den Grad n hat, gilt $\sum_{i=1}^m k_i = n$. Es sei

$$(b_1, \dots, b_n) := (b_1^{(1)}, \dots, b_{k_1}^{(1)}, b_1^{(2)}, \dots, b_{k_2}^{(2)}, \dots, b_1^{(m)}, \dots, b_{k_m}^{(m)}).$$

Wir wollen zeigen, dass $(b_1, \dots, b_n)$ linear unabhängig ist.
Seien dazu $\alpha_1^{(1)}, \dots, \alpha_{k_1}^{(1)}, \alpha_1^{(2)}, \dots, \alpha_{k_2}^{(2)}, \dots, \alpha_1^{(m)}, \dots, \alpha_{k_m}^{(m)} \in \mathbb{K}$ mit

$$\sum_{i=1}^{m} \sum_{j=1}^{k_i} \alpha_j^{(i)} b_j^{(i)} = 0.$$

Sei $v_i := \sum_{j=1}^{k_i} \alpha_j^{(i)} b_j^{(i)}$ für alle $i = 1, \ldots, m$. Dann gilt $v_i \in E_A(\lambda_i)$ für alle $i = 1, \ldots, m$ und $\sum_{i=1}^{m} v_i = 0$.
Da Eigenvektoren zu paarweise verschiedenen Eigenwerten linear unabhängig sind (Lemma VIII.1.7), folgt daraus $v_i = 0$ für alle $i = 1, \ldots, m$.
Also ist $\sum_{j=1}^{k_i} \alpha_j^{(i)} b_j^{(i)} = 0$ und somit (da $(b_1^{(i)}, \ldots, b_{k_i}^{(i)})$ linear unabhängig ist) $\alpha_j^{(i)} = 0$ für alle $j = 1, \ldots, k_i$ und alle $i = 1, \ldots, m$. Das zeigt die lineare Unabhängigkeit von $(b_1, \ldots, b_n)$.
Somit ist $(b_1, \ldots, b_n)$ also eine geordnete Basis des $\mathbb{K}^n$, die aus Eigenvektoren von A besteht. Nach Lemma VIII.2.3 ist A also diagonalisierbar. □

Beispiele:
1) Sei

$$A := \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Da A eine obere Dreiecksmatrix ist, folgt für das charakteristische Polynom leicht $\chi_A(t) = -(t-1)^2(t-2)$. Die Eigenwerte von A sind also 1 und 2 mit den algebraischen Vielfachheiten $k_A(1) = 2$ und $k_A(2) = 1$.
Mit Hilfe des Gaußschen Algorithmus bestimmt man die zugehörigen Eigenräume zu $E_A(1) = \text{span}\{e_1\}$ und $E_A(2) = \text{span}\{e_3\}$. Also gilt $g_A(2) = 1 = k_A(2)$ und $g_A(1) = 1 < 2 = k_A(1)$. Nach dem obigen Satz ist also A nicht diagonalisierbar.
2) Sei

$$B := \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Für das charakteristische Polynom ergibt sich wieder $\chi_B(t) = -(t-1)^2(t-2)$, die Eigenwerte sind also wieder 1 und 2 mit den algebraischen Vielfachheiten $k_B(1) = 2$ und $k_B(2) = 1$.
Mit dem Gaußschen Algorithmus bestimmt man die Eigenräume zu $E_B(1) = \text{span}\{e_1, e_2\}$ und $E_B(2) = \text{span}\{e_1 - e_3\}$. Also ist $g_B(1) = 2 = k_B(1)$ und $g_B(2) = 1 = k_B(2)$. Daher ist nach dem obigen Satz B diagonalisierbar.
Eine Basis des $\mathbb{R}^3$ aus Eigenvektoren von B erhält man, indem man Basen der einzelnen Eigenräume wählt und diese zusammenfügt (vgl. den Beweis des obigen Satzes). Also ist $\mathcal{B} := (e_1, e_2, e_1 - e_3)$ eine geordnete Basis des $\mathbb{R}^3$ aus Eigenvektoren von B.
Sei $T := T_{\mathcal{B}}^{\mathcal{A}}$, wobei $\mathcal{A}$ die kanonische Basis des $\mathbb{R}^3$ bezeichnet. Dann gilt

$$TAT^{-1} = \mathcal{M}_{\mathcal{B}}(F_A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Da $\mathcal{A}$ die kanonische Basis ist, ist

$$T^{-1} = T_{\mathcal{A}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Durch Inversenbildung mit dem Gaußschen Algorithmus erhält man

$$T = (T^{-1})^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = T^{-1}.$$

Als Nächstes wollen wir uns noch speziell mit der Diagonalisierbarkeit von sogenannten symmetrischen/hermiteschen Matrizen befassen. Hierzu zunächst die Definition.

