Skript: Lineare Algebra I

Prof. Dr. Vladimir Lazić

(nach dem Mitschrieb von Niklas Schneider im Wintersemester 2019/20)

Inhaltsverzeichnis

1 Mengen und Abbildungen **1**
Funktionen . . . 6
Relationen . . . 12

2 Gruppen und Symmetrie **15**
Gruppenhomomorphismen . . . 19
Kanonische Projektion . . . 26
Permutationsgruppen . . . 30
Signatur . . . 32

3 Ringe und Körper **37**
Komplexe Zahlen . . . 41
Polynome . . . 45
Division mit Rest . . . 47
Nullstellen . . . 49

4 Vektorräume **60**
Austauschsätze von Steinitz . . . 71
Schnitt und Summe von Unterräumen . . . 75
Direkte Summe . . . 79
Quotientenräume . . . 81
Austauschsätze von Steinitz . . . 84

5 Lineare Abbildungen und Matrizen **85**
Matrizen und Matrixdarstellungen von linearen Abbildungen . . . 89
Matrixalgebra . . . 95
Rang einer Matrix . . . 99
Invertierbare Matrizen . . . 106
Lineare Gleichungssysteme . . . 111

6 Determinanten **118**
Cramersche Regel . . . 133
Basiswechsel . . . 134

7 Eigenwerte und Eigenvektoren **138**
Diagonalisierung . . . 143

8 Skalarprodukte **148**
Die Cauchy-Schwarz-Ungleichung 151
Orthogonalität und Orthonormalität 153
Isometrien 162
Spektralsätze 166

1 Mengen und Abbildungen

Definition 1.1. Eine **Menge** M ist eine Ansammlung wohlbestimmter Objekte.

Beispiel 1.2.

(1) Die Menge aller Buchstaben im englischen Alphabet:

$$A := \{A, a, B, b, \dots, Z, z\}.$$

(Hier sprechen wir „:=“ aus als „ist definiert durch“.)

(2) Die Menge aller Hörer dieser Vorlesung:

$$B := \{\text{alle Hörer dieser Vorlesung}\}.$$

(3) Alle Primzahlen, die kleiner als zehn sind:

$$C := \{2, 3, 5, 7\}.$$

Bemerkung 1.3.

(1) Die Reihenfolge der Elemente spielt keine Rolle:

$$C = \{3, 2, 7, 5\} = \{2, 7, 5, 3\}.$$

(2) Die Objekte einer Menge M nennt man *Elemente von M*.

Beispiel 1.4. Wir haben die folgenden speziellen Mengen:

(1) Die *leere Menge* $\{\} = \emptyset$.

(2) Die *natürlichen Zahlen* $\mathbb{N} = \{0, 1, 2, 3, \dots\}$.

(3) Die *ganzen Zahlen* $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}$.

(4) Die *reellen Zahlen* $\mathbb{R} = \{\text{unendliche Dezimalzahlen}\}$.

(5) Die *komplexen Zahlen* $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$.

Bemerkung 1.5.

(1) Wenn ein Element a zu einer Menge A gehört, dann schreiben wir

$$a \in A \quad \text{oder} \quad A \ni a.$$

(2) Wenn a nicht zu einer Menge A gehört, dann schreiben wir

$$a \notin A \quad \text{oder} \quad A \not\ni a.$$

(3) Seien M und N zwei Mengen. Wenn jedes Element von N auch ein Element von M ist, dann sagen wir, dass N eine *Untermenge* oder *Teilmenge von M* ist und schreiben

$$N \subset M \quad \text{oder} \quad N \subseteq M.$$

(4) Eine Teilmenge $N \subseteq M$ heißt *echte Teilmenge*, wenn $N \neq \emptyset$ und $N \neq M$. Dann schreiben wir auch

$$N \subsetneq M.$$

(5) Wenn N keine Teilmenge von M ist, so schreiben wir

$$N \not\subseteq M.$$

Definition 1.6. Sei M eine Menge. Die Anzahl von Elementen von M bezeichnen wir mit

$$|M| \quad \text{oder} \quad \#M.$$

Diese Zahl heißt *Kardinalität* oder *Mächtigkeit* von M.

Beispiel 1.7.

(1) $|\emptyset| = 0$

(2) $|\{\text{Buchstaben des englischen Alphabets}\}| = 26$

(3) $|\{0\}| = 1$

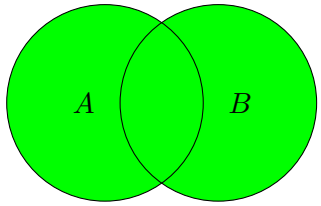
(4) Hat M unendlich viele Elemente, so schreiben wir $|M| = \infty$.

Definition 1.8. Seien A und B Mengen. Dann ist die *Vereinigung* von A und B die Menge

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

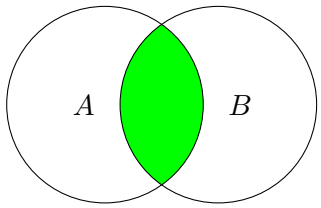
Für Mathematiker bedeutet „oder“ *nicht* „exklusiv oder“, d.h. es gilt:

$$A \cup B = \{x \text{ gehört zu } A, \text{ zu } B \text{ oder zu den beiden Mengen}\}.$$



Der *Durchschnitt* von A und B ist die Menge

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

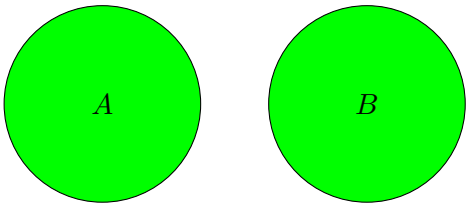


Gilt $A \cap B = \emptyset$, dann sind A und B *disjunkt*. Dann schreiben wir für $A \cup B$ auch

$$A \sqcup B,$$

und nennen diese Vereinigung *disjunkt*. Für eine disjunkte Vereinigung gilt

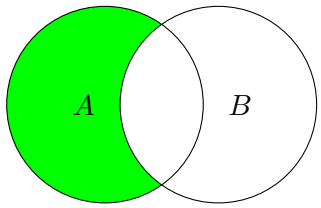
$$|A \sqcup B| = |A| + |B|.$$



Die *Differenz* von A und B ist die Menge

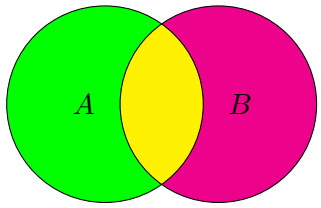
$$A \setminus B := \{x \mid x \in A \text{ und } x \notin B\}.$$

Wir sprechen dies aus als „A minus B“ und schreiben manchmal auch $A - B$.



Damit lässt sich die Vereinigung ausdrücken als

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



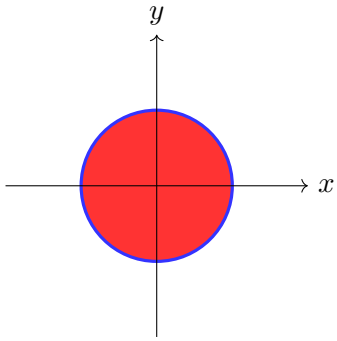
Beispiel 1.9. Betrachte die folgenden Mengen:

(1) $\delta_1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$,

(2) $\delta_2 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$,

(3) $\delta_3 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$.

Dann gilt $\delta_1 = \delta_2 \sqcup \delta_3$.



Lemma 1.10. *Seien A, B und C Mengen. Dann gelten die* Distributivgesetze*:*

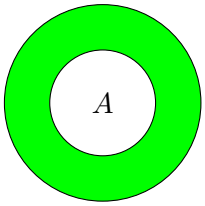
(1) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,

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

Definition 1.11. Betrachte eine Menge A, die Teilmenge einer bestimmten Menge M ist. Dann schreiben wir

$$\overline{A}$$

für die Menge $M \setminus A$. Dies sprechen wir aus als „A Komplement“ und schreiben auch A^c.



Lemma 1.12. *Seien A und B Teilmengen einer Menge M. Dann gelten die* Gesetze von De Morgan*:*

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B}, \quad \overline{\overline{A}} = A.$$

Definition 1.13. Seien A und B Mengen. Dann bezeichnet

$$\begin{aligned} A \times B :&= \{(a,b) \mid a \in A,\, b \in B\} \\ &= \{\text{alle geordneten Paare } (a,b) \text{ mit } a \in A,\, b \in B\} \end{aligned}$$

das *(kartesische) Produkt von A und B.*

Bemerkung 1.14. Wenn $A = B$ in der obigen Definition, dann schreiben wir auch

$$A^2 := A \times A,$$

und in ähnlicher Weise:

$$A^n := \underbrace{A \times A \times \cdots \times A}_{n \text{ mal}}.$$

Beispiel 1.15. $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(a,b) \mid a, b \in \mathbb{R}\}$.

Bemerkung 1.16. Probleme mit der Definition einer Menge:

(1) Sei L die Liste aller Sachen, die ich in meiner Tasche habe, zum Beispiel

$$L = \{\text{Kulis, Papier, Taschentücher}\}.$$

Wenn ich nun L in meine Tasche stecke, dann enthält die Tasche die Liste L, das heißt

$$L = \{\text{Kulis, Papier, Taschentücher, } L\}.$$

Das heißt, nach unserer Definition können Mengen sich selbst enthalten.

(2) *Russelsche Antimonie*
Sei M die Menge aller Mengen, die sich selbst nicht enthalten. Frage: Ist $M \in M$?
Wenn $M \in M$, dann gilt $M \notin M$ nach Definition von M. Wenn $M \notin M$, dann gilt $M \in M$ nach Definition von M. Fazit: M ist keine Menge. In der Mengenlehre löst man dieses Problem, indem man den Begriff *Familie* einführt.

Definition 1.17. Die Menge aller Teilmengen einer Menge M nennt man die *Potenzmenge von* M, und schreibt

$$\mathcal{P}(M).$$

Satz 1.18. *Sei M eine endliche Menge. Dann gilt*

$$|\mathcal{P}(M)| = 2^{|M|}.$$

Beweis. Übung! □

Funktionen

Definition 1.19. Seien M und N zwei Mengen. Eine *Abbildung* oder eine *Funktion* zwischen M und N ordnet *jedem* Element in M *genau ein* Element aus N zu. Wenn wir eine Zuordnung durch f bezeichnen, dann schreiben wir

$$f\colon M \to N, \quad x \mapsto f(x).$$

Hier heißt M *die Urmenge* und N *die Zielmenge* von f. Die Menge aller Abbildungen von M nach N bezeichnen wir mit $\mathrm{Abb}(M, N)$.

Beispiel 1.20. Seien $M = \{a, b, c, d\}$ und $N = \{1, 3\}$. Eine Funktion von M nach N können wir beispielsweise durch Zuordnung von Elementen definieren:

$$(a \mapsto 3,\ b \mapsto 3,\ c \mapsto 1,\ d \mapsto 3).$$

Eine Funktion kann auch durch eine Formel gegeben werden:

$$f\colon \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^2 + 1.$$

Notation 1.21. Sei M eine Menge. Die Abbildung

$$\mathrm{id}_M\colon M \to M, \quad m \to m$$

heißt die *Identität auf* M.

Bemerkung 1.22. Eine Funktion $f\colon M \to N$ heißt *wohldefiniert*, falls für jedes $m \in M$ gilt $f(m) \in N$ und für jedes $M \in M$ genau ein $n \in N$ existiert, sodass $n = f(m)$.

Beispiel 1.23. Es gibt keine Funktion $f\colon \mathbb{R} \to \mathbb{N}$ mit $f(x) = x^2 + 1$, da $f\left(\frac{5}{2}\right) \notin \mathbb{N}$.

Definition 1.24. Sei $f\colon M \to N$ eine Funktion und $A \subseteq M$ eine Teilmenge von M. Dann heißt die Menge

$$f(A) := \big\{x \in N \mid \exists\, y \in A \text{ mit } x = f(y)\big\} = \{f(x) \mid x \in A\}$$

das *Bild von* A.[1] Für eine Teilmenge $B \subseteq N$ heißt die Menge

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

das *Urbild von* B. Für eine einelementige Teilmenge $\{y\} \subseteq N$ schreiben wir auch $f^{-1}(y)$ statt $f^{-1}(\{y\})$, und nennen wir die Menge $f^{-1}(y)$ die *Faser von* f *über* y.

Beispiel 1.25. Im Beispiel 1.20 gilt:

$$f(M) = \{1, 3\}, \qquad f^{-1}(1) = \{c\}, \qquad f^{-1}(3) = \{a, b, d\}.$$

Definition 1.26. Sei $f\colon M \to N$ eine Abbildung und sei $A \subseteq M$ eine Teilmenge von M. Wir bezeichnen mit

$$f\,|_A\colon A \to N$$

(und lesen „f eingeschränkt auf A“) die Abbildung

$$A \ni x \mapsto f(x) \in N.$$

Definition 1.27. Sei $f\colon M \to N$ eine Abbildung. Die Menge

$$\Gamma_f \subseteq M \times N$$

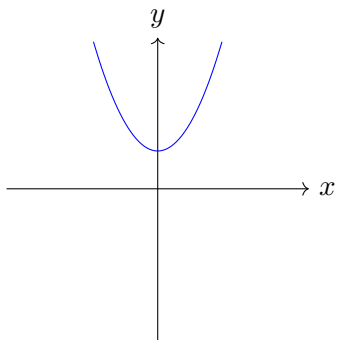
gegeben durch

$$\Gamma_f := \{(x, y) \in M \times N \mid y = f(x)\} = \{(x, f(x)) \mid x \in M\}$$

ist der *Graph von* f.

[1] Wir benutzen zwei *Quantoren*: $\forall$ = „für alle“ und $\exists$ = „es existiert“.

Beispiel 1.28. Der Graph von $f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto x^2 + 1$ ist



Definition 1.29. Sei $f\colon M \to N$ eine Abbildung. Dann heißt f:

(1) *injektiv*, falls für alle $x, y \in M$ mit $x \neq y$ gilt $f(x) \neq f(y)$;

(2) *surjektiv*, falls für jedes $n \in N$ ein $m \in M$ existiert, sodass $f(m) = n$;

(3) *bijektiv*, falls f surjektiv und injektiv ist.

Eine unendliche Menge M heißt *abzählbar*, falls es eine bijektive Abbildung $f\colon \mathbb{N} \to M$ gibt; ansonsten heißt M *überabzählbar*.

Beispiel 1.30. Im Beispiel 1.20 ist die Abbildung f weder injektiv noch surjektiv, da $f^{-1}(3) = \{a, b, d\}$ und $f(M) = \{1, 3\}$.

Lemma 1.31. *Die Menge $\mathbb{R}$ ist überabzählbar.*

Beweis. Angenommen, es existiert eine bijektive Abbildung $f\colon \mathbb{N} \to \mathbb{R}$. Wir werden einen Widerspruch herleiten. Die Abbildung f lässt sich darstellen als

$$\begin{aligned} f(0) &= n_0, d_{0,1}d_{0,2}d_{0,3}\dots \\ f(1) &= n_1, d_{1,1}d_{1,2}d_{1,3}\dots \\ f(2) &= n_2, d_{2,1}d_{2,2}d_{2,3}\dots \\ f(3) &= n_3, d_{3,1}d_{3,2}d_{3,3}\dots \\ &\dots \end{aligned}$$

wobei für alle $i, j \in \mathbb{N}$ wir haben $n_j \in \mathbb{Z}$ und $d_{j,i} \in \{0, 1, 2, \dots, 9\}$. Für jedes $i \in \mathbb{N}$ setze

$$r_i := \begin{cases} r_i = 5, & \text{falls } d_{i,i} = 4, \\ 4, & \text{sonst}, \end{cases}$$

und definiere

$$r := 0, r_1 r_2 r_3 \ldots$$

Dann ist r eine reelle Zahl, die nicht in der obigen Liste steht, denn r unterscheidet sich von $f(i)$ an der i-ten Stelle nach dem Komma, für jedes $i \in \mathbb{N}$. Das ist ein Widerspruch. $\square$

Bemerkung 1.32. Ein mögliches Problem im obigen Beweis: Eine reelle Zahl kann zwei unterschiedliche Darstellungen haben, zum Beispiel $0,945\overline{9} = 0,946$. Aber dies ist auch die einzige Möglichkeit, also wenn *fast alle Ziffern gleich* 9 *sind*. Im obigen Beweis tritt dieses Problem nicht auf, da r per Konstruktion nur aus 4 und 5 besteht.

Notation 1.33. Sei I eine Indexmenge, und für jedes $i \in I$ sei A_i eine Menge. Setze:

$$\bigcup_{i \in I} A_i := \{x \mid \exists i \in I \text{ mit } x \in A_i\},$$
$$\bigcap_{i \in I} A_i := \{x \mid \forall i \in I \text{ gilt } x \in A_i\}.$$

Satz 1.34. *Sei $f\colon M \to N$ eine Abbildung zwischen endlichen Mengen M und N mit $|M| = |N|$. Dann sind äquivalent:*

(a) f ist injektiv,

(b) f ist surjektiv,

(c) f ist bijektiv.

Beweis. Die Implikationen $(c) \Rightarrow (a)$ und $(c) \Rightarrow (b)$ folgen aus der Definition.

$(a) \Rightarrow (b)$: Sei f injektiv. Für alle $n_1, n_2 \in \mathbb{N}$ mit $n_1 \neq n_2$, die Mengen $f^{-1}(n_1)$ und $f^{-1}(n_2)$ sind disjunkte Teilmengen von M (per Definition einer Funktion); und in ähnlicher Weise gilt $\bigcup_{n \in N} f^{-1}(n) = M$. Daraus folgt:

$$M = \bigsqcup_{n \in N} f^{-1}(n).$$

Damit gilt

$$|M| = \sum_{n \in N} |f^{-1}(n)|,$$

und deswegen:

$$|M| = \sum_{n \in N} \underbrace{|f^{-1}(n)|}_{\leq 1} \overset{f \text{ inj.}}{\leq} \sum_{n \in N} 1 = |N|.$$

Aber wir wissen, dass $|M| = |N|$, und damit $\sum_{n \in N} |f^{-1}(n)| = \sum_{n \in N} 1$. Die impliziert $|f^{-1}(n)| = 1$ für alle $n \in N$, und damit ist f surjektiv.

$(a) \Leftarrow (b)$: Sei f surjektiv. Ähnlich wie oben haben wir:

$$|M| = \sum_{n \in N} \underbrace{|f^{-1}(n)|}_{\geq 0} \overset{f \text{ sur.}}{\geq} \sum_{n \in N} 1 = |N|.$$

Aber wir wissen, dass $|M| = |N|$, und damit $\sum_{n \in N} |f^{-1}(n)| = \sum_{n \in N} 1$. Dies impliziert $|f^{-1}(n)| = 1$ für alle $n \in N$, und damit ist f injektiv.

Dies zeigt $(a) \Leftrightarrow (b)$, und somit sind die fehlenden Implikationen bewiesen. □

Korollar 1.35. *Seien M und N endliche Mengen. Ist $f\colon M \to N$ eine injektive Abbildung, so gilt*

$$|M| \leq |N|.$$

Ist f eine surjektive Abbildung, so gilt

$$|M| \geq |N|.$$

Ist f eine bijektive Abbildung, so gilt

$$|M| = |N|.$$

Beweis. Folgt aus dem Beweis von Satz 1.34. □

Korollar 1.36 (Schubfachprinzip)**.** *Eine Abbildung $f\colon M \to N$ mit $|M| > |N|$ ist nicht injektiv.*

Beweis. Das ist die Kontraposition zum obigen Korollar. □

Beispiel 1.37. Gegeben seien $n^2 + 1$ Punkte im folgenden $n \times n$ Quadrat:

$$Q = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x, y \leq n\}.$$

Dann existieren zwei Punkte $a, b \in Q$, sodass der Abstand zwischen a und b kleiner oder gleich $\sqrt{2}$ ist.

Beweis. Wir unterteilen Q in n^2 Einheitsquadrate. Diese bezeichnen wir mit $Q_{i,j}$, wobei $1 \leq i, j \leq n$. Die $n^2 + 1$ Punkte bezeichnen wir mit P_i, wobei $1 \leq i \leq n^2 + 1$. Wir definieren eine Abbildung

$$f\colon \{1, 2, \ldots, n^2 + 1\} \to \{Q_{i,j} \mid 1 \leq i, j \leq n\}$$

wie folgt: Sei $k \in \{1, 2, \ldots, n^2 + 1\}$. Dann:

(a) falls es nur ein Quadrat $Q_{i,j}$ gibt, zu welchem der Punkt P_k gehört, dann setzen wir $f(k) := Q_{i,j}$,

(b) falls es mehrere Quadrate gibt, zu welchem der Punkt P_k gehört, dann wählen wir das Quadrat $f(k)$ so, dass es am linkesten und am untersten liegt.

Da

$$|\{1, 2, \ldots, n^2 + 1\}| = n^2 + 1 \quad \text{und} \quad |\{Q_{i,j} \mid 1 \leq i, j \leq n\}| = n^2,$$

so gibt es nach dem Schubfachprinzip zwei Werte $p, q \in \{1, 2, \ldots, n^2 + 1\}$, sodass $f(p) = f(q)$. In anderen Worten, die Punkte P_p und P_q liegen im selben Einheitsquadrat. Das impliziert, dass der Abstand zwischen P_p und P_q kleiner oder gleich der Länge der Diagonalen des Quadrats ist, also $\leq \sqrt{2}$. $\square$

Definition 1.38. Sei $f\colon M \to N$ eine bijektive Abbildung. Die Funktion

$$f^{-1}\colon N \to M, \quad n \mapsto m,$$

wobei m das eindeutige Element in M ist, für welches $f(m) = n$ gilt, heißt die *Umkehrfunktion* zu der Funktion f.

Definition 1.39. Seien $f\colon A \to B$ und $g\colon B \to C$ zwei Abbildungen. Die Funktion

$$g \circ f\colon A \to C, a \mapsto g(f(a))$$

(ausgesprochen „g nach f“) heißt die *Komposition* oder die *Verknüpfung* von f und g.

Proposition 1.40. *Seien $f\colon A \to B$ und $g\colon B \to C$ zwei Abbildungen. Dann gilt:*

(1) $g \circ f$ injektiv $\Longrightarrow f$ injektiv,

(2) $g \circ f$ surjektiv $\Longrightarrow g$ surjektiv.

Beweis.
(1) Seien $a_1, a_2 \in A$ mit $f(a_1) = f(a_2)$. Dann gilt

$$(g \circ f)(a_1) = g(f(a_1)) = g(f(a_2)) = (g \circ f)(a_2).$$

Da $g \circ f$ injektiv ist, so folgt $a_1 = a_2$, und daher ist f injektiv.

(2) Sei $c \in C$. Da $g \circ f$ surjektiv ist, so existiert ein $a \in A$ mit $(g \circ f)(a) = c$, und damit $g(f(a)) = c$. Definiere $b := f(a)$. Dann gilt $g(b) = c$, und deswegen ist g surjektiv. $\square$

Relationen

Definition 1.41. Seien A und B Mengen. Eine *Relation R zwischen A und B* ist eine Teilmenge von $A \times B$. Wenn $(a, b) \in R$, dann schreiben wir

$$a \, R \, b$$

(ausgesprochen: „a ist in der Relation R zu b“). Die Menge A ist die *Quellmenge* oder der *Vorbereich* von R. Die Menge B ist die *Zielmenge* oder der *Nachbereich* von R.

Beispiel 1.42. Sei $A = \{\text{Barack, Michelle}\} = B$. Betrachten wir die folgende Relation R in $A \times B$:

$$\forall (a, b) \in A \times B : \quad a \, R \, b \Longleftrightarrow a \text{ und } b \text{ sind verheiratet.}$$

Dann gilt: Barack R Michelle, aber Barack $\not R$ Barack.

Definition 1.43. Wenn in der obigen Definition $A = B$ (also die Ziel- und die Quellmenge sind gleich), dann heißt die Relation R *homogen*, und wir sprechen über eine *Relation auf der Menge A*. Homogene Relationen bezeichnen wir normalerweise mit $\sim$.

Eine homogene Relation $\sim$ ist:

(1) *reflexiv*, wenn $a \sim a$ für alle $a \in A$,

(2) *symmetrisch*, wenn aus $a \sim b$ folgt $b \sim a$ für alle $a, b \in A$,

(3) *transitiv*, wenn aus $a \sim b$ und $b \sim c$ folgt, dass $a \sim c$ für alle $a, b, c \in A$,

(4) *antisymmetrisch*, wenn aus $a \sim b$ und $b \sim a$ folgt, dass $a = b$ für alle $a, b \in A$,

(5) eine *Äquivalenzrelation*, wenn $\sim$ reflexiv, symmetrisch und transitiv ist.

Beispiel 1.44. Betrachte die folgenden Relationen $\sim$ auf $\mathbb{N}$:

$x \sim y$	reflexiv	symmetrisch	antisymmetrisch	transitiv
$x \leq y$	ja	nein	ja	ja
x teilt y	ja	nein	ja	ja
$x + y = 7$	nein	ja	nein	nein

Definition 1.45.

(1) Seien $a, b \in \mathbb{Z}$. Wir sagen, dass die Zahl a die Zahl b *teilt* und schreiben

$$a \mid b,$$

falls es ein $k \in \mathbb{Z}$ gibt, sodass $a = nk$.

(2) Sei $n \in \mathbb{N}$. Zwei Zahlen $a, b \in \mathbb{Z}$ heißen *kongruent modulo* n, und wir schreiben

$$a \equiv b \pmod{n},$$

falls $n \mid a - b$.

Proposition 1.46. *Sei $n \in \mathbb{N}$. Dann definiert*

$$a \sim b \quad \Longleftrightarrow \quad a \equiv b \pmod{n}$$

eine Äquivalenzrelation $\sim$ auf $\mathbb{Z}$.

Beweis.
Reflexivität: Es gilt $a \sim a$ für alle $a \in \mathbb{Z}$, weil $n \mid \underbrace{a - a}_{=0}$.
Symmetrie: Sei $a \sim b$ für $a, b \in \mathbb{Z}$. Dann gilt $n \mid a - b$, und damit $n \mid b - a$, also $b \sim a$.
Transitivität: Seien $a \sim b$ und $b \sim c$ für $a, b, c \in \mathbb{Z}$. Dann existieren $k, \ell \in \mathbb{Z}$, sodass $a - b = nk$ und $b - c = n\ell$, und damit:

$$a - c = (a - b) + (b - c) = nk + n\ell = n(k + \ell).$$

Es gilt also $n \mid a - c$, und damit $a \sim c$. □

Definition 1.47. Sei M eine Menge, sei $\sim$ eine Äquivalenzrelation auf M und sei $m \in M$. Die Teilmenge

$$[m] := \{x \in M \mid x \sim m\} \subseteq M$$

heißt die *Äquivalenzklasse von m unter der Relation $\sim$ auf M*. Manchmal bezeichnen wir diese Äquivalenzklasse auch mit $\overline{m}$. Wir bezeichnen mit

$$M/\sim$$

die Menge aller Äquivalenzklassen von $\sim$ auf M (ausgesprochen: „M modulo $\sim$“).

Satz 1.48. *Sei $\sim$ eine Äquivalenzrelation auf einer Menge M. Dann bilden die Äquivalenzklassen eine* Partition *von M, das heißt*

(1) für alle $x, y \in M$ gilt entweder $[x] = [y]$ oder $[x] \cap [y] = \emptyset$, und

(2) $\bigcup_{x \in M} [x] = M$.

Beweis.
(1) Angenommen, $[x] \cap [y] \neq \emptyset$. Dann müssen wir zeigen, dass $[x] = [y]$.
Fixiere ein z aus $[x] \cap [y]$. Sei $a \in [x]$ ein beliebiges Element. Dann gilt $a \sim x$ und $x \sim z$, und damit folgt aus der Transitivität, dass

$$a \sim z.$$

Außerdem folgt aus $z \in [y]$, dass $z \sim y$. Damit folgt wieder aus der Transitivität, dass $a \sim y$; also $a \in [y]$. Dies zeigt, dass

$$[x] \subseteq [y].$$

In ähnlicher Weise zeigen wir, dass $[y] \subseteq [x]$, und damit $[x] = [y]$.

(2) Für alle $x \in M$ gilt $x \in [x]$ (Reflexivität), und somit

$$M \supseteq \bigcup_{x \in M} [x] \supseteq \bigcup_{x \in M} \{x\} = M.$$

Dies zeigt die gewünschte Gleichheit. □

Beispiel 1.49.

(1) Eine Uhr ist ein Beispiel für eine Äquivalenzrelation auf $\mathbb{Z}$, und zwar, die Relation ist $\equiv \pmod{12}$.

(2) Sei $n \in \mathbb{N}$. Die Äquivalenzklassen der Relation $\equiv \pmod{n}$ auf $\mathbb{Z}$ sind

$$[0], [1], [2], \ldots, [n-1].$$

Und zwar, die Division mit Rest besagt, dass für jedes $a \in \mathbb{Z}$ eindeutige Zahlen $q \in \mathbb{Z}$ und $0 \leq r \leq n \leq n-1$ existieren, sodass $a = nq + r$. Damit ist $[a] = [r]$, also

$$\mathbb{Z} = \bigcup_{r=0}^{n-1} [r].$$

In ähnlicher Weise zeigt man, dass für alle $0 \leq r_1, r_2 \leq n-1$ mit $r_1 \neq r_2$ gilt $[r_1] \neq [r_2]$ (Übung!). Damit bilden die Klassen $[0], [1], [2], \ldots, [n-1]$ eine Partition von $\mathbb{Z}$ bezüglich der Relation $\equiv \pmod{n}$.

2 Gruppen und Symmetrie

Definition 2.1. Ein Paar $(G, \cdot)$, bestehend aus einer Menge G zusammen mit einer Verknüpfung

$$\cdot \colon G \times G \to G, \quad (a, b) \mapsto a \cdot b,$$

heißt eine *Gruppe* mit der *Gruppenoperation* $\cdot$, wenn die folgenden Axiome erfüllt sind:

(G1) (*Assoziativgesetz*) Für alle $a, b, c \in G$ gilt:

$$(a \cdot b) \cdot c = a \cdot (b \cdot c).$$

(G2) (Existenz eines *neutralen Elements*) Es existiert ein $e \in G$, sodass für alle $a \in G$ gilt:

$$a \cdot e = a.$$

Das Element e heißt ein *neutrales Element* oder ein *Neutral*.

(G3) (Existenz von *Inversen*) Für jedes $a \in G$ gibt es ein $a' \in G$, sodass:

$$a \cdot a' = e.$$

Angenommen, es gilt darüber hinaus noch:

(G4) (*Kommutativgesetz*) Für alle $a, b \in G$ gilt:

$$a \cdot b = b \cdot a.$$

Dann ist G eine *kommutative Gruppe* oder eine *abelsche Gruppe*.

Beispiel 2.2.

(1) $(\mathbb{Z}, +)$ ist eine abelsche Gruppe, mit dem Neutral 0, und die Inverse von m ist $-m$ für jedes $m \in \mathbb{Z}$.

(2) $(\mathbb{Q} \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe, mit dem Neutral 1, und die Inverse von m ist $\frac{1}{m}$ für jedes $m \in \mathbb{Q} \setminus \{0\}$.

(3) $(\mathbb{N}, +)$ ist keine Gruppe, da es keine Inversen gibt.

(4) $(\mathbb{Z} \setminus \{0\}, \cdot)$ ist keine Gruppe, da es keine Inversen gibt.

Definition 2.3. Sei $n \in \mathbb{N}_{>0}$ und $M = \{1, 2, \ldots, n\}$. Betrachte die Menge

$$S_n = \{\sigma \colon M \to M \mid \sigma \text{ ist eine Bijektion}\}$$

mit der Operation $\circ$ (also die Verknüpfung von Abbildungen). Dann ist $(S_n, \circ)$ eine Gruppe: und zwar, es gilt:

(G1) $\circ$ ist assoziativ;

(G2) die Abbildung $\mathrm{id}_M \colon M \to M,\ m \mapsto m$ ist das Neutral;

(G3) Für jedes $\sigma \in S_n$ ist die Abbildung σ^{-1} die Inverse von σ.

Die Elemente von S_n heißen *Permutationen* von M, und S_n heißt die *Permutationsgruppe* auf einer Menge von n Elementen. Die Elemente $\sigma \in S_n$ können wir auch durch *Wertetabellen* schreiben:

$$\sigma \colon \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

Notation 2.4. Das Neutral id_M wie in der Definition wird oft einfach mit id bezeichnet und heißt die *Identität in S_n*.

Beispiel 2.5. Betrachte die folgenden Elemente in S_4:

$$\sigma \colon \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}, \quad \tau \colon \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}.$$

Dann gilt:

$$\sigma^{-1} \colon \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}, \quad \tau^{-1} \colon \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}.$$

Ferner gilt:

$$\sigma \circ \tau \colon \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}, \quad \tau \circ \sigma \colon \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix},$$

und insbesondere $\sigma \circ \tau \neq \tau \circ \sigma$. Damit ist S_4 nicht kommutativ. Im Allgemeinen ist S_n für jedes $n \geq 3$ keine abelsche Gruppe.

Satz 2.6. *Für jedes $n \in \mathbb{N}$ gilt $|S_n| = n!$.*

Beweis. Man kann dies mithilfe von Induktion beweisen; ein anderer Beweis lautet:

Für $\sigma(1)$ haben wir n Möglichkeiten. Nachdem wir $\sigma(1)$ ausgewählt haben, haben wir für $\sigma(2)$ noch $n-1$ Möglichkeiten. Im Allgemeinen: Nachdem wir die ersten $k-1$ Werte ausgewählt haben, haben wir für $\sigma(k)$ noch $n-k+1$ Möglichkeiten. Damit ist $|S_n| = n \cdot (n-1) \cdot \ldots \cdot 1 = n!$. $\square$

Bemerkung 2.7. Sei M eine Menge. Die Menge $\mathrm{Abb}(M, M)$ aller Abbildungen zwischen M und M hat die Mächtigkeit $|M|^{|M|}$.

Satz 2.8. *Sei $(G, \cdot)$ eine Gruppe. Dann gelten die folgenden Aussagen.*

(a) Ein neutrales Element e erfüllt auch $e \cdot a = a$.

(b) Das neutrale Element ist eindeutig bestimmt.

(c) Ein inverses Element $a' \in G$ von $a \in G$ erfüllt auch $a' \cdot a = e$.

(d) Das inverse Element eines Elements $a \in G$ ist eindeutig bestimmt. Wir schreiben a^{-1} statt a'.[2]

(e) Für Elemente $a_1, \ldots, a_n \in G$ ist das Element $a_1 \cdot a_2 \cdot \ldots \cdot a_n \in G$ eindeutig bestimmt, egal wie man die Klammern in diesem Produkt setzt.

Beweis.
(c) Da a' ein inverses Element zu a ist, so gilt $a \cdot a' = e$. Ferner, nach $(G3)$ gibt es ein $a'' \in G$, sodass

$$a' \cdot a'' = e. \tag{1}$$

Es folgt:

$$\begin{aligned} a' \cdot a &\overset{(G2)}{=} (a' \cdot a) \cdot e \overset{(1)}{=} (a' \cdot a) \cdot (a' \cdot a'') \\ &\overset{(G1)}{=} a' \cdot (a \cdot (a' \cdot a'')) \overset{(G1)}{=} a' \cdot ((a \cdot a') \cdot a'') \overset{(G3)}{=} a' \cdot (e \cdot a'') \\ &\overset{(G1)}{=} (a' \cdot e) \cdot a'' \overset{(G2)}{=} a' \cdot a'' \overset{(1)}{=} e. \end{aligned}$$

(a) Sei a' ein inverses Element zu a. Dann:

$$e \cdot a \overset{(G3)}{=} (a \cdot a') \cdot a \overset{(G1)}{=} a \cdot (a' \cdot a) \overset{(c)}{=} a \cdot e \overset{(G2)}{=} a.$$

(b) Sei $\widetilde{e}$ ein weiteres neutrales Element. Dann gilt:

$$e \overset{(G2) \text{ für } \widetilde{e}}{=} e \cdot \widetilde{e} \overset{(a)}{=} \widetilde{e}.$$

(d) Sei a' ein inverses Element zu a, und sei $\widetilde{a}$ ein weiteres inverses Element zu a. Nach (c) gilt

$$\widetilde{a} \cdot \ a = e. \tag{2}$$

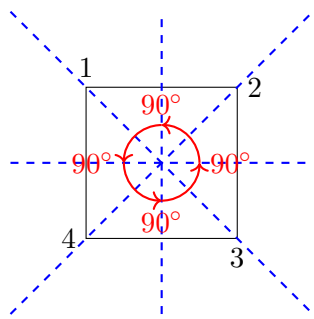
[2] Wenn die Operation auf G mit $+$ bezeichnet wird, dann schreiben wir $-a$ für die Inverse von a.

Dann folgt:

$$\widetilde{a} \overset{(G2)}{=} \widetilde{a} \cdot e \overset{(G3)}{=} \widetilde{a} \cdot (a \cdot a') \overset{(G1)}{=} (\widetilde{a} \cdot a) \cdot a' \overset{(2)}{=} e \cdot a' \overset{(a)}{=} a'.$$

(e) Übung! $\square$

Beispiel 2.9. Mit D_4 bezeichnen wir die folgende Symmetriegruppe des Quadrats:



Die Menge D_4 enthält vier Spiegelungen und vier Drehungen (0°, 90°, 180°, 270°). Das neutrale Element ist die Drehung um 0°.

Sei σ die Drehung um 90°:

$$\sigma\colon \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix},$$

und sei τ die Spiegelung an der y-Achse:

$$\tau\colon \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}.$$

Dann gilt

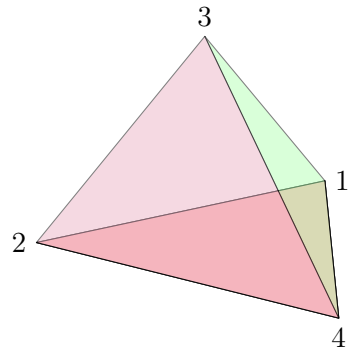
$$\tau \circ \sigma\colon \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix} = \text{ Spiegelung an der Diagonale } (1,3)$$

und

$$\sigma \circ \tau\colon \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} = \text{Spiegelung an der Diagonale } (2,4).$$

Insbesondere ist D_4 nicht abelsch. Die Gruppe D_4 ist eine *Untergruppe* von S_4. (Siehe Definition 2.15.)

Beispiel 2.10. Sei T ein Tetraeder:



Übung: Finde alle Elemente der Symmetriegruppe von T geometrisch, und zeige, dass diese Gruppe S_4 ist.

Beispiel 2.11. Gruppen lassen sich auch durch *Verknüpfungstafeln* darstellen. Als Beispiel: Sei $V_4 = \{e, a, b, c\}$ die folgende *Kleinsche Vierergruppe* mit der Verknüpfungstafel:

$\cdot$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Aus der Tafel kann man folgendes sehen:

- $x^2 = e$ für jedes $x \in V_4$;
- e ist das neutrale Element;
- V_4 ist abelsch.