Definition VIII.2.5. Sei $n \in \mathbb{N}$.
1) Eine Matrix $A \in M(n \times n, \mathbb{R})$ heißt *symmetrisch*, falls $A = A^T$ gilt.
2) Eine Matrix $A \in M(n \times n, \mathbb{C})$ heißt *hermitesch*, falls $A = \overline{A}^T$ gilt.
Anstelle von symmetrisch oder hermitesch sagt man auch, die Matrix A sei *selbstadjungiert*.

Selbstadjungierte Matrizen lassen sich wie folgt charakterisieren.

Lemma VIII.2.6. *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $A \in M(n \times n, \mathbb{K})$. Sei $\langle \cdot, \cdot \rangle$ das kanonische Skalarprodukt auf $\mathbb{K}^n$. Sei $A \in M(n \times n, \mathbb{K})$. Dann gilt: A ist selbstadjungiert genau dann, wenn $\langle Ax, y \rangle = \langle x, Ay \rangle$ für alle $x, y \in \mathbb{K}^n$ gilt.*

Beweis. Sei $A^* := A^T$, falls $\mathbb{K} = \mathbb{R}$ bzw. $A^* := \overline{A}^T$, falls $\mathbb{K} = \mathbb{C}$. Ist A selbstadjungiert, so ist $A^* = A$ und aus Lemma VII.1.10 folgt daher $\langle Ax, y \rangle = \langle x, Ay \rangle$ für alle $x, y \in \mathbb{K}^n$.
Gilt umgekehrt $\langle Ax, y \rangle = \langle x, Ay \rangle$ für alle $x, y \in \mathbb{K}^n$, so folgt aus Lemma VII.1.10 $\langle x, Ay \rangle = \langle x, A^* y \rangle$, also $\langle x, (A - A^*)y \rangle = 0$ für alle $x, y \in \mathbb{K}^n$. Insbesondere ist $\|(A - A^*)y\|^2 = \langle (A - A^*)y, (A - A^*)y \rangle = 0$ und somit $(A - A^*)y = 0$ für alle $y \in \mathbb{K}^n$.
Also ist $(A - A^*)e_i = 0$ für alle $i = 1, \ldots, n$. $(A - A^*)e_i$ ist aber gerade die i-te Spalte von $A - A^*$. Also ist $A = A^*$, d. h. A ist selbstadjungiert. □

Als Nächstes beobachten wir, dass sämtliche Eigenwerte einer hermiteschen Matrix reell sein müssen.

Lemma VIII.2.7. *Sei $A \in M(n \times n, \mathbb{C})$ hermitesch und $\lambda \in \mathbb{C}$ sei ein Eigenwert von A. Dann ist $\lambda \in \mathbb{R}$.*

Beweis. Sei $x \in \mathbb{C}^n$ mit $Ax = \lambda x$ und $x \neq 0$. Da A hermitesch ist, folgt

$$\lambda \|x\|^2 = \lambda \langle x, x \rangle = \langle \lambda x, x \rangle = \langle Ax, x \rangle = \langle x, Ax \rangle = \langle x, \lambda x \rangle = \overline{\lambda} \langle x, x \rangle = \overline{\lambda} \|x\|^2.$$

Wegen $x \neq 0$ ist $\|x\|^2 > 0$ und es folgt $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$. □

Es folgt die wichtige Erkenntnis, dass die Eigenvektoren zu paarweise verschiedenen Eigenwerten einer selbstadjungierten Matrix senkrecht aufeinander stehen.

Lemma VIII.2.8. *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $A \in M(n \times n, \mathbb{K})$ selbstadjungiert. Seien λ und μ Eigenwerte von A mit $\lambda \neq \mu$. Sei $x \in \mathbb{K}^n$ ein Eigenvektor von A zum Eigenwert λ und sei $y \in \mathbb{K}^n$ ein Eigenvektor von A zum Eigenwert μ. Dann gilt $x \perp y$ (bzgl. des kanonischen Skalarprodukts auf $\mathbb{K}^n$).*

Beweis. Wegen der Selbstadjungiertheit von A gilt

$$\lambda\langle x, y\rangle = \langle \lambda x, y\rangle = \langle Ax, y\rangle = \langle x, Ay\rangle = \langle x, \mu y\rangle = \mu\langle x, y\rangle,$$

wobei wir im letzten Schritt $\mu \in \mathbb{R}$ ausgenutzt haben.
Wegen $\lambda \neq \mu$ folgt aus der obigen Gleichung $\langle x, y\rangle = 0$. □

Als Nächstes definieren wir noch orthogonale und unitäre Matrizen.

Definition VIII.2.9. Für $n \in \mathbb{N}$ sei

$$O(n) := \left\{A \in M(n \times n, \mathbb{R}) : A \text{ ist invertierbar mit } A^{-1} = A^T\right\}$$

und

$$U(n) := \left\{A \in M(n \times n, \mathbb{C}) : A \text{ ist invertierbar mit } A^{-1} = \overline{A}^T\right\}.$$

Die Elemente von $O(n)$ heißen *orthogonale Matrizen*, die Elemente von $U(n)$ nennt man *unitäre Matrizen*.

Es ist leicht zu zeigen, dass $O(n)$ und $U(n)$ jeweils bzgl. der Matrixmultiplikation eine Gruppe bilden, die sogenannte *orthogonale* bzw. *unitäre Gruppe*.

Ist $A \in M(n \times n, \mathbb{R})$ und bezeichnen wir die Spalten von A mit $a_1, \ldots, a_n$ und die Zeilen mit $z_1, \ldots, z_n$, so gilt:

$$A \in O(n) \;\Leftrightarrow\; \langle a_i, a_j\rangle = \delta_{ij} \;\forall i, j = 1, \ldots, n \;\Leftrightarrow\; \langle z_i, z_j\rangle = \delta_{ij} \;\forall i, j = 1, \ldots, n$$

(Beweis als Übung). A ist also genau dann eine orthogonale Matrix, wenn ihre Spalten (bzw. Zeilen) eine Orthonormalbasis des $\mathbb{R}^n$ bzgl. des euklidischen Skalarprodukts bilden. Eine analoge Charakterisierung gilt für unitäre Matrizen.

Nun kommen wir zum entscheidenden Satz über die Diagonalisierbarkeit selbstadjungierter Matrizen, dem sogenannten *Satz von der Hauptachsentransformation*.

Satz VIII.2.10 (Satz von der Hauptachsentransformation)**.** *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und sei $A \in M(n \times n, \mathbb{K})$ selbstadjungiert. Dann existiert eine Orthonormalbasis des $\mathbb{K}^n$ (bzgl. des Standardskalarprodukts), die aus Eigenvektoren von A besteht. Insbesondere ist A diagonalisierbar.*

Beweis. Wir wollen zunächst argumentieren, dass A mindestens einen Eigenwert besitzt. Im Falle $\mathbb{K} = \mathbb{C}$ ist das wegen des Fundamentalsatzes der Algebra klar: Das charakteristische Polynom χ_A muss eine Nullstelle in $\mathbb{C}$ haben. Wegen Lemma VIII.2.7 muss diese aber sogar in $\mathbb{R}$ liegen. Im Falle $\mathbb{K} = \mathbb{R}$ können wir A auch künstlich als eine hermitesche Matrix in $M(n \times n, \mathbb{C})$ auffassen, deren Einträge nur "zufällig" alle reell sind. Dann argumentiert man wie eben, dass χ_A eine reelle Nullstelle, also A einen reellen Eigenwert besitzt.
Nun zeigen wir die eigentliche Behauptung des Satzes durch vollständige Induktion nach n. Für $n = 1$ ist die Aussage klar.
Angenommen nun die Behauptung stimmt für selbstadjungierte Matrizen vom Format $n \times n$ und es ist A eine selbstadjungierte $(n+1) \times (n+1)$-Matrix. Wie eben gezeigt, existiert ein reeller Eigenwert λ von A. Es sei $x_0 \in \mathbb{K}^{n+1}$ ein Eigenvektor von A zum Eigenwert λ. Indem man ggf. x_0 durch $x_0/\|x_0\|$ ersetzt, kann man $\|x_0\| = 1$ annehmen.
Wir setzen $U := \operatorname{span}\{x_0\}$ und

$$W := \left\{x \in \mathbb{K}^{n+1} : \langle x, x_0\rangle = 0\right\} = U^{\perp}.$$

Aus Satz VII.2.9 folgt $\dim(W) = n + 1 - \dim(U) = n$. Es sei $\mathcal{B} = (b_1, \ldots, b_n)$ eine geordnete ONB von W.
Ist $x \in W$, so folgt wegen der Selbstadjungiertheit von A und $Ax_0 = \lambda x_0$

$$\langle Ax, x_0\rangle = \langle x, Ax_0\rangle = \langle x, \lambda x_0\rangle = \lambda\langle x, x_0\rangle = 0,$$

also ist auch $Ax \in W$.
Daher ist die Abbildung $G : W \to W$ mit $G(x) := Ax$ für alle $x \in W$ wohldefiniert und natürlich ist G linear.
Es sei $B := \mathcal{M}_{\mathcal{B}}(G) \in M(n \times n, \mathbb{K})$.
Als Nächstes machen wir folgende Beobachtung: Es gilt

$$\langle x, y\rangle = \langle \Phi_{\mathcal{B}}(x), \Phi_{\mathcal{B}}(y)\rangle \quad \forall x, y \in \mathbb{K}^n. \tag{VIII.1}$$