Gruppenhomomorphismen

Definition 2.12. Seien $(G, \cdot_G)$ und $(H, \cdot_H)$ zwei Gruppen. Ein *Gruppenhomomorphismus* von G nach H ist eine Abbildung $\varphi\colon G \to H$, sodass gilt:

$$\forall a, b \in G: \quad \varphi(a \cdot_G b) = \varphi(a) \cdot_H \varphi(b).$$

In anderen Worten: φ ist verträglich mit den Gruppenoperationen $\cdot_G$ und $\cdot_H$ auf G und H.

Zusätzlich ist φ ein:

(a) *Monomorphismus*, wenn φ *injektiv* ist;

(b) *Epimorphismus*, wenn φ *surjektiv* ist;

(c) *Isomorphismus*, wenn φ *bijektiv* ist. In diesem Falle schreiben wir auch $G \cong H$ oder $G \simeq H$. (Ausgesprochen: „G ist isomorph zu H“.)

Satz 2.13. *Seien $(G, \cdot_G)$ und $(H, \cdot_H)$ zwei Gruppen und $\varphi\colon G \to H$ ein Gruppenhomomorphismus. Seien e_G bzw. e_H die neutralen Elemente in G bzw. H. Dann gilt:*

(a) $\varphi(e_G) = e_H$,

(b) $\varphi(a^{-1}) = \varphi(a)^{-1}$ *für jedes* $a \in G$.

Beweis.
(a) Es gilt

$$\varphi(e_G) = \varphi(e_G \cdot e_G) = \varphi(e_G) \cdot \varphi(e_G), \tag{3}$$

und damit

$$\begin{aligned} e_H = \varphi(e_G) \cdot \varphi(e_G)^{-1} &\overset{(3)}{=} \big(\varphi(e_G) \cdot \varphi(e_G)\big) \cdot \varphi(e_G)^{-1} \\ &= \varphi(e_G) \cdot (\underbrace{\varphi(e_G) \cdot \varphi(e_G)^{-1}}_{e_H}) = \varphi(e_G) \cdot e_H = \varphi(e_G). \end{aligned}$$

(b) Es gilt

$$\varphi(a) \cdot \varphi(a^{-1}) = \varphi(a \cdot a^{-1}) = \varphi(e_G) \overset{(a)}{=} e_H,$$

und damit $\varphi(a^{-1}) = \varphi(a)^{-1}$, wie gewünscht. □

Beispiel 2.14. Sei

$$\exp_a\colon (\mathbb{Z}, +) \to (\mathbb{R}_{>0}, \cdot), \quad n \mapsto a^n,$$

wobei $\mathbb{R}_{>0} = \{x \in \mathbb{R} \mid x > 0\}$ und $a \in \mathbb{R}_{>0} \setminus \{1\}$. Dann ist $\exp_a$ ein Gruppenhomomorphismus, da

$$\exp_a(n_1 + n_2) = a^{n_1+n_2} = a^{n_1} \cdot a^{n_2} = \exp_a(n_1) \cdot \exp_a(n_2).$$

(In der mathematischen Analysis wird dieser Homomorphismus zu einem Homomorphismus $\exp_a\colon (\mathbb{R}, +) \to (\mathbb{R}_{>0}, \cdot)$ fortgesetzt.)

Definition 2.15. Sei $(G, \cdot)$ eine Gruppe und H eine Teilmenge von G. Dann heißt H eine *Untergruppe* von G, wenn:

(1) für alle $a, b \in H$ gilt $a \cdot b \in H$, und

(2) $(H, \cdot)$ bildet eine Gruppe.

Wir schreiben dann auch $H \leq G$.

Beispiel 2.16.

(a) Die Menge aller geraden Zahlen (häufig auch bezeichnet mit $2\mathbb{Z}$) ist eine Untergruppe von $(\mathbb{Z}, +)$.

(b) $(\mathbb{Z}, +)$ ist eine Untergruppe von $(\mathbb{R}, +)$.

(c) Seien H_1 und H_2 zwei Untergruppen einer Gruppe G. Dann ist auch $H_1 \cap H_2$ eine Untergruppe von G. (Übung!)

(d) Sei G eine Gruppe und $A \subseteq G$ eine Teilmenge. Dann ist

$$H = \bigcap_{A \subseteq \widetilde{H} \leq G} \widetilde{H}$$

eine Untergruppe von G (Übung!). Dies ist auch die kleinste Untergruppe, die A enthält. Sie nennen wir auch die *von A erzeugte Untergruppe*, und schreiben

$$H := \langle A \rangle.$$

Wenn $A = \{g_1, \ldots, g_n\}$, dann schreiben wir für diese Gruppe auch $\langle g_1, \ldots, g_n \rangle$.

Beispiel 2.17. Seien $(G, \cdot_G)$ und $(H, \cdot_H)$ zwei Gruppen. Dann ist $G \times H$ eine Gruppe mit der wie folgend definierten Operation $\cdot$ auf $G \times H$:

$$(a, b) \cdot (c, d) := (a \cdot_G c, b \cdot_H d) \quad \text{für alle } a, c \in G, b, d \in H.$$

Das neutrale Element in $G \times H$ ist (e_G, e_H), und es gilt $(a, b)^{-1} = (a^{-1}, b^{-1})$ für alle $(a, b) \in G \times H$.

Satz 2.18. *Sei $(G, \cdot)$ eine Gruppe. Eine nichtleere Teilmenge $H \subseteq G$ ist eine Untergruppe von G genau dann, wenn*

$$\forall a, b \in H : \quad a \cdot b^{-1} \in H. \tag{4}$$

Beweis.
$\Longrightarrow$: Sei H eine Untergruppe von G. Dann ist die Abbildung

$$\theta\colon H \to G, \quad h \mapsto h$$

ein Homomorphismus. Deswegen stimmen, nach Satz 2.13, die neutralen sowie die inversen Elemente in diesen Gruppen überein. Es gilt also:

$$a, b \in H \Longrightarrow a, b^{-1} \in H,$$

und damit ist $a \cdot b^{-1} \in H$ nach der Definition einer Untergruppe.

$\Longleftarrow$: Die Menge H ist nicht leer, und fixiere ein $h \in H$. Damit ist nach (4) (für $a = b = h$):

$$e_G = h \cdot h^{-1} \in H. \tag{5}$$

Für jedes $g \in H$ gilt nach (4) und (5):

$$g^{-1} = e_G \cdot g^{-1} \in H. \tag{6}$$

Seien nun $c, d \in H$. Nach (6) gilt dann $c, d^{-1} \in H$, und dann folgt aus (4):

$$c \cdot (d^{-1})^{-1} \in H.$$

Es gilt also $c \cdot d \in H$, welches zeigt, dass H eine Untergruppe von G ist. $\square$

Satz/Definition 2.19. *Sei $\varphi\colon G \to H$ ein Gruppenhomomorphismus. Der* Kern *von φ ist die Menge*

$$\ker\varphi := \{x \in G \mid \varphi(x) = e_H\} = \varphi^{-1}(e_H),$$

und das Bild von φ ist die Menge $\operatorname{Im}\varphi := \varphi(G) \subseteq H$. *Es gilt:*

(a) der Kern $\ker\varphi$ *ist eine Untergruppe von G;*

(b) das Bild $\operatorname{Im}\varphi$ *ist eine Untergruppe von H.*

Beweis.
(a) Sei $b \in \ker\varphi$; äquivalent: $\varphi(b) = e_H$. Dann folgt:

$$\varphi(b^{-1}) \overset{\text{Satz 2.13}}{=} \varphi(b)^{-1} = e_H^{-1} = e_H,$$

und damit

$$b^{-1} \in \ker\varphi. \tag{7}$$

Seien nun $a, b \in \ker \varphi$. Nach (7) gilt dann $a, b^{-1} \in \ker \varphi$, also $\varphi(a) = \varphi(b^{-1}) = e_H$. Es folgt:

$$\varphi(a \cdot b^{-1}) = \varphi(a) \cdot \varphi(b^{-1}) = e_H \cdot e_H = e_H,$$

also $a \cdot b^{-1} \in \ker \varphi$. Nach Satz 2.18 zeigt dies, dass $\ker \varphi$ eine Untergruppe von G ist.

(b) Seien $a, b \in \operatorname{Im} \varphi$. Dann existieren $x, y \in G$ mit $\varphi(x) = a$ und $\varphi(y) = b$, und damit

$$ab^{-1} = \varphi(x)\varphi(y)^{-1} \overset{\text{Satz 2.13}}{=} \varphi(xy^{-1}).$$

In anderen Worten, $ab^{-1} \in \operatorname{Im} \varphi$, welches nach Satz 2.18 zeigt, dass $\operatorname{Im} \varphi$ eine Untergruppe von H ist. □

Definition 2.20. Sei $(G, \cdot)$ eine Gruppe und sei H eine Untergruppe von G. Sei $a \in G$. Die Menge

$$aH := \{a \cdot x \mid x \in H\}$$

heißt die *linke Nebenklasse* von H bezüglich a. Die Menge

$$Ha := \{x \cdot a \mid x \in H\}$$

heißt *rechte Nebenklasse* von H bezüglich a. Die Menge aller linken Nebenklassen von H wird mit

$$G/H$$

bezeichnet; die Menge aller rechten Nebenklassen wird mit

$$H \backslash G$$

bezeichnet. (Ausgesprochen: „G modulo H“.)

Lemma 2.21. *Sei $(G, \cdot)$ eine Gruppe und sei H eine Untergruppe von G. Seien $a, b \in G$. Dann gilt:*

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

Beweis.
$\Longrightarrow$: Es gilt:

$$a = ae_G \in aH = bH = \{bx \mid x \in H\}.$$

Damit existiert $x \in H$ mit $a = bx$. Es folgt:

$$b^{-1}a = b^{-1}(bx) = (b^{-1}b)x = x \in H.$$

$\Longleftarrow$: Es reicht zu zeige, dass $aH \subseteq bH$ und $bH \subseteq aH$.

Sei $x := b^{-1}a \in H$. In anderen Worten, $a = bx$. Es folgt:

$$\begin{aligned} aH &= \{ay \mid y \in H\} = \{(bx)y \mid y \in H\} = \{b\underbrace{(xy)}_{\in H} \mid x, y \in H\} \\ &\subseteq \{bz \mid z \in H\} = bH, \end{aligned}$$

also $aH \subseteq bH$. Um die umgekehrte Inklusion zu zeigen, bemerken wir, dass $(b^{-1}a)^{-1} \in H$, da H eine Gruppe ist. Aus $(b^{-1}a)^{-1} = a^{-1}b$ folgt also, dass

$$a^{-1}b \in H.$$

Dann zeigen wir, dass $bH \subseteq aH$, indem wir im obigen Beweis a und b vertauschen. $\square$

Lemma 2.22. *Sei $(G, \cdot)$ eine Gruppe und sei H eine Untergruppe von G. Seien $a, b \in G$. Dann gilt:*

$$aH = bH \Longleftrightarrow aH \cap bH \neq \emptyset.$$

Beweis.
$\Longrightarrow$: Trivial.

$\Longleftarrow$: Sei $y \in aH \cap bH$. Dann existieren $x \in H$ und $z \in H$, sodass

$$y = ax \quad \text{und} \quad y = bz.$$

Es folgt also

$$ax = bz,$$

und damit

$$b^{-1}a = b^{-1}(ax)x^{-1} = b^{-1}(bz)x^{-1} = zx^{-1} \in H.$$

Es folgt aus Satz 2.21, dass $aH = bH$. $\square$

Lemma 2.23. *Sei $(G, \cdot)$ eine endliche Gruppe und sei H eine Untergruppe von G. Dann gilt für alle $a, b \in G$:*

$$|aH| = |bH|.$$

Beweis. Betrachte die Abbildung

$$\varphi\colon aH \to bH, \quad y \mapsto ba^{-1}y.$$

Zuerst zweigen wir, dass die Abbildung φ wohldefiniert ist. Und zwar, da $y \in aH$, so existiert $x \in H$ mit $y = ax$, und damit $ba^{-1}y = bx \in bH$.

Ferner ist φ bijektiv, da die Abbildung

$$\psi\colon bH \to aH, \quad z \mapsto ab^{-1}z$$

die inverse Abbildung zu φ ist. Daraus folgt, dass $|aH| = |bH|$. □

Definition 2.24. Sei $(G, \cdot)$ eine Gruppe und sei H eine Untergruppe von G. Die Mächtigkeit der Menge G/H heißt der *Index* von H und wird mit

$$[G : H]$$

bezeichnet.

Satz 2.25 (Satz von Lagrange)**.** *Sei $(G, \cdot)$ eine endliche Gruppe und sei H eine Untergruppe von G. Dann gilt*

$$|G| = |H| \cdot [G : H].$$

Beweis. Nach Satz 2.22 bilden die Elemente von G/H eine Partition von G, d.h.

$$G = \bigsqcup_{aH \in G/H} aH.$$

Damit gilt

$$|G| = \sum_{aH \in G/H} |aH|.$$

Diese Summe hat $|G/H|$ Summanden und jeder Summand hat $|H|$ Elemente nach Satz 2.23. Es folgt:

$$|G| = |G/H| \cdot |H| = [G : H] \cdot |H|,$$

wie gewünscht. □

Korollar 2.26. *Sei $(G, \cdot)$ eine endliche Gruppe und sei H eine Untergruppe von G. Dann teilt die Zahl $|H|$ die Zahl $|G|$.*

Definition 2.27. Sei $(G, \cdot)$ eine Gruppe. Die Zahl $|G|$ heißt die *Ordnung* von G. Sei zudem $g \in G$ und sei e das neutrale Element von G. Die Zahl

$$\inf\{k \in \mathbb{N}_{>0} \mid g^k = e\}^3$$

heißt die *Ordnung* von g in G, und wird mit $\operatorname{ord}(g)$ bezeichnet. Es gilt also

$$\operatorname{ord}(g) \in \mathbb{N}_{>0} \cup \{\infty\}.$$

Kanonische Projektion

Seien $(G, \cdot)$ und $(H, \cdot)$ Gruppen und sei $\theta\colon G \to H$ ein Gruppenhomomorphismus. Dann ist $\ker\theta$ eine Untergruppe von G nach Satz 2.19, also ist die Menge $G/\ker\theta$ wohldefiniert. Wir werden zeigen, dass diese Menge tatsächlich eine Gruppe ist.

Wenn $g \in G$, dann schreiben wir $[g]$ für das Element $g \cdot \ker\theta \in G/\ker\theta$.

Zuerst brauchen wir das folgende Lemma.

Lemma 2.28. *Seien $(G, \cdot)$ und $(H, \cdot)$ Gruppen und sei $\theta\colon G \to H$ ein Gruppenhomomorphismus. Für jedes $g \in G$ gilt*

$$g \cdot \ker\theta = \ker\theta \cdot g.$$

Beweis. Betrachte die Menge

$$M := g \cdot \ker\theta \cdot g^{-1} := \{gcg^{-1} \mid c \in \ker\theta\}.$$

Es reicht zu zeigen (Übung!), dass $M = \ker\theta$.

$\subseteq$: Sei $z \in M$. Dann gibt es ein $c \in \ker\theta$ mit $z = gcg^{-1}$, und es gilt:

$$\theta(z) = \theta(gcg^{-1}) = \theta(g)\underbrace{\theta(c)}_{=e_H}\theta(g^{-1}) = \theta(g)\theta(g^{-1}) = \theta(g)\theta(g)^{-1} = e_H,$$

damit $z \in \ker\theta$. Es gilt also $M \subseteq \ker\theta$.

$\supseteq$: Sei $c \in \ker\theta$. Setze $z := g^{-1}cg$, und bemerke, dass $z \in \ker\theta$ (wie im ersten Teil des Beweises). Aber dann $c = gzg^{-1} \in M$, welches zeigt, dass $\ker\theta \subseteq M$. $\square$

[3] Hier ist $g^k = \underbrace{g \cdot \ldots \cdot g}_{k \text{ Mal}}$

Wir definieren nun auf $G/\ker\theta$ eine Gruppenoperation. Und zwar, seien $[x], [y] \in G/\ker\theta$, und setze

$$\underbrace{[x]\cdot[y]}_{\text{auf } G/\ker\theta} := [\underbrace{x\cdot y}_{\text{auf } G}].$$

Diese Verknüpfung ist wohldefiniert, d.h. sie ist unabhängig von der Wahl der Elemente x und y in G. Und zwar, seien $x, x', y, y' \in G$ mit $[x] = [x']$ und $[y] = [y']$. Dann müssen wir zeigen, dass $[x]\cdot[y] = [x']\cdot[y']$, d.h., dass

$$[xy] = [x'y'].$$

Wir zeigen dies nun. Aus $[x] = [x']$ und $[y] = [y']$ folgt, dass es $a, b \in \ker\theta$ gibt mit $x' = xa$ und $y' = yb$ (dies folgt aus Lemma 2.21). Damit gilt

$$x'y' = (xa)(yb) = x(ay)b. \tag{8}$$

Da $ay \in \ker\theta\cdot y$, und da $y\cdot\ker\theta = \ker\theta\cdot y$ nach Lemma 2.28, so gibt es $a' \in \ker\theta$ mit

$$ay = ya'.$$

Es folgt dann aus (8):

$$(xy)^{-1}x'y' = (xy)^{-1}x(ay)b = (xy)^{-1}x(ya')b = (xy)^{-1}xya'b = a'b \in \ker\theta.$$

Nach Lemma 2.21 zeigt dies, dass $[xy] = [x'y']$, wie gewünscht.

Wir haben also gezeigt, dass $G/\ker\theta$ mit der oben definierten Operation eine Gruppe ist. Es gibt den *kanonischen* oder *natürlichen* Gruppenhomomorphismus

$$\pi\colon G \to G/\ker\theta, \quad g \mapsto [g].$$

Diese Abbildung π heißt auch die *Projektion* von G auf $G/\ker\theta$.

Satz 2.29 (Homomorphiesatz für Gruppen)**.** *Sei $\theta\colon G \to H$ ein Gruppenhomomorphismus. Dann ist $G/\ker\theta$ eine Gruppe mit der kanonischen Projektion $\pi\colon G \to G/\ker\theta$. Es gibt einen wohl-definierten Gruppenhomomorphismus*

$$\xi\colon G/\ker\theta \to H, \quad [g] \mapsto \theta(g),$$

und das Diagramm

$$\begin{tikzcd} G \arrow[rr, "\theta"] \arrow[dr, "\pi"'] & & H \\ & G/\ker\theta \arrow[ur, "\xi"'] & \end{tikzcd}$$

kommutiert, d.h. $\theta = \xi \circ \pi$. Die Abbildung ξ induziert den Isomorphismus

$$\rho\colon G/\ker\theta \to \operatorname{Im}\theta, \quad [g] \mapsto \xi([g]) = \theta(g).$$

Beweis. Wir haben bereits gezeigt, dass $G/\ker\theta$ eine Gruppe ist.

Wir zeigen als Nächstes, dass die Abbildung ξ wohldefiniert ist: in anderen Worten, wenn $[g] = [g']$ für $g, g' \in G$, dann $\xi([g]) = \xi([g'])$, oder äquivalent: $\theta(g) = \theta(g')$. Und zwar, nach Lemma 2.21 gilt $g^{-1}g' \in \ker\theta$, oder äquivalent,

$$e_H = \theta(g^{-1}g') = \theta(g)^{-1}\theta(g').$$

Es gilt also $\theta(g) = \theta(g')$, wie gewünscht.

Per Definition von ξ folgt $\theta = \xi \circ \pi$.

Zuletzt müssen wir zeigen, dass ρ ein Isomorphismus ist. Injektivität: Angenommen, $\rho([g]) = \rho([g'])$ für $g, g' \in G$. Dann gilt $\theta(g) = \theta(g')$ per Definition von ρ, und damit wie oben:

$$e_H = \theta(g^{-1}g').$$

Dies ist äquivalent zu $g^{-1}g' \in \ker\theta$, und dann $[g] = [g']$ nach Lemma 2.21. Surjektivität: Sei $x \in \operatorname{Im}\theta$. Dann existiert $z \in G$ mit $\theta(z) = x$, und damit

$$x = \theta(z) = (\xi \circ \pi)(z) = \xi(\pi(z)) = \xi([z]).$$

Der Satz ist bewiesen. □

Beispiel 2.30.

(a) Die Menge $\mathbb{Z}/\sim$, wobei $\sim$ die Relation $\equiv \pmod{n}$ ist, ist eine Gruppe mit der Operation

$$[a] + [b] := [a + b] \quad \text{für alle } a, b \in \mathbb{Z}$$

(Übung!). Diese Gruppe wird mit $\mathbb{Z}_n$ oder mit $\mathbb{Z}/n\mathbb{Z}$ bezeichnet (siehe auch (b) unten). Die Elemente von $\mathbb{Z}_n$ sind die Klassen $[0], [1], \ldots, [n-1]$.

(b) Betrachte die Abbildung

$$\varphi\colon \mathbb{Z} \to \mathbb{Z}_n, \quad a \mapsto [a].$$

Dann ist φ ein Gruppenepimorphismus und

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

Nach dem Homomorphiesatz folgt, dass

$$\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n.$$

Wenn wir den Beweis von Satz 2.29 noch einmal lesen, dann sehen wir, dass die (fast) einzige Eigenschaft von $\ker\theta$, die wir benutzt haben, ist

$$g \cdot \ker\theta = \ker\theta \cdot g \quad \text{für alle } g \in G.$$

Dies führt zur folgenden Definition.

Definition 2.31. Sei $(G, \cdot)$ eine Gruppe und H eine Untergruppe von G. Dann ist H ein *Normalteiler* oder eine *normale Untergruppe*, falls für alle $g \in G$ gilt

$$gH = Hg,$$

oder äquivalent

$$gHg^{-1} = H.$$

Man zeigt (Übung!), dass G/H eine Gruppe ist, mit der Operation

$$(g_1H) \cdot (g_2H) := (g_1g_2)H \quad \text{für alle } g_1, g_2 \in G.$$

Diese Gruppe heißt die *Quotientengruppe* von G und H.

Beispiel 2.32.

(a) Der Kern eines beliebigen Gruppenhomomorphismus ist ein Normalteiler.

(b) Jede Untergruppe einer abelschen Gruppe ist normal.

Beispiel 2.33. Sei G eine Gruppe und sei $g \in G$. Welche Gruppe ist $\langle g \rangle$ (bis auf Isomorphie)? Wir unterscheiden zwei Fälle:

(a) $\operatorname{ord}(g) = \infty$.
Betrachte die Abbildung

$$\varphi\colon (\mathbb{Z}, +) \to (\langle g \rangle, \cdot), \quad k \mapsto g^k.$$ [4]

Die Abbildung ist ein Gruppenhomomorphismus (Übung!), und sie ist offensichtlich surjektiv. Außerdem ist sie injektiv: für $k \neq \ell$ gilt $g^k \neq g^\ell$. Und zwar, ansonsten wäre

$$g^{k-\ell} = g^{-\ell}g^k = g^{-\ell}g^\ell = e_G,$$

und damit $k = \ell$, da $\operatorname{ord}(g) = \infty$. Die Abbildung φ ist daher ein Isomorphismus.

[4] Hier bedeutet $g^0 := e_G$ und $g^{-k} := (g^k)^{-1} = \underbrace{g^{-1} \cdot \ldots \cdot g^{-1}}_{k \text{ mal}}$.

(b) $n := \operatorname{ord}(g) \in \mathbb{N}_{>0}$.
Betrachte wieder die Abbildung

$$\varphi\colon (\mathbb{Z}, +) \to (\langle g \rangle, \cdot), \quad k \mapsto g^k.$$

Dann ähnlich wie oben kann man zeigen, dass $\ker \varphi = n\mathbb{Z}$ (Übung!). Es folgt nach dem Homomorphiesatz, dass

$$\langle g \rangle \cong \mathbb{Z}/n\mathbb{Z}.$$

Lemma 2.34. *Sei G eine endliche Gruppe und sei $g \in G$. Dann gilt*

$$\operatorname{ord}(g) \mid |G|.$$

Beweis. Sei $n := \operatorname{ord}(g)$. Nach Beispiel 2.32 gilt

$$|\langle g \rangle| = |\mathbb{Z}/n\mathbb{Z}| = n,$$

und nach dem Satz von Lagrange haben wir $|\langle g \rangle| \mid |G|$. Das Lemma folgt. $\square$

Permutationsgruppen

Definition 2.35. Sei $n \in \mathbb{N}_{>0}$ und seien $i_1, i_2, \ldots, i_k \in \{1, \ldots, n\}$ paarweise verschieden. Dann bezeichnet

$$\tau := \begin{pmatrix} i_1 & i_2 & \cdots & i_k \end{pmatrix}$$

die Permutation $\tau \in S_n$ gegeben durch

$$\tau(m) = \begin{cases} i_{j+1}, & \text{wenn } m = i_j \text{ für ein } 1 \leq j \leq k-1, \\ i_1, & \text{wenn } m = i_k, \\ m, & \text{sonst.} \end{cases}$$

Dann ist τ ein *k-Zyklus* oder *Zyklus der Länge k*.

Beispiel 2.36.

(a) $S_5 \ni \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$.

(b) $S_5 \ni \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 \end{pmatrix}$.

(c) $S_5 \ni \rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 5 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \circ \begin{pmatrix} 4 & 5 \end{pmatrix} = \begin{pmatrix} 4 & 5 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$
$\Rightarrow \rho$ ist *kein* Zyklus.

Im Allgemeinen gilt: Für disjunkte Teilmengen $\{i_1, \dots, i_k\}$ und $\{j_1, \dots, j_l\}$ von $\{1, \dots, n\}$ gilt

$$\begin{pmatrix} i_1 & \cdots & i_k \end{pmatrix} \circ \begin{pmatrix} j_1 & \cdots & j_l \end{pmatrix} = \begin{pmatrix} j_1 & \cdots & j_l \end{pmatrix} \circ \begin{pmatrix} i_1 & \cdots & i_k \end{pmatrix},$$

d.h. zwei disjunkte Zyklen kommutieren. (Übung!)

Definition 2.37. Eine *Transposition* in S_n ist ein 2-Zyklus. In anderen Worten, $\tau \in S_n$ ist eine Transposition genau dann, wenn es zwei verschiedene $i, j \in \{1, \dots, n\}$ gibt, sodass

$$\tau(i) = j, \quad \tau(j) = i \quad \text{und} \quad \tau(k) = k \text{ für } k \neq i, j.$$

Lemma 2.38. *Ist $n \geq 2$, so ist jede Permutation in S_n das Produkt von Transpositionen. In anderen Worten, für jedes $\sigma \in S_n$ existieren Transpositionen $\tau_1, \dots \tau_k \in S_n$, sodass*

$$\sigma = \tau_1 \circ \cdots \circ \tau_k.$$

Beweis. Sei $\sigma \in S_n$. Angenommen, $\sigma = \mathrm{id}$. Sei τ eine beliebige Transposition. Dann gilt $\sigma = \tau \circ \tau$. Wir dürfen also annehmen, dass $\sigma \neq \mathrm{id}$.

Dann gibt es $i_1 \in \{1, \dots, n\}$ mit der Eigenschaft, dass

$$\sigma(i) = i \text{ für alle } i < i_1 \quad \text{und} \quad \sigma(i_1) \neq i_1.$$

Insbesondere gilt $\sigma(i_1) > i_1$. Definiere die Transposition

$$\tau_1 := \begin{pmatrix} i_1 & \sigma(i_1) \end{pmatrix} \in S_n,$$

und bezeichne

$$\sigma_1 := \tau_1 \circ \sigma.$$

Dann gilt für jedes $i < i_1$:

$$\sigma_1(i) = (\tau_1 \circ \sigma)(i) = \tau_1(\sigma(i)) = \tau_1(i) = i,$$

und es gilt

$$\sigma_1(i_1) = (\tau_1 \circ \sigma)(i_1) = \tau_1(\sigma(i_1)) = i_1.$$

Damit fixiert die Permutation $\sigma_1 \in S_n$ mindestens die Werte $1, 2, \dots, i_1$.

Angenommen, $\sigma_1 = \mathrm{id}$. Dann gilt

$$\tau_1^{-1} = \tau_1^{-1} \circ \mathrm{id} = \tau_1^{-1} \circ \sigma_1 = \tau_1^{-1} \circ \tau_1 \circ \sigma = \sigma.$$

Da $\tau_1^{-1} = \tau_1$ eine Transposition ist, so folgt, dass dann σ auch eine Transposition ist.

Wir dürfen also annehmen, dass $\sigma_1 \neq \mathrm{id}$. Nach der Diskussion oben gibt es $i_2 > i_1$, sodass

$$\sigma_1(i) = i \text{ für alle } i < i_2 \quad \text{und} \quad \sigma_1(i_2) \neq i_2.$$

Insbesondere gilt $\sigma_1(i_2) > i_2$. Definiere die Transposition

$$\tau_2 := \begin{pmatrix} i_2 & \sigma_1(i_2) \end{pmatrix} \in S_n,$$

und bezeichne

$$\sigma_2 := \tau_2 \circ \sigma_1.$$

Wie oben zeigt man, dass $\sigma_2(i) = i$ für alle $i \leq i_2$.

Angenommen, $\sigma_2 = \mathrm{id}$. Da $\sigma_2 = \tau_2 \circ \sigma_1 = \tau_2 \circ \tau_1 \circ \sigma$, so zeigt man ähnlich wie oben, dass dann

$$\sigma = (\tau_2 \circ \tau_1)^{-1} = \tau_1^{-1} \circ \tau_2^{-1} = \tau_1 \circ \tau_2.$$

Wir dürfen also annehmen, dass $\sigma_2 \neq \mathrm{id}$. Wir setzen diesen Prozess fort. Der Prozess muss aufhören, da $i_1 < i_2 < \cdots \leq n$. $\square$

Signatur

Definition 2.39. Sei $\sigma \in S_n$. Die *Signatur* oder das *Signum* von σ ist definiert durch die Formel

$$\mathrm{sign}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Satz 2.40. *Sei $n \in \mathbb{N}_{>0}$. Dann gilt $\mathrm{sign}(\sigma) \in \{-1, 1\}$ für jede Permutation $\sigma \in S_n$. Ferner ist die Abbildung*

$$\mathrm{sign}\colon (S_n, \circ) \to (\{-1, 1\}, \cdot)$$

ein Gruppenhomomorphismus. In anderen Worten, für alle $\sigma, \tau \in S_n$ gilt

$$\mathrm{sign}(\sigma \circ \tau) = \mathrm{sign}(\sigma) \cdot \mathrm{sign}(\tau).$$

Beweis. Es gilt

$$\operatorname{sign}(\sigma) = \prod_{1\le i<j\le n} \frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\prod\limits_{1\le i<j\le n}(\sigma(j)-\sigma(i))}{\prod\limits_{1\le i<j\le n}(j-i)}.$$

Da σ eine Permutation ist, kommen im Nenner und im Zähler die gleichen Zahlen vor, nur eventuell mit anderem Vorzeichen. Damit ist $\operatorname{sign}(\sigma) \in \{-1, 1\}$.

Nun beweisen wir die zweite Aussage. Für alle $\sigma, \tau \in S_n$ gilt

$$\begin{aligned}
\operatorname{sign}(\sigma\circ\tau) &= \prod_{1\le i<j\le n} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{j-i} \\
&= \frac{\prod\limits_{1\le i<j\le n}(\sigma(\tau(j))-\sigma(\tau(i)))}{\prod\limits_{1\le i<j\le n}(j-i)} \cdot \underbrace{\frac{\prod\limits_{1\le i<j\le n}(\tau(j)-\tau(i))}{\prod\limits_{1\le i<j\le n}(\tau(j)-\tau(i))}}_{=1} \\
&= \underbrace{\frac{\prod\limits_{1\le i<j\le n}(\sigma(\tau(j))-\sigma(\tau(i)))}{\prod\limits_{1\le i<j\le n}(\tau(j)-\tau(i))}}_{=:C} \cdot \underbrace{\frac{\prod\limits_{1\le i<j\le n}(\tau(j)-\tau(i))}{\prod\limits_{1\le i<j\le n}(j-1)}}_{=:D}.
\end{aligned}$$

Per Definition ist $D = \operatorname{sign}(\tau)$. Es bleibt zu zeigen, dass $C = \operatorname{sign}(\sigma)$.

Und zwar, es gilt

$$\begin{aligned}
C &= \frac{\prod\limits_{1\le i<j\le n}(\sigma(\tau(j))-\sigma(\tau(i)))}{\prod\limits_{1\le i<j\le n}(\tau(j)-\tau(i))} \\
&= \prod_{\substack{1\le i<j\le n\\ \tau(i)<\tau(j)}} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} \cdot \prod_{\substack{1\le i<j\le n\\ \tau(j)<\tau(i)}} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} \\
&= \prod_{\substack{1\le i<j\le n\\ \tau(i)<\tau(j)}} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} \cdot \prod_{\substack{1\le j<i\le n\\ \tau(i)<\tau(j)}} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)}.
\end{aligned}$$

Da $\tau\colon \{1,\dots,n\} \to \{1,\dots,n\}$ eine Bijektion ist, so ist es einfach zu sehen, dass das letzte Produkt gleich

$$\prod_{1\le k<\ell\le n} \frac{\sigma(\ell)-\sigma(k)}{\ell-k},$$

und das ist genau die Signatur von σ, welches es zu beweisen gab. $\square$

Bemerkung 2.41. In dieser Bemerkung zeigen wir, wie man die Signatur effektiv berechnen kann.

Sei $\sigma \in S_n$ eine Permutation. Ein Paar $(i,j) \in \{1,\dots,n\} \times \{1,\dots,n\}$ heißt ein *Fehlstand* von σ, falls

$$i < j \qquad \text{und} \qquad \sigma(i) > \sigma(j).$$

Sei k die Anzahl der Fehlstände von σ. Dann behaupten wir, dass

$$\operatorname{sign}(\sigma) = (-1)^k. \tag{9}$$

In anderen Worten:

$$\operatorname{sign}(\sigma) = \begin{cases} 1, & \text{falls } \sigma \text{ eine gerade Anzahl der Fehlstände hat,} \\ -1, & \text{falls } \sigma \text{ eine ungerade Anzahl der Fehlstände hat.} \end{cases}$$

Um (9) zu zeigen, berechnen wir:

$$\prod_{1\le i<j\le n} (\sigma(j)-\sigma(i)) = \underbrace{\prod_{\substack{1\le i<j\le n \\ \sigma(j)>\sigma(i)}} (\sigma(j)-\sigma(i))}_{=:A} \cdot \underbrace{\prod_{\substack{1\le i<j\le n \\ \sigma(j)<\sigma(i)}} (\sigma(j)-\sigma(i))}_{=:B},$$

und bemerken, dass

$$A = \prod_{1\le i<j\le n} \underbrace{(\sigma(j)-\sigma(i))}_{>0} = \prod_{1\le i<j\le n} |\sigma(j)-\sigma(i)|$$

und

$$\begin{aligned} B = \prod_{1\le i<j\le n} \underbrace{(\sigma(j)-\sigma(i))}_{<0} &= \prod_{1\le i<j\le n} (-1)\cdot|\sigma(j)-\sigma(i)| \\ &= (-1)^k \cdot \prod_{1\le i<j\le n} |\sigma(j)-\sigma(i)|. \end{aligned}$$

Damit gilt:

$$\begin{aligned} \prod_{1\le i<j\le n} (\sigma(j)-\sigma(i)) = A\cdot B &= (-1)^k \cdot \prod_{1\le i<j\le n} |\sigma(j)-\sigma(i)| \\ &= (-1)^k \cdot \prod_{1\le i<j\le n} (j-i), \end{aligned}$$

wobei die letzte Gleichung folgt, da σ eine Permutation ist. Deswegen:

$$\operatorname{sign}(\sigma) = \frac{\prod\limits_{1\le i<j\le n} (\sigma(j) - \sigma(i))}{\prod\limits_{1\le i<j\le n} (j - i)} = (-1)^k,$$

wie gewünscht.

Beispiel 2.42. Die Permutation $S_4 \ni \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$ hat 4 Fehlstände, und damit ist $\operatorname{sign}(\sigma) = 1$. (Die Fehlstände sind $(1,3), (1,4), (2,3), (2,4)$.)

Korollar 2.43. *Sei $n \ge 2$.*

(a) Für jedes $\sigma \in S_n$ gilt $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$.

(b) Für jede Transposition $\tau \in S_n$ gilt $\operatorname{sign}(\tau) = -1$.

(c) Sei $\sigma := \tau_1 \circ \cdots \circ \tau_k \in S_n$ das Produkt von k Transpositionen $\tau_1, \dots \tau_k \in S_n$. Dann gilt $\operatorname{sign}(\sigma) = (-1)^k$.

Beweis.
Der Teil (c) folgt aus (b) nach Lemma 2.38 und Satz 2.40.

(a) Nach Satz 2.40 gilt

$$1 = \operatorname{sign}(\mathrm{id}) = \operatorname{sign}(\sigma \circ \sigma^{-1}) = \operatorname{sign}(\sigma)\operatorname{sign}(\sigma^{-1}),$$

und damit $\operatorname{sign}(\sigma)^{-1} = \operatorname{sign}(\sigma)$, da beide zur Menge $\{\pm 1\}$ gehören.

(b) Setze $\tau_0 := \begin{pmatrix} 1 & 2 \end{pmatrix}$. Dann hat τ_0 nur einen Fehlstand, und damit ist $\operatorname{sign}(\tau_0) = -1$.

Sei $\tau = \begin{pmatrix} k & l \end{pmatrix} \in S_n$ nun eine beliebige Transposition, wobei $k, \ell \in \{1, \dots, n\}$ mit $k < \ell$. Sei $\sigma \in S_n$ eine beliebige Permutation mit $\sigma(1) = k$ and $\sigma(2) = \ell$ und setze

$$\tau' := \sigma \circ \tau_0 \circ \sigma^{-1}.$$

Wir behaupten, dass

$$\tau = \tau'. \tag{10}$$

Angenommen, (10) ist bewiesen. Dann gilt nach Satz 2.40 und nach (a):

$$\begin{aligned} \operatorname{sign}(\tau) &= \operatorname{sign}(\sigma \circ \tau_0 \circ \sigma^{-1}) = \operatorname{sign}(\sigma)\operatorname{sign}(\tau_0)\operatorname{sign}(\sigma^{-1}) \\ &= \operatorname{sign}(\sigma)\operatorname{sign}(\sigma^{-1})\operatorname{sign}(\tau_0) = 1 \cdot \operatorname{sign}(\tau_0) = -1, \end{aligned}$$

wie gewünscht.