Beweis dazu: Für $x = (x_1 \ \ldots \ x_n)^T$ und $y = (y_1 \ \ldots \ y_n)^T$ gilt

$$\begin{aligned}\langle \Phi_{\mathcal{B}}(x), \Phi_{\mathcal{B}}(y)\rangle &= \left\langle \sum_{i=1}^{n} x_i b_i, \sum_{j=1}^{n} y_j b_j \right\rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\overline{y_j}\langle b_i, b_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\overline{y_j}\delta_{ij} \\ &= \sum_{i=1}^{n} x_i\overline{y_i} = \langle x, y\rangle.\end{aligned}$$

Sind nun $v, w \in \mathbb{K}^n$, so gilt mit $a := \Phi_{\mathcal{B}}(v)$ und $b := \Phi_{\mathcal{B}}(w)$

$$\langle Bv, w\rangle = \langle \mathcal{M}_{\mathcal{B}}(G)\mathcal{K}_{\mathcal{B}}(a), \mathcal{K}_{\mathcal{B}}(b)\rangle = \langle \mathcal{K}_{\mathcal{B}}(G(a)), \mathcal{K}_{\mathcal{B}}(b)\rangle = \langle G(a), b\rangle,$$

wobei wir im letzten Schritt (VIII.1) benutzt haben. Es folgt

$$\begin{aligned}&\langle Bv, w\rangle = \langle Aa, b\rangle = \langle a, Ab\rangle = \langle a, G(b)\rangle = \langle \mathcal{K}_{\mathcal{B}}(a), \mathcal{K}_{\mathcal{B}}(G(b))\rangle \\ &= \langle \mathcal{K}_{\mathcal{B}}(a), \mathcal{M}_{\mathcal{B}}(G)\mathcal{K}_{\mathcal{B}}(b)\rangle = \langle v, Bw\rangle,\end{aligned}$$

wobei wir die Selbstadjungiertheit von A und erneut (VIII.1) ausgenutzt haben.
Also it B selbstadjungiert und somit existiert nach Induktionsvoraussetzung eine geordnete ONB $\mathcal{A} = (a_1, \ldots, a_n)$ des $\mathbb{K}^n$, die aus Eigenvektoren von B besteht. Sei λ_i der zu a_i gehörige Eigenwert von B für alle $i = 1, \ldots, n$.
Wir setzen $x_i := \Phi_{\mathcal{B}}(a_i) \in W$ für alle $i = 1, \ldots, n$.
Es folgt $\mathcal{K}_{\mathcal{B}}(G(x_i)) = \mathcal{M}_{\mathcal{B}}(G)\mathcal{K}_{\mathcal{B}}(x_i) = Ba_i = \lambda_i a_i = \lambda_i \mathcal{K}_{\mathcal{B}}(x_i)$, also $Ax_i = G(x_i) = \lambda_i x_i$, d. h. x_i ist ein Eigenvektor von A für alle $i = 1, \ldots, n$.
Ferner gilt wegen (VIII.1) auch $\langle x_i, x_j\rangle = \langle a_i, a_j\rangle = \delta_{ij}$ für alle $i, j = 1, \ldots, n$.
Außerdem ist x_0 ein Eigenvektor von A mit $\|x_0\| = 1$ und wegen $x_1, \ldots, x_n \in W$ gilt $\langle x_i, x_0\rangle = 0$ für alle $i = 1, \ldots, n$. Also ist $(x_0, x_1, \ldots, x_n)$ eine ONB des $\mathbb{K}^{n+1}$, die aus Eigenvektoren von A besteht. □

Als Korollar erhält man folgendes Resultat.

Korollar VIII.2.11. *Ist A eine symmetrische bzw. hermitesche $(n \times n)$-Matrix, so existieren eine Diagonalmatrix D und eine Matrix $T \in O(n)$ bzw. $T \in U(n)$ mit $TAT^{-1} = D$.*

Beweis. Nach dem Satz über die Hauptachsentransformation gibt es eine geordnete ONB $\mathcal{B} = (b_1, \ldots, b_n)$ des $\mathbb{K}^n$, die aus Eigenvektoren von A besteht. Es sei $\mathcal{A} = (e_1, \ldots, e_n)$ die kanonische Basis des $\mathbb{K}^n$ und $T := T_{\mathcal{B}}^{\mathcal{A}}$. Dann ist $TAT^{-1} = T_{\mathcal{B}}^{\mathcal{A}} A T_{\mathcal{A}}^{\mathcal{B}} = \mathcal{M}_{\mathcal{B}}(F_A) =: D$ eine Diagonalmatrix.
Ferner sind die Spalten von $T^{-1} = T_{\mathcal{A}}^{\mathcal{B}}$ gerade die Vektoren $b_1, \ldots, b_n$, die eine ONB des $\mathbb{K}^n$ bilden. Also ist T^{-1} und damit auch T eine orthogonale bzw. unitäre Matrix. □

VIII.3 Anwendung: Der PageRank

In diesem letzten Abschnitt wollen wir eine Anwendung der Eigenwerttheorie diskutieren, den sogenannten *PageRank*, den die Google-Gründer Sergei Brin und Larry Page zur Bewertung der Relevanz von Internetseiten bei der Suche im Web eingeführt haben.[1]

Es bezeichne N die Gesamtzahl aller Webseiten im Internet (die von $1, \ldots, N$ durchnummeriert werden). Es bezeichne weiter c_j die Anzahl der von Seite j ausgehenden Links. Ist $c_j > 0$, so setzen wir $a_{ij} := 1/c_j$, falls es einen Link von Seite j auf Seite i gibt und $a_{ij} = 0$, falls das nicht der Fall ist.

[1] Die Originalpublikation ist: S. Brin, L. Page, *The Anatomy of a Large-Scale Hypertextual Web Search Engine*, Computer Networks and ISDN Systems 30, 107–117 (1998).

Ist $c_j = 0$, so setzen wir $a_{ij} = 1/N$ für alle $i = 1, \ldots, N$. Außerdem fixieren wir noch eine Zahl $0 < d < 1$.

Nun betrachten wir ein Modell eines Zufallssurfers, der bei einer beliebig gewählten Webseite startet und sich gemäß des folgenden Prinzips zufällig von Seite zu Seite weiter klickt. Befindet sich der Surfer gerade auf Seite j, so führt er mit Wahrscheinlichkeit d folgenden Schritt aus: Falls es mindestens einen von Seite j ausgehenden Link gibt (d. h. $c_j > 0$), so wählt er zufällig einen dieser Links aus und klickt zur entsprechenden Seite weiter. Die Wahrscheinlichkeit, dass er dabei auf Seite i landet ist gerade a_{ij}. Falls es keinen von Seite j ausgehenden Link gibt ($c_j = 0$), so gibt er zufällig eine neue Webaddresse in der Browserleiste ein. Die Wahrscheinlichkeit, dass er dabei auf Seite i gelangt ist gerade $1/N = a_{ij}$.

Mit Wahrscheinlichkeit $1 - d$ gibt er dagegen direkt zufällig eine neue Webaddresse ein, unabhängig davon, ob es von Seite j ausgehende Links gibt oder nicht. Die Wahrscheinlichkeit, dabei auf Seite i zu landen ist wieder $1/N$.

Dieses wahrscheinlichkeitstheoretische Modell ist ein Beispiel für eine sogenannte *Markov-Kette* (die genaue Definition einer solchen wollen wir hier nicht extra anführen).

Jeder Seite soll nun eine gewisse positive Zahl, ihr *PageRank*, zugeordnet werden. Wir bezeichnen die PageRanks der Seiten $1, \ldots, N$ mit $p_1, \ldots, p_N$. Diese Zahlen sollen eine *stationäre Verteilung* für die obige Markov-Kette bilden, d. h. es soll $\sum_{j=1}^{N} p_j = 1$ und

$$p_i = \sum_{j=1}^{N} (da_{ij} + (1-d)/N) p_j \quad \forall i = 1, \ldots, N \tag{VIII.2}$$

gelten.

Setzt man

$$G_{ij} = \frac{1-d}{N} + da_{ij}$$

für alle $i, j = 1, \ldots, N$, $G = (G_{ij})_{i,j=1}^{N,N}$ und $p = (p_1 \ \ldots \ p_n)^T$, so ist (VIII.2) äquivalent zu $Gp = p$. Es handelt sich also um ein Eigenwertproblem für die Matrix G (die sogenannte *Google-Matrix*). Der Parameter d wird auch *Dämpfungsfaktor* genannt. Typischerweise setzt man $d = 0,85$ (wir werden unten sehen, wo wir die Voraussetzung $d < 1$ brauchen).