Es bleibt also zu zeigen (10). Zuerst haben wir:

$$\tau'(k) = \sigma(\tau_0(\sigma^{-1}(k))) = \sigma(\tau_0(1)) = \sigma(2) = \ell = \tau(k),$$
$$\tau'(\ell) = \sigma(\tau_0(\sigma^{-1}(\ell))) = \sigma(\tau_0(2)) = \sigma(1) = k = \tau(\ell).$$

Sei nun $m \in \{1, \ldots, n\} \setminus \{k, \ell\}$. Dann ist $\sigma^{-1}(m) \notin \{1, 2\}$, denn ansonsten wäre $m = \sigma(1)$ oder $m = \sigma(2)$, was ein Widerspruch wäre. Es gilt also

$$\tau'(m) = \sigma(\tau_0(\sigma^{-1}(m))) = \sigma(\sigma^{-1}(m)) = m = \tau(m),$$

und dies zeigt (10). Das Korollar ist bewiesen. □

Definition 2.44. Der Kern $A_n := \ker(\mathrm{sign})$ der Abbildung $\mathrm{sign}\colon S_n \to \{\pm 1\}$ heißt *alternierende Gruppe vom Grad n*. In anderen Worten,

$$A_n := \{\sigma \in S_n \mid \mathrm{sign}(\sigma) = 1\}.$$

Lemma 2.45. *Sei $\tau \in S_n$ eine beliebige Transposition. Dann gilt:*

$$S_n = A_n \sqcup \tau A_n.$$

Beweis. Die Vereinigung $A_n \cup \tau A_n$ ist disjunkt, da alle Permutationen aus A_n die Signatur 1 und alle Permutationen aus τA_n die Signatur -1 haben, nach Satz 2.40 und Korollar 2.43.

Sei nun $\sigma \in S_n$. Dann $\mathrm{sign}(\sigma) = \pm 1$. Wenn $\mathrm{sign}(\sigma) = 1$, dann $\sigma \in A_n$ per Definition. Wenn $\mathrm{sign}(\sigma) = -1$, dann $\sigma = \tau \circ (\tau^{-1}\sigma)$, und es gilt $\tau^{-1}\sigma \in A_n$ nach Satz 2.40 und Korollar 2.43. □

Korollar 2.46. *Sei $n \geq 2$. Dann gilt $|A_n| = n!/2$.*

Beweis. Nach Lemma 2.23 gilt $|A_n| = |\tau A_n|$. Lemma 2.45 ergibt:

$$|S_n| = |A_n| + |\tau A_n| = 2|A_n|,$$

und das Korollar folgt, da $|S_n| = n!$. □

3 Ringe und Körper

Definition 3.1. Ein *Körper* ist ein Tripel $(K, +, \cdot)$ aus einer Menge K und zwei Verknüpfungen

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

sodass die folgenden Axiome erfüllt sind:

(K1) $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0, d.h.:

- (K1.1) $(a+b)+c = a+(b+c)$ für alle $a, b, c \in K$,
- (K1.2) $a+b = b+a$ für alle $a, b \in K$,
- (K1.3) $0+a = a+0 = a$ für alle $a \in K$,
- (K1.4) für jedes $a \in K$ gibt es $-a \in K$, sodass $a+(-a) = (-a)+a = 0$.

(K2) $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element 1, und es gilt ferner:

- (K2.1) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ für alle $a, b, c \in K$,
- (K2.2) $a \cdot b = b \cdot a$ für alle $a, b \in K$,
- (K2.3) $1 \cdot a = a \cdot 1 = a$ für alle $a \in K$,
- (K2.4) für jedes $a \in K \setminus \{0\}$ gibt es $a^{-1} \in K \setminus \{0\}$, sodass $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

(K3) Es gelten die folgenden *Distributivgesetze*: für alle $a, b, c \in K$ gilt

$$(a+b) \cdot c = a \cdot c + b \cdot c \quad \text{und} \quad a \cdot (b+c) = a \cdot b + a \cdot c.$$

Beispiel 3.2. Bekannte Körper sind $(\mathbb{Q}, +, \cdot), (\mathbb{R}, +, \cdot)$ und $(\mathbb{C}, +, \cdot)$. Die Struktur $(\mathbb{Z}, +, \cdot)$ ist kein Körper, da es keine multiplikativen Inversen gibt.

Definition 3.3. Ein *Ring* ist ein Tripel $(R, +, \cdot)$ aus einer Menge R und zwei Verknüpfungen $+$ und $\cdot$, für welches die Axiome

$$(K1.1) - (K1.4), \qquad (K2.1) \qquad \text{und} \qquad (K3)$$

erfüllt sind.
Falls zusätzlich das Axiom $(K2.2)$ erfüllt ist, so ist R ein *kommutativer* Ring.
Falls zusätzlich das Axiom $(K2.3)$ erfüllt ist, so ist R ein kommutativer Ring *mit Eins*.

Beispiel 3.4.

(a) $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer Ring mit 1.

(b) Jeder Körper ist ein kommutativer Ring mit 1.

(c) $(2\mathbb{Z}, +, \cdot)$ ist ein kommutativer Ring ohne 1.

Lemma 3.5. *Jeder Körper hat mindestens zwei Elemente:* 0 *und* 1*. Der kleinste Körper hat zwei Elemente: das ist* $(\mathbb{F}_2, +\cdot)$ *mit* $\mathbb{F}_2 := \{0, 1\}$ *und den Verknüpfungstafeln:*

$+$	0	1
0	0	1
1	1	0

und

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

Insbesondere gilt: Die Gruppe $(\mathbb{F}_2, +)$ *ist* $(\mathbb{Z}_2, +)$*. Die zwei Operationen oben sind* $+$ *und* $\cdot$ *modulo* 2*.*

Beispiel 3.6. Für jede positive ganze Zahl n ist $(\mathbb{Z}_n, +, \cdot)$ ein Ring (Übung!), wobei

$$[a] + [b] := [a + b] \qquad \text{und} \qquad [a] \cdot [b] := [ab].$$

Ab jetzt schreiben wir in einem Körper:

$$a - b := a + (-b) \qquad \text{und} \qquad \frac{a}{b} := ab^{-1}.$$

Lemma 3.7. *Sei* $(K, +, \cdot)$ *ein Körper. Dann gilt:*

(1) 0 *und* 1 *sind eindeutig bestimmt,*

(2) das Negative $-a$ *zu* a *und das Inverse* a^{-1} *zu* a *sind eindeutig bestimmt,*

(3) (a) $0 \cdot a = 0$ *für alle* $a \in K$*,*

(b) $(-1)(-1) = 1$*,*

(c) aus $ab = 0$ *und* $a \neq 0$ *folgt, dass* $b = 0$*.*

(4) In endlichen Summen und Produkten kommt es nicht auf die Reihenfolge an. Das heißt: Seien $a_1, \ldots, a_n \in K$ *und sei* $\sigma \in S_n$*. Dann gilt*

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} a_{\sigma(i)} \quad \textit{und} \quad \prod_{i=1}^{n} a_i = \prod_{i=1}^{n} a_{\sigma(i)}.$$

Beweis.
(1) und (2) folgen aus der Tatsache, dass $(K, +)$ und $(K, \cdot)$ Gruppen sind.

(4) folgt aus den Assoziativ- und Kommutativgesetzen.

(3)(a) Es gilt

$$0 \cdot a \overset{(K1.3)}{=} (0+0) \cdot a \overset{(K3)}{=} 0 \cdot a + 0 \cdot a,$$

und damit

$$0 = 0 \cdot a - 0 \cdot a = 0 \cdot a + 0 \cdot a - 0 \cdot a = 0 \cdot a.$$

(3)(b) Es gilt

$$1 + (-1) = 0,$$

und damit

$$(1 + (-1)) \cdot (-1) = 0 \cdot (-1).$$

Dann folgt aus (K3) und aus (3)(a):

$$-1 + (-1)(-1) = 0,$$

und damit

$$(-1)(-1) = 1.$$

(3)(c) Aus $ab = 0$ folgt, nach (3)(a):

$$b = \underbrace{a^{-1}a}_{=1} b = a^{-1}(ab) = a^{-1} \cdot 0 = 0,$$

wie gewünscht. □

Definition 3.8. Sei $(K, +, \cdot)$ ein Körper. Ein *Unterkörper* oder *Teilkörper* $(L, +, \cdot)$ von $(K, +, \cdot)$ besteht aus einer Teilmenge $L \subseteq K$, die bezüglich der bestehenden Verknüpfungen $+$ und $\cdot$ zu einem Körper wird.

Analog dazu ist ein *Unterring* oder *Teilring* definiert.

Definition 3.9. Seien $(R_1, +, \cdot)$ und $(R_2, +, \cdot)$[5] zwei Ringe mit Eins. Eine Abbildung $\varphi\colon R_1 \to R_2$ ist ein *Ringhomomorphismus*, falls:

(1) $\varphi(1_{R_1}) = 1_{R_2}$,

(2) $\varphi(a + b) = \varphi(a) + \varphi(b)$ für alle $a, b \in R_1$,

(3) $\varphi(ab) = \varphi(a)\varphi(b)$ für alle $a, b \in R_1$.

[5] Wir bezeichnen die Operationen in R_1 und in R_2 mit denselben Symbolen.

Falls R_1 und R_2 Körper sind, so ist φ ein *Körperhomomorphismus.*

Bemerkung 3.10. Sei $\varphi\colon R_1 \to R_2$ ein Ringhomomorphismus. Dann gilt

$$\varphi(0_{R_1}) = 0_{R_2},$$

denn φ ist ein Gruppenhomomorphismus zwischen $(R_1, +)$ und $(R_2, +)$.

Lemma 3.11. *Sei $(K, +, \cdot)$ ein Körper und L eine nichtleere Teilmenge von K. Dann ist $(L, +, \cdot)$ ein Teilkörper von K genau dann, wenn:*

(a) $0, 1 \in L$,

(b) für alle $a, b \in L$ gilt $a + b \in L$ und $ab \in L$,

(c) für alle $a \in L$ gilt $-a \in L$ und $a^{-1} \in L$.

Beweis. Übung! (Ähnlich wie für Gruppen und Untergruppen.) □

Beispiel 3.12.

(1) Die Abbildung
$$\varphi\colon (\mathbb{Z}, +, \cdot) \to (\mathbb{F}_2, +, \cdot),\ n \mapsto [n]$$
ist ein Ringhomomorphismus. (Übung!)

(2) $(\mathbb{Q}, +, \cdot)$ ist ein Teilkörper von $(\mathbb{R}, +, \cdot)$.

(3) $(\mathbb{R}, +, \cdot)$ ist ein Teilkörper von $(\mathbb{C}, +, \cdot)$.

Notation 3.13. Sei R ein Ring mit 1, Seien $n \in \mathbb{Z}$ und $a \in R$. Dann bezeichnen wir

$$n \cdot a := \begin{cases} \underbrace{a + \cdots + a}_{n \text{ Mal}}, & \text{wenn } n > 0, \\ 0, & \text{wenn } n = 0, \\ \underbrace{-a - \cdots - a}_{(-n) \text{ Mal}}, & \text{wenn } n < 0. \end{cases}$$

Beispiel 3.14.

(1) Für $1 \in \mathbb{Z}$ gilt $n \cdot 1 \neq 0$ für alle $n \in \mathbb{Z} \setminus \{0\}$.

(2) Für jedes $n \in \mathbb{N}_{>0}$ gilt im Ring $\mathbb{Z}_n$: $n \cdot [1] = [n] = [0]$.

Definition 3.15. Sei R ein Ring mit 1. Die *Charakteristik* von R, geschrieben $\operatorname{char}(R)$, ist definiert durch

$$\operatorname{char}(R) := \begin{cases} 0, & \text{falls } n \cdot 1 \neq 0 \text{ für alle } n \in \mathbb{Z} \setminus \{0\}, \\ \min\{n \in \mathbb{N}_{>0} \mid n \cdot 1 = 0\}, & \text{sonst.} \end{cases}$$

Lemma 3.16. *Sei K ein Körper. Dann ist* $\operatorname{char}(K)$ *eine Primzahl oder* 0.

Beweis. Sei $n := \operatorname{char}(K)$. Angenommen, $n \neq 0$. Wir müssen zeigen, dass n eine Primzahl ist.

Angenommen, n ist keine Primzahl. Dann gibt es ganze Zahlen $1 < k, \ell < n$ mit $n = k\ell$, und es folgt:

$$\begin{aligned}(k \cdot 1) \cdot (\ell \cdot 1) &= (\underbrace{1 + \cdots + 1}_{k \text{ Mal}})(\underbrace{1 + \cdots + 1}_{\ell \text{ Mal}}) \overset{(K3)}{=} \underbrace{1 \cdot 1 + \cdots + 1 \cdot 1}_{k \cdot \ell \text{ Mal}} \\ &= \underbrace{1 + \cdots + 1}_{n \text{ Mal}} = n \cdot 1 = 0.\end{aligned}$$

Dann folgt aus Lemma 3.7, dass $k \cdot 1 = 0$ oder $\ell \cdot 1 = 0$. Insbesondere ist $n = \operatorname{char}(K) \leq k$ oder $n = \operatorname{char}(K) \leq \ell$, welches ein Widerspruch ist. $\square$

Komplexe Zahlen

Das Problem mit dem Körper $\mathbb{R}$ ist, dass die Gleichung

$$x^2 + 1 = 0$$

keine Lösung in $\mathbb{R}$ hat. Betrachte nun das Tripel $(\mathbb{R}^2, +, \cdot)$, wobei für alle $a, b, c \in \mathbb{R}$ gilt:

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

Dann checkt man einfach, dass $(\mathbb{R}^2, +, \cdot)$ ein Körper ist. Diesen Körper bezeichnen wir mit $(\mathbb{C}, +, \cdot)$. Die neutralen Elemente sind $(0, 0)$ (bezüglich der Operation $+$) und $(1, 0)$ (bezüglich der Operation $\cdot$).

Sei nun $(a, b) \neq (0, 0)$. Dann sind die zu (a, b) inversen Elemente:

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

Lemma 3.17.

(a) Die Teilmenge $K := \{(a,0) \mid a \in \mathbb{R}\} \subseteq \mathbb{C}$ *ist ein Teilkörper von* $\mathbb{C}$.

(b) Die Abbildung $\varphi\colon \mathbb{R} \to K, \quad a \mapsto (a,0)$ *ist ein Körperisomorphismus.*

Beweis. Übung! □

Wegen dieses Lemmas identifizieren wir $\mathbb{R}$ mit K, und wir nennen K den *Körper der reellen Zahlen.* Für $(a,0)$ schreiben wir wie üblich einfach a.

Bemerkung 3.18. Definiere

$$i := (0,1) \in \mathbb{C}.$$

Dann gilt $i^2 = (0,1)\cdot(0,1) = (-1,0) = -1$, und damit ist i eine Lösung der Gleichung $x^2 + 1 = 0$ in $\mathbb{C}$. (Die andere Lösung ist $-i$.) Des Weiteren gilt:

$$(a,b) = (a,0) + (0,b) = (a,0) + (b,0)\cdot(0,1) = a + bi.$$

Definition 3.19. Betrachte eine komplexe Zahl $z = x + iy$, wobei $x, y \in \mathbb{R}$.

(a) Die reelle Zahl x ist der *Realteil* von z, und wir schreiben

$$\mathrm{Re}(z) := x.$$

(b) Die reelle Zahl y ist der *Imaginärteil* von z, und wir schreiben

$$\mathrm{Im}(z) := y.$$

(c) Die komplexe Zahl $\overline{z} := x - iy$ ist die *zu z konjugierte* komplexe Zahl.

(d) Die reelle Zahl $\sqrt{x^2 + y^2}$ ist der *Absolutbetrag* von z, und wir schreiben

$$|z| := \sqrt{x^2 + y^2}.$$

(e) Die komplexe Zahl z heißt *reell*, falls $\mathrm{Im}(z) = 0$. Sie heißt *imaginär*, falls $\mathrm{Re}(z) = 0$.

Proposition 3.20. *Für jede komplexe Zahl* $z \in \mathbb{C}$ *gilt:*

(a) $z + \overline{z} = 2\,\mathrm{Re}(z)$,

(b) $z - \overline{z} = 2\,\mathrm{Im}(z)$,

(c) $z \cdot \overline{z} = |z|^2$.

Beweis. Einfach (Übung!). $\square$

Proposition 3.21. *Für komplexe Zahlen* $z_1, z_2 \in \mathbb{C}$ *gilt:*

(a) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$,

(b) $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$,

(c) $\overline{\left(\frac{1}{z_1}\right)} = \frac{1}{\overline{z_1}}$,

(d) $\overline{\overline{z_1}} = z_1$,

(e) $|\overline{z_1}| = |z_1|$.

Beweis. Einfach (Übung!). $\square$

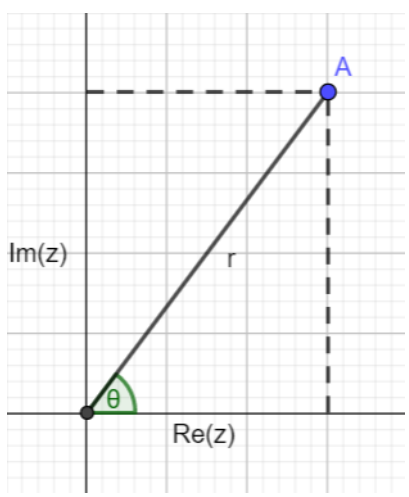
Bemerkung 3.22. Sei $z \neq 0$ eine komplexe Zahl. Dann gilt

$$\frac{1}{z} = \frac{\overline{z}}{z\overline{z}} \overset{3.20(c)}{=} \frac{\overline{z}}{|z|^2},$$

und damit

$$\operatorname{Re}\left(\frac{1}{z}\right) = \frac{1}{|z|^2}\operatorname{Re}(\overline{z}) \quad \text{und} \quad \operatorname{Im}\left(\frac{1}{z}\right) = \frac{1}{|z|^2}\operatorname{Im}(\overline{z}).$$

Bemerkung 3.23 (Geometrische Darstellung komplexer Zahlen)**.**
Sei $z = a + bi$ eine komplexe Zahl. Dann können wir z in der Zahlenebene $\mathbb{R}^2$ geometrisch darstellen:



Auf dem Bild ist der Punkt $A := (a, b)$ und die Länge r des Segments zwischen 0 und A ist der Betrag $|z|$. Der Winkel θ nennen wir das *Argument von z*. Dann gilt:

$$a = r\cos\theta, \quad b = r\sin\theta, \quad \frac{b}{a} = \tan\theta.$$

Insbesondere gilt

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

Wir schreiben auch

$$z := e^{i\theta} := \cos\theta + i\sin\theta.$$

Lemma 3.24. *Betrachte zwei komplexen Zahlen:*

$$z_1 = r_1(\cos\theta_1 + i\sin\theta_1) \quad \textit{und} \quad z_2 = r_2(\cos\theta_2 + i\sin\theta_2).$$

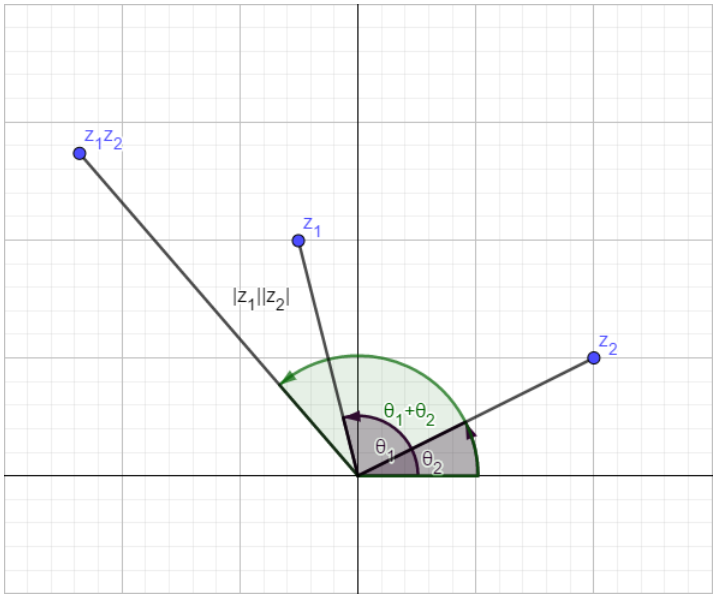
Dann gilt das Gesetz von De Moivre:

$$z_1 z_2 = r_1 r_2\big(\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)\big).$$

Insbesondere gilt

$$z_1^n = r_1^n\big(\cos(n\theta_1) + i\sin(n\theta_1)\big) \quad \textit{für alle } n \in \mathbb{N}_{>0}.$$

In anderen Worten: Bei der Multiplikation komplexer Zahlen multiplizieren sich die Beträge und addieren sich die Argumente.



Beweis. Es gilt:

$$\begin{aligned} z_1 z_2 &= r_1(\cos\theta_1 + i\sin\theta_1) r_2(\cos\theta_2 + i\sin\theta_2) \\ &= r_1 r_2(\cos\theta_1\cos\theta_2 + i\cos\theta_1\sin\theta_2 + i\sin\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) \\ &= r_1 r_2\big(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)\big), \end{aligned}$$

welches die erste Formel beweist. Die zweite Formel folgt aus dieser per Induktion nach n. □

Polynome

Es gibt zwei mögliche naive Definitionen von Polynomen.

(1) Ein *Polynom* (über einem Körper K) ist eine Formel

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

wobei $n \in \mathbb{N}$ und a_i sind Elemente aus K. Hier ist x eine *Unbekannte* oder eine *Variable*. Problem: Wie definiert man, was eine Unbekannte ist?

(2) Ein *Polynom* (über einem Körper K) ist eine Abbildung

$$f\colon K \to K, \quad x \mapsto a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

wobei $n \in \mathbb{N}$ und a_i sind Elemente aus K. Problem: Dies Funktioniert nicht über endlichen Körpern. Betrachte zum Beispiel $K := \mathbb{F}_2$ mit zwei Abbildungen:

$$\begin{aligned} f\colon \mathbb{F}_2 \to \mathbb{F}_2, &\quad x \mapsto x, \\ g\colon \mathbb{F}_2 \to \mathbb{F}_2, &\quad x \mapsto x^2. \end{aligned}$$

Diese zwei Funktionen sind gleich.

Die echte Definition eines Polynoms ist wie folgt.

Definition 3.25. Sei R ein kommutativer Ring mit 1. Bezeichne

$$\begin{aligned} R^{(\mathbb{N})} &:= \{f\colon \mathbb{N} \to \mathbb{R} \mid \text{ alle bis auf endlich viele } f(i) \text{ sind } 0\} \\ &= \{f\colon \mathbb{N} \to \mathbb{R} \mid \exists n \in \mathbb{N} : f(i) = 0 \ \forall i > n\}. \end{aligned}$$

Ein Element $f \in R^{(\mathbb{N})}$ können wir f tabellarisch schreiben als

$$f = (a_0, a_1, a_2, a_3, \dots), \quad \text{wobei } a_i = f(i) \text{ für jedes } i.$$

Auf der Menge $R^{(\mathbb{N})}$ definieren wir zwei Verknüpfungen $+$ und $\cdot$ wie folgt. Seien $f = (a_0, a_1, \dots)$ und $g = (b_0, b_1, \dots)$ zwei Elemente aus $R^{(\mathbb{N})}$. Dann definieren wir

$$f + g := (a_0 + b_0, a_1 + b_1, \dots),$$

und

$$f \cdot g := (c_0, c_1, \dots), \quad \text{wobei } c_k := \sum_{i+j=k} a_i b_j.$$

Man prüft einfach, dass $(R^{(\mathbb{N})}, +, \cdot)$ ein kommutativer Ring mit Eins ist. Diesen Ring nennen wir den *Polynomring in einer Variable über R* oder den *Polynomring in einer Variable mit Koeffizienten in R.*

Die Null in diesem Ring ist das Element $(0, 0, 0, \dots)$. Die Eins in diesem Ring ist das Element $(1, 0, 0, \dots)$. Die *Unbekannte* oder die *Variable* in diesem Ring ist das Element $x := (0, 1, 0, 0, \dots)$. Daraus folgt

$$x^n = (\underbrace{0}_{\text{0-te Stelle}}, \dots, 0, \underbrace{1}_{n\text{-te Stelle}}, 0, \dots).$$

Diesen Ring bezeichnen wir üblicherweise mit

$$R[x].$$

Der Ring R ist isomorph zum Unterring

$$\{(a, 0, 0, \dots) \mid a \in R\} \subseteq R[x].$$

Aus den Definitionen von $+$ und $\cdot$ folgt sofort, dass ein Element

$$f := (a_0, a_1, a_2, \dots)$$

geschrieben werden kann als

$$f = a_0 + a_1 x + a_2 x^2 + \dots.$$

Die Elemente $a_0, a_1, \dots \in R$ nennen wir *Koeffizienten* von f. Aus der Definition von $R[x]$ folgt, dass für nur endlich viele $i \in \mathbb{N}$ gilt $a_i \neq 0$. Wenn $f \neq 0$, dann können wir f also schreiben als

$$f = a_0 + a_1 x + \dots + a_n x^n \quad \text{für ein } n \in \mathbb{N} \text{ mit } a_n \neq 0.$$

Diese Zahl n nennen wir den *Grad von f* und wir schreiben

$$\deg(f) := n.$$

Der Koeffizient a_n heißt der *Leitkoeffizient*. Wenn $f = 0$ ist, dann setzen wir

$$\deg(f) := -\infty.$$

In anderen Worten,

$$\deg(f) = \begin{cases} -\infty, & \text{wenn } f = 0, \\ \max\{k \in \mathbb{N} \mid a_k \neq 0\}, & \text{sonst.} \end{cases}$$

Ein Polynom $f \neq 0$ heißt *normiert* oder *monisch*, falls sein Leitkoeffizient gleich 1 ist.

Bemerkung 3.26. Sei K ein Körper. Für alle $f, g \in K[x]$ gilt

$$\deg(fg) = \deg f + \deg g.$$

(Hier gelten die Regeln $(-\infty) + n = n + (-\infty) = -\infty$.) Dies ist einfach zu zeigen: Wir dürfen annehmen, dass $f \neq 0$ und $g \neq 0$. Wir können dann schreiben

$$f\colon a_0 + a_1 x + \cdots + a_n x^n \text{ mit } a_n \neq 0,$$

und

$$g\colon b_0 + b_1 x + \cdots + b_m x^m \text{ mit } b_m \neq 0.$$

Dann ist $\deg f = n$ und $\deg g = m$, und man checkt einfach, dass der Leitkoeffizient von fg gleich $a_n b_m$ (bemerke, dass $a_n b_m \neq 0$, da $a_n \neq 0$ und $b_m \neq 0$: hier benutzen wir die Voraussetung, dass K ein Körper ist). Es gilt also $\deg(fg) = n + m$.

Insbesondere zeigt dasselbe Argument:

Lemma 3.27. *Sei K ein Körper. Für alle $f, g \in K[x]$ gilt:*

$$fg = 0 \quad \Longrightarrow \quad f = 0 \text{ oder } g = 0.$$

Division mit Rest

In einem Polynomring gilt Division mit Rest, ähnlich wie im Ring $\mathbb{Z}$.

Satz 3.28. *Sei K ein Körper und seien $g, f \in K[x]$ mit $g \neq 0$. Dann gibt es eindeutig bestimmte Polynome $q, r \in K[x]$ mit*

$$f = gq + r, \quad \textit{wobei } \deg r < \deg g.$$

Beweis. Zuerst beweisen wir die Eindeutigkeit. Angenommen, es gibt Polynome $q_1, q_2, r_1, r_2 \in K[x]$ mit

$$f = gq_1 + r_1 \quad \text{und} \quad f = gq_2 + r_2, \quad \text{wobei } \deg r_1, \deg r_2 < \deg g.$$

Dann gilt:

$$0 = f - f = (gq_1 + r_1) - (gq_2 + r_2) = g(q_1 - q_2) + (r_1 - r_2),$$

und damit

$$r_2 - r_1 = g(q_1 - q_2).$$

Also gilt:

$$\deg g > \deg(r_2 - r_1) = \deg(g(q_1 - q_2)) = \deg g + \deg(q_1 - q_2).$$

Das ist möglich nur, wenn $\deg(q_1 - q_2) = -\infty$, also $q_1 = q_2$. Dann folgt einfach, dass $r_1 = r_2$.

Nun zeigen wir die Existenz per Induktion nach $\deg f$. Die Aussage ist einfach, wenn $f = 0$; wir nehmen daher an, dass $f \neq 0$. Seien $n := \deg f$ und $m := \deg g$. Dann können wir schreiben

$$f = a_n x^n + \cdots + a_0 \quad \text{und} \quad g = b_m x^m + \cdots + b_0, \text{wobei } a_n, b_m \neq 0.$$

Angenommen zuerst, dass $n < m$. Dann ist $q = 0$ und $r = f$, da $f = g \cdot 0 + f$.

Nun nehmen wir an, dass $n = m$. Setze $q := \frac{a_n}{b_n} \in K$ und $r := f - qg$. Dann checkt man einfach, dass

$$r = \sum_{i=0}^{n-1} (a_i - qb_i) x^i,$$

und daher ist r ein Polynom des Grades $\leq n - 1$.

Zuletzt nehmen wir an, dass $n > m$. Setze $q_1 := \frac{a_n}{b_m} x^{n-m} \in K[x]$ und $r_1 := f - gq_1$. Dann berechnet man einfach, dass

$$r_1 = \sum_{i=0}^{n-1} \left(a_i - \frac{a_n}{b_m} b_{i-m} \right) x^i.$$

Daher ist r_1 ein Polynom des Grades $\leq n - 1 < \deg f$. Damit können wir die Induktion anwenden auf die Polynome r_1 und g: Es gibt $q_2, r_2 \in K[x]$ mit $r_1 = gq_2 + r_2$, wobei $\deg r_2 < \deg g$. Damit gilt:

$$f = gq_1 + r_1 = gq_1 + (gq_2 + r_2) = g(q_1 + q_2) + r_2.$$

Setze nun $q := q_1 + q_2$ und $r := r_2$. Das sind die gewünschten Polynome. $\square$

Beispiel 3.29. Sei $K = \mathbb{R}$ und seien $f, g \in \mathbb{R}[x]$ die Polynome

$$f = 3x^3 + 2x + 1 \qquad \text{und} \qquad g = 4x^2 - 4.$$

Dann gilt:[6]

$$\begin{aligned} &(3x^3 + 2x + 1) : (4x^2 - 4) = \frac{3}{4}x + 0 \\ -&\underline{(3x^2 - 3x)} \\ &\qquad\quad 5x + 1 \end{aligned}$$

Setze $q := \frac{3}{4}x$ und $r := 5x + 1$. Man prüft, dass $f = gq + r$.

Nullstellen

Definition 3.30. Sei K ein Körper und seien $a_0, \ldots, a_n \in K$. Dann ist die Abbildung

$$p\colon K \to K, \quad x \mapsto a_0 + a_1 x + \cdots + a_n x^n$$

eine *Polynomfunktion auf* K. Es gibt dann die Abbildung

$$\widetilde{\cdot}\colon K[X] \to \mathrm{Abb}(K, K),$$

die zu einem Polynom $f = a_0 + a_1 X + \cdots + a_n X^n \in K[X]$ die folgende Funktion assoziiert:

$$\widetilde{f}\colon K \to K, \quad \widetilde{f}(z) := a_0 + a_1 z + \cdots + a_n z^n.$$

Bemerkung 3.31. Im Allgemeinen ist die Abbildung $\widetilde{\cdot}$ in der Definition nicht injektiv: Ein Gegenbeispiel für $K = \mathbb{F}_2$ ist gegeben oben in (2).

Definition 3.32. Sei K ein Körper und sei $f \in K[x]$ mit $f = a_n x^n + \cdots + a_0$. Ein Element $\lambda \in K$ ist eine *Nullstelle* von f, falls

$$a_n \lambda^n + \cdots + a_0 = 0.$$

In anderen Worten:

$$\lambda \text{ ist eine Nullstelle von } f \quad \Longleftrightarrow \quad \widetilde{f}(\lambda) = 0.$$

Wenn λ eine Nullstelle von f ist, dann schreiben wir auch $f(\lambda) = 0$.

[6] $3x^2 - 3x = \frac{3}{4}x \cdot (4x^2 - 4)$

Satz 3.33. *Sei K ein Körper und sei $f \in K[x]$. Ist $\lambda \in K$ eine Nullstelle von f, so gibt es ein eindeutig bestimmtes Polynom $g \in K[x]$ mit den folgenden Eigenschaften:*

(a) $f = (x - \lambda)g$,

(b) $\deg g = \deg f - 1$.

Beweis.
(a) Wir wenden die Division mit Rest auf f und $x - \lambda$ an: Es gibt Polynome $q, r \in K[x]$ mit

$$f = (x - \lambda)q + r, \quad \text{wobei } \deg r < \deg(x - \lambda) = 1.$$

Damit ist r ein Element in K. Wir ersetzen nun x durch λ:

$$\underbrace{f(\lambda)}_{0} = \underbrace{(\lambda - \lambda)}_{0} \cdot q(\lambda) + \underbrace{r(\lambda)}_{r \in K},$$

und damit $r = 0$.

(b) Es gilt:

$$\deg f \overset{(a)}{=} \deg\big((x - \lambda)g\big) = \deg(x - \lambda) + \deg g = 1 + \deg g,$$

woraus die Aussage folgt. □

Korollar 3.34. *Sei K ein Körper, sei $f \in K[x]$ und sei k die Anzahl an verschiedenen Nullstellen von f. Ist $f \neq 0$, so gilt*

$$k \leq \deg f.$$

Beweis. Der Beweis ist per Induktion nach $\deg f$.

Wenn $\deg f = 0$, dann ist $f \in K \setminus \{0\}$, und damit hat f keine Nullstelle. Es folgt: $k = 0 \leq \deg f$.

Wir dürfen also annehmen, dass $n := \deg f \geq 1$, und dass die Aussage des Satzes gilt für alle Polynome des Grades $< n$. Wir dürfen auch annehmen, dass f eine Nullstelle $\lambda \in K$ hat (ansonsten ist das Korollar trivial). Nach Satz 3.33 gibt es $g \in K[x]$, sodass $f = (x - \lambda)g$ und $\deg g = n - 1$.

Wenn f hat keine weitere Nullstelle hat, dann gilt $k = 1 \leq \deg f$.

Wir nehmen also an, dass f weitere Nullstelle hat. Sei λ' eine weitere Nullstelle von f mit $\lambda' \neq \lambda$. Dann gilt

$$\underbrace{f(\lambda')}_{=0} = \underbrace{(\lambda' - \lambda)}_{\neq 0} \cdot g(\lambda'),$$

und damit $g(\lambda') = 0$. Also, λ' ist eine Nullstelle von g. Per Induktion folgt, dass g nur ℓ verschiedene Nullstellen hat, wobei $\ell \leq \deg g$. Es folgt:

$$k = \ell + 1 \leq \deg g + 1 = \deg f,$$

wie gewünscht. □

Korollar 3.35. *Ist ein Körper K unendlich, so ist die Funktion $\widetilde{\cdot}\colon K[x] \to \mathrm{Abb}(K, K)$ aus Definition 3.30 injektiv.*

Beweis. Angenommen, es gibt zwei Polynome $f, g \in K[x]$ mit $\widetilde{f} = \widetilde{g}$. Wir müssen zeigen, dass $f = g$.

Betrachte $h := f - g \in K[x]$. Dann gilt

$$\widetilde{h} = \widetilde{f - g} = \widetilde{f} - \widetilde{g} = 0.^{7}$$

Insbesondere ist jedes Element in K eine Nullstelle von h. Damit hat h unendlich viele Nullstellen, da K unendlich ist. Nach Korollar 3.34 ist dann $h = 0$, also $f = g$. □

Bemerkung 3.36. Korollar 3.35 besagt, dass, wenn K unendlich ist, es keinen Unterschied zwischen den Polynomen in $K[x]$ und den Polynomfunktionen über K gibt.

Satz 3.37 (Fundamentalsatz der Algebra)**.** *Jedes Polynom $f \in \mathbb{C}[x]$ mit $\deg f \geq 1$ hat in $\mathbb{C}$ mindestens eine Nullstelle.*

Bemerkung 3.38. Der Satz wurde zunächst 1799 von Gauß bewiesen. Es gibt keinen rein algebraischen Beweis. Ein Beweis wird in der Vorlesung *Algebra* gegeben.

Korollar 3.39. *Jedes Polynom $f \in \mathbb{C}[x]$ zerfällt in* Linearfaktoren*. In anderen Worten, wenn $n = \deg f$, dann gibt es $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ und $a \in \mathbb{C}$, sodass*

$$f = a(x - \lambda_1) \cdots (x - \lambda_n).$$

Hier sind $\lambda_1, \ldots, \lambda_n$ nicht notwendigerweise verschieden.

Beweis. Wenn $n = 1$ ist die Aussage klar. Wir nehmen also an, dass $n > 1$.

Nach dem Fundamentalsatz der Algebra gibt es eine Nullstelle $\lambda_n \in \mathbb{C}$ von f. Nach Satz 3.33 gilt

$$f = (x - \lambda_n)g \quad \text{mit} \quad \deg g = n - 1.$$

[7] Hier bezeichnet 0 die *Nullfunktion*, d.h. die Abbuldung, die alle Elemente aus K nach 0 schickt.

Per Induktion nach n gibt es $\lambda_1, \ldots, \lambda_{n-1} \in \mathbb{C}$ und $a \in \mathbb{C}$, sodass

$$g = a(x - \lambda_1) \cdots (x - \lambda_{n-1}),$$

und damit

$$f = a(x - \lambda_1) \cdots (x - \lambda_{n-1})(x - \lambda_n),$$

wie gewünscht. □

Definition 3.40. Sei K ein Körper, sei $f \in K[x]$ und sei $\lambda \in K$. Dann heißt

$$\mu(f, \lambda) := \max\{r \in \mathbb{N} \mid f = (x - \lambda)^r g \text{ für ein } g \in K[x]\}$$

die *Vielfachheit* oder die *Multiplizität* von λ in f.

Bemerkung 3.41. Es gilt $\mu(f, \lambda) > 0$ genau dann, wenn λ eine Nullstelle von f ist.

Lemma 3.42. *Sei K ein Körper, seien $f, g \in K[x]$ und sei $\lambda \in K$. Dann gilt*

$$\mu(fg, \lambda) = \mu(f, \lambda) + \mu(g, \lambda).$$

Beweis. Setze $k := \mu(f, \lambda)$, $\ell := \mu(g, \lambda)$ und $n := \mu(fg, \lambda)$.

Per Definition der Multiplizität gibt es Polynome f' und g', sodass

$$f = (x - \lambda)^k f' \qquad g = (x - \lambda)^\ell g',$$

und damit

$$fg = (x - \lambda)^{k+\ell} f' g'.$$

Daraus folgt sofort, dass $k + \ell \leq n$. Angenommen, $k + \ell \neq n$, oder äquivalent,

$$n - (k + \ell) > 0.$$

Wir werden einen Widerspruch herleiten.