Wir suchen also einen Eigenvektor für die Matrix G zum Eigenwert 1, der die zusätzlichen Bedingungen erfüllt, dass alle seine Koordinaten positiv sind und sich zu 1 aufsummieren.

Um zu beweisen, dass dieses Problem tatsächlich eine eindeutige Lösung besitzt, ist folgende Eigenschaft der Matrix G entscheidend: G ist eine *strikt positive, spaltenstochastische Matrix*, d. h. es gilt $G_{ij} > 0$ für alle $i, j = 1, \ldots, N$ und $\sum_{i=1}^{N} G_{ij} = 1$ für alle $j = 1, \ldots, N$.

Das sieht man wie folgt ein: Zunächst gilt (wegen $d > 0$ und $a_{ij} \geq 0$) $G_{ij} \geq (1-d)/N > 0$ (hier brauchen wir $d < 1$).

Weiter gilt für $j \in \{1, \ldots, N\}$

$$\sum_{i=1}^{N} G_{ij} = \sum_{i=1}^{N}((1-d)/N + da_{ij}) = \sum_{i=1}^{N}(1-d)/N + d\sum_{i=1}^{N} a_{ij}$$
$$= N(1-d)/N + d\sum_{i=1}^{N} a_{ij} = 1 - d + d\sum_{i=1}^{N} a_{ij}.$$

Ist $c_j > 0$, so gilt $\sum_{i=1}^{N} a_{ij} = \sum_{i \in L_j} 1/c_j$, wobei L_j die Menge aller $i \in \{1, \ldots, N\}$ bezeichnet, für die es einen Link von Seite j auf Seite i gibt. Die Anzahl der Elemente von L_j ist also gerade die Gesamtzahl der von Seite j ausgehenden Links, also c_j. Es folgt $\sum_{i=1}^{N} a_{ij} = c_j(1/c_j) = 1$.

Ist $c_j = 0$, so folgt $\sum_{i=1}^{N} a_{ij} = \sum_{i=1}^{N} 1/N = N/N = 1$.

Insgesamt folgt also $\sum_{i=1}^{N} G_{ij} = 1 - d + d\sum_{i=1}^{N} a_{ij} = 1 - d + d = 1$ für alle $j = 1, \ldots, N$.

Für strikt positive, spaltenstochastische Matrizen gibt es den folgenden *Satz von Perron-Frobenius.*

Satz VIII.3.1 (Satz von Perron-Frobenius)**.** *Es sei $S = (s_{ij})_{i,j=1}^{N,N}$ eine strikt positive, spaltenstochastische Matrix. Dann gilt:*

(a) *1 ist ein Eigenwert von S.*

(b) *Der Eigenraum $E_S(1)$ ist eindimensional.*

(c) *Es gibt genau ein $y = (y_1 \ \ldots \ y_n)^T \in E_S(1)$ mit $y_j > 0$ für alle $j = 1, \ldots, N$ und $\sum_{j=1}^{N} y_j = 1$.*

Beweis. (a) Es sei $v := (1 \ \ldots \ 1)^T$. Dann gilt nach Voraussetzung für den j-ten Eintrag von $S^T v$

$$(S^T v)_j = \sum_{i=1}^{N} s_{ij} = 1.$$

Also ist $S^T v = v$ und daher ist 1 ein Eigenwert von S^T, d. h. $\chi_{S^T}(1) = 0$. Es folgt $\chi_S(1) = \det(S - E_n) = \det((S - E_n)^T) = \det(S^T - E_n) = \chi_{S^T}(1) = 0$. Also ist 1 auch ein Eigenwert von S.

(b) und (c): Es sei $x = (x_1 \ \ldots \ x_n)^T \in E_S(1) \setminus \{0\}$ beliebig. Angenommen es gäbe $k, l \in \{1, \ldots, N\}$ mit $x_k > 0$ und $x_l < 0$. Dann gilt für alle $i = 1, \ldots, N$

$$|x_i| = \left| \sum_{j=1}^{N} s_{ij} x_j \right| < \sum_{j=1}^{N} s_{ij} |x_j|$$