Nun, per Definition der Multiplizität gibt es ein Polynom h, sodass

$$f = (x - \lambda)^n h.$$

Dann gilt:

$$\begin{aligned} 0 = f - f &= (x - \lambda)^n h - (x - \lambda)^{k+\ell} f' g' \\ &= \underbrace{(x - \lambda)^{k+\ell}}_{\neq 0} \big((x - \lambda)^{n-(k+\ell)} h - f' g'\big), \end{aligned}$$

und es folgt dann aus Lemma 3.27, dass

$$(x-\lambda)^{n-(k+\ell)}h - f'g' = 0. \tag{11}$$

Wenn wir nun $x = \lambda$ einsetzen, bekommen wir:

$$\underbrace{(\lambda-\lambda)^{n-(k+\ell)}}_{=0,\text{ da } n-(k+\ell)>0} h(\lambda) - f'(\lambda)g'(\lambda) = 0.$$

Es folgt wieder aus Lemma 3.27, dass $f'(\lambda)g'(\lambda) = 0$, und daher $f'(\lambda) = 0$ oder $g'(\lambda) = 0$. Ohne Beschränkung der Allgemeinheit dürfen wir annehmen, dass $f'(\lambda) = 0$. Aus Satz 3.33 folgt, dass dann $f' = (x-\lambda)f''$ für ein Polynom $f'' \in K[x]$. Aber dann ist

$$f = (x-\lambda)^k f' = (x-\lambda)^k (x-\lambda) f'' = (x-\lambda)^{k+1} f''.$$

welches bedeuten würde, dass $\mu(f,\lambda) \geq k+1$. Widerspruch. □

Lemma 3.43. *Sei K ein Körper, sei $f \in K[x]$ und seien $\lambda_1, \ldots, \lambda_k \in K$ verschiedene Nullstellen von f mit jeweiligen Multiplizitäten $r_1, \ldots, r_k$. Dann gilt:*

(a) $r_1 + \cdots + r_k \leq \deg f$,

(b) es existiert ein Polynom $h \in K[x]$, sodass

$$f = (x-\lambda_1)^{r_1} \cdots (x-\lambda_k)^{r_k} h.$$

Beweis. Setze $n := \deg f$ und

$$g := (x-\lambda_1)^{r_1} \cdots (x-\lambda_k)^{r_k}.$$

(a) Wir zeigen (a) per Induktion nach n. Für $n = 1$ ist die Aussage trivial.

Sei nun $n \geq 2$. Wir argumentieren per Widerspruch und nehmen an, dass $r_1 + \cdots + r_k > n$. Wie wenden die Division mit Rest auf die Polynome g und f: Es gibt Polynome $q, r \in K[x]$ mit

$$g = fq + r, \quad \text{wobei } \deg r < \deg f = n. \tag{12}$$

Per Definition der Multiplizität gibt es ein Polynom h_1, sodass

$$f = (x-\lambda_1)^{r_1} h_1. \tag{13}$$

Aus (12) und (13) folgt:

$$\begin{aligned} r = g - fq &= (x-\lambda_1)^{r_1} \cdots (x-\lambda_k)^{r_k} - (x-\lambda_1)^{r_1} h_1 q \\ &= (x-\lambda_1)^{r_1} \big((x-\lambda_2)^{r_2} \cdots (x-\lambda_k)^{r_k} - h_1 q \big). \end{aligned}$$

Insbesondere, per Definition der Multiplizität folgt:

$$\mu(r, \lambda_1) \geq r_1.$$

Analog zeigt man, dass

$$\mu(r, \lambda_i) \geq r_i \quad \text{für jedes } i = 1, \ldots, k,$$

und insbesondere:

$$\sum_{i=1}^{k} \mu(r, \lambda_i) \geq \sum_{i=1}^{k} r_i > n. \tag{14}$$

Andererseits folgt per Induktionsschritt, dass

$$\sum_{i=1}^{k} \mu(r, \lambda_i) \leq \deg r < n,$$

ein Widerspruch zu (14). Dies zeigt (a).

(b) Wir zeigen (b) per Induktion nach n. Für $n = 1$ ist die Aussage trivial.

Sei nun $n \geq 2$. Wie wenden die Division mit Rest auf die Polynome f und g: Es gibt Polynome $\widehat{q}, \widehat{r} \in K[x]$ mit

$$f = g\widehat{q} + \widehat{r}, \quad \text{wobei } \deg \widehat{r} < \deg g = r_1 + \cdots + r_k \leq n, \tag{15}$$

wobei die letzte Ungleichung aus (a) folgt. Ähnlich wie in (a) zeigt man, dass

$$\mu(\widehat{r}, \lambda_i) \geq r_i \quad \text{für jedes } i = 1, \ldots, k.$$

Insbesondere, per Induktionsschritt existiert ein Polynom $\widehat{h}$, sodass

$$\widehat{r} = (x-\lambda_1)^{r_1} \cdots (x-\lambda_k)^{r_k} \widehat{h},$$

oder umgeschrieben,

$$\widehat{r} = g\widehat{h}.$$

Da $\deg \widehat{r} < \deg g$, so ist dies nur möglich, wenn $\widehat{h} = 0$, und damit $\widehat{r} = 0$, welches, nach (15), zeigt (b). $\square$

Um zu untersuchen, wie eine ähnliche Aussage zum Korollar 3.39 für reelle Polynome aussehen könnte, brauchen wir zuerst ein Lemma.

Lemma 3.44. *Sei f ein Polynom in $\mathbb{R}[x]$ und $\lambda \in \mathbb{C}$ eine nicht-reelle Nullstelle. Sei*

$$g := (x - \lambda)(x - \overline{\lambda}) \in \mathbb{C}[x].$$

Dann gilt:

(a) $g \in \mathbb{R}[x]$,

(b) $\overline{\lambda}$ *ist auch eine Nullstelle von* f,

(c) es gibt $q \in \mathbb{R}[x]$ *mit* $f = gq$,

(d) für $k := \mu(f, \lambda)$ *gibt es* $h \in \mathbb{R}[x]$ *mit* $f = g^k h$.

Das Lemma besagt also: Wenn λ eine Nullstelle von f ist, dann ist es auch $\overline{\lambda}$.

Beweis. (a) Sei $\lambda = a + bi$. Dann ist $\overline{\lambda} = a - bi$, und damit

$$\begin{aligned} g &= (x - \lambda)(x - \overline{\lambda}) = \big(x - (a + bi)\big)\big(x - (a - bi)\big) \\ &= x^2 - x(a - bi + a + bi) + (a + bi)(a - bi) \\ &= x^2 - 2ax + (a^2 + b^2) \in \mathbb{R}[x]. \end{aligned}$$

(b) Wir schreiben

$$f = \sum_{i=0}^{n} a_i x^i \in \mathbb{R}[x],$$

wobei $a_i = \overline{a_i}$ für alle $i \geq 0$, da alle a_i reell sind. Damit gilt:

$$f(\overline{\lambda}) = \sum_{i=0}^{n} a_i \overline{\lambda}^i = \sum_{i=0}^{n} \overline{a_i}\,\overline{\lambda^i} = \sum_{i=0}^{n} \overline{a_i \lambda^i} = \overline{\sum_{i=0}^{n} a_i \lambda^i} = \overline{f(\lambda)} = \overline{0} = 0$$

(c) Division mit Rest ergibt, dass es $q, r \in \mathbb{R}[x]$ gibt, sodass

$$f = gq + r \quad \text{und} \quad \deg r < \deg g = 2.$$

Daraus folgt:

$$0 = f(\lambda) = \underbrace{g(\lambda)}_{=0} q(\lambda) + r(\lambda) = r(\lambda)$$

und

$$0 \overset{\text{(b)}}{=} f(\overline{\lambda}) = \underbrace{g(\overline{\lambda})}_{=0} q(\overline{\lambda}) + r(\overline{\lambda}) = r(\overline{\lambda}).$$

Das Polynom r hat also den Grad ≤ 1 und mindestens zwei Nullstellen (λ und $\overline{\lambda}$). Es folgt aus Korollar 3.34, dass $r = 0$.

(d) Wir führen den Beweis per Induktion nach k. Die Aussage ist trivial für $k = 0$.

Sei jetzt die Aussage wahr für alle Zahlen, die kleiner als k sind. Wir müssen zeigen, dass die Aussage gilt auch für k.

Nach (c) gibt es $q \in \mathbb{R}[x]$ mit $f = gq = (x - \lambda)(x - \overline{\lambda})q$. Nach Lemma 3.42 gilt:

$$k = \mu(f, \lambda) = \mu(gq, \lambda) = \mu(g, \lambda) + \mu(q, \lambda) = 1 + \mu(q, \lambda).$$

Also, $\mu(q, \lambda) = k - 1$, sodass, per Induktion, ein Polynom $w \in \mathbb{R}[x]$ existiert mit

$$q = g^{k-1} w.$$

Daher:

$$f = gq = g \cdot g^{k-1} w = g^k w,$$

wie gewünscht. □

Satz 3.45. *Sei $f \in \mathbb{R}[x]$ und $\lambda \in \mathbb{C}$ eine Nullstelle von f. Dann ist $\overline{\lambda}$ auch eine Nullstelle von f und es gilt*

$$\mu(f, \lambda) = \mu(f, \overline{\lambda}).$$

Beweis. Die erste Aussage wurde in Lemma 3.44 gezeigt.

Wir zeigen nun die zweite Aussage. Wenn $\lambda \in \mathbb{R}$ ist, dann ist die Aussage trivial. Wir dürfen also annehmen, dass $\lambda \notin \mathbb{R}$. Sei $k := \mu(f, \lambda)$. Dann gibt es, nach Lemma 3.44(d), ein $h \in \mathbb{R}[x]$ mit

$$f = (x - \lambda)^k (x - \overline{\lambda})^k h.$$

Per Definition der Multiplizität folgt daraus, dass

$$\mu(f, \overline{\lambda}) \geq k = \mu(f, \lambda).$$

Analog zeigen wir, dass $\mu(f, \lambda) \geq \mu(f, \overline{\lambda})$, und damit $\mu(f, \overline{\lambda}) = \mu(f, \lambda)$. □

Korollar 3.46. *Jedes Polynom $f \in \mathbb{R}[x]$ mit $n := \deg f \geq 1$ lässt sich schreiben als*

$$f = a(x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_r) g_1 g_2 \cdots g_m,$$

wobei $a \in \mathbb{R} \setminus \{0\}$, $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$, und $g_1, \ldots, g_m \in \mathbb{R}[x]$ sind monische Polynome vom Grad 2, die keine reellen Nullstellen haben. Insbesondere gilt

$$n = r + 2m.$$

Beweis. Aus dem Fundamentalsatz der Algebra folgt, dass existieren $a \in \mathbb{R} \setminus \{0\}$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ mit $f = a(x - \lambda_1) \ldots (x - \lambda_n)$. Ohne Beschränkung der Allgemeinheit dürfen wir annehmen, dass $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ und $\lambda_{r+1}, \ldots, \lambda_n \notin \mathbb{R}$.

Übung: Vervollständigen Sie den Beweis mithilfe vom Satz 3.45. □

Bemerkung 3.47. Jedes Polynom $f \in \mathbb{R}[x]$ von ungeradem Grad hat mindestens eine reelle Nullstelle. Und zwar, in der Notation von Korollar 3.46 gilt $n = r + 2m$. Also folgt, wenn n ungerade ist, dass $r \neq 0$.

Bemerkung 3.48. Der Beweis des Fundamentalsatzes der Algebra funktioniert aber andersrum: man beweist *zuerst* die Bemerkung und nutzt dies, um den Fundamentalsatz der Algebra zu zeigen. Die Bemerkung folgt aus dem Zwischenwertsatz der mathematischen Analysis.

Satz 3.49. *Sei $f = ax^2 + bx + c \in \mathbb{C}[x]$. Dann sind*

$$\lambda_1 := \frac{-b + \sqrt{b^2 - 4ac}}{2a} \qquad \textit{und} \qquad \lambda_2 := \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

Nullstellen von f. Hier ist $\sqrt{b^2 - 4ac}$ eine komplexe Zahl z mit der Eigenschaft, dass

$$z^2 = b^2 - 4ac.$$

Definition 3.50. Sei $a \in \mathbb{C}$. Eine *n-te Wurzel von a* ist eine Lösung $z \in \mathbb{C}$ der Gleichung

$$z^n = a.$$

Beispiel 3.51. Sei $n \in \mathbb{N}_{>0}$. Wir wollen die Gleichung $z^n = 1$ lösen in $\mathbb{C}$. Und zwar, sei $z = r(\cos\theta + i \sin\theta)$ eine Lösung. Dann gilt

$$1 \cdot (\cos 0 + i \sin 0) = 1 = z^n \overset{\text{De Moivre}}{=} r^n(\cos n\theta + i \sin n\theta),$$

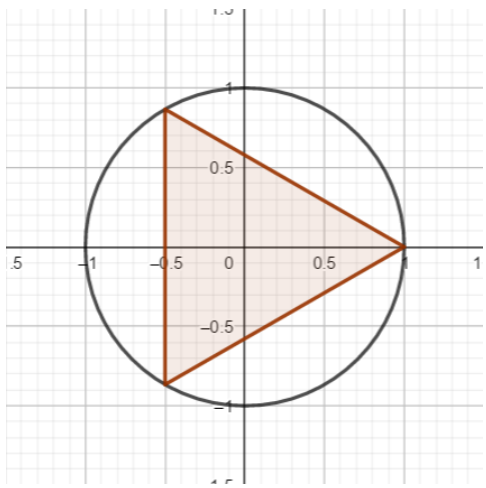
und damit

$$r^n = 1 \quad \text{und} \quad n\theta = 2k\pi, \text{ wobei } k \in \mathbb{Z}.$$

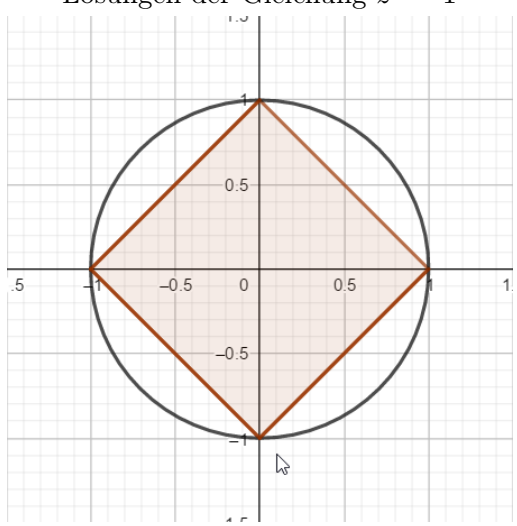
Daraus folgt, dass

$$r = 1 \quad \text{und} \quad \theta = \frac{2k\pi}{n}, \text{ wobei } k \in \mathbb{Z}.$$

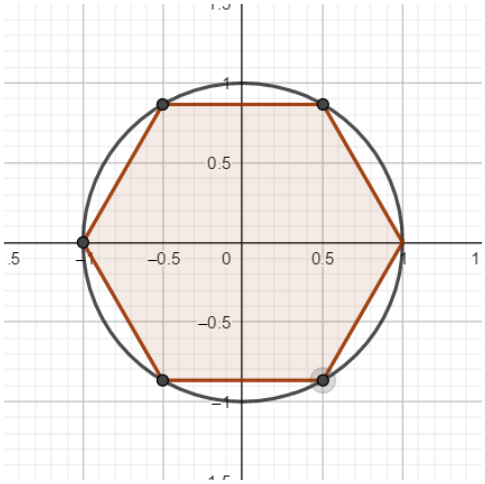
Die Lösungen der Gleichung sind also die Eckpunkte der regulären n-gons. Übung: Was sind dann die Lösungen der Gleichung $z^n = a$, wobei $a \in \mathbb{C}$?



Lösungen der Gleichung $z^3 = 1$



Lösungen der Gleichung $z^4 = 1$



Lösungen der Gleichung $z^6 = 1$

4 Vektorräume

Definition 4.1. Sei $(V, +)$ eine abelsche Gruppe und $(K, +, \cdot)$ ein Körper. Wir nennen V einen *Vektorraum über* K oder einen K-*Vektorraum*, falls es eine Abbildung

$$K \times V \to V, \quad (\alpha, v) \mapsto \alpha v \tag{16}$$

gibt mit den folgenden Eigenschaften.

(V1) $(\alpha + \beta)v = \alpha v + \beta v$ für alle $\alpha, \beta \in K$ und $v \in V$,

(V2) $(\alpha \cdot \beta)v = \alpha(\beta v)$ für alle $\alpha, \beta \in K$ und $v \in V$,

(V3) $1v = v$ für alle $v \in V$,

(V4) $\alpha(v + w) = \alpha v + \alpha w$ für alle $\alpha \in K$ und $v, w \in V$.

Die Elemente von V heißen *Vektoren* und die Elemente von K *Skalare*. Die Verknüpfung $+$ heißt *Vektoraddition* und die Abbildung (16) heißt *Skalarmultiplikation*. Ist $K = \mathbb{R}$ bzw. $K = \mathbb{C}$, so heißt V ein *reeller* bzw. *komplexer* Vektorraum.

Beispiel 4.2.

(1) Die „klassischen“ Vektoren sind die n-Tupel aus $\mathbb{R}^n$. Dann ist $\mathbb{R}^n$ ein $\mathbb{R}$-Vektorraum, indem wir definieren

$$\alpha(x_1, \ldots, x_n) := (\alpha x_1, \ldots, \alpha x_n)$$

für alle $\alpha \in \mathbb{R}$ und $(x_1, \ldots, x_n) \in \mathbb{R}^n$. (Übung!)

(2) Ähnlich wie in (1): Sei K ein Körper. Dann ist K^n ein K-Vektorraum.

(3) Sei K ein Körper. Die Menge $\{0\}$ ist immer ein K-Vektorraum.

(4) Sei K ein Körper. Die Menge $\mathrm{Abb}(K, K)$ ist ein K-Vektorraum auf folgende Weise: Wir definieren die Addition von zweier Elementen $f, g \in \mathrm{Abb}(K, K)$ durch

$$(f + g)(x) := f(x) + g(x) \quad \text{für alle } x \in K,$$

und die Skalarmultiplikation von Elementen $\alpha \in K$ und $f \in \mathrm{Abb}(K, K)$ durch

$$(\alpha f)(x) := \alpha f(x) \qquad \forall x \in K.$$

Übung: Diese Operationen machen $\mathrm{Abb}(K, K)$ zu einem K-Vektorraum.

Bemerkung 4.3. Seien K ein Körper und $L \subseteq K$ ein Unterkörper. Ein K-Vektorraum ist immer auch ein L-Vektorraum.

Proposition 4.4. *Sei V ein K-Vektorraum.*[8] *Dann gilt:*

(a) $\underbrace{0}_{in\ K} \cdot v = \underbrace{0}_{in\ V}$ *für alle* $v \in V$,

(b) $(-1) \cdot v = -v$ *für alle* $v \in V$.

Beweis.
(a) Es gilt

$$0v = (0+0)v \overset{(V1)}{=} 0v + 0v,$$

und damit $0v = 0$.

(b) Es gilt

$$v + (-1)v \overset{(V3)}{=} 1v + (-1)v \overset{(V1)}{=} (1 + (-1))v = 0v \overset{(a)}{=} 0,$$

und damit $(-1)v = -v$. □

Definition 4.5. Sei V ein K-Vektorraum. Eine Teilmenge $W \subseteq V$ heißt *Unterraum* von V, falls W zusammen mit der Vektoraddition von V und der Skalarmultiplikation von K ein K-Vektorraum wird.

Bemerkung 4.6. Implizit in der Definition ist es, dass $(W, +)$ eine abelsche Gruppe ist. Ferner gilt: W ist genau dann ein Unterraum, wenn:

(1) W ist abgeschlossen bezüglich der Vektoraddition, d.h.:

$$\text{für alle } v, w \in W \text{ gilt } v + w \in W, \text{ und}$$

(2) W ist abgeschlossen bezüglich der Skalarmultiplikation, d.h.:

$$\text{für alle } v \in W \text{ und } \alpha \in K \text{ gilt } \alpha v \in W.$$

Beispiel 4.7 (Hyperebene)**.**

(a) Sei K ein Körper, sei $n \in \mathbb{N}_{>0}$ und seien $\lambda_1, \ldots, \lambda_n$ feste Elemente in K, die nicht alle verschwinden. Die Menge

$$H := \{(x_1, \ldots, x_n) \in K^n \mid \lambda_1 x_1 + \cdots + \lambda_n x_n = 0\}$$

ist eine *Hyperebene* in K^n. Eine Hyperebene ist ein Unterraum von K^n, denn es gilt:

[8] Dieser Satz impliziert immer implizit, dass K ein Körper ist.

(1) H ist abgeschlossen bezüglich der Vektoraddition: Und zwar, seien $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in H$. Dann gilt

$$\sum_{i=1}^{n} \lambda_i x_i = 0 \quad \text{und} \quad \sum_{i=1}^{n} \lambda_i y_i = 0,$$

und damit

$$0 = \sum_{i=1}^{n} \lambda_i x_i + \sum_{i=1}^{n} \lambda_i y_i = \sum_{i=1}^{n} \lambda_i (x_i + y_i),$$

also $(x_1 + y_1, \ldots, x_n + y_n) \in H$.

(2) H ist abgeschlossen bezüglich der Skalarmultiplikation: Sei $\alpha \in K$ und sei $(x_1, \ldots, x_n) \in H$. Dann gilt

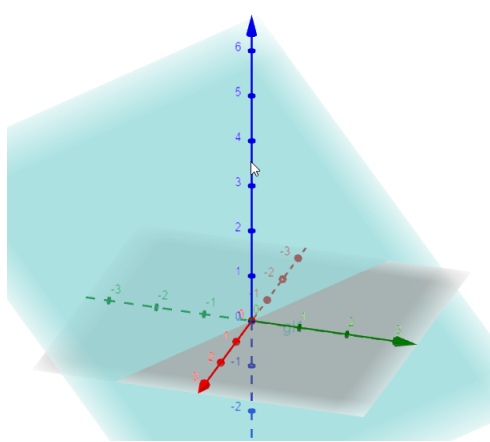
$$\sum_{i=1}^{n} \lambda_i x_i = 0,$$

und damit

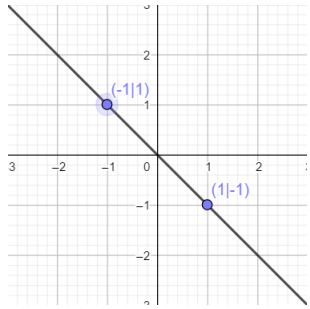
$$0 = \alpha \sum_{i=1}^{n} \lambda_i x_i = \sum_{i=1}^{n} \lambda_i (\alpha x_i),$$

also $(\alpha x_1, \ldots, \alpha x_n) \in H$.

(b) $H = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y + 3z = 0\} \subseteq \mathbb{R}^3$:



(c) $H = \{(x, y) \in \mathbb{R}^2 \mid x + y = 0\} \subseteq \mathbb{R}^2$:



Proposition 4.8. *Sei V ein K-Vektorraum und sei $\{W_i\}_{i \in I}$ eine Familie von Unterräumen von V. Dann ist*

$$\bigcap_{i \in I} W_i$$

auch ein Unterraum von V.

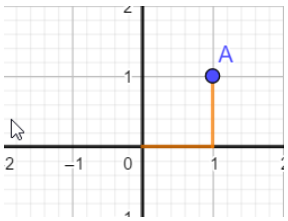
Beweis. Abgeschlossenheit der Vektoraddition: Seien $v, w \in \bigcap_{i \in I} W_i$. Dann gilt $v, w \in W_i$ für jedes $i \in I$, und da jedes W_i ein Unterraum ist, so gilt $v + w \in W_i$ für jedes $i \in I$, und damit $v + w \in \bigcap_{i \in I} W_i$.

Abgeschlossenheit der Skalarmultiplikation: Sei $\alpha \in K$ und sei $v \in \bigcap_{i \in I} W_i$. Dann gilt $v \in W_i$ für jedes $i \in I$, und da jedes W_i ein Unterraum ist, so gilt $\alpha v \in W_i$ für jedes $i \in I$, und damit $\alpha v \in \bigcap_{i \in I} W_i$. $\square$

Warnung. Die Vereinigung von Unterräumen ist nicht zwingend ein Unterraum! Gegenbeispiel: Sei

$$V = \mathbb{R}^2, \quad W_1 = \{(x, 0) \in \mathbb{R} \mid x \in \mathbb{R}\}, \quad W_2 = \{(0, y) \in \mathbb{R}^2 \mid y \in \mathbb{R}\}.$$

Dann ist $W_1 \cup W_2$ nicht abgeschlossen bezüglich der Vektoraddition, da $(1, 0) \in W_1$ und $(0, 1) \in W_2$, aber $(1, 0) + (0, 1) = (1, 1) \notin W_1 \cup W_2$.



Notation 4.9. Ein n-Tupel $(x_1, \ldots, x_n) \in K^n$ schreiben wir üblicherweise als *Spaltenvektor*:

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

Definition 4.10. Sei V ein K-Vektorraum und seien $v_1, \ldots, v_n \in V$ Vektoren. Eine *lineare Kombination* von $v_1, \ldots, v_n$ ist ein Vektor der Form

$$k_1 v_1 + \cdots + k_n v_n, \quad \text{wobei } k_1, \ldots, k_n \in K.$$

Beispiel 4.11. Sei $V = \mathbb{R}^2$. Dann ist $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ eine lineare Kombination von $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Definition 4.12. Sei V ein K-Vektorraum und $v_1, \ldots, v_n \in V$. Das *Erzeugnis* oder die *lineare Hülle* oder der (*lineare*) *Spann* dieser Vektoren ist die Menge

$$\langle v_1, \ldots, v_n \rangle := \left\{ \sum_{i=1}^{n} k_i v_i \,\middle|\, k_i \in K \right\}.$$

Notation 4.13. In der Regel schreiben wir $\langle \emptyset \rangle := \{0\}$.

Beispiel 4.14. Seien $V = \mathbb{R}^3$ und seien v_1, v_2 zwei *nicht parallele* Vektoren (d.h. es gibt keinen Skalar $\alpha \in K$, sodass $v_1 = \alpha v_2$ oder $v_2 = \alpha v_1$). Dann ist

$$\langle v_1, v_2 \rangle = \{k_1 v_1 + k_2 v_2 \mid k_1, k_2 \in \mathbb{R}\}$$

eine Ebene durch den Nullpunkt.

Proposition 4.15. *Sei V ein K-Vektorraum und seien $v_1, \ldots, v_n \in V$. Dann ist $\langle v_1, \ldots, v_n \rangle$ ein Unterraum von V.*

Beweis. Seien

$$\sum_{i=1}^{n} k_i v_i, \sum_{i=1}^{n} l_i v_i \in \langle v_1, \ldots, v_n \rangle.$$

Dann gilt

$$\sum_{i=1}^{n} k_i v_i + \sum_{i=1}^{n} l_i v_i = \sum_{i=1}^{n} (k_i + l_i) v_i \in \langle v_1, \ldots, v_n \rangle,$$

und für $\alpha \in K$ gilt

$$\alpha \sum_{i=1}^{n} k_i v_i = \sum_{i=1}^{n} (\alpha k_i) v_i \in \langle v_1, \ldots, v_n \rangle.$$

Die Proposition ist bewiesen. $\square$

Definition 4.16. Ein K-Vektorraum V ist *endlich-dimensional*, falls es endlich viele $v_1, \ldots, v_n \in V$ gibt, sodass

$$V = \langle v_1, \ldots, v_n \rangle.$$

Ansonsten ist V *unendlich-dimensional*.

Beispiel 4.17.

(a) Der K-Vektorraum K^n ist immer endlich-dimensional, weil K^n das Erzeugnis der Vektoren

$$\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{-te Stelle}$$

ist, für $1 \leq i \leq n$.

(b) Der Vektorraum $\{0\}$ ist endlich-dimensional, da $\{0\} = \langle \emptyset \rangle$.

(c) Der $\mathbb{R}$-Vektorraum $\mathbb{R}[x]$ ist unendlich-dimensionaler. Und zwar, nehmen wir an, dass es endlich viele Vektoren $v_1, \ldots, v_n \in \mathbb{R}[x]$ gibt, sodass $\mathbb{R}[x] = \langle v_1, \ldots, v_n \rangle$. Dann ist jedes Element $w \in \mathbb{R}[x]$ der Form $w = \sum_{i=1}^{n} \alpha_i v_i$, wobei $\alpha_i \in \mathbb{R}$. Sei $n_i := \deg v_i$. Dann gilt

$$\deg w \leq \max\{n_i \mid 1 \leq i \leq n\}.$$

Im Allgemeinen aber kann w einen beliebigen Grad haben. Widerspruch.

Definition 4.18. Sei V ein K-Vektorraum. Vektoren $v_1, \ldots, v_n \in V$ heißen *linear unabhängig*, bzw. eine Menge $\{v_1, \ldots, v_n\}$ heißt *linear unabhängig*, falls für $\alpha_i \in K$ gilt

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0 \quad \Longrightarrow \quad \alpha_i = 0 \text{ für alle } i.$$

Ansonsten heißen sie *linear abhängig*.

Beispiel 4.19.

(1) In $\mathbb{R}^3$ sind die Vektoren $\begin{pmatrix}1\\0\\0\end{pmatrix}$, $\begin{pmatrix}0\\1\\0\end{pmatrix}$ und $\begin{pmatrix}0\\0\\1\end{pmatrix}$ linear unabhängig, weil aus

$$\underbrace{\alpha_1\begin{pmatrix}1\\0\\0\end{pmatrix}+\alpha_2\begin{pmatrix}0\\1\\0\end{pmatrix}+\alpha_3\begin{pmatrix}0\\0\\1\end{pmatrix}}_{=\begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix}}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

folgt, dass $\alpha_1=\alpha_2=\alpha_3=0$.

(2) In $\mathbb{R}^3$ sind die Vektoren $\begin{pmatrix}1\\1\\0\end{pmatrix}$ und $\begin{pmatrix}1\\-1\\0\end{pmatrix}$ linear unabhängig: Sei

$$\underbrace{\alpha_1\begin{pmatrix}1\\1\\0\end{pmatrix}+\alpha_2\begin{pmatrix}1\\-1\\0\end{pmatrix}}_{=\begin{pmatrix}\alpha_1+\alpha_2\\\alpha_1-\alpha_2\\0\end{pmatrix}}=\begin{pmatrix}0\\0\\0\end{pmatrix}.$$

Daraus folgt

$$\alpha_1+\alpha_2=0 \quad \text{und} \quad \alpha_1-\alpha_2=0,$$

und die einzige Lösung dieses Gleichungssystems ist $\alpha_1=\alpha_2=0$.

(3) In $\mathbb{R}^2$ sind $\begin{pmatrix}1\\0\end{pmatrix}$, $\begin{pmatrix}0\\1\end{pmatrix}$ und $\begin{pmatrix}1\\1\end{pmatrix}$ linear abhängig, denn es gilt

$$-1\begin{pmatrix}1\\1\end{pmatrix}+1\begin{pmatrix}1\\0\end{pmatrix}+1\begin{pmatrix}0\\1\end{pmatrix}=0.$$

(4) In $\mathbb{R}[x]$ sind die Vektoren $x^4+x^2+2, x^4+\frac{1}{2}x^2, x^2+1$ und 1 linear abhängig, denn es gilt

$$-1\left(x^4+x^2+2\right)+1\left(x^4+\frac{1}{2}x^2\right)+\frac{1}{2}\left(x^2+1\right)+\frac{3}{2}\cdot 1=0.$$

(5) In einem beliebigen K-Vektorraum V ist $\{0\}$ linear abhängig, da für alle $\alpha \in K$ gilt $\alpha \cdot 0 = 0$.

Proposition 4.20. *Sie V ein K-Vektorraum. Vektoren $v_1, \ldots, v_n \in V$ sind genau dann linear unabhängig, wenn sich kein Vektor v_i als lineare Kombination von den Vektoren $v_1, \ldots, v_{i-1}$ schreiben lässt.*

Beweis.
$\Longrightarrow$: Seien $v_1, \ldots, v_n$ linear unabhängig. Angenommen, ein Vektor v_i ist eine lineare Kombination der Vektoren $v_1, \ldots, v_{i-1}$. Dann gibt es $\alpha_1, \ldots, \alpha_{i-1} \in K$, sodass

$$v_i = \alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1},$$

und damit

$$0 = \alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + (-v_i) + 0 v_{i+1} + \cdots + 0 v_n.$$

Die Vektoren $v_1, \ldots, v_n$ sind also linear abhängig. Widerspruch.

$\Longleftarrow$: Wir zeigen die Aussage durch die Kontraposition.

Seien $v_1, \ldots, v_n$ linear abhängig. Dann gibt es $\alpha_1, \ldots, \alpha_n \in K$, die nicht alle verschwinden, sodass

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0.$$

Sei $\ell := \max\{1 \leq j \leq n \mid \alpha_j \neq 0\}$. Dann folgt:

$$\alpha_1 v_1 + \cdots + \alpha_{\ell-1} v_{\ell-1} + \alpha_\ell v_\ell + 0 v_{l+1} + \cdots + 0 v_n = 0,$$

und damit

$$v_\ell = -\frac{1}{\alpha_\ell}(\alpha_1 v_1 + \cdots + \alpha_{\ell-1} v_{\ell-1}).$$

Dies zeigt, dass v_ℓ eine lineare Kombination von $v_1, \ldots, v_{\ell-1}$ ist. $\square$

Proposition 4.21. *Sei V ein K-Vektorraum, seien $v_1, \ldots, v_n \in V$ linear unabhängige Vektoren und sei $v \in \langle v_1, \ldots, v_n \rangle$. Dann sind die Skalare $\alpha_1, \ldots, \alpha_n \in K$ in der Formel*

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

eindeutig bestimmt.

Beweis. Seien $k_1, \ldots, k_n \in K$ und $\ell_1, \ldots, \ell_n \in K$ mit

$$v = k_1 v_1 + \cdots + k_n v_n \quad \text{und} \quad v = \ell_1 v_1 + \cdots + \ell_n v_n.$$

Daraus folgt:

$$\begin{aligned} 0 = v - v &= (k_1 v_1 + \cdots + k_n v_n) - (\ell_1 v_1 + \cdots + \ell_n v_n) \\ &= (k_1 - \ell_1) v_1 + \cdots + (k_n - \ell_n) v_n. \end{aligned}$$

Da $v_1, \ldots, v_n$ linear unabhängig sind, so folgt $k_i - \ell_i = 0$ für alle i, und damit $k_i = \ell_i$ für alle i. □

Definition 4.22. Sei V ein endlich-dimensionaler K-Vektorraum und seien $v_1, \ldots, v_n \in V$. Die Menge $\{v_1, \ldots, v_n\}$ ist ein *Erzeugendensystem* von V, falls $V = \langle v_1, \ldots, v_n \rangle$. Wenn zusätzlich die Vektoren $v_1, \ldots, v_n$ linear unabhängig sind, dann ist die Menge $\{v_1, \ldots, v_n\}$ eine *Basis* von V.

Definition 4.23. Sei V ein endlich-dimensionaler K-Vektorraum und sei $\mathcal{B} = \{v_1, \ldots, v_n\}$ eine Basis von V. Nach Proposition 4.21 sind für jedes $v \in V$ die Skalare $\alpha_1, \ldots, \alpha_n \in K$ in der Formel

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

eindeutig bestimmt. Der Vektor

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in K^n$$

heißt der *Koordinatenspaltenvektor von v bezüglich der Basis $\mathcal{B}$*, oder die *Koordinaten von v in der Basis $\mathcal{B}$*.

Beispiel 4.24.

(a) Die Menge $\{e_i \mid 1 \leq i \leq n\}$, wobei

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

heißt die *kanonische Basis* von K^n.

(b) Sei $\mathbb{R}[x]^{\leq n}$ die Menge aller Polynome in $\mathbb{R}[x]$ vom Grad $\leq n$. Dies ist ein $\mathbb{R}$-Vektorraum und eine Basis von $\mathbb{R}[x]^{\leq n}$ ist die Menge

$$\{1, x, x^2, \ldots, x^n\}.$$

Diese Menge nennen wir *kanonische Basis* von $\mathbb{R}[x]^{\leq n}$.

Lemma 4.25. *Sei V ein endlich dimensionaler K-Vektorraum und sei $M = \{v_1, \ldots, v_n\}$ eine nichtleere Teilmenge von V.*

(i) Die Menge M ist genau dann eine Basis von V, wenn sie eine maximale linear unabhängige Teilmenge von V ist. (In anderen Worten: Keine echte Obermenge von M ist linear unabhängig.)

(ii) Die Menge M ist genau dann eine Basis von V, wenn sie eine minimale erzeugende Teilmenge von V ist. (In anderen Worten: Keine echte Teilmenge von M erzeugt V.)

Beweis.
(i)

$\Longrightarrow$: Sei M eine Basis von V. Angenommen, M ist nicht eine maximal linear unabhängige Teilmenge von V. Dann gibt es $v \in V \setminus M$, sodass die Menge $M \cup \{v\} = \{v_1, \ldots, v_n, v\}$ linear unabhängig ist. Da M eine Basis ist, so gibt es $\beta_1, \ldots, \beta_n \in K$ mit

$$\beta_1 v_1 + \cdots + \beta_n v_n = v.$$

Damit sind $v, v_1, \ldots, v_n$ sind nicht linear unabhängig. Widerspruch.

$\Longleftarrow$: Angenommen, M ist eine maximale linear unabhängige Teilmenge von V. Dann ist die Menge $M \cup \{v\}$ linear abhängig für alle $v \in V \setminus M$. In anderen Worten, es gibt $\alpha, \alpha_1, \ldots, \alpha_n \in K$, die nicht alle verschwinden, sodass

$$\alpha_1 v_1 + \cdots + \alpha_n v_n + \alpha v = 0.$$

Nun: Wenn $\alpha = 0$, daraus folgt

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0,$$

welches besagt, dass $v_1, \ldots, v_n$ linear abhängig sind, ein Widerspruch. Wenn $\alpha \neq 0$, dann haben wir

$$v = -\frac{1}{\alpha}(\alpha_1 v_1 + \cdots + \alpha_n v_n),$$

und damit $v \in \langle v_1, \ldots, v_n \rangle$ für alle $v \in V$. In anderen Worten, M ist ein Erzeugendensystem von V. Da M auch linear unabhängig ist, so ist M eine Basis.

(ii)

$\Longrightarrow$: Sei M eine Basis von V. Angenommen, es existiert eine echte Teilmenge von M, die V erzeugt. Ohne Beschränkung der Allgemeinheit: $\{v_1, \ldots, v_\ell\}$ erzeugt V, wobei $\ell < n$. Dann gibt es $\alpha_1, \ldots, \alpha_l \in K$, sodass

$$v_n = \alpha_1 v_1 + \cdots + \alpha_\ell v_\ell,$$

oder äquivalent,

$$\alpha_1 v_1 + \cdots + \alpha_\ell v_\ell - v_n = 0.$$

Also, die Menge $\{v_1, \ldots, v_\ell, v_n\}$ ist nicht linear unabhängig, ein Widerspruch.

$\Longleftarrow$: Sei M eine minimale erzeugende Teilmenge von V. Wir müssen zeigen, dass M linear unabhängig ist. Angenommen, M ist linear abhängig. Dann gibt es $\gamma_1, \ldots, \gamma_n \in K$, die nicht alle verschwinden, sodass

$$\gamma_1 v_1 + \cdots + \gamma_n v_n = 0.$$

Ohne Beschränkung der Allgemeinheit dürfen wir annehmen, dass $\gamma_n \neq 0$, und damit

$$v_n = \sum_{i=1}^{n-1} \left(-\frac{\gamma_i}{\gamma_n} \right) v_i.$$

Nun: Da M ein Erzeugendensystem ist, für jedes $v \in V$ existieren $\alpha_1, \ldots, \alpha_n \in K$, sodass

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

Es folgt:

$$\begin{aligned} v &= \alpha_1 v_1 + \cdots + \alpha_{n-1} v_{n-1} + \alpha_n \sum_{i=1}^{n-1} \left(-\frac{\gamma_i}{\gamma_n} \right) v_i \\ &= \sum_{i=1}^{n-1} \left(\alpha_i + \alpha_n \left(-\frac{\gamma_i}{\gamma_n} \right) \right) v_i, \end{aligned}$$

und damit ist jedes $v \in V$ eine lineare Kombination von $v_1, \ldots, v_{n-1}$. Also, $V = \langle v_1, \ldots, v_{n-1} \rangle$, welches der Minimalität von M widerspricht.

□

Satz 4.26. *Sei V ein endlich-dimensionaler K-Vektorraum. Dann hat V eine Basis.*

Beweis. Sei E_0 eine endliche erzeugende Teilmenge von V. Falls E_0 eine minimale erzeugende Teilmenge ist, dann ist sie eine Basis von V nach Lemma 4.25. Wir dürfen also annehmen, dass E_0 nicht eine minimale erzeugende Teilmenge von V ist. Es gibt also $E_1 \subsetneqq E_0$, sodass E_1 erzeugende Teilmenge ist. Nun ist E_1 entweder eine minimale erzeugende Teilmenge von V oder nicht. Wir führen diesen Prozess iterativ weiter; der Prozess muss aufhören, da E_0 endlich ist. □

Austauschsätze von Steinitz

Lemma 4.27 (Austauschlemma)**.** *Sei V ein K-Vektorraum mit einer Basis $\mathcal{B} = \{v_1, \dots, v_n\}$, sei*

$$w = \lambda_1 v_1 + \cdots + \lambda_n v_n \tag{17}$$

eine lineare Kombination von $v_1, \dots, v_n$ mit $\lambda_k \neq 0$ für ein $k \in \{1, \dots, n\}$. Dann ist die Menge

$$\mathcal{B}' = \{v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_n\}$$

eine Basis von V. (Wir können also v_k gegen w austauschen.)

Beweis. Ohne Beschränkung der Allgemeinheit dürfen wir annehmen, dass $k = 1$. Wir müssen zeigen, dass $\mathcal{B}'$ ein Erzeugendensystem von V und linear unabhängig ist.

Sei $v \in V$. Dann gibt es $\mu_1, \dots, \mu_n \in K$, sodass

$$v = \mu_1 v_1 + \cdots + \mu_n v_n. \tag{18}$$

Aus (17) folgt

$$v_1 = -\frac{\lambda_2}{\lambda_1} v_2 - \cdots - \frac{\lambda_n}{\lambda_1} v_n + \frac{1}{\lambda_1} w. \tag{19}$$

Aus (18) und (19) folgt

$$\begin{aligned} v &= \mu_1 \left(-\frac{\lambda_2}{\lambda_1} v_2 - \cdots - \frac{\lambda_n}{\lambda_1} v_n + \frac{1}{\lambda_1} w \right) + \mu_2 v_2 + \cdots + \mu_n v_n \\ &= \frac{\mu_1}{\lambda_1} w + \left(\mu_2 - \frac{\mu_1 \lambda_2}{\lambda_1} \right) v_2 + \cdots + \left(\mu_n - \frac{\mu_1 \lambda_n}{\lambda_1} \right) v_n. \end{aligned}$$

Also, v ist eine lineare Kombination von $w, v_2, \ldots, v_n$, sodass $\mathcal{B}'$ ein Erzeugendensystem von V ist.

Jetzt zeigen wir, dass $\mathcal{B}'$ linear unabhängig ist. Angenommen, das ist nicht der Fall: dann gibt es $\theta_1, \ldots, \theta_n \in K$, die nicht alle verschwunden, sodass

$$\theta_1 w + \theta_2 v_2 + \cdots + \theta_n v_n = 0.$$

Aus (17) folgt

$$\theta_1(\lambda_1 v_1 + \cdots + \lambda_n v_n) + \theta_2 v_2 + \cdots + \theta_n v_n = 0,$$

und damit

$$\theta_1\lambda_1 v_1 + (\theta_1\lambda_2 + \theta_2)v_2 + \cdots + (\theta_1\lambda_n + \theta_n)v_n = 0.$$

Da $v_1, \ldots, v_n$ linear unabhängig sind, so folgt:

$$\begin{aligned} \theta_1\lambda_1 = 0 \quad &\overset{\lambda_1 \neq 0}{\Longrightarrow} \quad \theta_1 = 0 \\ \theta_1\lambda_2 + \theta_2 = 0 \quad &\Longrightarrow \quad \theta_2 = 0 \\ &\vdots \\ \theta_1\lambda_n + \theta_n = 0 \quad &\Longrightarrow \quad \theta_n = 0 \end{aligned}$$

Das ist ein Widerspruch und damit ist $\mathcal{B}'$ linear unabhängig. □

Satz 4.28 (Austauschsatz)**.** *Sei V ein K-Vektorraum, sei $\mathcal{B} = \{v_1, \ldots, v_n\}$ eine Basis von V und sei $\{w_1, \ldots, w_m\}$ eine linear unabhängige Teilmenge von V. Dann gelten die folgenden Aussagen.*

(i) $m \leq n$.

(ii) Es gibt Indizes $i_1, \ldots, i_m \in \{1, \ldots, n\}$ derart, dass man nach dem Austausch (in $\mathcal{B}$) von v_{i_1} gegen w_1, v_{i_2} gegen w_2, usw. wieder eine Basis von V gewinnt.

Beweis.
Der Beweis ist per Induktion nach m, gleichzeitig für (i) und (ii).

Für $m = 0$ ist nichts zu beweisen.

Sei nun $m \geq 1$ und sei die Aussage bewiesen für alle Zahlen $< m$. Bemerke, dass die Menge $\{w_1, \ldots, w_{m-1}\}$ linear unabhängig ist, weil die Menge $\{w_1, \ldots, w_m\}$ linear unabhängig ist. Nach dem Induktionsschritt für (i) gilt $m - 1 \leq n$, also $m \leq n + 1$.

Angenommen, $m \not\leq n$. Dann gilt $m = n + 1$, also $n = m - 1$. Nun, nach dem Induktionsschritt für (ii) können wir alle Elemente von $\mathcal{B}$ gegen $w_1, \ldots, w_{m-1}$ austauschen, sodass dann die Menge $\{w_1, \ldots, w_{m-1}\}$ wieder eine Basis von V ist. Aber das widerspricht Lemma 4.25(i), da die Obermenge $\{w_1, \ldots, w_m\}$ linear unabhängig ist. Dies zeigt (i).

Es gilt also $m \leq n$. Wieder nach dem Induktionsschritt für (ii) können wir $m-1$ Elemente von $\mathcal{B}$ gegen $w_1, \ldots, w_{m-1}$ austauschen; ohne Beschränkung der Allgemeinheit dürfen wir annehmen, dass

$$\mathcal{B}' := \{w_1, \ldots, w_{m-1}, v_m, \ldots, v_n\}$$

eine Basis von V ist. Dann gibt es $\lambda_1, \ldots, \lambda_n \in K$, sodass

$$w_m = \lambda_1 w_1 + \cdots + \lambda_{m-1} w_{m-1} + \lambda_m v_m + \cdots + \lambda_n v_n.$$

Wäre $\lambda_m = \cdots = \lambda_n = 0$, dann hätten wir

$$w_m = \lambda_1 w_1 + \cdots + \lambda_{m-1} w_{m-1}.$$

Das ist ein Widerspruch, da $w_1, \ldots, w_m$ linear unabhängig sind. Es gibt also $i \geq m$ mit $\lambda_i \neq 0$; ohne Beschränkung der Allgemeinheit dürfen wir annehmen, dass $i = m$. Nach Lemma 4.27 können wir dann in der Basis $\mathcal{B}'$ das Element v_m gegen w_m austauschen. Also, $\{w_1, \ldots, w_m, v_{m+1}, \ldots, v_n\}$ ist eine Basis von V. Dies zeigt (ii). □

Korollar 4.29. *Sei V ein endlich-dimensionaler K-Vektorraum. Dann haben jede zwei Basen von V dieselbe Mächtigkeit.*

Beweis. Seien $\mathcal{B}$ und $\mathcal{B}'$ zwei Basen von V. Sei $n := |\mathcal{B}|$ und $m := |\mathcal{B}'|$. Nach Satz 4 gilt dann $m \leq n$ und $m \geq n$, also $m = n$. □

Definition 4.30. Sei V ein endlich-dimensionaler K-Vektorraum. Die *Dimension von V* ist die Anzahl der Elemente in einer beliebigen Basis von V.[9] Diese Zahl bezeichnen wir mit

$$\dim V.$$

Beispiel 4.31.

(a) Per Definition ist $\dim V = 0$ genau dann, wenn $V = \{0\}$.

(b) $\dim K^n = n$ nach Beispiel 4.24.

[9] Dies ist unabhängig von der Wahl des Basis von V nach Korollar 4.29.

(c) $\dim \mathbb{R}[x]^{\leq n} = n + 1$ nach Beispiel 4.24.

Korollar 4.32. *Sei V ein endlich-dimensionaler K-Vektorraum mit $n = \dim V$.*

(1) Jede linear unabhängige endliche Teilmenge von V hat höchstens n Elemente. Hat sie genau n Elemente, so ist diese Menge eine Basis von V.

(2) Jede erzeugende endliche Teilmenge von V hat mindestens n Elemente. Hat sie genau n Elemente, so ist diese Menge eine Basis von V.

Beweis. Folgt aus Lemma 4.25 und Korollar 4.29. □

Lemma 4.33 (Dimension von Unterräumen)**.** *Sei V ein endlich-dimensionaler K-Vektorraum. Dann ist jeder Unterraum $U \subseteq V$ auch endlich-dimensional und es gilt*

$$\dim U \leq \dim V.$$

Es gilt $\dim U = \dim V$ genau dann, wenn $U = V$.

Beweis. Seien $n := \dim V$ und seien $v_1, \ldots, v_m \in U$ linear unabhängige Elemente in U. Dann sind $v_1, \ldots, v_m$ auch linear unabhängig in V, und damit gilt $m \leq n$ nach Satz 4.28. Also, jede linear unabhängige Menge in U hat maximal n Elemente, und damit ist U auch endlich-dimensional. Sei $\{w_1, \ldots, w_{m'}\}$ eine maximale linear unabhängige Teilmenge von U. Nach Lemma 4.25 ist dann $\{w_1, \ldots, w_{m'}\}$ ist eine Basis von U, und es folgt:

$$\dim U = m' \leq n = \dim V.$$

Wenn $\dim U = \dim V$, also $m' = n$, und $\{w_1, \ldots, w_{m'}\}$ linear unabhängig in V. Nach Satz 4.28 und Lemma 4.25 ist dann $\{w_1, \ldots, w_{m'}\}$ eine Basis von V. Folglich gilt: $U = \langle w_1, \ldots, w_{m'} \rangle = V$, wie gewünscht. □

Beispiel 4.34. Da $\dim \mathbb{R}^3 = 3$, so haben alle echten Unterräume von $\mathbb{R}^3$ entweder Dimension 1 oder Dimension 2. Die Unterräume der Dimension 1 sind Geraden durch den Nullpunkt; die Unterräume der Dimension 2 sind Ebenen durch den Nullpunkt.

Lemma 4.35 (Basisergänzung)**.** *Sei V ein endlich-dimensionaler K-Vektorraum mit $\dim V = n$, sei W ein Unterraum von V und sei $\{v_1, \ldots, v_m\}$ eine Basis von W. Dann gibt es Elemente $v_{m+1}, \ldots, v_n \in V \setminus W$, sodass $\{v_1, \ldots, v_n\}$ eine Basis von V ist.*

Beweis. Die Aussage ist trivial, wenn $W = V$. Wir dürfen also annehmen, dass $W \neq V$. Wähle $v_{m+1} \in V \setminus W$.

Wir behaupten, dass die Menge $\{v_1, \ldots, v_{m+1}\}$ linear unabhängig ist. Und zwar, seien $\lambda_1, \ldots, \lambda_{m+1} \in K$ mit

$$\lambda_1 v_1 + \cdots + \lambda_{m+1} v_{m+1} = 0.$$

Falls $\lambda_{m+1} \neq 0$, dann

$$v_{m+1} = -\frac{1}{\lambda_{m+1}}(\lambda_1 v_1 + \cdots + \lambda_m v_m) \in \langle v_1, \ldots, v_m \rangle,$$

Widerspruch. Es ist also $\lambda_{m+1} = 0$. Aber dann gilt

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0,$$

und daraus folgt, dass auch $\lambda_1 = \cdots = \lambda_m = 0$, da $v_1, \ldots, v_m$ linear unabhängig sind. Die Behauptung ist bewiesen.

Setze $W_{m+1} := \langle v_1, \ldots, v_{m+1} \rangle$. Falls $W_{m+1} = V$, dann sind wir fertig. Ansonsten wählen wir ein $v_{m+2} \in V \setminus W_{m+1}$ und setzen diesen Prozess iterativ fort. Der Prozess muss aufhören, da V endlich-dimensional ist. □

Schnitt und Summe von Unterräumen

Definition 4.36. Seien W_1 und W_2 zwei Unterräume eines K-Vektorraumes V. Die *Summe von W_1 und W_2* ist definiert durch

$$W_1 + W_2 := \{w_1 + w_2 \mid w_1 \in W_1, w_2 \in W_2\}.$$

Proposition 4.37. *Seien W_1 und W_2 zwei Unterräume eines K-Vektorraumes V. Dann ist $W_1 + W_2$ auch ein Unterraum von V.*

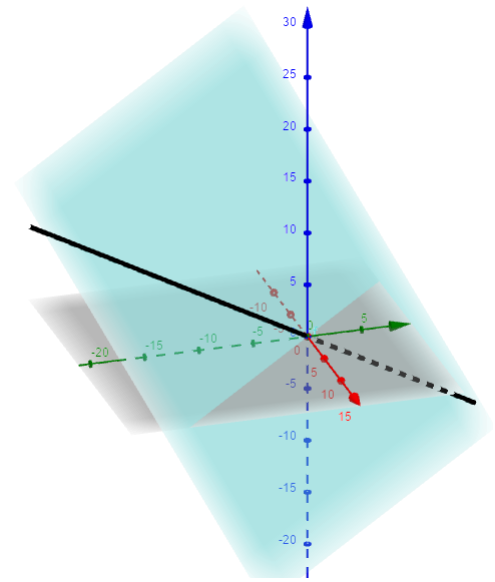
Beweis. Seien $w_1, w_1' \in W_1$ und $w_2, w_2' \in W_2$. Dann gilt:

$$(w_1 + w_2) + (w_1' + w_2') = \underbrace{(w_1 + w_1')}_{\in W_1} + \underbrace{(w_2 + w_2')}_{\in W_2} \in W_1 + W_2,$$

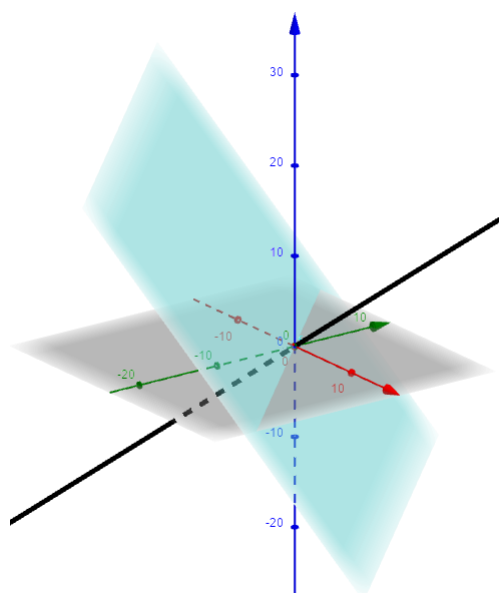
und damit ist $W_1 + W_2$ abgeschlossen bezüglich der Vektoraddition. Analog zeigen wir, dass $W_1 + W_2$ abgeschlossen ist bezüglich der Skalarmultiplikation. □

Beispiel 4.38. Sei W_1 eine Gerade in $\mathbb{R}^3$ und sei W_2 eine Ebene in $\mathbb{R}^3$. Es gibt dann zwei Fälle.

<u>Fall 1</u>: $W_1 \subseteq W_2$. Dann ist $\dim W_1 \cap W_2 = \dim W_1 = 1$. Ferner ist $W_1 + W_2 = W_2$, und damit $\dim W_1 + W_2 = \dim W_2 = 2$. Bemerke, dass $1 + 2 = 3$.



<u>Fall 2</u>: $W_1 \not\subseteq W_2$. Dann ist $W_1 \cap W_2 = \{0\}$, und damit $\dim W_1 \cap W_2 = 0$. Ferner ist $W_1 + W_2 = \mathbb{R}^3$, und damit $\dim W_1 + W_2 = \dim \mathbb{R}^3 = 3$. Bemerke, dass $0 + 3 = 3$.



Im Allgemeinen gilt:

Satz 4.39 (Dimensionssatz/Dimensionsformel für Unterräume)**.** *Seien W_1 und W_2 zwei Unterräume eines endlich-dimensionalen K-Vektorraumes V. Dann gilt*

$$\dim W_1 + \dim W_2 = \dim (W_1 \cap W_2) + \dim (W_1 + W_2).$$

Beweis. Seien $r := \dim W_1 \cap W_2$, $s := \dim W_1$ und $t := \dim W_2$, und sei $\{v_1, \ldots, v_r\}$ eine Basis von $W_1 \cap W_2$.

Da $W_1 \cap W_2$ ein Unterraum von W_1 ist, so existieren nach Satz 4.35 Elemente $w_{r+1}, \ldots, w_s \in W_1 \setminus (W_1 \cap W_2)$, sodass $\{v_1, \ldots, v_r, w_{r+1}, \ldots, w_s\}$ eine Basis von W_1 ist.

In ähnlicher Weise gibt es Elemente $z_{r+1}, \ldots, z_t \in W_2 \setminus (W_1 \cap W_2)$, sodass $\{v_1, \ldots, v_r, z_{r+1}, \ldots, z_t\}$ eine Basis von W_2 ist.

Wir müssen zeigen, dass

$$\dim(W_1 + W_2) = t + s - r. \tag{20}$$

Wir zeigen dies in zwei Schritten.

Schritt 1. Setze

$$U := \langle v_1, \ldots, v_r, w_{r+1}, \ldots, w_s, z_{r+1}, \ldots, z_t \rangle.$$

Dann behaupten wir, dass

$$W_1 + W_2 = U.$$

Und zwar, seien $w_1 \in W_1$ und $w_2 \in W_2$. Dann können wir schreiben

$$w_1 = \alpha_1 v_1 + \cdots + \alpha_r v_r + \alpha_{r+1} w_{r+1} + \cdots + \alpha_s w_s$$

und

$$w_2 = \beta_1 v_1 + \cdots + \beta_r v_r + \beta_{r+1} z_{r+1} + \cdots + \beta_t z_t,$$

und folglich:

$$w_1 + w_2 = \sum_{i=1}^{r} (\alpha_i + \beta_i) v_i + \sum_{i=r+1}^{s} \alpha_i w_i + \sum_{i=r+1}^{t} \beta_i z_i \in U.$$

Also,

$$W_1 + W_2 \subseteq U.$$

Sei umgekehrt $w \in U$. Dann können wir schreiben

$$w = \underbrace{\sum_{i=1}^{r} \alpha_i v_i + \sum_{i=r+1}^{s} \beta_i w_i}_{\in W_1} + \underbrace{\sum_{i=r+1}^{t} \gamma_i z_i}_{\in W_2} \in W_1 + W_2,$$

und damit

$$U \subseteq W_1 + W_2.$$

Dies zeigt die Behauptung.

Schritt 2. Um (20) zu zeigen, reicht es zu zeigen, dass die Menge

$$\mathcal{B} := \{v_1, \ldots, v_r, w_{r+1}, \ldots, w_s, z_{r+1}, \ldots, z_t\}$$

eine Basis von $W_1 + W_2$ ist. Schritt 1 zeigt, dass $\mathcal{B}$ ein Erzeugnis von $W_1 + W_2$ ist, und es bleibt zu beweisen, dass die Menge $\mathcal{B}$ linear unabhängig ist.

Seien $\alpha_i, \beta_j, \gamma_k \in K$ mit

$$\sum_{i=1}^{r} \alpha_i v_i + \sum_{j=r+1}^{s} \beta_j w_j + \sum_{k=r+1}^{t} \gamma_k z_k = 0. \tag{21}$$

Daraus folgt

$$\underbrace{\sum_{k=r+1}^{t} \gamma_k z_k}_{\in W_2} = \underbrace{-\sum_{i=1}^{r} \alpha_i v_i - \sum_{j=r+1}^{s} \beta_j w_j}_{\in W_1},$$

und insbesondere:

$$\sum_{k=r+1}^{t} \gamma_k z_k \in W_1 \cap W_2 = \langle v_1, \ldots, v_r \rangle.$$

Es existieren dann Skalare $\delta_1, \ldots, \delta_r \in K$, sodass

$$\sum_{k=r+1}^{t} \gamma_k z_k = \sum_{i=1}^{r} \delta_i v_i,$$

oder äquivalent:

$$\sum_{i=1}^{r} \delta_i v_i - \sum_{k=r+1}^{t} \gamma_k z_k = 0.$$

Da die Menge $\{v_1, \ldots, v_r, z_{r+1}, \ldots, z_t\}$ linear unabhängig ist, so folgt, dass

$$\delta_i = \gamma_k = 0 \quad \text{für alle } i \text{ und } k.$$

Wenn wir dies einsetzen in (21), bekommen wir:

$$\sum_{i=1}^{r} \alpha_i v_i + \sum_{j=r+1}^{s} \beta_j w_j = 0.$$

Da die Menge $\{v_1, \ldots, v_r, w_{r+1}, \ldots, w_s\}$ linear unabhängig ist, so folgt, dass

$$\alpha_i = \beta_j = 0 \quad \text{für alle } i \text{ und } j.$$

Es gilt also $\alpha_i = \beta_j = \gamma_k = 0$ für alle i, j und k, und damit ist die Menge $\mathcal{B}$ linear unabhängig, wie gewünscht. □

Direkte Summe

Definition 4.40. Seien W_1 und W_2 zwei Unterräume eines K-Vektorraums V. Die Summe $W_1 + W_2$ heißt *direkt*, falls jedes $w \in W_1 + W_2$ eindeutig geschrieben werden kann als

$$w = w_1 + w_2, \quad \text{wobei } w_1 \in W_1 \text{ und } w_2 \in W_2.$$

Dann bezeichnen wir die Summe $W_1 + W_2$ mit

$$W_1 \oplus W_2.$$

Beispiel 4.41. Sei $V = \mathbb{R}^2$, sei W_1 die x-Achse und sei W_2 die y-Achse. Dann ist $W_1 + W_2$ direkt. Und zwar, sei $w \in W_1 + W_2$. Dann ist $w = (a, b) \in \mathbb{R}^2$ und die Summe $w = w_1 + w_2 = (a, 0) + (0, b)$ ist eindeutig bestimmt.

Lemma 4.42. *Seien W_1 und W_2 zwei Unterräume eines K-Vektorraums V. Dann ist die Summe $W_1 + W_2$ genau dann direkt, wenn*

$$W_1 \cap W_2 = \{0\}.$$

Beweis.
$\Longrightarrow$: Sei $w \in W_1 \cap W_2$. Dann ist

$$w = \underbrace{0}_{\in W_1} + \underbrace{w}_{\in W_2} = \underbrace{w}_{\in W_1} + \underbrace{0}_{\in W_2}.$$

Aus der Eindeutigkeit folgt, dass $w = 0$.

$\Longleftarrow$: Angenommen, die Summe $W_1 + W_2$ ist nicht direkt. Dann gibt es $w \in W_1 + W_2$, welches nicht eindeutig geschrieben werden kann als Summe von Elementen aus W_1 und W_2. D.h. es gibt verschiedene $y_1, y_2 \in W_1$ und verschiedene $z_1, z_2 \in W_2$ mit

$$w = y_1 + z_1 = y_2 + z_2.$$

Dann gilt:

$$\underbrace{y_1 - y_2}_{\in W_1} = \underbrace{z_2 - z_1}_{\in W_2},$$

und damit

$$y_1 - y_2 \in W_1 \cap W_2 = \{0\},$$

welches ein Widerspruch ist. $\square$

Bemerkung 4.43. Nach Satz 4.39 gilt:

$$\dim(W_1 \oplus W_2) = \dim W_1 + \dim W_2 - \underbrace{\dim(W_1 \cap W_2)}_{=0} = \dim W_1 + \dim W_2.$$

Definition 4.44. Sei V ein K-Vektorraum. Zwei Unterräume W_1 und W_2 von V heißen *komplementär*, falls

$$V = W_1 \oplus W_2.$$

Wir nennen dann W_1 das *Komplement* von W_2, und umgekehrt.

Lemma 4.45. *Sei V ein endlich-dimensionaler K-Vektorraum und sei W ein Unterraum von V. Dann existiert ein Komplement von W in V.*

Beweis. Sei $\{v_1, \ldots, v_r\}$ eine Basis von W. Nach Satz 4.35 existieren Vektoren $v_{r+1}, \ldots, v_n \in V$, sodass $\{v_1, \ldots, v_n\}$ eine Basis von V ist. Setze

$$Z := \langle v_{r+1}, \ldots, v_n \rangle.$$

Wir behaupten, dass Z ein Komplement von W ist. Wir zeigen zuerst, dass $V = W + Z$. Die Inklusion $W + Z \subseteq V$ ist trivial. Sei umgekehrt $v \in V$. Dann gibt es $\alpha_1, \ldots, \alpha_n \in K$ mit

$$v = \sum_{i=1}^{n} \alpha_i v_i = \underbrace{\sum_{i=1}^{r} \alpha_i v_i}_{\in W} + \underbrace{\sum_{i=r+1}^{n} \alpha_i v_i}_{\in Z} \in W + Z,$$

und damit $V \subseteq W + Z$.

Wir zeigen nun, dass die Summe $W + Z$ direkt ist. Nach Lemma 4.42 reicht es zu zeigen, dass $W \cap Z = \{0\}$. Und zwar, sei $w \in W \cap Z$. Dann existieren $\lambda_1, \ldots, \lambda_r \in K$ und $\mu_{r+1}, \ldots, \mu_n \in K$ mit

$$w = \lambda_1 v_1 + \cdots + \lambda_r v_r \quad \text{und} \quad w = \mu_{r+1} v_{r+1} + \cdots + \mu_n v_n,$$

und folglich

$$0 = w - w = \sum_{i=1}^{r} \lambda_i v_i - \sum_{i=r+1}^{n} \mu_i v_i.$$

Da die Menge $\{v_1, \ldots, v_n\}$ eine Basis von V ist, so folgt $\lambda_i = \mu_i = 0$ für alle i, und damit $w = 0$, wie gewünscht. $\square$

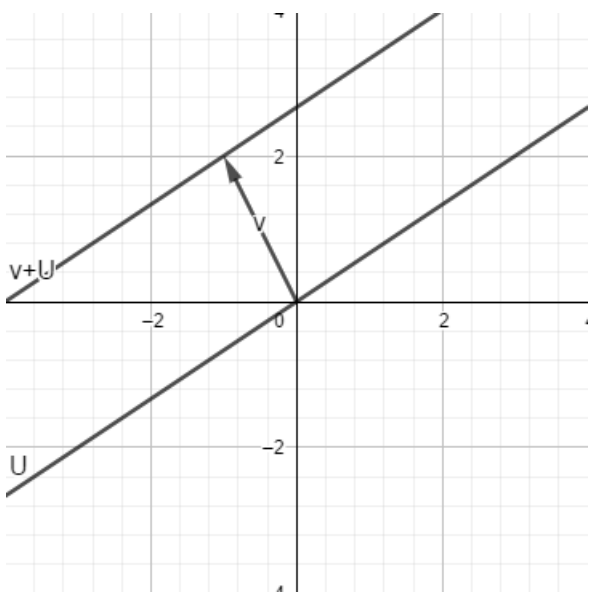
Quotientenräume

Definition 4.46. Sei V ein K-Vektorraum, sei U ein Unterraum von V und sei $v \in V$. Die Menge

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

heißt die *Nebenklasse von U durch v*. Der Vektor v heißt ein *Repräsentant* der Klasse $v + U$.

Beispiel 4.47. Sei $V = \mathbb{R}^2$, sei U eine Gerade durch den Nullpunkt und sei $v \in V$. Dann ist $v + U$ die Gerade, die v enthält und parallel zu U verläuft.



Proposition 4.48. *Sei V ein K-Vektorraum, sei U ein Unterraum von V und seien $v, w \in V$. Dann ist $v + U = w + U$ genau dann, wenn $v - w \in U$.*

Beweis. Folgt sofort aus Lemma 2.21, da $(V, +)$ eine Gruppe und U eine Untergruppe von V ist. □

Lemma/Definition 4.49. *Sei V ein K-Vektorraum und sei U ein Unterraum von V. Die Menge*

$$V/U := \{v + U \mid v \in V\}$$

aller Nebenklassen von U heißt der Quotientenraum *und ist ein K-Vektorraum bezüglich der folgenden Operationen:*

$$\begin{aligned} (v + U) + (w + U) &:= (v + w) + U \quad \textit{für alle } v, w \in V, \\ \alpha(v + U) &:= \alpha v + U \quad \textit{für alle } \alpha \in K,\ v \in V. \end{aligned}$$

Beweis. Die Struktur $(V/U, +)$ ist eine Gruppe (Übung! Siehe Definition 2.31). Es bleibt zu zeigen, dass die Skalarmultiplikation wohldefiniert ist.

Seien $v, w \in V$, sodass $v + U = w + U$. Dann ist $v - w \in U$ nach Proposition 4.48. Damit ist $\alpha(v - w) \in U$, somit nach Proposition 4.48:

$$\alpha v + U = \alpha w + U.$$

Aber dies zeigt genau, dass $\alpha(v + U) = \alpha(w + U)$, wie gewünscht. □

Lemma 4.50. *Sei V ein K-Vektorraum und sei U ein Unterraum von V. Wenn V endlich-dimensional ist, dann gilt*

$$\dim V/U = \dim V - \dim U.$$

Beweis. Sei $\{v_1, \ldots, v_r\}$ eine Basis von U. Nach Satz 4.35 existieren Vektoren $v_{r+1}, \ldots, v_n \in V$, sodass $\{v_1, \ldots, v_n\}$ eine Basis von V ist. Wir behaupten, dass die Menge

$$\mathcal{B} := \{v_{r+1} + U, \ldots, v_n + U\}$$

eine Basis von V/U ist. Das Lemma folgt sofort aus dieser Behauptung.

Zuerst zeigen wir, dass die Menge $\mathcal{B}$ den Vektorraum V/U erzeugt. Sei $v \in V$, und wir möchten zeigen, dass

$$v + U \in \langle v_{r+1} + U, \ldots, v_n + U \rangle.$$

Und zwar, es gibt $\lambda_1, \ldots, \lambda_n \in K$, sodass $v = \sum\limits_{i=1}^{n} \lambda_i v_i$, und damit

$$v + U = \left(\sum_{i=1}^{n} \lambda_i v_i \right) + U. \tag{22}$$

Nun: es gilt

$$\sum_{i=1}^{n} \lambda_i v_i - \sum_{i=r+1}^{n} \lambda_i v_i = \sum_{i=1}^{r} \lambda_i v_i \in U,$$

und damit, nach Proposition 4.48:

$$\left(\sum_{i=1}^{n} \lambda_i v_i \right) + U = \left(\sum_{i=r+1}^{n} \lambda_i v_i \right) + U.$$

Zusammen mit (22) ergibt dies:

$$v + U = \left(\sum_{i=r+1}^{n} \lambda_i v_i \right) + U = \sum_{i=r+1}^{n} \lambda_i (v_i + U) \in \langle v_{r+1} + U, \ldots, v_n + U \rangle,$$

wie gewünscht.

Es bleibt zu zeigen, dass die Menge $\mathcal{B}$ linear unabhängig ist. Und zwar, seien $\nu_{r+1}, \ldots, \nu_n \in K$ mit

$$\sum_{i=r+1}^{n} \nu_i (v_i + U) = 0 + U = U.$$

Das ist äquivalent zu

$$\left(\sum_{i=r+1}^{n} \nu_i v_i \right) + U = 0 + U,$$

und nach Proposition 4.48, dies ist äquivalent zu

$$\sum_{i=r+1}^{n} \nu_i v_i \in U.$$

Dann existieren $\theta_1, \ldots, \theta_r \in K$, sodass

$$\sum_{i=r+1}^{n} \nu_i v_i = \sum_{i=1}^{r} \theta_i v_i,$$

also

$$\sum_{i=1}^{r} \theta_i v_i - \sum_{i=r+1}^{n} \nu_i v_i = 0.$$

Da $v_1, \dots, v_n$ linear unabhängig sind, so folgt $\theta_i = \nu_i = 0$ für alle i. Das zeigt, dass die Menge $\mathcal{B}$ linear unabhängig ist. $\square$

Ausblick auf unendlich-dimensionale Vektorräume

Viele Dinge kann man auch für unendlich dimensionale Vektorräume definieren und beweisen. Man benutzt dafür das *Auswahlaxiom* oder das *Zornsche Lemma* der mathematischen Logik.

Definition 4.51. Sei V ein K-Vektorraum.

(1) Das *Erzeugnis* einer Teilmenge $S \subseteq V$ ist die Menge

$$\langle S \rangle := \left\{ \sum_{i=1}^{n} k_i v_i \mid n \in \mathbb{N}, k_i \in K, v_i \in S \right\}.$$

In anderen Worten, $\langle S \rangle$ ist die Menge aller endlichen linearen Kombinationen von Vektoren in S.

(2) Eine Teilmenge $S \subseteq V$ heißt *linear unabhängig*, falls für $v_1, \dots, v_n \in S$ und $k_1, \dots, k_n \in K$ gilt:

$$\sum_{i=1}^{n} k_i v_i = 0 \quad \Longrightarrow \quad k_i = 0 \text{für alle } i.$$

(3) Eine Teilmenge $S \subseteq V$ heißt eine *Basis* von V, falls S linear unabhängig ist und $V = \langle S \rangle$.

Beispiel 4.52. $\{1, x, x^2, x^3, \dots\}$ ist eine Basis von $\mathbb{R}[x]$.

Dann kann man beweisen:

Satz 4.53. *Jeder K-Vektorraum hat eine Basis.*

5 Lineare Abbildungen und Matrizen

Definition 5.1. Seien V und W zwei K-Vektorräume. Eine Abbildung

$$T\colon V \to W$$

ist ein *Vektorraumhomomorphismus* oder *K-lineare Abbildung*, falls:

(a) $T(v_1 + v_2) = T(v_1) + T(v_2)$ für alle $v_1, v_2 \in V$,

(b) $T(\alpha v) = \alpha T(v)$ für alle $v \in V, \alpha \in K$.

Ferner ist T ein Vektorraum*mono*morphismus, -*epi*morphismus bzw. -*iso*morphismus, falls T injektiv, surjektiv bzw. bijektiv ist. Falls T ein Isomorphismus ist, dann sagen wir, dass V und W *isomorph* sind und schreiben

$$V \cong W \quad \text{oder} \quad V \simeq W.$$

Bemerkung 5.2. Bemerke, dass, in der Notation wie oben, T ein Gruppenhomomorphismus zwischen den Gruppen $(V, +)$ und $(W, +)$ ist.

Beispiel 5.3.

(1) Seien V und W zwei K-Vektorräume und sei $T\colon V \to W$ ein Homomorphismus. Dann ist T die *Nullabbildung*, falls $T(v) = 0$ für alle $v \in V$.

(2) Sei W ein K-Vektorraum und sei $V \subseteq W$ ein Unterraum. Dann ist die Abbildung

$$T\colon V \to W, \quad v \mapsto v$$

linear und heißt *Identitätsabbildung auf V*.

(3) Betrachte zwei Abbildungen $T_1\colon \mathbb{R}^2 \to \mathbb{R}$ und $T_2\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$T_1\begin{pmatrix} x \\ y \end{pmatrix} = x, \qquad T_2\begin{pmatrix} x \\ y \end{pmatrix} = y.$$

Diese Abbildungen sind linear und heißen *Projektionen* auf die x- bzw. y-Achse.

(4) Sei $w \neq 0$ eine feste komplexe Zahl. Eine *Drehstreckung* ist die Abbildung

$$T\colon \mathbb{C} \to \mathbb{C}, \quad z \mapsto zw.$$

(5) Die Abbildung $T\colon \mathbb{R}[x] \to \mathbb{R}[x],\ p \mapsto p'$ (wobei p' die Ableitung von p ist) ist eine lineare Abbildung.

Definition 5.4. Seien V und W zwei K-Vektorräume und sei $T\colon V \to W$ eine lineare Abbildung.

(1) Der *Kern* oder *Nullraum von* T ist die Teilmenge

$$\ker T := T^{-1}(0).$$

(2) Die *Bildmenge* oder das *Bild von* T ist die Menge

$$\operatorname{Im} T := T(V).$$

Proposition 5.5. *Seien V und W zwei K-Vektorräume und sei $T\colon V \to W$ eine lineare Abbildung. Dann gilt:*

(a) $\ker T$ *ist ein Unterraum von* V*,*

(b) $\operatorname{Im} T$ *ist ein Unterraum von* W*.*

Beweis. Die Mengen $\ker T$ und $\operatorname{Im} T$ sind Untergruppen von V bzw. W, nach Satz 2.19. Es bleibt, die Abgeschlossenheit bezüglich Skalarmultiplikation zu zeigen.

Und zwar, seien $v \in \ker T$ und $\alpha \in K$. Dann gilt:

$$T(\alpha v) = \alpha T(v) = 0,$$

und damit $\alpha v \in \ker T$. Dies zeigt (a).

Ferner, seien $w \in \operatorname{Im} T$ und $\beta \in K$. Dann existiert $v \in V$ mit $T(v) = w$, und damit:

$$\alpha w = T(\alpha v),$$

also $\alpha w \in \operatorname{Im} T$. Dies zeigt (b). □

Satz 5.6 (Homomorphiesatz für Vektorräume)**.** *Seien V und W zwei K-Vektorräume und sei $T\colon V \to W$ eine lineare Abbildung. Dann ist die Abbildung*

$$\theta\colon V/\ker T \to \operatorname{Im} T, \quad [v] \mapsto T(v)$$

ein Vektorraumisomorphismus.