(Ersteres wegen $Sx = x$, Letzteres wegen unserer Annahme und der Ungleichung $|a + b| < |a| + |b|$ für $a < 0 < b$, die Sie leicht als Übung beweisen können). Es folgt

$$\sum_{i=1}^{N} |x_i| < \sum_{i=1}^{N} \sum_{j=1}^{N} s_{ij}|x_j| = \sum_{j=1}^{N} \sum_{i=1}^{N} s_{ij}|x_j| = \sum_{j=1}^{N} |x_j| \sum_{i=1}^{N} s_{ij} = \sum_{j=1}^{N} |x_j|,$$

was natürlich ein Widerspruch ist.
Also gilt $x_j \geq 0$ für alle $j = 1, \ldots, N$ oder $x_j \leq 0$ für alle $j = 1, \ldots, N$.
Wegen $x \neq 0$ existiert weiterhin ein $j_0 \in \{1, \ldots, N\}$ mit $x_{j_0} \neq 0$. Da alle Koordinaten von x dasselbe Vorzeichen haben, folgt

$$|x_i| = \sum_{j=1}^{N} s_{ij}|x_j| \geq s_{ij_0}|x_{j_0}| > 0$$

für alle $i = 1, \ldots, N$.
Also gilt sogar $x_j > 0$ für alle $j = 1, \ldots, N$ oder $x_j < 0$ für alle $j = 1, \ldots, N$.
Insbesondere folgt, dass es einen Vektor $u = (u_1 \ \ldots \ u_n)^T \in E_S(1)$ mit $u_j > 0$ für alle $j = 1, \ldots, N$ gibt. Es sei $s := \sum_{j=1}^{N} u_j$ und $y = (y_1 \ \ldots \ y_n)^T := u/s \in E_S(1)$. Dann ist $y_j = u_j/s > 0$ für alle $j = 1, \ldots, N$ und $\sum_{j=1}^{N} y_j = 1$.
Ist $w = (w_1 \ \ldots \ w_n)^T \in E_S(1)$ beliebig, so ist auch $z := y_1 w - w_1 y \in E_S(1)$ und die erste Koordinate von z ist $y_1 w_1 - w_1 y_1 = 0$. Nach unserer obigen Überlegung muss also $z = 0$ und damit $w = (w_1/y_1)y$ gelten.
Es gilt also $\text{span}\{y\} = E_S(1)$ und somit $\dim(E_S(1)) = 1$.
Es bleibt nur noch die Eindeutigkeit von y zu zeigen. Ist auch $y' = (y'_1 \ \ldots \ y'_n)^T$ ein Element von $E_S(1)$ mit $y'_j > 0$ für alle $j = 1, \ldots, N$ und $\sum_{j=1}^{N} y'_j = 1$, so existiert ein $\lambda \in \mathbb{R}$ mit $y' = \lambda y$. Es folgt $1 = \sum_{j=1}^{N} y'_j = \lambda \sum_{j=1}^{N} y_j = \lambda$, also $y' = y$. □

Aus dem Satz von Perron-Frobenius folgt, dass der PageRank-Vektor existiert und eindeutig bestimmt ist. Offen bleibt damit allerdings noch das Problem der konkreten Berechnung des PageRanks. Es gibt hunderte Millionen von Seiten im Web, die Google-Matrix ist also viel zu groß, als dass eine Berechnung mit dem Gaußschen Algorithmus noch praktikabel wäre. Stattdessen bedient man sich numerischer Methoden zur approximativen (näherungsweisen) Lösung linearer Gleichungssysteme, siehe z. B. [7].

A Anhang

A.1 Logiksymbole

Wir wollen hier kurz die am häufigsten verwendeten Logiksymbole zusammenstellen und ihre Bedeutung klären. Dabei sollen $\mathcal{A}$ und $\mathcal{B}$ stets zwei mathematische Aussagen bezeichnen. Wir verwenden dann folgende Schreibweisen:

(i) $\mathcal{A} \wedge \mathcal{B}$ steht für die Aussage "$\mathcal{A}$ und $\mathcal{B}$".

(ii) $\mathcal{A} \vee \mathcal{B}$ steht für die Aussage "$\mathcal{A}$ oder $\mathcal{B}$" (im Sinne eines einschließenden oders, d. h. es gilt mindestens eine der beiden Aussagen $\mathcal{A}$, $\mathcal{B}$, eventuell auch beide).

(iii) $\mathcal{A} \Rightarrow \mathcal{B}$ steht für die Aussage "aus $\mathcal{A}$ folgt $\mathcal{B}$".

(iv) $\mathcal{A} \Leftrightarrow \mathcal{B}$ steht für die Aussage "$\mathcal{A}$ ist äquivalent zu $\mathcal{B}$" (auch gelesen als "$\mathcal{A}$ genau dann, wenn $\mathcal{B}$"). Das bedeutet definitionsgemäß "$\mathcal{A} \Rightarrow \mathcal{B}$ und $\mathcal{B} \Rightarrow \mathcal{A}$".

(v) $\neg\, \mathcal{A}$ steht für die Verneinung von $\mathcal{A}$ (gelesen als "nicht $\mathcal{A}$").