Beweis. Die Abbildung θ ist ein Gruppenisomorphismus ist nach dem Homomorphiesatz für Gruppen, Satz 2.29. Es bleibt noch zu zeigen, dass für alle $\alpha \in K$ gilt

$$\theta(\alpha[v]) = \alpha\theta([v]).$$

Und zwar:

$$\theta(\alpha[v]) = \theta([\alpha v]) = T(\alpha v) = \alpha T(v) = \alpha\theta([v]),$$

wie gewünscht. □

Definition 5.7. Seien V und W zwei endlich-dimensionale K-Vektorräume und sei $T\colon V \to W$ eine lineare Abbildung. Der *Rang von T* ist

$$\operatorname{rang} T := \dim \operatorname{Im} T,$$

und der *Defekt von T* ist

$$\operatorname{def} T := \dim \ker T.$$

Satz 5.8. *Seien V und W zwei K-Vektorräume und sei $T\colon V \to W$ eine lineare Abbildung. Dann gelten die folgenden Aussagen.*

(a) Wenn Vektoren $v_1, \dots, v_n \in V$ linear abhängig sind, so sind die Vektoren $T(v_1), \dots, T(v_n) \in W$ linear abhängig.

(b) Wenn Vektoren $T(v_1), \dots, T(v_n) \in W$ linear unabhängig sind, so sind Vektoren $v_1, \dots, v_n \in V$ linear unabhängig.

Beweis.
Seien $v_1, \dots, v_n \in V$ linear abhängig. Dann existieren $\lambda_i \in K$, die nicht alle verschwinden, sodass

$$\sum_{i=1}^{n} \lambda_i v_i = 0.$$

Daraus folgt:

$$0 = T(0) = T\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i T(v_i),$$

und damit sind $T(v_1), \dots, T(v_n)$ linear abhängig. Dies zeigt (a), und (b) ist die logische Kontraposition von (a). □

Korollar 5.9. *Sei $T\colon V \to W$ ein Isomorphismus zwischen zwei endlich-dimensionalen K-Vektorräumen V und W. Dann*

$$\dim V = \dim W.$$

Beweis. Sei $n := \dim V$ und sei $\{v_1, \dots, v_n\}$ eine Basis von V. Wir behaupten, dass die Vektoren $T(v_1), \dots, T(v_n)$ linear unabhängig sind. Wir argumentieren per Widerspruch und nehmen an, dass sie linear abhängig sind. Nun können wir Satz 5.8(a) auf den Isomorphismus $T^{-1}\colon W \to V$ anwenden und schließen, dass die Vektoren $v_1, \dots, v_n$ linear abhängig sind (da $T^{-1}\big(T(v_i)\big) = v_i$ für alle i), ein Widerspruch.

Also, die Vektoren $T(v_1), \dots, T(v_n)$ sind linear unabhängig, und damit, nach Satz 4.28, $\dim V = n \leq \dim W$. Die umgekehrte Ungleichung folgt, wenn wir die Abbildung $T^{-1}\colon W \to V$ betrachten. □

Satz 5.10 (Dimensionssatz)**.** *Sei $T\colon V \to W$ eine lineare Abbildung zwischen zwei endlich-dimensionalen K-Vektorräumen V und W. Dann gilt*

$$\dim V = \operatorname{def} T + \operatorname{rang} T.$$

Beweis. Nach dem Homomorphiesatz gilt $\operatorname{Im} T \cong V/\ker T$, und damit, nach Korollar 5.9,

$$\dim \operatorname{Im} T = \dim(V/\ker T).$$

Es folgt:

$$\begin{aligned}\operatorname{rang} T &= \dim \operatorname{Im} T = \dim\left(V/\ker T\right)\\ &\overset{\text{Satz 4.50}}{=} \dim V - \dim \ker T = \dim V - \operatorname{def} T,\end{aligned}$$

wie gewünscht. □

Korollar 5.11. *Sei $T\colon V \to W$ eine lineare Abbildung zwischen zwei endlich-dimensionalen K-Vektorräumen V und W. Wenn $\dim V = \dim W$, dann gilt:*

$$T \text{ ist surjektiv} \quad \Longleftrightarrow \quad T \text{ ist injektiv.}$$

Beweis. Es gilt:

$$\begin{aligned}T \text{ surjektiv} \quad &\Longleftrightarrow \quad T(V) = W \quad \Longleftrightarrow \quad \operatorname{rang} T = \dim W\\ &\overset{\text{Dimensionssatz}}{\Longleftrightarrow} \quad \dim V - \operatorname{def} T = \dim W\\ &\overset{\dim W = \dim V}{\Longleftrightarrow} \quad \operatorname{def} T = 0\\ &\Longleftrightarrow \quad \ker T = \{0\} \quad \Longleftrightarrow \quad T \text{ injektiv},\end{aligned}$$

wie gewünscht. □

Proposition 5.12. *Sei $T\colon V \to W$ eine lineare Abbildung zwischen zwei endlich-dimensionalen K-Vektorräumen V und W. Sei $\{e_1, \ldots, e_n\}$ eine Basis von V. Dann gilt*

$$\operatorname{Im} T = \langle T(e_1), \ldots, T(e_n)\rangle.$$

In anderen Worten, die Abbildung T ist eindeutig definiert durch Vektoren $T(e_1), \ldots, T(e_n)$.

Beweis. Sei $v \in \langle T(e_1), \ldots, T(e_n)\rangle$. Dann gibt es $\lambda_i \in K$, sodass $v = \sum\limits_{i=1}^{n} \lambda_i T(e_i)$, und damit

$$v = \sum_{i=1}^{n} \lambda_i T(e_i) = T\left(\sum_{i=1}^{n} \lambda_i e_i\right) \in \operatorname{Im} T.$$

Es gilt also $\langle T(e_1), \ldots, T(e_n)\rangle \subseteq \operatorname{Im} T$.

Umgekehrt, sei $w \in \operatorname{Im} T$. Dann gibt es $v \in V$ mit $T(v) = w$, und dann gibt es $\mu_i \in K$, sodass $v = \sum\limits_{i=1}^{n} \mu_i e_i$. Daraus folgt:

$$w = T(v) = T\left(\sum_{i=1}^{n} \mu_i e_i\right) = \sum_{i=1}^{n} \mu_i T(e_i) \in \langle T(e_1), \ldots, T(e_n)\rangle,$$

welches die Inklusion $\operatorname{Im} T \subseteq \langle T(e_1), \ldots, T(e_n)\rangle$ zeigt. □

Matrizen und Matrixdarstellungen von linearen Abbildungen

Definition 5.13. Sei K ein Körper. Eine *Matrix* der Dimension $m \times n$ ist eine Abbildung

$$\{1, \ldots, m\} \times \{1, \ldots, n\} \to K, \quad (i, j) \mapsto a_{ij} \in K.$$

Eine Matrix schreiben wir als Tabelle:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

oder kürzer:

$$(a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} \quad \text{oder} \quad (a_i j).$$

Eine Matrix hat *Werte* oder *Einträge* in K und heißt auch Matrix *über* K. Die Menge aller Matrizen über K bezeichnet man in der Literatur mit

$$K^{m\times n}, \qquad K^{m,n}, \qquad M_{m\times n}(K), \qquad M_{m,n}(K);$$

wir verwenden hier die Schreibweise $M_{m,n}(K)$. Wenn $m = n$, dann schreiben wir $M_n(K)$ statt $M_{m,n}(K)$.

Lemma/Definition 5.14. *Sei K ein Körper. Wir definieren die Addition und die Skalarmultiplikation auf $M_{m,n}(K)$ wie folgt:*

$$(a_{ij}) + (b_{ij}) := (a_{ij} + b_{ij}) \quad \textit{(komponentenweise Addition)}$$

und, für $\alpha \in K$,

$$\alpha(a_{ij}) := (\alpha a_{ij}) \quad \textit{(komponentenweise Skalarmultiplikation).}$$

Dann wird $M_{m,n}(K)$ zu einem K-Vektorraum bezüglich dieser Operationen.

Beweis. Übung! □

Definition 5.15. Sei $T\colon V \to W$ eine lineare Abbildung zwischen zwei endlich-dimensionalen K-Vektorräumen V und W. Sei $\{e_1, \dots, e_n\}$ eine Basis von V und sei $\{f_1, \dots, f_m\}$ als Basis von W. Dann lässt sich jedes $T(e_i) \in W$ eindeutig schreiben als lineare Kombination von $f_1, \dots, f_m$: In anderen Worten, es eindeutige gibt $a_{ij} \in K$, sodass

$$\begin{aligned} T(e_1) &= a_{11} f_1 + \cdots + a_{m1} f_m, \\ &\vdots \\ T(e_n) &= a_{1n} f_1 + \cdots + a_{mn} f_m. \end{aligned}$$

Die Matrix

$$A := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

heißt die *Darstellungsmatrix von T* bezüglich der Basen $\{e_1, \dots, e_n\}$ von V und $\{f_1, \dots, f_m\}$ von W.

Bemerkung 5.16. Nach Proposition 5.12 ist die Abbildung T eindeutig bestimmt durch die Wahl der Skalare $a_{ij} \in K$, also durch die Matrix A.

Beispiel 5.17.

(a) Sei $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ die lineare Abbildung

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y - z \\ 2x + z \end{pmatrix}.$$

Wir finden nun die Darstellungsmatrix A von T bezüglich der kanonischen Basen von $\mathbb{R}^3$ und $\mathbb{R}^2$. Und zwar, es gilt:

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

$$T\begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} = 1\begin{pmatrix}1\\0\end{pmatrix} + 0\begin{pmatrix}0\\1\end{pmatrix},$$

$$T\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}-1\\1\end{pmatrix} = -1\begin{pmatrix}1\\0\end{pmatrix} + 1\begin{pmatrix}0\\1\end{pmatrix},$$

und damit $A = \begin{pmatrix}1 & 1 & -1\\2 & 0 & 1\end{pmatrix}$.

(b) Sei T wie in (a). Wir finden nun die Darstellungsmatrix B von T bezüglich der Basen $\left\{\begin{pmatrix}1\\0\\-1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ von $\mathbb{R}^3$, und $\left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right\}$ von $\mathbb{R}^2$. Und zwar, es gilt:

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

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

$$T\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}-1\\1\end{pmatrix} = 1\begin{pmatrix}1\\1\end{pmatrix} - 2\begin{pmatrix}1\\0\end{pmatrix},$$

und damit $B = \begin{pmatrix}1 & 3 & 1\\1 & -2 & -2\end{pmatrix}$.

(c) Sei $n \in \mathbb{N}$ und $T\colon \mathbb{R}[x]^{\leq n} \to \mathbb{R}[x]^{\leq n}$ die Ableitung. Wir finden nun die Darstellungsmatrix C von T bezüglich der Standardbasis von $\mathbb{R}[x]^{\leq n}$.

Und zwar, es gilt:

$$\begin{aligned}
T(1) &= 0 = 0 \cdot 1 + 0 \cdot x + \cdots + 0 \cdot x^n, \\
T(x) &= 1 = 1 \cdot 1 + 0 \cdot x + \cdots + 0 \cdot x^n, \\
&\vdots \\
T(x^i) &= ix^{i-1} = 0 \cdot 1 + \cdots + i \cdot x^{i-1} + \cdots + 0 \cdot x^n, \\
&\vdots \\
T(x^n) &= nx^{n-1} = 0 \cdot 1 + \cdots + n \cdot x^{n-1} + 0 \cdot x^n,
\end{aligned}$$

und damit $C = \begin{pmatrix} 0 & 1 & 0 & \cdots & & & 0 \\ 0 & 0 & 2 & 0 & \cdots & & 0 \\ \vdots & \vdots & & \ddots & & & \vdots \\ 0 & 0 & \cdots & 0 & i & 0 & \vdots \\ \vdots & \vdots & & & & \ddots & \vdots \\ 0 & 0 & \cdots & & \cdots & 0 & n \\ 0 & 0 & \cdots & & \cdots & 0 & 0 \end{pmatrix}$.

Definition 5.18. Seien V und W zwei K-Vektorräume. Definiere

$$\mathrm{Hom}_K(V, W) := \{f \colon V \to W \mid f \text{ ist eine } k\text{-lineare Abbildung}\}.$$

Wenn es aus dem Kontext klar ist, um welchen Körper K es sich hier handelt, dann schreiben wir einfach

$$\mathrm{Hom}(V, W).$$

Bemerkung 5.19.

(1) $\mathrm{Hom}_K(V, W) \subseteq \mathrm{Abb}(V, W)$.

(2) Wir definieren die Addition von $f, g \in \mathrm{Hom}_K(V, W)$ durch

$$(f + g)(x) := f(x) + g(x) \quad \text{für alle } x \in K,$$

und die Skalarmultiplikation von $\alpha \in K$ und $f \in \mathrm{Hom}_K(V, W)$ durch

$$(\alpha f)(x) := \alpha f(x) \quad \text{für alle } x \in K.$$

Übung: Diese Operationen machen $\mathrm{Hom}_K(V, W)$ zu einem K-Vektorraum.

(3) Wir definieren

$$\mathrm{Iso}_K(V,W) := \{f\colon V \to W \mid f \text{ ist ein Vektorraumisomorphismus}\}.$$

(4) Wenn $V = W$, dann schreiben wir

$$\mathrm{End}(V) := \mathrm{Hom}_K(V,V)$$

und nennen deren Elemente *Endomorphismen*. Sind $F, G \in \mathrm{End}(V)$, so ist $F \circ G \in \mathrm{End}(V)$ (Übung!).

(5) $\big(\mathrm{Iso}(V,V), \circ\big)$ ist eine Gruppe.

Lemma 5.20.

(1) Sei V ein n-dimensionaler K-Vektorraum. Dann

$$V \cong K^n.$$

(2) Seien V und W zwei K-Vektorräume mit $n = \dim V$ und $m = \dim W$. Sei $\mathcal{B}_V$ eine Basis von V und sei $\mathcal{B}_W$ von W . Dann ist die Abbildung

$$\theta\colon \mathrm{Hom}_K(V,W) \to M_{m,n}(K),$$

die zu jedem $F \in \mathrm{Hom}_K(V,W)$ seine Darstellungsmatrix bezüglich $\mathcal{B}_V$ und $\mathcal{B}_W$ zuordnet, ein Vektorraumisomorphismus.

(3) Mit der Notation aus (2), so ordnet die inverse Abbildung θ^{-1} jeder Matrix $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in M_{m,n}(K)$ die lineare Abbildung

$$F\colon V \to W, \quad v \to w,$$

auf folgende Weise: Seien $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ die Koordinaten von v in der Basis $\mathcal{B}_V$. Dann sind

$$\begin{pmatrix} a_{11}\lambda_1 + \cdots + a_{1n}\lambda_n \\ \vdots \\ a_{m1}\lambda_1 + \cdots + a_{mn}\lambda_n \end{pmatrix}$$

die Koordinaten von w in der Basis $\mathcal{B}_W$.

Beweis.
(1) Sei $\{e_1, \dots, e_n\}$ eine Basis von V und sei $\{f_1, \dots, f_m\}$ die kanonische Basis von K^n. Definiere die lineare Abbildung

$$T\colon V \to K^n, \quad e_i \mapsto f_i.$$

Diese Abbildung ist offensichtlich surjektiv, da $T(V) = \langle f_1, \dots, f_n \rangle = K^n$. Nach Korollar 5.11 ist T dann auch ist injektiv, und damit bijektiv.

(2) Fixiere Basen $\mathcal{B}_1$ von V und $\mathcal{B}_2$ von W, und setze

$$\theta\colon \operatorname{Hom}(V, W) \to M_{m,n}(K), \quad T \to A(T),$$

wobei $A(T)$ die Darstellungsmatrix von T bezüglich $\mathcal{B}_1$ und $\mathcal{B}_2$ ist. Es ist einfach zu zeigen, dass

$$A(T_1 + T_2) = A(T_1) + A(T_2) \quad \text{und} \quad A(\lambda T) = \lambda A(T) \text{ für } \lambda \in K.$$

Es ist auch einfach zu zeigen, dass θ ein Isomorphismus ist (Übung!).

(3) folgt einfach aus der Definition der Darstellungsmatrix (Übung!). □

Korollar 5.21.

(1) Zwei endlich-dimensionale K-Vektorräume V und W mit derselben Dimension sind isomorph.

(2) Für zwei endlich dimensionale K-Vektorräume V und W mit $n = \dim V$ und $m = \dim W$ ist die Menge $\operatorname{Hom}_K(V, W)$ auch ein endlich dimensionaler K-Vektorraum der Dimension $n \cdot m$.

Beweis.
(a) Nach Lemma 5.20 gibt es Isomorphismen $T_1\colon V \to K^n$ und $T_2\colon W \to K^n$. Dann ist

$$T_2^{-1} \circ T_1\colon V \to W$$

ein Isomorphismus.

(b) Aus Lemma 5.20 folgt, dass $\operatorname{Hom}_K(V, W) \cong M_{m,n}(K)$, und damit

$$\dim \operatorname{Hom}_K(V, W) = \dim M_{m,n}(K)$$

nach Korollar 5.9. Die Matrizen

$$E_{ij} = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \leftarrow \text{die 1 ist an der Stelle } (i, j),$$

bilden eine Basis von $M_{m,n}(K)$ (für $i = 1, \ldots, n$ und $j = 1, \ldots, m$), und damit $\dim M_{m,n}(K) = mn$. $\square$

Matrixalgebra

Seien X, Y und Z drei K-Vektorräume mit

$$n := \dim X, \qquad m := \dim Y, \qquad \ell := \dim Z.$$

Seien

$$\mathcal{B}_X = \{x_1, \ldots, x_n\}, \quad \mathcal{B}_Y = \{y_1, \ldots, y_m\}, \quad \mathcal{B}_Z = \{z_1, \ldots, z_\ell\}$$

fixierte Basen von X bzw. Y bzw. Z.

Sei $S\colon X \to Y$ eine lineare Abbildung mit der Darstellungsmatrix $M(S)$ bezüglich $\mathcal{B}_X$ und $\mathcal{B}_Y$. Sei $T\colon Y \to Z$ eine lineare Abbildung mit der Darstellungsmatrix $M(T)$ bezüglich $\mathcal{B}_Y$ und $\mathcal{B}_Z$.

In anderen Worten, wenn

$$M(S) = \begin{pmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & & \vdots \\ s_{m1} & \cdots & s_{mn} \end{pmatrix} \in M_{m,n}(K), \quad \text{und}$$

$$M(T) = \begin{pmatrix} t_{11} & \cdots & s_{1m} \\ \vdots & & \vdots \\ t_{\ell 1} & \cdots & s_{\ell m} \end{pmatrix} \in M_{\ell,m}(K),$$

dann gilt

$$S(x_j) = \sum_{k=1}^{m} s_{kj} y_k \qquad \text{für alle } j = 1, \ldots, n, \quad \text{und} \tag{23}$$

$$T(y_j) = \sum_{k=1}^{\ell} t_{kj} z_k \qquad \text{für alle } j = 1, \ldots, m. \tag{24}$$

Lemma 5.22. *In der Notation wie oben, sei $M(T \circ S)$ der Darstellungsmatrix von $T \circ S$ bezüglich $\mathcal{B}_X$ und $\mathcal{B}_Z$. Dann lautet der Eintrag (i, j) der Matrix $M(T \circ S)$:*

$$\sum_{k=1}^{m} t_{ik} s_{kj}.$$

Beweis. Sei $R := T \circ S$ und

$$M(R) = \begin{pmatrix} r_{11} & \cdots & r_{1m} \\ \vdots & & \vdots \\ r_{\ell 1} & \cdots & r_{\ell m} \end{pmatrix} \in M_{l,m}(K).$$

Es gilt

$$\begin{aligned} R(x_j) &= (T \circ S)(x_j) = T(S(x_j)) \\ &\overset{(23)}{=} T\left(\sum_{k=1}^{m} s_{kj} y_k\right) = \sum_{k=1}^{m} s_{kj} T(y_k) \\ &\overset{(24)}{=} \sum_{k=1}^{m} s_{kj} \sum_{i=1}^{\ell} t_{ik} z_i = \sum_{k=1}^{m} \sum_{i=1}^{\ell} t_{ik} s_{kj} z_i \\ &= \sum_{i=1}^{\ell} \sum_{k=1}^{m} t_{ik} s_{kj} z_i = \sum_{i=1}^{\ell} \left(\sum_{k=1}^{m} t_{ik} s_{kj}\right) z_i. \end{aligned}$$

Damit gilt $r_{ij} = \sum\limits_{k=1}^{r} t_{ik} s_{kj}$, wie gewünscht. □

Definition 5.23. Sei K ein Körper, und seien

$$A = (a_{ij})_{\substack{i=1,\dots,\ell \\ j=1,\dots,m}} \in M_{\ell,m}(K)$$

und

$$B = (b_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in M_{m,n}(K).$$

Das *Produkt* $A \cdot B$ ist die Matrix

$$C = (c_{ij})_{\substack{i=1,\dots,\ell \\ j=1,\dots,n}} \in M_{\ell,n}(K),$$

die durch die folgenden Formel definiert ist:

$$c_{ij} := \sum_{k=1}^{m} a_{ik} b_{kj}.$$

Wir bekommen also den Eintrag c_{ij}, indem wir die *jeweiligen Einträge der i-ten Zeile von A* mit den *jeweiligen Einträge der j-ten Spalte von B* ausmultiplizieren und dann summieren.

Bemerkung 5.24. Mit der Notation wie am Anfang des Abschnitts: Die Matrixmultiplikation ist so definiert, dass die Matrix $M(T) \cdot M(S)$ die Darstellungsmatrix der Abbildung $T \circ S$ ist. In anderen Worten, es gilt

$$M(T) \cdot M(S) = M(T \circ S).$$

$$\begin{tikzcd} X \arrow[r, "M(S)"] \arrow[rr, bend right, "M(T\circ S)=M(T)\cdot M(S)"'] & Y \arrow[r, "M(T)"] & Z \end{tikzcd}$$

Beispiel 5.25. Es gilt:

$$\underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}}_{3\times 2} \cdot \underbrace{\begin{pmatrix} 6 & 5 & 4 & 3 \\ 2 & 1 & 0 & -1 \end{pmatrix}}_{2\times 4}$$

$$= \underbrace{\begin{pmatrix} 1\cdot 6 + 2\cdot 2 & 1\cdot 5 + 2\cdot 1 & 1\cdot 4 + 2\cdot 0 & 1\cdot 3 + 2\cdot(-1) \\ 3\cdot 6 + 4\cdot 2 & 3\cdot 5 + 4\cdot 1 & 3\cdot 4 + 4\cdot 0 & 3\cdot 3 + 4\cdot(-1) \\ 5\cdot 6 + 6\cdot 2 & 5\cdot 5 + 6\cdot 1 & 5\cdot 4 + 6\cdot 0 & 5\cdot 3 + 6\cdot(-1) \end{pmatrix}}_{3\times 4}.$$

Lemma 5.26.

(a) Seien $A, B \in M_{\ell,m}(K)$ und $C \in M_{m,n}(K)$. Dann gilt

$$(A + B)C = AC + BC.$$

(b) Seien $A \in M_{\ell,m}(K)$ und $B, C \in M_{m,n}(K)$. Dann gilt

$$A(B + C) = AB + BC.$$

(c) Seien $A \in M_{\ell,m}(K)$, $B \in M_{m,n}(K)$ und $C \in M_{n,p}(K)$. Dann gilt

$$A(BC) = (AB)C.$$

Beweis. Wir zeigen nur (a); (b) und (c) folgen auf ähnliche Weise.

Man prüft entweder direkt aus den Formeln, dass (a) stimmt, oder man geht folgendermaßen vor. Sei $F\colon K^m \to K^\ell$ die lineare Abbildung, wessen Darstellungsmatrix bezüglich der kanonischen Basen von K^m und K^ℓ genau A ist; ähnlich für Abbildungen $G\colon K^m \to K^\ell$, die zu B korrespondiert, und $H\colon K^n \to K^m$, die zu C korrespondiert. Dann korrespondiert die Abbildung

$$(F + G) \circ H\colon K^n \to K^\ell$$

zur Matrix $(A+B)C$. Andererseits korrespondiert die Abbildung

$$F \circ H + G \circ H$$

zur Matrix $AC + BC$. Aber diese zwei Abbildungen sind gleich, die korrespondierende Matrizen sind also auch gleich. $\square$

Bemerkung 5.27. Die Matrixmultiplikation ist *nicht* kommutativ. Zum Beispiel gilt:

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

Notation 5.28. Sei K ein Körper und $A = (a_{ij}) \in M_{m,n}(K)$. Wir bezeichnen die *Spalten* von A mit

$$c_1^A, \ldots, c_n^A \qquad (\text{„}c\text{“ für „column“}),$$

und die *Zeilen* mit

$$r_1^A, \ldots, r_m^A \qquad (\text{„}r\text{“ für „row“}).$$

Also gilt

$$c_i^A = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} \qquad \text{und} \qquad r_i^A = \begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix}$$

Dann können wir A schreiben als

$$A = \begin{pmatrix} r_1^A \\ \vdots \\ r_m^A \end{pmatrix} \quad \text{oder} \quad A = \begin{pmatrix} c_1^A & \cdots & c_n^A \end{pmatrix}.$$

Bemerkung 5.29. Seien $A = (a_{ij}) \in M_{\ell,m}(K)$ und $B = (b_{ij}) \in M_{m,n}$, und sei $C = (c_{ij}) := AB \in M_{\ell,n}$. Dann folgt aus der Definition der Matrixmultiplikation, dass

$$c_{ij} = \underbrace{r_i^A}_{\in M_{1,m}(K)} \cdot \underbrace{c_j^B}_{\in M_{m,1}(K)} \in M_{1,1}(K) \cong K.$$

Man checkt dann einfach, dass

$$c_j^C = \sum_{k=1}^{m} b_{kj} c_k^A \qquad \text{für alle } j = 1, \ldots, \ell,$$

und

$$r_j^C = \sum_{k=1}^{m} a_{jk} r_k^B \qquad \text{für alle } j = 1, \ldots, n.$$

Definition 5.30. Sei $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in M_{m,n}(K)$. Dann ist die *zu A transponierte Matrix* definiert als

$$A^T := (a_{ij}^T)_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \in M_{n,m}(K)$$

durch die Formel

$$a_{ij}^T := a_{ji}.$$

Bildlich wird dies dargestellt als:

$$A = \begin{pmatrix} \rightarrow \\ \vdots \\ \rightarrow \end{pmatrix} \quad \rightsquigarrow \quad A^T = \begin{pmatrix} \downarrow & \cdots & \downarrow \end{pmatrix}$$

Bemerkung 5.31. Es gilt

$$(A^T)^T = A.$$

Proposition 5.32. *Seien $A \in M_{m,n}(K)$ und $B \in M_{n,\ell}(K)$. Dann gilt*

$$(AB)^T = B^T A^T.$$

Beweis. Schreibe $C = (c_{ij}) := AB$ und $D = (d_{ij}) := B^T A^T$. Dann gilt

$$c_{ij} = r_i^A c_j^B,$$

und daher

$$c_{ij}^T = r_j^A c_i^B = r_i^{B^T} c_j^{A^T} = d_{ij}.$$

Das zeigt die Proposition. □

Rang einer Matrix

Definition 5.33. Sie K ein Körper und sei $A \in M_{m,n}(K)$.

(1) Der *Spaltenraum von A* ist der Unterraum $\langle c_1^A, \dots, c_n^A \rangle$ von K^m, und seine Dimension ist der *Spaltenrang von A*.

(2) Der *Zeilenraum von A* ist der Unterraum $\langle r_1^A, \dots, r_m^A \rangle$ von K^n, und seine Dimension ist der *Zeilenrang von A*.

Bemerkung 5.34. Der Spalten- bzw. Zeilenrang ist die maximale Anzahl an linear unabhängigen Spalten bzw. Zeilen.

Beispiel 5.35. Sei

$$A = \begin{pmatrix} 1 & -2 & 0 & 4 \\ 3 & 1 & 1 & 0 \\ -1 & -5 & -1 & 8 \\ 3 & 8 & 2 & -12 \end{pmatrix}.$$

Der Zeilenrang von A ist 2, da r_1^A und r_2^A linear abhängig sind, und es gilt

$$r_3^A = 2r_1^A - r_2^A, \quad r_4^A = -3r_1^A + 2r_2^A.$$

Der Spaltenrang von A ist 2, da c_3^A und c_4^A linear unabhängig sind, und es gilt

$$c_1^A = \frac{1}{4}c_4^A + 3c_3^A, \quad c_2^A = \frac{1}{2}c_4^A + c_3^A.$$

Satz 5.36. *Sei $A \in M_{m,n}(K)$. Dann ist der Spaltenrang von A gleich dem Zeilenrang von A.*

Beweis. Sei s der Spaltenrang von A, und sei $\{f_1, \dots, f_s\}$ eine Basis vom Spaltenraum von A. Dann gibt es für jede Spalte c_j^A von A Skalare λ_{kj}, sodass

$$c_j^A = \sum_{k=1}^{s} \lambda_{kj} f_k.$$

Definiere

$$L := (\lambda_{kj})_{\substack{k=1,\dots,s \\ j=1,\dots,n}} \in M_{s,n}(K)$$

und

$$F = (f_{ik})_{\substack{i=1,\dots,m \\ k=1,\dots,s}} := \begin{pmatrix} f_1 & \cdots & f_s \end{pmatrix} \in M_{m,s}(K).$$

Dann prüft man einfach (siehe Bemerkung 5.29), dass

$$A = F \cdot L,$$

und daher, wieder nach Bemerkung 5.29,

$$r_i^A = \sum_{k=1}^{s} f_{ik} r_k^L.$$

Also, jede Zeile von A kann geschrieben werden als Summe von s Vektoren $r_1^L, \dots, r_s^L$, und daher gilt:

$$\text{Zeilenrang von } A \leq s = \text{Spaltenrang von } A. \tag{25}$$

Analog zeigt man, dass

$$\text{Zeilenrang von } A^T \leq \text{Spaltenrang von } A^T. \tag{26}$$

Aber es gilt offensichtlich, dass

$$\text{Zeilenrang von } A^T = \text{Spaltenrang von } A$$

und

$$\text{Spaltenrang von } A^T = \text{Zeilenenrang von } A.$$

Es folgt dann aus (26), dass

$$\text{Spaltenrang von } A \leq \text{Zeilenrang von } A.$$

Dies, zusammen mit (25), zeigt den Satz. □

Definition 5.37. Der *Rang einer Matrix* A ist der Zeilenrang von A (nach Satz 5.36 äquivalent zu: der Spaltenrang von A). Wir schreiben

$$\operatorname{rang} A.$$

Bemerkung 5.38. Da Zeilenrang von A^T = Spaltenrang von A, so gilt

$$\operatorname{rang} A = \operatorname{rang} A^T.$$

Proposition 5.39. *Sei A eine Matrix mit m Zeilen, bei der die ersten r Hauptdiagonaleneinträge von 0 verschieden, die letzten $m-r$ Zeilen und alle Einträge unter der Hauptdiagonalen identisch 0 sind. Dann ist* $\operatorname{rang} A = r$.

$$A = \left(\begin{array}{cccccc} a_{11} & & & & & \\ 0 & \ddots & & * & & \\ \vdots & \ddots & \ddots & & & \\ 0 & \cdots & 0 & a_{rr} & & \\ \hdashline 0 & \cdots & & & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdots & & & \cdots & 0 \end{array}\right)$$

Beweis. Übung! □

Definition 5.40. Unter einer *elementaren Zeilenumformung* einer Matrix versteht man eine der folgenden Operationen:

(a) Vertauschen zweier Zeilen,

(b) Multiplikation einer Zeile mit einem Skalar $\lambda \neq 0$,

(c) Addition eines beliebigen Vielfachen einer Zeile zu einer anderen Zeile.

Die *elementaren Spaltenumformungen* definiert man ähnlich.

Proposition 5.41. *Die elementaren Zeilen- und Spaltenumformungen ändern den Rang einer Matrix nicht.*

Beweis. Übung! □

Beispiel 5.42. Sei

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

Wir werden Propositionen 5.39 und 5.41 benutzen, um den Rang von A zu berechnen:

$$\begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & 1 \\ 0 & 2 & -1 \\ 1 & 1 & 4 \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 3 \\ 0 & 2 & -1 \\ 1 & 1 & 4 \end{pmatrix} \xrightarrow[r_4 \to r_4 - r_1]{r_2 \to r_2 - 2r_1} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 2 & -1 \\ 0 & 1 & 3 \end{pmatrix}$$

$$\xrightarrow[r_4 \to r_4 + r_2]{r_3 \to r_3 + 2r_2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 4 \end{pmatrix} \xrightarrow{r_4 \to r_4 - 4r_3} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Es gilt also $\operatorname{rang} A = 3$.

Lemma 5.43. *Seien M_1 und M_2 zwei Matrizen über einem Körper K. Dann gilt*

$$\operatorname{rang} \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix} = \operatorname{rang} M_1 + \operatorname{rang} M_2,$$

$$\operatorname{rang} \begin{pmatrix} M_1 & * \\ 0 & M_2 \end{pmatrix} \geq \operatorname{rang} M_1 + \operatorname{rang} M_2.$$

Beweis. Übung! □

Notation 5.44. Sei $n \in \mathbb{N}_{>0}$. Die Matrix

$$I_n := \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in M_n(K)$$

ist die *Identitätsmatrix* der Dimension $n \times n$. Es gilt

$$I_n \cdot A = A \cdot I_n = A$$

für alle $A \in M_n(K)$.

Satz 5.45 (Rangungleichung von Sylvester)**.** *Sei K ein Körper, und seien $A \in M_{\ell,m}(K)$ und $B \in M_{m,n}(K)$. Dann gilt*

$$\operatorname{rang} A + \operatorname{rang} B \leq \operatorname{rang} AB + m.$$

Beweis. Seien $A = (a_{ij})$ und $B = (b_{ij})$. Die elementaren Zeilen- und Spaltenumformungen ergeben:

$$\begin{pmatrix} I_m & 0 \\ 0 & AB \end{pmatrix} \xrightarrow[\substack{r_{m+\ell} \to r_{m+\ell} + a_{\ell i} r_i}]{\substack{r_{m+1} \to r_{m+1} + a_{1i} r_i \\ \cdots}} \begin{pmatrix} I_m & 0 \\ A & AB \end{pmatrix} \xrightarrow[\substack{c_{m+\ell} \to c_{m+\ell} - b_{j\ell} c_j}]{\substack{c_{m+1} \to c_{m+1} - b_{j1} c_j \\ \cdots}} \begin{pmatrix} I_m & -B \\ A & 0 \end{pmatrix}$$
$$\xrightarrow{\text{Spaltentausch}} \begin{pmatrix} -B & I_m \\ 0 & A \end{pmatrix} \xrightarrow{c_j \to -c_j} \begin{pmatrix} B & I_m \\ 0 & A \end{pmatrix}. \tag{27}$$

Es gilt also

$$\begin{aligned} m + \operatorname{rang} AB &= \operatorname{rang} I_m + \operatorname{rang} AB \overset{5.43}{=} \operatorname{rang} \begin{pmatrix} I_m & 0 \\ 0 & AB \end{pmatrix} \\ &\overset{(27)}{=} \operatorname{rang} \begin{pmatrix} B & I_m \\ 0 & A \end{pmatrix} \overset{5.43}{\geq} \operatorname{rang} B + \operatorname{rang} A, \end{aligned}$$

welches zu zeigen war. □

Definition 5.46. Eine $m \times n$ Matrix A hat *Zeilenstufenform*, wenn die folgenden Bedingungen erfüllt sind.

(1) Die letzten $m - r$ Zeilen sind identisch 0 ($0 \leq r \leq m$).

(2) Ist a_{ij_i}, der erste Eintrag in der i-ten Zeile, der nicht 0 ist, so gilt

$$0 \leq j_1 < j_2 < \cdots < j_r \leq n.$$

Also A hat dann die folgende Gestalt:

$$\begin{pmatrix} 0 & \cdots & 0 & a_{1j_1} & & & & \\ \vdots & & & 0 & a_{2j_2} & & * & \\ & & & & 0 & & & \\ \vdots & & & & & 0 & a_{rj_r} & \\ 0 & \cdots & & & & \cdots & 0 & \end{pmatrix}$$

Proposition 5.47. *Der Rang einer Matrix in Zeilenstufenform ist die Anzahl der Zeilen, die von* 0 *verschieden sind; also in der Notation aus Definition 5.46:*

$$\operatorname{rang} A = r.$$

Beweis. Übung! □

Lemma 5.48 (Gauß-Algorithmus)**.** *Jede Matrix kann durch elementare Zeilenumformungen in Zeilenstufenform gebracht werden.*

Beweis. Sei $A \in M_{m,n}(K)$. Falls die Zeilen von A alle 0 sind, dann sind wir fertig.

Wir dürfen also annehmen, dass es eine Zeile von A gibt, die von 0 verschieden ist. Sei j_1 der kleinste Index, sodass $c^A_{j_1} \neq 0$:

$$A = \left(\begin{array}{cc|c|c} 0 & \cdots & & \\ \vdots & & c^A_{j_1} & * \\ 0 & \cdots & & \end{array}\right)$$

Durch Vertauschen der Zeilen können wir erreichen, dass der erste Eintrag in $c^A_{j_1}$ ungleich null ist:

$$A = \begin{pmatrix} 0 & \cdots & 0 & a_{1j_1} & \\ \vdots & & \vdots & \vdots & * \\ 0 & \cdots & 0 & \vdots & \end{pmatrix}, \quad \text{wobei } a_{1j_1} \neq 0.$$

Dann erreicht man durch die folgenden elementaren Zeilenumformungen:

$$\xrightarrow{r_k \to r_k - \frac{a_{kj_1}}{a_{1j_1}} r_1 \ (\forall k > 1)} \left(\begin{array}{cccc|c} 0 & \cdots & 0 & a_{1j_1} & * \\ \hline 0 & \cdots & 0 & 0 & \\ \vdots & & \vdots & \vdots & A_1 \\ 0 & \cdots & 0 & 0 & \end{array}\right).$$

Der Rang von A_1 ist kleiner als der Rang von A. Per Induktion nach dem Rang kann dann A_1 durch elementare Zeilenumformungen in Zeilenstufenform gebracht werden. Deswegen gilt dies auch für A. $\square$

Beispiel 5.49.

$$A = \begin{pmatrix} 1 & 3 & 2 & 4 \\ 0 & 2 & 2 & 2 \\ 4 & 2 & 3 & 1 \\ 6 & 8 & 0 & 1 \end{pmatrix} \xrightarrow[\substack{r_3 \to r_3 - 4r_1 \\ r_4 \to r_4 - 6r_1}]{r_2 \to r_2} \begin{pmatrix} 1 & 3 & 2 & 4 \\ 0 & 2 & 2 & 2 \\ 0 & -10 & -5 & -15 \\ 0 & -10 & -12 & -23 \end{pmatrix}$$

$$\xrightarrow[\substack{r_4 \to r_4 + 5r_2}]{r_3 \to r_3 + 5r_2} \begin{pmatrix} 1 & 3 & 2 & 4 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 5 & -5 \\ 0 & 0 & -2 & -13 \end{pmatrix}$$

$$\xrightarrow{r_4 \to r_4 + \frac{2}{5} r_3} \begin{pmatrix} 1 & 3 & 2 & 4 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 5 & -5 \\ 0 & 0 & 0 & -15 \end{pmatrix}$$

Bemerkung 5.50. Sei K ein Körper. Wir betrachten im Folgenden die Vektorräume K^n und K^m mit jeweils kanonischen Basen. Wir wissen aus Lemma 5.20, dass die Abbildung

$$\theta\colon \operatorname{Hom}(K^n, K^m) \to M_{m,n}(K), \quad T \mapsto \text{Darstellungsmatrix von } T$$

ein Isomorphismus ist, wessen inverse Abbildung ist:

$$\theta^{-1}(A) = \text{Abbildung } T \in \operatorname{Hom}(K^n, K^m) \text{ mit } T(x) = Ax \text{ für alle } x \in K^n.$$

In anderen Worten, zu jeder Abbildung $T\colon K^n \to K^m$ kann man eine Matrix $A \in M_{m,n}(K)$ assoziieren und umgekehrt. Wir sagen, dass A *injektiv* bzw. *surjektiv* bzw. *bijektiv* ist, wenn dies entsprechend für die assoziierte Abbildung gilt. Wenn A eine quadratische Matrix ist, dann gilt

$$A \text{ injektiv} \quad \Longleftrightarrow \quad A \text{ surjektiv}.$$

Sei $\{e_1, \ldots, e_n\}$ die kanonische Basis von K^n. Dann gilt

$$\begin{aligned} \operatorname{Im} T &= \langle T(e_1), \ldots, T(e_n) \rangle = \langle Ae_1, \ldots, Ae_n \rangle \\ &= \langle c_1^A, \ldots, c_n^A \rangle = \text{Spaltenraum von } A. \end{aligned}$$

Damit gilt

$$\operatorname{rang} T = \dim \operatorname{Im} T = \dim(\text{Spaltenraum von } A) = \operatorname{rang} A;$$

in anderen Worten, unsere zwei Definitionen vom Rang sind äquivalent. Analog zu diesem, schreiben wir

$$\ker A, \quad \operatorname{Im} A, \quad \operatorname{rang} A, \quad \operatorname{def} A,$$

wenn wir eigentlich

$$\ker T, \quad \operatorname{Im} T, \quad \operatorname{rang} T, \quad \operatorname{def} T$$

meinen. Bemerke, dass nach dem Dimensionssatz gilt

$$\operatorname{def} A = \underbrace{n}_{=\operatorname{rang} K^n} - \operatorname{rang} A.$$

Invertierbare Matrizen

Definition 5.51. Sei K ein Körper. Eine Matrix $A \in M_n(K)$ heißt *invertierbar* oder *regulär*, falls es eine Matrix $B \in M_n(K)$ gibt mit

$$B \cdot A = A \cdot B = I_n.$$

In diesem Fall ist B eindeutig bestimmt (Übung!) und heißt die *Inverse von* A. Sie wird mit A^{-1} bezeichnet.

Proposition/Definition 5.52. *Sei K ein Körper. Die Menge aller invertierbaren Matrizen in $M_n(K)$ bildet eine Gruppe bezüglich der Matrixmultiplikation mit dem neutralen Element I_n. Diese Gruppe heißt* allgemeine lineare Gruppe vom Grad n über K *und wird mit*

$$\mathrm{GL}(n, K)$$

bezeichnet.

Beweis. Übung! □

Bemerkung 5.53. (1) Das neutrale Element in $\mathrm{GL}(n, K)$ ist eindeutig bestimmt.

(2) Das inverse Element eines Elements aus $\mathrm{GL}(n, K)$ ist eindeutig bestimmt.

(3) Es gilt $(AB)^{-1} = B^{-1}A^{-1}$ für alle $A, B \in \mathrm{GL}(n, K)$.

Beispiel 5.54. Es gilt (Übung!):

$$\mathrm{GL}(2,K) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(K) \;\middle|\; ad - bc \neq 0 \right\},$$

und für jede Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2,K)$ gilt

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Proposition 5.55. *Sei K ein Körper. Für jede lineare Abbildung $T\colon K^n \to K^n$ sei $A(T)$ die Darstellungsmatrix von T bezüglich der Standardbasis von K^n. Dann ist die Abbildung $T \mapsto A(T)$ ein Gruppenisomorphismus zwischen*

$$(\mathrm{Iso}(K^n, K^n), \circ) \quad \textit{und} \quad (\mathrm{GL}(n,K), \cdot).$$

Insbesondere ist die Matrix $A \in M_n(K)$ invertierbar genau dann, wenn die assoziierte lineare Abbildung $x \mapsto Ax$ bijektiv ist. In diesem Fall ist die Darstellungsmatrix der inversen Abbildung die Matrix A^{-1}.

Beweis. Übung! □

Proposition 5.56. *Sei $A \in M_n(K)$. Dann sind die folgenden Aussagen äquivalent.*

(1) A ist invertierbar.

(2) A ist injektiv.

(3) A ist surjektiv.

(4) A ist bijektiv.

(5) $\operatorname{rang} A = n$.

(6) A^T ist invertierbar.

Beweis. Sei $T\colon K^n \to K^n$ die zu A assoziierte Abbildung. Dann gilt

$$\begin{array}{rcl} A \text{ ist invertierbar} & \overset{5.55}{\Longleftrightarrow} & T \text{ ist bijektiv} \qquad \Longleftrightarrow A \text{ ist bijektiv} \\ & & \Updownarrow 5.11 \\ A \text{ ist injektiv} & \Longleftrightarrow & T \text{ ist injektiv} \\ & & \Updownarrow 5.11 \\ A \text{ ist surjektiv} & \Longleftrightarrow & T \text{ ist surjektiv} \end{array}$$

Dies zeigt (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4).

Ferner ist $\operatorname{rang} A = n$ äquivalent zu $\operatorname{rang} T = n$, und dies ist äquivalent zu $\operatorname{Im} T = K^n$, also äquivalent zur Aussage, dass T surjektiv ist. Dies zeigt (5) $\Leftrightarrow$ (3).

Außerdem ist $\operatorname{rang} A = n$ äquivalent zu $\operatorname{rang} A^T = n$, nach Bemerkung 5.38. Da wir bereits gezeigt haben, dass (1) $\Leftrightarrow$ (5), so ist dies äquivalent zur Aussage, dass A^T invertierbar ist. Dies zeigt (6) $\Leftrightarrow$ (5). $\square$

Lemma 5.57. *Eine Matrix $A \in M_n(K)$ ist genau dann invertierbar, wenn sie durch elementare Zeilenumformungen zur Identitätsmatrix I_n gebracht werden kann.*

Beweis. Nach Proposition 5.56 ist A genau dann invertierbar, wenn $\operatorname{rang} A = n$. Damit folgt die Implikation „$\Longleftarrow$“ aus Proposition 5.41.

Die Implikation „$\Longrightarrow$“: Nach dem Gauß-Algorithmus kann die Matrix A durch elementare Zeilenumformungen zu einer Matrix

$$A' = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} \qquad \text{mit } a_{11} \neq 0, \dots, a_{nn} \neq 0,$$

gebracht werden. Nun formen wir diese folgendermaßen um:

$$\begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} \xrightarrow{r_n \to \frac{1}{a_{nn}} r_n} \begin{pmatrix} a_{11} & & & * & \\ & \ddots & & & \\ 0 & & & a_{n-1,n-1} & \\ 0 & \cdots & & \cdots & 1 \end{pmatrix}$$

$$\xrightarrow{r_j \to r_j - a_{jn} r_n \ (\forall j = 1, \dots, n-1)} \left(\begin{array}{cccc|c} a_{11} & & & * & 0 \\ & \ddots & & & \vdots \\ 0 & & & a_{n-1,n-1} & 0 \\ \hline 0 & \cdots & & \cdots & 1 \end{array}\right)$$

Wenn wir diesen Prozess iterativ wiederholen, so bekommen wir die Identitätsmatrix $\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$. $\square$

Definition 5.58. Eine *elementare Matrix* ist das Ergebnis einer elementaren Umformung von I_n. In anderen Worten, eine elementare Matrix ist eine der folgenden Matrizen.

(1) Die elementaren Umformungen

$$(c_i \leftrightarrow c_j) \quad \text{oder} \quad (r_i \leftrightarrow r_j)$$

ergeben die Matrizen

$$V_{ij} = \begin{array}{c} \left(\begin{array}{cccccccccccc}
1 & \cdots & 0 & \cdots & 0 & \cdots & & & & & & \\
 & \ddots & & & & & & & & & & \\
0 & \cdots & 1 & 0 & 0 & \cdots & & & & & & \\
\hline
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & & \\
\hline
 & & \cdots & 0 & 1 & 0 & \cdots & & & & & \\
 & & & & & \ddots & & & & & & \\
 & & & & \cdots & 0 & 1 & 0 & \cdots & & & \\
\hline
 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & & \\
\hline
 & & & & & & \cdots & 0 & 1 & 0 & \cdots \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & \cdots & 0 & 1
\end{array}\right) \end{array} \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \\ \\ \end{array}$$

(2) Für $\lambda \in K \setminus \{0\}$, die elementaren Umformungen

$$(c_i \to \lambda c_i) \quad \text{oder} \quad (r_i \to \lambda r_i)$$

ergeben die Matrizen

$$M_i^\lambda = \left(\begin{array}{ccccccc}
1 & \cdots & 0 & \cdots & & & \\
 & \ddots & & & & & \\
0 & \cdots & 1 & 0 & \cdots & & \\
\hline
 & \cdots & 0 & \lambda & 0 & \cdots & \\
\hline
 & & \cdots & 0 & 1 & 0 & \cdots \\
 & & & & & \ddots & \\
 & & & & \cdots & 0 & 1
\end{array}\right) \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \end{array}$$

(3) Für $\lambda \in K$, die elementaren Umformungen

$$(c_i \to c_i + \lambda c_j) \quad \text{oder} \quad (r_i \to r_i + \lambda r_j)$$

ergeben die Matrizen

$$A_{ij}^{\lambda} = \begin{array}{c} \begin{pmatrix} 1 & \cdots & 0 & \cdots & & & & \\ & \ddots & & & & & & \\ \hline 0 & \cdots & 1 & \cdots & & \lambda & \cdots & \\ \hline & \cdots & 0 & 1 & & 0 & \cdots & \\ & & & & \ddots & & & \\ & & \cdots & 0 & \cdots & 1 & \cdots & \\ & & & & & & \ddots & \\ & & & & & \cdots & 0 & 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \\ \end{array}$$

Lemma 5.59. *Die elementaren Matrizen sind invertierbar und es gilt:*

(1) $(V_{ij})^{-1} = V_{ij}$,

(2) $(M_i^{\lambda})^{-1} = M_i^{\frac{1}{\lambda}}$,

(3) $(A_{ij}^{\lambda})^{-1} = A_{ij}^{-\lambda}$.

Beweis. Übung! □

Proposition 5.60.

(1) Das Ergebnis einer elementaren Zeilenumformung einer $m \times n$-Matrix A erhält man durch Linksmultiplikation *von A mit der entsprechenden elementaren $m \times m$-Matrix.*

(2) Das Ergebnis einer elementaren Spaltenumformung einer $m \times n$-Matrix A erhält man durch Rechtsmultiplikation *von A mit der entsprechenden elementaren $n \times n$-Matrix.*

Beweis. Übung! □

Korollar 5.61. *Sei $A \in M_n(K)$ eine invertierbare Matrix. Nach Lemma 5.57 kann A durch elementare Zeilenumformungen zu I_n gebracht werden, die den elementaren Matrizen $E_1, \ldots, E_p$ entsprechen. Dann gilt*

$$A^{-1} = E_p \cdot E_{p-1} \cdot \ldots \cdot E_1. \tag{28}$$

In anderen Worten: Die elementaren Umformungen, die die Matrix A zu I_n überführen, überführen die Matrix I_n zu A^{-1}.

Beweis. Nach Proposition 5.60 gilt

$$E_p \cdot E_{p-1} \cdot \ldots \cdot E_1 \cdot A = I_n.$$

Wenn wir diese Gleichung von rechts mit A^{-1} multiplizieren, bekommen wir (28). $\square$

Beispiel 5.62. Wir werden Korollar 5.61 anwenden, und die inverse Matrix der Matrix

$$A = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & 1 \\ 0 & 2 & -1 \end{pmatrix} \in M_3(\mathbb{R})$$

zu berechnen. Und zwar:

$$\left(\begin{array}{ccc|ccc} 2 & -1 & 3 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 2 & -1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{r_1 \leftrightarrow r_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 2 & -1 & 3 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{r_2 \to r_2 - 2r_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & -2 & 0 \\ 0 & 2 & -1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{r_3 \to r_3 + 2r_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & -2 & 0 \\ 0 & 0 & 1 & 2 & -4 & 1 \end{array}\right)$$

$$\xrightarrow[r_1 \to r_1 - r_3]{r_2 \to r_2 - r_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 5 & -1 \\ 0 & -1 & 0 & -1 & 2 & -1 \\ 0 & 0 & 1 & 2 & -4 & 1 \end{array}\right)$$

$$\xrightarrow{r_2 \to -r_2} \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & -2 & 5 & -1 \\ 0 & 1 & 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & 2 & -4 & 1 \end{array}\right)$$

Damit ist

$$A^{-1} = \begin{pmatrix} -2 & 5 & -1 \\ 1 & -2 & 1 \\ 2 & -4 & 1 \end{pmatrix}.$$

Lineare Gleichungssysteme

Definition 5.63. Ein *lineares Gleichungssystem über einem Körper K* ist eine Kollektion von m linearen Gleichungen mit n Unbestimmten, die man

lösen muss:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \tag{29}$$

Die Matrix

$$A := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in M_{m,n}(K)$$

ist die *Koeffizientenmatrix* des Systems und

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

ist der *Vektor der Unbekannten*. Setze

$$b := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m.$$

Dann lässt sich das Gleichungssystem (29) kürzer schreiben als Matrizengleichung

$$Ax = b.$$

Das System heißt *homogen*, wenn $b = 0$, ansonsten heißt es *inhomogen*. Die Matrix

$$\left(\begin{array}{c|c} A & b \end{array} \right) \in M_{m,n+1}(K)$$

ist die *erweiterte Koeffizientenmatrix* des Systems.

Bemerkung 5.64. In der Notation aus der Definition gilt offensichtlich $\operatorname{rang} A \leq \operatorname{rang} \left(\begin{array}{c|c} A & b \end{array} \right)$.

Proposition 5.65. *Ein lineares Gleichungssystem $Ax = b$ hat mindestens eine Lösung genau dann, wenn* $\operatorname{rang} A = \operatorname{rang} \left(\begin{array}{c|c} A & b \end{array} \right)$.

Beweis. Wir können das System $Ax = b$ umschreiben als

$$x_1 c_1^A + \cdots + x_n c_n^A = b.$$

Damit hat das System mindestens eine Lösung genau dann, wenn b zum Spaltenraum von A gehört, oder äquivalent, wenn

$$\text{Spaltenrang von } \left(\begin{array}{c|c} A & b \end{array} \right) = \text{Spaltenrang von } A.$$

Die linke Seite dieser Gleichung ist $\operatorname{rang}\left(\begin{array}{c|c} A & b \end{array} \right)$ und die rechte Seite ist $\operatorname{rang} A$. Damit ist die Proposition bewiesen. $\square$

Definition 5.66. Hat ein lineares Gleichungssystem $Ax = b$ keine Lösung, also ist

$$\operatorname{rang}\left(\begin{array}{c|c} A & b \end{array} \right) > \operatorname{rang} A,$$

so heißt das System *inkonsistent*.

Proposition 5.67.

(1) Die Lösungsmenge eines homogenen Gleichungssystems $Ax = 0$, wobei $A \in M_n(K)$, ist der Unterraum

$$\ker A \subseteq K^n$$

und hat die Dimension $n - \operatorname{rang} A$.

(2) Sei $A \in M_n(K)$ und $b \in K^n$. Angenommen, das inhomogene Gleichungssystem $Ax = b$ hat eine Lösung $x_0 \in K^n$. Dann ist die Lösungsmenge des Systems die Nebenklasse $x_0 + \ker A$.

Beweis.
(1) Sei T die zu A assoziierte Abbildung

$$T\colon K^n \to K, \quad x \mapsto Ax.$$

Dann ist

$$Ax = 0 \quad \Longleftrightarrow \quad T(x) = 0.$$

Deswegen ist die Lösungsmenge $\ker T = \ker A$. Nach dem Dimensionssatz gilt $\dim(\ker A) = n - \operatorname{rang} A$.

(2) Da $x_0 \in K^n$ eine Lösung des Systems $Ax = b$ ist, so gilt $Ax_0 = b$. Sei y eine weitere Lösung. Dann gilt

$$0 = Ay - Ax_0 = A(y - x_0),$$

und somit ist $y - x_0$ eine Lösung des homogenen Gleichungssystems $Az = 0$. Nach (1) ist also $y - x_0 \in \ker A$, oder äquivalent: $y \in x_0 + \ker A$.

Umgekehrt: Sei $y \in x_0 + \ker A$. Dann gibt es $t \in \ker A$ mit $y = x_0 + t$, und damit

$$Ay = Ax_0 + At = Ax_0 = b.$$

Also, y ist eine Lösung des Systems $Ax = b$. $\square$

Proposition 5.68. *Sei $A \in M_n(K)$ und $b \in K^n$. Angenommen, die Matrix $\left(\begin{array}{c|c} A & b \end{array}\right)$ kann durch elementare Zeilenumformungen zu einer Matrix $\left(\begin{array}{c|c} A' & b' \end{array}\right)$ gebracht werden. Dann haben die linearen Gleichungssysteme*

$$Ax = b \quad \textit{und} \quad A'x = b'$$

dieselbe Lösungsmenge.

Beweis. Nach Proposition 5.60 gibt es elementare Matrizen $E_1, \ldots, E_p$, sodass

$$E_p \cdot \ldots \cdot E_1 \cdot \left(\begin{array}{c|c} A & b \end{array}\right) = \left(\begin{array}{c|c} A' & b' \end{array}\right).$$

In anderen Worten gilt:

$$E_p \cdot \ldots \cdot E_1 \cdot A = A' \quad \text{und} \quad E_p \cdot \ldots \cdot E_1 \cdot b = b'.$$

Das heißt, das System $A'x = b'$ ist äquivalent zum System

$$E_p \cdot \ldots \cdot E_1 \cdot Ax = E_p \cdot \ldots \cdot E_1 \cdot b.$$

Da die Matrizen $E_1, \ldots, E_p$ invertierbar sind, ist dieses System äquivalent zum System $Ax = b$. $\square$

Bemerkung 5.69. Sei $A \in M_n(K)$ und $b \in K^n$. Wir können das System $Ax = b$ umschreiben als

$$x_1 c_1^A + \cdots + x_n c_n^A = b.$$

Die elementare Spaltenumformung $c_i \leftrightarrow c_j$ (wobei $1 \leq i, j \leq n$) der Matrix A transformiert dieses System zum System

$$x_1 c_1^A + \cdots + x_i c_j^A + \cdots + x_j c_i^A + \cdots + x_n c_n^A = b,$$

und entspricht der Umbenennung der Variablen x_i und x_j.

Beispiel 5.70. Sei $A \in M_n(K)$ und $b \in K^n$. Angenommen, die Matrix $\left(\begin{array}{c|c} A & b \end{array}\right) \in M_{m,n+1}(K)$ ist in Zeilenstufenform:

$$\left(\begin{array}{c|c} A & b \end{array}\right) = \left(\begin{array}{ccccc|c} * & & & & * & \\ & \ddots & & & & \\ & & * & & & b \\ & & & \ddots & & \\ 0 & & & & & \end{array}\right).$$

Angenommen,

$$r := \operatorname{rang}\left(\begin{array}{c|c} A & b \end{array}\right) = \operatorname{rang} A.$$

Durch Umbenennung der Variablen (dies erreichen wir durch Spaltenvertauschungen) können wir annehmen, dass $\left(\begin{array}{c|c} A & b \end{array}\right)$ die folgende Form hat:

$$\left(\begin{array}{c|c} A & b \end{array}\right) = \left(\begin{array}{ccccc|c} a_{11} & & & & & b_1 \\ & \ddots & & & & \vdots \\ & & \ddots & & * & \\ & & & a_{rr} & & b_r \\ \hline & & & 0 & & 0 \end{array}\right)$$

Bezeichne

$$A_1 := \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{rr} \end{pmatrix} \in M_r(K).$$

Dann ist $\operatorname{rang} A_1 = r$ und damit ist A_1 invertierbar. Dann gilt

$$\left(\begin{array}{c|c} A & b \end{array}\right) = \left(\begin{array}{c|c} A_1 & A_2 & \begin{matrix} b_1 \\ \vdots \\ b_r \end{matrix} \\ \hline 0 & & 0 \end{array}\right)$$

für eine Matrix $A_2 \in M_{r,n-r}(K)$. Dann ist unser System äquivalent zu

$$A_1 \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} + A_2 \begin{pmatrix} x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix}.$$

Wenn wir diese Gleichung von links mit A^{-1} multiplizieren und umorganisieren, so bekommen wir

$$\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = A_1^{-1} \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix} - A_1^{-1} A_2 \begin{pmatrix} x_{r+1} \\ \vdots \\ x_n \end{pmatrix}.$$

Die ersten r Unbekannten sind also von den anderen $n-r$ Unbekannten abhängig.

Beispiel 5.71 (Das Lösen eines linearen Gleichungssystems)**.** Sei das folgende lineare Gleichungssystem über $\mathbb{R}$ gegeben:

$$\begin{cases} x & +3y & +2z & = k \\ 2x & +ky & +4z & = 12 \\ kx & +2y & -4z & = 20 \end{cases}.$$

Wir werden es nun durch Gauß-Algorithmus lösen:

$$\left(\begin{array}{ccc|c} 1 & 3 & 2 & k \\ 2 & k & 4 & 12 \\ k & 2 & -4 & 20 \end{array}\right) \xrightarrow[r_3 \to r3 - kr_1]{r_2 \to r_2 - 2r_1} \left(\begin{array}{ccc|c} 1 & 3 & 2 & k \\ 0 & k-6 & 0 & 12-2k \\ 0 & 2-3k & -4-2k & 20-k^2 \end{array}\right) =: M.$$

Wir müssen nun zwischen verschiedenen Fällen unterscheiden.

(a) Angenommen, $k = 6$, sodass

$$M = \left(\begin{array}{ccc|c} 1 & 3 & 2 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & -16 & -16 & -16 \end{array}\right) \xrightarrow[r_3 \leftrightarrow r_2]{r_3 \to -\frac{1}{16} r_3} \left(\begin{array}{ccc|c} 1 & 3 & 2 & 6 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Das System wird also zu

$$\begin{aligned} x + 3y + 2z &= 6 \\ y + z &= 1. \end{aligned}$$

Die Lösungsmenge ist also

$$\{(x, y, z) \in \mathbb{R}^3 \mid y = 1 - z,\ x = 3 + z\} = \{(3 + z, 1 - z, z) \mid z \in \mathbb{R}\}.$$

Diese Menge können wir auch schreiben wie folgt:

$$(3, 1, 0) + \{z \cdot (1, -1, 1) \mid z \in \mathbb{R}\}.$$

(b) Wir nehmen nun an, dass $k \neq 6$. Dann

$$M = \left(\begin{array}{ccc|c} 1 & 3 & 2 & k \\ 0 & k-6 & 0 & 12-2k \\ 0 & 2-3k & -4-2k & 20-k^2 \end{array}\right)$$

$$\xrightarrow{r_2 \to \frac{1}{k-6} r_2} \left(\begin{array}{ccc|c} 1 & 3 & 2 & k \\ 0 & 1 & 0 & -2 \\ 0 & 2-3k & -4-2k & 20-k^2 \end{array}\right)$$

$$\xrightarrow{r_3 \to (2-3k) r_2} \left(\begin{array}{ccc|c} 1 & 3 & 2 & k \\ 0 & 1 & 0 & -2 \\ 0 & 0 & -4-2k & 24-6k-k^2 \end{array}\right) =: N.$$

Hier müssen wir wieder zwischen verschiedenen Fällen unterscheiden.

(i) Angenommen, $k = -2$. Dann

$$N = \left(\begin{array}{ccc|c} 1 & 3 & 2 & -2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 32 \end{array}\right),$$

und damit ist das System inkonsistent.

(ii) Wir nehmen nun an, dass $k \neq -2$. Dann hat das System genau eine Lösung

$$\left(\frac{36+2k}{2+k}, -2, \frac{24-6k-k^2}{-4-2k}\right).$$

6 Determinanten

In diesem Kapitel möchten wir wollen eine Abbildung $\det\colon M_n(K) \to K$ finden und untersuchen, die die folgende Bedingung erfüllt:

$$\det A \neq 0 \quad \Longleftrightarrow \quad A \text{ ist invertierbar.}$$

Notation 6.1. In diesem Kapitel werden wir die Zeilen einer Matrix $A = (a_{ij}) \in M_n(K)$ mit

$$a_1, \ldots, a_n$$

bezeichnen. In anderen Worten, wir schreiben

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Definition 6.2 (Axiome von Weierstraß)**.** Sei K ein Körper und sei $n \in \mathbb{N}_{>0}$. Eine Abbildung

$$\det\colon M_n(K) \to K, \quad A \mapsto \det A$$

heißt *Determinante*, falls die folgenden Axiome erfüllt sind.

(D1) (*Multilinearität*) Seien $a_i', a_i'' \in K^n$. Dann gilt für alle $\lambda, \mu \in K$:

$$\det \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i' + \mu a_i'' \\ \vdots \\ a_n \end{pmatrix} = \lambda \ \det \begin{pmatrix} a_1 \\ \vdots \\ a_i' \\ \vdots \\ a_n \end{pmatrix} + \mu \ \det \begin{pmatrix} a_1 \\ \vdots \\ a_i'' \\ \vdots \\ a_n \end{pmatrix}.$$

In anderen Worten, die Determinante ist eine lineare Abbildung in jeder Zeile.

(D2) Die Funktion det ist *alternierend*, d.h. wenn A zwei gleiche Zeilen hat, so ist

$$\det A = 0.$$

(D3) Die Funktion det ist *normiert*, d.h.

$$\det I_n = 1.$$

Der nächste wichtige Satz zeigt welche andere Eigenschaften eine Determinante erfüllt, falls sie existiert. Diese Eigenschaften sind in Praxis unverzichtbar, um die Determinante einer Matrix konkret zu berechnen.

Satz 6.3. *Sei* $\det\colon M_n(K) \to K$ *eine Determinante. Dann hat sie die folgenden weiteren Eigenschaften.*

(D4) Für alle $\lambda \in K$ *gilt*

$$\det \lambda A = \lambda^n \det A.$$

(D5) Ist eine Zeile von A *gleich* 0*, so ist* $\det A = 0$.

(D6) Seien

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} \begin{array}{l} \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \\ \end{array}, \qquad B = \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} \begin{array}{l} \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \\ \end{array}.$$

Dann gilt

$$\det B = -\det A.$$

(D7) Sei $\lambda \in K$ *und seien*

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}, \qquad B = \begin{pmatrix} a_1 \\ \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}.$$

Dann gilt

$$\det B = \det A.$$

(D8) Ist A eine obere Dreiecksmatrix*, d.h.*

$$A = \begin{pmatrix} \lambda_{11} & & * \\ & \ddots & \\ 0 & & \lambda_{nn} \end{pmatrix}.$$

Dann gilt

$$\det A = \lambda_{11} \cdots \lambda_{nn}.$$

(D9) Sei $n \geq 2$ und sei $A \in M_n(K)$ eine Matrix, die die folgende Gestalt hat:

$$A = \left(\begin{array}{c|c} A_1 & C \\ \hline 0 & A_2 \end{array}\right)$$

wobei A_1, A_2 und C Matrizen sind, sodass A_1 und A_2 quadratisch sind. Dann gilt

$$\det A = \det A_1 \cdot \det A_2.$$

(D10) Sei $A \in M_n(K)$. Dann gilt

$$\det A = 0 \quad \Longleftrightarrow \quad \operatorname{rang} A < n \quad \Longleftrightarrow \quad A \text{ ist nicht invertierbar.}$$

(D11) Für alle $A, B \in M_n(K)$ gilt

$$\det (AB) = \det A \cdot \det B.$$

Insbesondere gilt

$$\det A^{-1} = \frac{1}{\det A}.$$

Beweis. Bemerke zuerst, dass die folgende Formel ein Spezialfall von (D1) ist:

$$\det \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i \\ \vdots \\ a_n \end{pmatrix} = \lambda \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix}. \tag{30}$$

(D4) folgt nun, wenn wir die Formel (30) n-mal anwenden.

(D5) folgt aus (30) mit $\lambda = 0$.

(D6) Wir berechnen:

$$\det A + \det B = \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix}$$

$$= \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} + \underbrace{\det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix}}_{\overset{(D2)}{=} 0} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \underbrace{\det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}}_{\overset{(D2)}{=} 0}$$

$$\overset{(D1)}{=} \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j + a_i \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_j + a_i \\ \vdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \vdots \\ a_j + a_i \\ \vdots \\ a_j + a_i \\ \vdots \\ a_n \end{pmatrix} \overset{(D2)}{=} 0,$$

und damit $\det B = -\det A$.

(D7) Wir berechnen:

$$\det B = \det \begin{pmatrix} a_1 \\ \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} \overset{(D1)}{=} \underbrace{\det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}}_{=\det A} + \lambda \underbrace{\det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}}_{\overset{(D2)}{=} 0} = \det A.$$

(D8) Wir unterscheiden zwei Fälle. Wir nehmen zuerst an, dass für alle $i = 1, \ldots, n$ gilt $\lambda_{ii} \neq 0$. Durch elementare Zeilenumformungen der Form A^{μ}_{ij} können wir die Matrix A zu der Form

$$A' = \begin{pmatrix} \lambda_{11} & & 0 \\ & \ddots & \\ 0 & & \lambda_{nn} \end{pmatrix}$$

bringen (siehe den Beweis von Lemma 5.57). Dann gilt:

$$\det A \overset{(D7)}{=} \det A' \overset{(D1)}{=} \lambda_1 \ldots \lambda_n \det I_n \overset{(D3)}{=} \lambda_1 \ldots \lambda_n.$$

Wir nehmen nun an, dass es ein $1 \leq i \leq n$ gibt, sodass $\lambda_{ii} = 0$. Wir dürfen annehmen, dass $\lambda_{jj} \neq 0$ für $j > i$. Dann:

$$A = \begin{pmatrix} \lambda_{11} & & & & * \\ & \ddots & & & \\ & & 0 & & \\ & & & \ddots & \\ 0 & & & & \lambda_{nn} \end{pmatrix}$$

$$\xrightarrow[\forall j = i+1, \ldots, n]{r_i \to r_i - \frac{\lambda_{ij}}{\lambda_{jj}} r_j} \begin{pmatrix} \lambda_{11} & & & & & * \\ & \ddots & & & & \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ & & & \lambda_{i+1,i+1} & & \\ & & & & \ddots & \\ 0 & & & & & \lambda_{nn} \end{pmatrix} =: A'.$$

Dann gilt:

$$\det A \overset{(D7)}{=} \det A' \overset{(D5)}{=} 0.$$

(D9) Durch wiederholte Anwendung von Zeilenoperationen des Typs V_{ij} oder A_{ij}^{λ} können wir die Matrizen A_1 und A_2 in die Zeilenstufenform bringen (mithilfe des Gauß-Algorithmus). Damit umformen wir:

$$A_1 \rightsquigarrow B_1, \quad A_2 \rightsquigarrow B_2,$$

wobei B_1 und B_2 in der Zeilenstufenform sind. Seien k bzw. ℓ die Anzahl an Vertauschungen von Zeilen in diesen Operationen. Dann gilt nach (D6) und (D7):

$$\det A_1 = (-1)^k \det B_1, \quad \det A_2 = (-1)^\ell \det B_2. \tag{31}$$

Dieselben Operationen transformieren die Matrix

$$\left(\begin{array}{c|c} A_1 & C \\ \hline 0 & A_2 \end{array}\right) \rightsquigarrow \left(\begin{array}{c|c} B_1 & C' \\ \hline 0 & B_2 \end{array}\right),$$

und es gilt wieder nach (D6) und (D7):

$$\det A = (-1)^{k+\ell} \left(\begin{array}{c|c} B_1 & C' \\ \hline 0 & B_2 \end{array}\right).$$

Da die Matrix $\left(\begin{array}{c|c} B_1 & C' \\ \hline 0 & B_2 \end{array}\right)$ eine obere Dreiecksmatrix ist, so gilt nach D8, dass

$$\det \left(\begin{array}{c|c} B_1 & C' \\ \hline 0 & B_2 \end{array}\right) = \det B_1 \det B_2,$$

also

$$\det A = (-1)^{k+\ell} \det B_1 \det B_2.$$

Dies zusammen mit (31) zeigt (D9).

(D10) Durch elementare Zeilenumformungen des Typs V_{ij} oder A_{ij}^{λ} können wir die Matrix A in Zeilenstufenform bringen (mithilfe des Gauß-Algorithmus). Dann ist B eine obere Dreiecksmatrix:

$$B = \begin{pmatrix} \lambda_{11} & & * \\ & \ddots & \\ 0 & & \lambda_{nn} \end{pmatrix},$$

und es gilt

$$\operatorname{rang} A = \operatorname{rang} B \quad \text{und} \quad \det A = \pm \det B.$$

Damit gilt:

$$\begin{aligned} \operatorname{rang} B = n \quad &\Longleftrightarrow \quad \lambda_{ii} \neq 0 \text{ für alle } i = 1, \ldots, n \\ &\overset{(D8)}{\Longleftrightarrow} \quad \det B = \lambda_{11} \cdots \lambda_{nn} \neq 0. \end{aligned}$$

Aus diesen Äquivalenzen folgt die Aussage.

(D11) Wir nehmen zuerst an, dass $\operatorname{rang} A < n$. Nach (D10) ist dies äquivalent zu $\det A = 0$, und wir müssen zeigen, dass $\det AB = 0$.

Bezeichne mit T_A bzw. T_B die zu A bzw. B assoziierten linearen Abbildungen. Dann gilt:

$$\operatorname{rang} A < n \quad \Longleftrightarrow \quad \dim \big(T_A(K^n)\big) < n,$$

und damit

$$\operatorname{rang} AB = \dim \left((T_A \circ T_B)(K^n)\right) = \dim T_A(\underbrace{T_B(K^n)}_{\subseteq K^n}) \leq \dim T_A(K^n) < n.$$

Dann ist $\det AB = 0$ nach (D10), wie gewünscht.

Es bleibt (D11) zu zeigen, wenn $\operatorname{rang} A = n$. Nach (D10) ist dann A ist invertierbar, und nach Lemma 5.57 und Proposition 5.60 gibt es elementare Matrizen $E_1, \ldots, E_p$, sodass

$$A = E_p \cdot \ldots \cdot E_1.$$

Folglich ist

$$AB = E_p \cdot \ldots \cdot E_1 \cdot B.$$

Aus diesen zwei Formeln folgt, dass es ausreicht, die Formel

$$\det AB = \det A \cdot \det B$$

zu zeigen, wenn A eine elementare Matrix ist. (Übung!)

Wir haben drei Fälle. Sei $A := V_{ij}$. Da die Matrix V_{ij} bzw. $V_{ij}B$ der Vertauschung der i-ten und j-ten Zeile von I_n bzw. B korrespondiert, so gilt:

$$\det(V_{ij}B) \overset{(D6)}{=} -\det B \overset{(D3)}{=} -\det I_n \cdot \det B \overset{(D6)}{=} \det V_{ij} \cdot \det B,$$

wie gewünscht.

Sei nun $A := A_{ij}^{\lambda}$ für ein $\lambda \in K$. Ähnlich wie oben gilt:

$$\det\left(A_{ij}^{\lambda} B\right) \overset{(D7)}{=} \det B \overset{(D3)}{=} \det I_n \cdot \det B \overset{(D7)}{=} \det A_{ij}^{\lambda} \cdot \det B,$$

wie gewünscht.

Sei letztlich $A := M_i^{\lambda}$. Ähnlich wie oben gilt:

$$\det\left(M_i^{\lambda} B\right) \overset{(30)}{=} \lambda \det B \overset{(D3)}{=} \lambda \det I_n \cdot \det B \overset{(30)}{=} \det M_i^{\lambda},$$

und der Beweis ist komplett. $\square$

Beispiel 6.4. Wir werden die obigen Regeln benutzen, um die Determinante der folgenden Matrix zu berechnen (falls sie existiert):

$$\det \begin{pmatrix} 0 & 1 & 2 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix} \overset{r_1 \leftrightarrow r_3}{=} -\det \begin{pmatrix} 1 & 1 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \overset{r_2 \to r_2 - 3r_1}{=} -\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 2 & 1 \end{pmatrix}$$
$$\overset{r_3 \to r_3 + r_2}{=} -\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 3 \end{pmatrix} = -1 \cdot (-1) \cdot 3 = 3.$$

Bemerkung 6.5 (Anwendungen in Gleichungssystemen)**.** Sei K ein Körper und sei $A \in M_n(K)$. Dann gilt:

$$\begin{aligned} &\text{das lineare Gleichungssystem } Ax = 0 \text{ hat genau eine Lösung} \\ \overset{5.67}{\iff}\ & \ker A = \{0\} \iff A \text{ ist injektiv} \overset{5.56}{\iff} A \text{ ist bijektiv} \\ \overset{5.56}{\iff}\ & A \text{ ist invertierbar} \overset{(D10)}{\iff} \det A \neq 0. \end{aligned}$$

Notation 6.6. Seien $i, j \in \{1, \ldots, n\}$. Dann ist das *Kronecker-Symbol* oder *Kronecker-Delta* zum Paar (i, j) definiert als

$$\delta_{ij} = \begin{cases} 0, & \text{wenn } i \neq j, \\ 1, & \text{wenn } i = j. \end{cases}$$

Also lässt sich die Identitätsmatrix I_n schreiben als

$$I_n = (\delta_{ij})_{i,j=1,\ldots,n}.$$

Satz 6.7 (Satz von Leibnitz)**.** *Sei K ein Körper und $n \in \mathbb{N}_{>0}$. Dann gibt es genau eine Determinante*

$$\det\colon M_n(K) \to K.$$

Für eine Matrix $A = (a_{ij}) \in M_n(K)$ ist die Determinante gegeben durch die Formel

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)}. \tag{32}$$

Beweis. Wir zeigen zuerst die Eindeutigkeit. Dazu müssen wir zeigen, dass es maximal eine Funktion gibt, die die Axiome (D1)–(D3) erfüllt. Für eine Matrix $A = (a_{ij}) \in M_n(K)$ schreiben wir jede Zeile a_i in der Form

$$a_i = a_{i1}e_1 + \cdots + a_{in}e_n,$$

wobei $\{e_1, \ldots, e_n\}$ die kanonische Basis von K^n ist. Dann gilt:

$$\begin{aligned}
\det A = \det \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} &= \det \begin{pmatrix} a_{11}e_1 + \cdots + a_{1n}e_n \\ \vdots \\ a_n \end{pmatrix} \\
&\overset{(D2)}{=} a_{11} \det \begin{pmatrix} e_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \cdots + a_{1n} \det \begin{pmatrix} e_n \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \\
&= \sum_{i=1}^{n} a_{1i} \det \begin{pmatrix} e_i \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.
\end{aligned}$$

Wenn wir diesen Prozess weiterführen für jede weitere Zeile, dann bekommen wir am Ende die Formel:

$$\det A = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} a_{1i_1} \cdot \ldots \cdot a_{n,i_n} \det \begin{pmatrix} e_{i_1} \\ \vdots \\ e_{i_n} \end{pmatrix}.$$

Dann ist die Eindeutigkeit aus dieser Formel offensichtlich.