Die Symbole $\wedge, \vee, \neg, \Rightarrow, \Leftrightarrow$ werden auch *Junktoren* genannt. Hinzu kommen noch die sogenannten *Quantoren* $\forall$ und $\exists$, die wie folgt erklärt sind:

(vi) $\forall x\ \mathcal{A}$ bedeutet "für alle x gilt $\mathcal{A}$".

(vii) $\exists x\ \mathcal{A}$ bedeutet "es existiert (mindestens) ein x, für das $\mathcal{A}$ gilt".

Das Symbol $\forall$ heißt *Allquantor*, das Symbol $\exists$ wird *Existenzquantor* genannt. Häufig verwendet man diese Symbole auch in folgender Weise (wobei M eine vorgegebene Menge ist): $\forall x \in M\ \mathcal{A}$ bedeutet "für alle Elemente x der Menge M gilt $\mathcal{A}$" und $\exists x \in M\ \mathcal{A}$ steht für "es existiert (mindestens) ein Element $x \in M$, für welches $\mathcal{A}$ gilt".

Die Symbole $\wedge, \vee$ und $\neg$ werden wir in dieser Vorlesung eher selten oder gar nicht gebrauchen (stattdessen schreiben wir "und", "oder", "nicht" einfach aus), die Zeichen $\Rightarrow, \Leftrightarrow, \forall$ und $\exists$ werden wir dagegen des Öfteren zur Abkürzung verwenden.

Zum Abschluss ein paar konkrete Beispiele für den Gebrauch der Logiksymbole:

(i) $a > 0 \Leftrightarrow -a < 0$ bedeutet "$a > 0$ ist äquivalent zu $-a < 0$".

(ii) $((a < b) \wedge (b < c)) \Rightarrow (a < c)$ bedeutet "aus $a < b$ und $b < c$ folgt $a < c$".

(iii) $x \neq 0 \Rightarrow (x > 0 \vee x < 0)$ bedeutet "aus $x \neq 0$ folgt $x > 0$ oder $x < 0$".

(iv) $\forall x \in \mathbb{R}\ x^2 \geq 0$ bedeutet "für alle reellen Zahlen x ist $x^2 \geq 0$".

(v) $\forall x \in \mathbb{R}\, \exists n \in \mathbb{N}\ n > x$ bedeutet "für alle reellen Zahlen x existiert eine natürliche Zahl n mit $n > x$".

(vi) $\neg\, (\exists x \in \mathbb{Q}\ x^2 = 2)$ bedeutet "es existiert keine rationale Zahl x mit $x^2 = 2$".

A.2 Das griechische Alphabet

In der Mathematik (und auch in der Physik) werden häufig neben den lateinischen auch griechische Buchstaben zur Bezeichnung mathematischer (physikalischer) Größen verwendet. Das griechische Alphabet lautet wie folgt:

Name	Großbuchstabe	Kleinbuchstabe
Alpha	A	α
Beta	B	β
Gamma	Γ	γ
Delta	Δ	δ
Epsilon	E	ε oder ϵ
Zeta	Z	ζ
Eta	H	η
Theta	Θ	θ oder ϑ
Iota	I	ι
Kappa	K	κ
Lambda	Λ	λ
My	M	μ
Ny	N	ν
Xi	Ξ	ξ
Omikron	O	o
Pi	Π	π
Rho	P	ρ
Sigma	Σ	σ
Tau	T	τ
Ypsilon	Υ	υ
Phi	Φ	φ oder ϕ
Chi	X	χ
Psi	Ψ	ψ
Omega	Ω	ω

Literaturhinweise

Es gibt diverse einführende Lehrbücher zur Linearen Algebra. Hier eine kleine Auswahl:

[1] A. Beutelspacher, *Lineare Algebra*, Springer, Heidelberg, 2014. (8.Auflage).

[2] S. Bosch, *Lineare Algebra*, Springer, Heidelberg, 2014. (5.Auflage).

[3] G. Fischer, *Lineare Algebra*, Springer, Heidelberg, 2014. (18.Auflage).

[4] K. Jänich, *Lineare Algebra*, Springer, Berlin, 2008. (11.Auflage).

[5] H.-J. Kowalsky und G. O. Michler, *Lineare Algebra*, Walter de Gruyter, Berlin, 2003. (12.Auflage).

Zum Aufbau der Zahlenbereiche siehe etwa:

[6] H.-D. Ebbinghaus, H. Hermes, F. Hirzebruch, M. Koecher, K. Mainzer, J. Neukirch, A. Prestel, R. Remmert, *Zahlen*, Springer, Berlin–Heidelberg, 1992. (3.Auflage).

Zur numerischen Linearen Algebra:

[7] F. Bornemann, *Numerische Lineare Algebra*, Springer Spektrum, Wiesbaden, 2016.