Nun zeigen wir die Existenz. Dazu müssen wir zeigen, dass die in (32) gegebene Formel eine Determinante definiert, d.h. dass diese Abbildung die Axiome (D1)–(D3) erfüllt.

(D1) Seien $\lambda, \mu \in K$. Dann gilt:

$$\begin{aligned}
\det \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i' + \mu a_i'' \\ \vdots \\ a_n \end{pmatrix} &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdots (\lambda a_{i,\sigma(i)}' + \mu a_{i,\sigma(i)}'') \cdots a_{n,\sigma(n)} \\
&= \lambda \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdots a_{i,\sigma(i)}' \cdots a_{n,\sigma(n)} \\
&\quad + \mu \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdots a_{i,\sigma(i)}'' \cdots a_{n,\sigma(n)} \\
&= \lambda \det \begin{pmatrix} a_1 \\ \vdots \\ a_i' \\ \vdots \\ a_n \end{pmatrix} + \mu \det \begin{pmatrix} a_1 \\ \vdots \\ a_i'' \\ \vdots \\ a_n \end{pmatrix}.
\end{aligned}$$

(D2) Sei A eine $n \times n$-Matrix, in der die k-te und die ℓ-te Zeile gleich sind (mit $k < \ell$). Setze

$$\tau := \begin{pmatrix} k & \ell \end{pmatrix} \in S_n.$$

Nach Lemma 2.45 gilt

$$S_n = A_n \sqcup A_n \tau,$$

und für jedes $\sigma \in A_n$ gilt $\operatorname{sign}(\sigma) = 1$ und $\operatorname{sign}(\sigma \circ \tau) = -1$. Es folgt:

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)} = \sum_{\sigma \in A_n} \prod_{i=1}^n a_{i,\sigma(i)} - \sum_{\sigma \in A_n} \prod_{i=1}^n a_{i,\sigma(\tau(i))},$$

wobei

$$\begin{aligned}
\prod_{i=1}^n a_{i,\sigma(\tau(i))} &= a_{1,\sigma(\tau(1))} \cdots a_{k,\sigma(\tau(k))} \cdots a_{\ell,\sigma(\tau(\ell))} \cdots a_{n,\sigma(\tau(n))} \\
&= a_{1,\sigma(1)} \cdots \underbrace{a_{k,\sigma(\tau(k))}}_{=a_{k,\sigma(\ell)}=a_{\ell,\sigma(\ell)}} \cdots \underbrace{a_{\ell,\sigma(\tau(\ell))}}_{=a_{\ell,\sigma(k)}=a_{k,\sigma(k)}} \cdots a_{n,\sigma(n)} \\
&= a_{1,\sigma(1)} \cdots a_{l,\sigma(l)} \cdots a_{k,\sigma(k)} \cdots a_{n,\sigma(n)},
\end{aligned}$$

da die k-te und die ℓ-te Zeile gleich sind. Damit tauchen alle Summanden in der zweiten Summe auch in der ersten Summe auf, und damit ist $\det A = 0$.

(D3) Da $I_n = (\delta_{ij})_{i,j=1,\dots,n}$, und da

$$\prod_{i=1}^{n} \delta_{i,\sigma(i)} = \begin{cases} 1, & \text{wenn } \sigma = \mathrm{id}, \\ 0, & \text{wenn } \sigma \neq \mathrm{id}, \end{cases}$$

so gilt

$$\det I_n = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} \delta_{i,\sigma(i)} = \delta_{11} \cdots \delta_{nn} = 1,$$

wie gewünscht. □

Beispiel 6.8.

(1) $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$.

(2) $\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$
$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{12}a_{21}a_{33}$.

Satz 6.9. *Für $A \in M_n(K)$ gilt*

$$\det A^T = \det A.$$

Beweis. Sei $A = (a_{ij})$, sodass $A^T = (a_{ji})$. Dann gilt:

$$\det A^T = \sum_{\sigma \in S_n} \underbrace{\operatorname{sign}(\sigma)}_{=\operatorname{sign}(\sigma^{-1})} \underbrace{\prod_{i=1}^{n} a_{\sigma(i),i}}_{=\prod_{i=1}^{n} a_{i,\sigma^{-1}(i)}} = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma^{-1}) \prod_{i=1}^{n} a_{i,\sigma^{-1}(i)} = \det A,$$

wie gewünscht. □

Bemerkung 6.10. Dieser Satz zeigt, dass wir die Axiome (D1)–(D3) auch bezüglich der Spalten formulieren können.

Definition 6.11. Sei K ein Körper und sei $A = (a_{k\ell}) \in M_n(K)$. Seien $i, j \in \{1, \dots, n\}$ und setze

$$A_{ij} := \left(\begin{array}{ccc|c|ccc} a_{11} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ \hline 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \hline a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn} \end{array}\right);$$

in anderen Worten, wir bekommen die Matrix A_{ij} aus der Matrix A, indem wir in der i-ten Zeile und der j-ten Spalte alle Einträge außer (i, j) durch 0 ersetzen, und den (i, j)-Eintrag durch 1 ersetzen. Setze

$$a_{ij}^{\#} := \det A_{ji}.$$

(Achtung: die Indizes sind vertauscht!) Die *zu A komplementäre Matrix* ist definiert durch

$$A^{\#} := (a_{ij}^{\#}).$$

Notation 6.12. Sei K ein Körper und sei $A = (a_{k\ell}) \in M_n(K)$. Seien $i, j \in \{1, \dots, n\}$ und setze

$$A'_{ij} := \begin{pmatrix} a_{11} & \cdots & \cancel{a_{1j}} & \cdots & a_{1n} \\ \vdots & & \cancel{\vdots} & & \vdots \\ \cancel{a_{i1}} & \cancel{\cdots} & \cancel{a_{ij}} & \cancel{\cdots} & \cancel{a_{in}} \\ \vdots & & \cancel{\vdots} & & \vdots \\ a_{n1} & \cdots & \cancel{a_{nj}} & \cdots & a_{nn} \end{pmatrix} \in M_{n-1}(K).$$

Also die Matrix A'_{ij} entsteht aus A durch das Streichen der i-ten Zeile und j-ten Spalte. Wir werden die Spalten von A von nun an mit $a^1, \dots, a^n$ bezeichnen; es gilt also

$$A = \begin{pmatrix} a^1 & \cdots & a^n \end{pmatrix}.$$

Wir bezeichnen den i-ten Einheitsspaltenvektor mit e^i, also

$$e^i := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{-te Stelle}$$

Lemma 6.13. *Mit der Notation wie oben gilt:*

(1) $\det A_{ij} = (-1)^{i+j} \det A'_{ij}$,

(2) $\det A_{ij} = \det \begin{pmatrix} a^1 & \cdots & a^{j-1} & e^i & a^{j+1} & \cdots & a^n \end{pmatrix}$.

Beweis.
(1) Durch $i-1$ Vertauschungen von Zeilen und $j-1$ Vertauschungen von Spalten können wir die Matrix A_{ij} in die Form

$$B = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A'_{ij} & \\ 0 & & & \end{array}\right)$$

bringen. Dann gilt nach (D6):

$$\det B = (-1)^{i-1}(-1)^{j-1} \det A_{ij} = (-1)^{i+j} \det A_{ij}.$$

Andererseits gilt

$$\det B \overset{(D9)}{=} \det(1) \cdot \det A'_{ij} = \det A'_{ij},$$

und (1) folgt.

(2) Es gilt

$$\begin{aligned} &\begin{pmatrix} a^1 & \cdots & a^{j-1} & e^i & a^{j+1} & \cdots & a^n \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & 1 & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix} \leftarrow i\text{-te Zeile} \end{aligned}$$

Wir können Spaltenumformungen der Form A_{ij}^{λ} benutzen, um alle Einträge der i-ten Zeile dieser Matrix zu null zu machen außer den Eintrag (i,j). Daher wird die Matrix oben durch diese Operationen zur Matrix A_{ij}, und damit sind ihre Determinanten gleich. $\square$

Satz 6.14. *Sei K ein Körper. Ist $A \in M_n(K)$, dann gilt*

$$A \cdot A^{\#} = A^{\#} \cdot A = \det A \cdot I_n.$$

Insbesondere, wenn $A \in \mathrm{GL}(n,K)$, dann

$$A^{-1} = \frac{1}{\det A} A^{\#}.$$

Beweis. Wir werden alle Einträge der Matrix $A^{\#} \cdot A$ berechnen; der Beweis für $A \cdot A^{\#}$ ist ähnlich. Der (i,k)-ter Eintrag von $A^{\#} \cdot A$ lautet:

$$\begin{aligned}
\sum_{j=1}^{n} a_{ij}^{\#} a_{jk} &= \sum_{j=1}^{n} a_{jk} \det A_{ji} \\
&\overset{6.13(2)}{=} \sum_{j=1}^{n} a_{jk} \det \begin{pmatrix} a^1 & \cdots & a^{i-1} & e^j & a^{i+1} & \cdots & a^n \end{pmatrix} \\
&\overset{(D1)}{=} \det \begin{pmatrix} a^1 & \cdots & a^{i-1} & \underbrace{\sum_{j=1}^{n} a_{jk} e^j}_{=a^k} & a^{i+1} & \cdots & a^n \end{pmatrix} \\
&= \det \begin{pmatrix} a^1 & \cdots & a^{i-1} & a^k & a^{i+1} & \cdots & a^n \end{pmatrix} \\
&\overset{(D2)}{=} \begin{cases} \det A, & \text{wenn } k = i, \\ 0, & \text{sonst} \end{cases} \\
&= \delta_{ik} \det A.
\end{aligned}$$

Daher gilt $A^{\#} A = \det A \cdot I_n$. $\square$

Das folgende Korollar ist eine Umformulierung der letzten Aussage von Satz 6.14.

Korollar 6.15. *Sei K ein Körper und sei $A \in \mathrm{GL}(n,K)$. Definiere die Matrix $C = (c_{ij}) \in M_n(K)$ durch*

$$c_{ij} := (-1)^{i+j} \det A'_{ij}.$$

Dann gilt

$$A^{-1} = \frac{1}{\det A} C^T.$$

Satz 6.16 (Entwicklungssatz von Laplace)**.** *Sei K ein Körper, sei $n \geq 2$ und sei $A = (a_{ij}) \in M_n(K)$. Für jedes $i \in \{1, \ldots, n\}$ gilt*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A'_{ij}$$

(Entwicklung nach der i-ten Zeile). Für jedes $j \in \{1, \ldots, n\}$ gilt

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A'_{ij}$$

(Entwicklung nach der j-ten Spalte).

Beweis. Nach Satz 6.14 ist $\det A$ der Eintrag auf der Hauptdiagonale der Matrix $A \cdot A^{\#}$. Also für jedes fixierte $i \in \{1, \ldots, n\}$ gilt:

$$\det A = \sum_{j=1}^{n} a_{ij} a^{\#}_{ji} = \sum_{j=1}^{n} a_{ij} \det A^{\#}_{ij} \overset{6.13(1)}{=} \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A'_{ij}.$$

Dies zeigt die erste Formel. Die zweite beweist man ähnlich, indem man stattdessen das Produkt $A^{\#} \cdot A$ betrachtet. □

Beispiel 6.17. In diesem Beispiel berechnen wir die Determinante der Matrix $\begin{pmatrix} 0 & 1 & 2 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ auf zwei verschiedene Weise. Zuerst entwickeln wir nach der ersten Zeile:

$$\begin{aligned} \det \begin{pmatrix} \boxed{0 \quad 1 \quad 2} \\ 3 \quad 2 \quad 1 \\ 1 \quad 1 \quad 0 \end{pmatrix} &= 0 \cdot \det \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} \\ &= 0 - (-1) + 2 \cdot 1 = 3. \end{aligned}$$

Nun entwickeln wir nach der ersten Spalte:

$$\begin{aligned} \det \begin{pmatrix} \boxed{\begin{matrix} 0 \\ 3 \\ 1 \end{matrix}} & \begin{matrix} 1 \\ 2 \\ 1 \end{matrix} & \begin{matrix} 2 \\ 1 \\ 0 \end{matrix} \end{pmatrix} &= 0 \cdot \det \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} - 3 \cdot \det \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} + 1 \cdot \det \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \\ &= 0 - 3 \cdot (-2) + 1 \cdot (-3) = 3. \end{aligned}$$

Bemerkung 6.18. Den Faktor $(-1)^{i+j}$ können wir uns nach dem „Schachbrettmuster“ merken:

+	-	+	-	$\cdots$
-	+	-	+	
+	-	+	-	
-	+	-	+	
$\vdots$				$\ddots$

Beispiel 6.19. Aus Satz 6.14 können wir berechnen:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Cramersche Regel

Sei $Ax = b$ ein lineares Gleichungssystem, wobei $A \in \mathrm{GL}(n, k)$ und $b \in K^n$. Dann gilt

$$x = A^{-1}b,$$

und dies ist die einzige Lösung des Systems.

Man kann die Lösung aber auch finden, ohne die Matrix A^{-1} berechnen zu müssen.

Satz 6.20 (Cramersche Regel)**.** *Sei $A = \begin{pmatrix} a^1 & \cdots & a^n \end{pmatrix} \in \mathrm{GL}(n, K)$, wobei $a^1, \ldots, a^n$ die Spalten der Matrix A sind. Sei $b \in K^n$ und sei $y = \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix}^T \in K^n$ die eindeutige Lösung des Systems $Ax = b$. Für jedes $i \in \{1, \ldots, n\}$ gilt:*

$$y_i = \frac{\det \begin{pmatrix} a^1 & \cdots & a^{i-1} & b & a^{i+i} & \cdots & a^n \end{pmatrix}}{\det A}.$$

Beweis. Sei $A^{-1} := (c_{ij})$ und $b = \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix}^T$. Nach Satz 6.14 gilt, für alle $i, j \in \{1, \ldots, n\}$:

$$c_{ij} = \frac{a_{ij}^{\#}}{\det A} = \frac{\det A_{ji}}{\det A} \overset{6.13(2)}{=} \frac{\det \begin{pmatrix} a^1 & \cdots & a^{i-1} & e^j & a^{i+1} & \cdots & a^n \end{pmatrix}}{\det A}. \quad (33)$$

Da $y = A^{-1}b$, so folgt:

$$\begin{aligned} y_i &= \sum_{j=1}^{n} c_{ij} b_j \overset{(33)}{=} \sum_{j=1}^{n} b_j \frac{\det\begin{pmatrix} a^1 & \cdots & a^{i-1} & e^j & a^{i+1} & \cdots & a^n \end{pmatrix}}{\det A} \\ &\overset{(D1)}{=} \frac{\det\begin{pmatrix} a^1 & \cdots & a^{i-1} & \sum_{j=1}^{n} b_j e^j & a^{i+1} & \cdots & a^n \end{pmatrix}}{\det A} \\ &= \frac{\det\begin{pmatrix} a^1 & \cdots & a^{i-1} & b & a^{i+i} & \cdots & a^n \end{pmatrix}}{\det A}, \end{aligned}$$

welches zu beweisen war. □

Beispiel 6.21. Betrachte das folgende Gleichungssystem über $\mathbb{R}$:

$$\begin{array}{rcrcrl} x_1 & + & x_2 & & & = 1 \\ & & x_2 & + & x_3 & = 1 \\ 3x_1 & + & 2x_2 & + & x_3 & = 0. \end{array}$$

Dann ist die Koeffizientenmatrix

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix},$$

und da $\det A = 2 \neq 0$, so ist $A \in \mathrm{GL}(3, \mathbb{R})$. Nun folgt aus der Cramerschen Regel, dass die eindeutige Lösung des Systems ist:

$$y_1 = \frac{1}{2} \det \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} = -1,$$

$$y_2 = \frac{1}{2} \det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 3 & 0 & 1 \end{pmatrix} = 2,$$

$$y_3 = \frac{1}{2} \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 3 & 2 & 0 \end{pmatrix} = -1.$$

Basiswechsel

Definition 6.22. Sei K ein Körper, sei V ein n-dimensionaler K-Vektorraum und seien

$$\mathcal{A} := \{v_1, \ldots, v_n\} \quad \text{und} \quad \mathcal{B} := \{w_1, \ldots, w_n\}$$

zwei Basen von V. Die Darstellungsmatrix der Abbildung

$$\mathrm{id}_V \colon V \to V \text{ bezüglich der Basen } \mathcal{A} \text{ und } \mathcal{B}$$

nennt man die *Basiswechselmatrix von $\mathcal{A}$ zu $\mathcal{B}$*, und bezeichnet mit

$$T_{\mathcal{B}}^{\mathcal{A}}.$$

In anderen Worten: Es gibt Skalare $a_{ij} \in K$, sodass

$$\begin{aligned} \mathrm{id}_V(v_1) = v_1 &= a_{11}w_1 + a_{21}w_2 + \cdots + a_{n1}w_n, \\ &\vdots \\ \mathrm{id}_V(v_n) = v_n &= a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{nn}w_n. \end{aligned}$$

Dann ist

$$T_{\mathcal{B}}^{\mathcal{A}} := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Es ist klar, dass $T_{\mathcal{B}}^{\mathcal{A}} \in \mathrm{GL}(n, K)$, und es gilt

$$\left(T_{\mathcal{B}}^{\mathcal{A}}\right)^{-1} = T_{\mathcal{A}}^{\mathcal{B}}.$$

Satz 6.23. *Seien V und W zwei K-Vektorräume mit $n = \dim V$ und $m = \dim W$, und sei $F \colon V \to W$ eine lineare Abbildung. Seien $\mathcal{A}$ und $\mathcal{A}'$ zwei Basen von V, und seien $\mathcal{B}$ und $\mathcal{B}'$ zwei Basen von W. Sei weiterhin*

$$A := \textit{Darstellungsmatrix von } F \textit{ bezüglich } \mathcal{A} \textit{ und } \mathcal{B},$$

und sei

$$B := \textit{Darstellungsmatrix von } F \textit{ bezüglich } \mathcal{A}' \textit{ und } \mathcal{B}'.$$

Dann gilt

$$B = T_{\mathcal{B}'}^{\mathcal{B}} \cdot A \cdot (T_{\mathcal{A}'}^{\mathcal{A}})^{-1}.$$

(Bemerke, dass $T_{\mathcal{A}'}^{\mathcal{A}} \in \mathrm{GL}(n, K)$ und $T_{\mathcal{B}'}^{\mathcal{B}} \in \mathrm{GL}(m, K)$.)

Beweis. Sei R die Darstellungsmatrix von F bezüglich der Basen $\mathcal{A}$ und $\mathcal{B}'$. Betrachte das folgende kommutative Diagramm:

$$\begin{tikzcd} \mathcal{A} \arrow[r, phantom, "\subseteq"] & V \arrow[r, "\mathrm{id}_V"] \arrow[d, "F"] \arrow[dr, "F"] & V \arrow[r, phantom, "\supseteq"] \arrow[d, "F"] & \mathcal{A}' \\ \mathcal{B} \arrow[r, phantom, "\subseteq"] & W \arrow[r, "\mathrm{id}_W"] & W \arrow[r, phantom, "\supseteq"] & \mathcal{B}'. \end{tikzcd}$$

Aus dem oberen Dreieck haben wir $F = F \circ \mathrm{id}_V$, und damit gilt, nach Bemerkung 5.24:

$$R = B \cdot T^{\mathcal{A}}_{\mathcal{A}'}.$$

Aus dem unteren Dreieck haben wir $F = \mathrm{id}_W \circ F$, und damit gilt, wieder nach Bemerkung 5.24:

$$R = T^{\mathcal{B}}_{\mathcal{B}'} \cdot A.$$

Aus diesen zwei Gleichungen folgt

$$B \cdot T^{\mathcal{A}}_{\mathcal{A}'} = T^{\mathcal{B}}_{\mathcal{B}'} \cdot A,$$

und der Satz ist bewiesen. □

Korollar 6.24. *Seien V ein K-Vektorraum mit $n = \dim V$, und sei $F\colon V \to V$ ein Endomorphismus. Seien $\mathcal{A}$ und $\mathcal{B}$ zwei Basen von V. Sei weiterhin*

$$A := \textit{Darstellungsmatrix von } F \textit{ bezüglich } \mathcal{A} \textit{ und } \mathcal{A},$$

und sei

$$B := \textit{Darstellungsmatrix von } F \textit{ bezüglich } \mathcal{B} \textit{ und } \mathcal{B}.$$

Dann gilt

$$B = T^{\mathcal{A}}_{\mathcal{B}} \cdot A \cdot (T^{\mathcal{A}}_{\mathcal{B}})^{-1}.$$

Insbesondere gilt

$$\det B = \det A.$$

Beweis. Die erste Aussage folgt aus Satz 6.23. Dann folgt:

$$\begin{aligned}\det B &= \det\left(T^{\mathcal{A}}_{\mathcal{B}} \cdot A \cdot (T^{\mathcal{A}}_{\mathcal{B}})^{-1}\right) = \det T^{\mathcal{A}}_{\mathcal{B}} \cdot \det A \cdot \det\left((T^{\mathcal{A}}_{\mathcal{B}})^{-1}\right)\\ &= \det T^{\mathcal{A}}_{\mathcal{B}} \cdot \det A \cdot \frac{1}{\det T^{\mathcal{A}}_{\mathcal{B}}} = \det A,\end{aligned}$$

wie gewünscht. □

Definition 6.25. Sei K ein Körper. Zwei Matrizen $A, B \in M_n(K)$ sind *ähnlich*, falls es eine Matrix $P \in \mathrm{GL}(n, K)$ gibt mit

$$B = PAP^{-1}.$$

Bemerkung 6.26. *Ähnlichkeit* ist eine Äquivalenzrelation auf $M_n(K)$. (Übung!)

Definition 6.27 (Determinante eines Endomorphismus)**.** Sei V ein n-dimensionaler K-Vektorraum und sei $F\colon V \to V$ ein Endomorphismus. Seien $\mathcal{A}$ eine Basis von V und sei

$$A := \text{Darstellungsmatrix von } F \text{ bezüglich } \mathcal{A} \text{ und } \mathcal{A}.$$

Nach Korollar 6.24 ist $\det A$ unabhängig von der Wahl der Basis $\mathcal{A}$. Definiere

$$\det F := \det A.$$

7 Eigenwerte und Eigenvektoren

Definition 7.1. Sei V ein K-Vektorraum und sei $T\colon V \to V$ ein Endomorphismus. Gibt es einen Skalar $\lambda \in K$ und einen Vektor $v \in V \setminus \{0\}$ mit

$$T(v) = \lambda v,$$

so heißt λ *Eigenwert von* T, und v heißt *Eigenvektor von* T *zu* λ. Wenn λ ein Eigenwert von T ist, dann heißt die Menge

$$\mathrm{Eig}(T, \lambda) := \{v \in V \mid T(v) = \lambda v\}$$

Eigenraum von T *zum Eigenwert* λ. Es gilt

$$\mathrm{Eig}(T, \lambda) = \ker\left(T - \lambda \cdot \mathrm{id}_V\right)$$

(siehe Bemerkung 7.2 unten), und damit ist $\mathrm{Eig}(T, \lambda)$ ein Unterraum von V. Seine Dimension heißt *geometrische Vielfachheit* von λ.

Bemerkung 7.2. Sei λ ein Eigenwert von T. Dann gilt:

$$\begin{aligned} v \in \mathrm{Eig}(T, \lambda) \quad &\iff \quad T(v) = \lambda v \quad \iff \quad T(v) - \lambda v = 0 \\ &\iff \quad T(v) - \lambda \cdot \mathrm{id}_V(v) = 0 \quad \iff \quad (T - \lambda \cdot \mathrm{id}_V)(v) = 0 \\ &\iff \quad v \in \ker\left(T - \lambda \cdot \mathrm{id}_V\right). \end{aligned}$$

Beispiel 7.3.

(1) Sei V ein K-Vektorraum. Dann ist 1 der einzige Eigenwert von id_V und es gilt

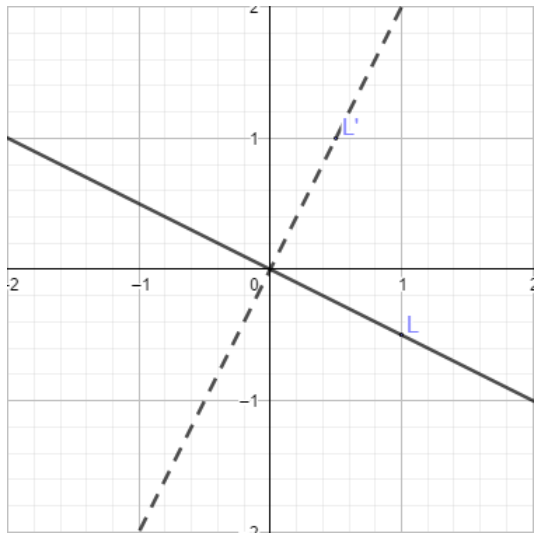
$$\mathrm{Eig}(\mathrm{id}_V, 1) = V.$$

(2) Sei V ein K-Vektorraum. Jede nicht-injektive lineare Abbildung $T\colon V \to V$ hat den Eigenwert 0 und es gilt

$$\mathrm{Eig}(T, 0) = \ker T.$$

(3) Sei $S\colon \mathbb{R}^2 \to \mathbb{R}^2$ eine Spiegelung an einer Geraden L in $\mathbb{R}$. Die Eigenwerte von S sind 1 und -1, und es gilt

$$\begin{aligned} \mathrm{Eig}(S, 1) &= L, \\ \mathrm{Eig}(S, -1) &= L' := \text{die senkrechte Gerade zu } L \text{ durch den Punkt } (0, 0). \end{aligned}$$



(4) Sei $R_\theta \colon \mathbb{R}^2 \to \mathbb{R}^2$ die Drehung um einen Winkel θ. Dann hat R_θ mindestens einen Eigenwert genau dann, wenn $\theta = \pi k$ für ein $k \in \mathbb{Z}$.

Proposition 7.4. *Sei V ein endlich-dimensionaler K-Vektorraum und sei $T \colon V \to V$ ein Endomorphismus. Dann ist $\lambda \in K$ ein Eigenwert von T genau dann, wenn* $\det(T - \lambda \cdot \mathrm{id}_V) = 0$.

Beweis. Ein $\lambda \in K$ ist ein Eigenwert von T genau dann, wenn

$$\mathrm{Eig}(T, \lambda) \neq \{0\}.$$

Nach Bemerkung 7.2 ist dies äquivalent zu

$$\ker(T - \lambda \cdot \mathrm{id}_V) \neq \{0\},$$

oder äquivalent:

$$\mathrm{def}(T - \lambda \cdot \mathrm{id}_V) > 0.$$

Nach dem Dimensionssatz ist dies genau dann der Fall, wenn

$$\mathrm{rang}(T - \lambda \cdot \mathrm{id}_V) < \dim V,$$

und dies ist äquivalent zu $\det(T - \lambda \cdot \mathrm{id}_V) = 0$ nach (D10). □

Definition 7.5. Sei V ein n-dimensionaler K-Vektorraum und sei $T \colon V \to V$ ein Endomorphismus. Die Abbildung

$$\chi_T \colon K \to K, \quad \lambda \mapsto \det(T - \lambda \cdot \mathrm{id}_V)$$

heißt *charakteristische Funktion von T*.

Bemerkung 7.6. Sei V ein n-dimensionaler K-Vektorraum und sei $T\colon V \to V$ ein Endomorphismus. Sei $\mathcal{A}$ eine beliebige Basis von V, und sei $A = (a_{ij})$ die Darstellungsmatrix von T bezüglich der Basen $\mathcal{A}$ und $\mathcal{A}$. Dann ist

$$\chi_T(\lambda) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix}.$$

Wir können diese Determinante berechnen nach dem Entwicklungssatz von Laplace, und daraus folgt, dass es $b_0, \ldots, b_{n-1}, b_n \in K$ gibt, sodass

$$\chi_T(\lambda) = b_n \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_0.$$

Man sieht sofort, dass

$$b_n = (-1)^n \quad \text{und} \quad b_0 = \chi_T(0) = \det T.$$

Damit ist χ_T ein *Polynom vom Grad n in λ mit Koeffizienten in K*. Deswegen nennen wir χ_T auch das *charakteristische Polynom von T*.

Korollar 7.7. *In der Notation aus Definition 7.5 hat die Abbildung T höchstens n verschiedene Eigenwerte.*

Beweis. Nach Proposition 7.4 sind die Eigenwerte von T Nullstellen des Polynoms χ_T, und dieses Polynom hat maximal n verschiedene Nullstellen. □

Beispiel 7.8. Wir berechnen nun alle Eigenwerte und Eigenräume der Matrix

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix} \in M_3(\mathbb{Q}).$$

Es gilt

$$\chi_A(\lambda) = \det(A - \lambda I_3) = \det \begin{pmatrix} 3-\lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3-\lambda \end{pmatrix} = -(\lambda+1)^2(\lambda-8),$$

und damit sind -1 und 8 die Eigenwerte von A. Die Eigenräume sind:

$$\begin{aligned}
\operatorname{Eig}(A,-1) &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{Q}^3 \;\middle|\; \underbrace{\begin{pmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{pmatrix}}_{=A-(-1)I_3} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \right\} \\
&= \left\{ \begin{pmatrix} -\frac{1}{2}\mu - \lambda \\ \mu \\ \lambda \end{pmatrix} \;\middle|\; \mu, \lambda \in \mathbb{Q} \right\} = \left\langle \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\rangle,
\end{aligned}$$

$$\begin{aligned}
\operatorname{Eig}(A,8) &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{Q}^3 \;\middle|\; \underbrace{\begin{pmatrix} -5 & 2 & 4 \\ 2 & -3 & 2 \\ 4 & 2 & -5 \end{pmatrix}}_{=A-8I_3} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \right\} \\
&= \left\{ \begin{pmatrix} \lambda \\ \frac{1}{2}\lambda \\ \lambda \end{pmatrix} \;\middle|\; \lambda \in \mathbb{Q} \right\} = \left\langle \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix} \right\rangle.
\end{aligned}$$

Bemerkung 7.9. Es kann passieren, dass eine lineare Abbildung keine Eigenwerte hat. (Siehe auch Beispiel 7.3(4).)

(1) Die Matrix

$$B = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$$

hat die Eigenwerte $\pm\sqrt{2} \in \mathbb{R}$. Aber die Matrix

$$B' = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{Q})$$

hat keine Eigenwerte in $\mathbb{Q}$, da $\pm\sqrt{2} \notin \mathbb{Q}$.

(2) Die Matrix

$$C = \begin{pmatrix} 0 & 4 \\ -1 & 0 \end{pmatrix} \in M_2(\mathbb{C})$$

hat die Eigenwerte $\pm 2i$. Aber die Matrix

$$C' = \begin{pmatrix} 0 & 4 \\ -1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$$

hat keine Eigenwerte in $\mathbb{R}$.

Lemma 7.10. *Sei V ein endlich-dimensionaler komplexer Vektorraum und sei $T\colon V \to V$ ein Endomorphismus. Dann hat T mindestens einen Eigenwert.*

Beweis. Folgt aus Proposition 7.4 und dem Fundamentalsatz der Algebra. □

Lemma 7.11. *Sei V ein endlich-dimensionaler K-Vektorraum und sei $T\colon V \to V$ eine lineare Abbildung. Seien $\lambda_1, \dots, \lambda_n$ verschiedene Eigenwerte von T und seien $v_1, \dots, v_n$ entsprechende Eigenvektoren. Dann sind $v_1, \dots, v_n$ linear unabhängig.*

Beweis. Der Beweis ist per Induktion nach n. Für $n = 1$ ist die Aussage offensichtlich. Sei nun $n \geq 2$.

Seien $\alpha_1, \dots, \alpha_n \in K$ Skalare mit

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0. \tag{34}$$

Wenn wir diese Gleichung mit λ_1 multiplizieren, bekommen wir

$$\lambda_1 \alpha_1 v_1 + \cdots + \lambda_1 \alpha_n v_n = 0. \tag{35}$$

Andererseits, wenn wir auf die Gleichung (34) die Abbildung T anwenden, bekommen wir

$$\begin{aligned} 0 = T(0) &= T(\alpha_1 v_1 + \cdots + \alpha_n v_n) \\ &= \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n) = \alpha_1 \lambda_1 v_1 + \cdots + \alpha_n \lambda_n v_n. \end{aligned}$$

Wenn wir jetzt von der letzten Gleichung die Gleichung (35) subtrahieren, bekommen wir

$$0 = \sum_{i=2}^{n} (\lambda_i \alpha_i - \lambda_1 \alpha_i) v_i = \sum_{i=2}^{n} \alpha_i (\lambda_i - \lambda_1) v_i.$$

Per Induktion folgt nun, dass für alle $2 \leq i \leq n$ gilt

$$\alpha_i (\lambda_i - \lambda_1) = 0.$$

Da $\lambda_1 \neq \lambda_i$ für alle $i \geq 2$, so folgt, dass $\alpha_i = 0$ für alle $i \geq 2$. Nun folgt aus (34), dass $\alpha_1 v_1 = 0$, und damit $\alpha_1 = 0$. Also, $v_1, \dots, v_n$ sind linear unabhängig. □

Korollar 7.12. *Sei V ein endlich-dimensionaler K-Vektorraum und seien $\lambda_1, \dots, \lambda_n$ verschiedene Eigenwerte einer linearen Abbildung $T\colon V \to V$. Sei weiterhin $E_i := \operatorname{Eig}(T, \lambda_i)$ für jedes i. Dann ist die Summe*

$$E_1 + \cdots + E_n$$

direkt.

Beweis. Folgt direkt aus Lemma 7.11. (Übung!) □

Diagonalisierung

Definition 7.13. Sei K ein Körper und sei V ein endlich-dimensionaler K-Vektorraum.

(1) Eine Matrix $M \in M_n(K)$ heißt *diagonal* oder *Diagonalmatrix*, falls es $m_{11}, \dots, m_{nn} \in K$ gibt, sodass

$$M = \begin{pmatrix} m_{11} & & 0 \\ & \ddots & \\ 0 & & m_{nn} \end{pmatrix}.$$

Eine solche Matrix bezeichnen wir auch mit

$$\operatorname{diag}(m_{11}, \dots, m_{nn}).$$

(2) Eine Matrix $A \in M_n(K)$ heißt *diagonalisierbar*, wenn es eine Diagonalmatrix in $M_n(K)$ gibt, die zu A ähnlich ist.

(3) Eine lineare Abbildung $T\colon V \to V$ heißt *diagonalisierbar*, falls es eine Basis von V gibt, sodass die Darstellungsmatrix von T bezüglich dieser Basis eine Diagonalmatrix ist.

Proposition 7.14. *Sei V ein endlich-dimensionaler K-Vektorraum. Ein Endomorphismus $T\colon V \to V$ ist diagonalisierbar genau dann, wenn es eine Basis von V gibt, die aus Eigenvektoren von T besteht.*

Beweis. Die Abbildung T ist diagonalisierbar genau dann, wenn es eine Basis $\mathcal{B} = \{v_1, \dots, v_n\}$ von V gibt, sodass die Darstellungsmatrix von T bezüglich $\mathcal{B}$ eine Diagonalmatrix

$$\operatorname{diag}(\lambda_{11}, \dots, \lambda_{nn})$$

ist. In anderen Worten, es gilt

$$T(v_i) = \lambda_{ii} v_i \quad \text{für alle } i = 1, \dots, n.$$

Aber das bedeutet genau, dass $\mathcal{B}$ aus Eigenvektoren von T besteht. □

Bemerkung 7.15. Eine Matrix $A \in M_n(K)$ ist diagonalisierbar genau dann, wenn die assoziierte lineare Abbildung

$$T\colon K^n \to K^n, \quad x \mapsto Ax$$

diagonalisierbar ist. (Übung: Basiswechsel!). Folglich ist nach Proposition 7.14 eine Matrix $A \in M_n(K)$ genau dann diagonalisierbar, wenn es eine Basis von K^n aus Eigenvektoren von A gibt.

Beispiel 7.16. Aus Beispiel 7.8: die Matrix $A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix} \in M_3(\mathbb{Q})$ hat die Eigenwerte -1 und 8, und

$$\operatorname{Eig}(A, -1) = \left\langle \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\rangle, \quad \operatorname{Eig}(A, 8) = \left\langle \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix} \right\rangle.$$

Man sieht einfach, dass die Vektoren $\begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix}$ linear unabhängig sind, also

$$\mathcal{B} := \left\{ \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix} \right\}$$

ist eine Basis von $\mathbb{Q}^3$. Damit ist die Matrix A nach Proposition 7.14 diagonalisierbar. Sei

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

die kanonische Basis von $\mathbb{Q}^3$. Die assoziierte Abbildung

$$T\colon \mathbb{Q}^3 \to \mathbb{Q}^3, \quad x \mapsto Ax$$

hat die Darstellungsmatrix A bezüglich der Basis $\mathcal{A}$. Die Darstellungsmatrix von T bezüglich der Basis $\mathcal{B}$ ist die Matrix

$$\operatorname{diag}(-1, -1, 8).$$

Nach Korollar 6.24 gilt:

$$T_{\mathcal{B}}^{\mathcal{A}} \cdot A \cdot (T_{\mathcal{B}}^{\mathcal{A}})^{-1} = \operatorname{diag}(-1, -1, 8).$$

Korollar 7.17. *Sei V ein endlich-dimensionaler K-Vektorraum und seien $\lambda_1, \ldots, \lambda_n$ alle verschiedenen Eigenwerte einer linearen Abbildung $T\colon V \to V$. Dann gilt*

$$\sum_{i=1}^{n} \dim \operatorname{Eig}(T, \lambda_i) \leq \dim V,$$

mit Gleichheit genau dann, wenn T diagonalisierbar ist.

Beweis. Setze $E_i := \operatorname{Eig}(T, \lambda_i)$. Dann ist, nach Korollar 7.12, der Vektorraum $\bigoplus_{i=1}^{n} E_i$ ein Unterraum von V. Damit ist nach Bemerkung 4.43:

$$\sum_{i=1}^{n} \dim E_i = \dim \left(\bigoplus_{i=1}^{n} E_i \right) \leq \dim V.$$

Die Gleichheit gilt genau dann, wenn $\bigoplus_{i=1}^{n} E_i = V$, oder äquivalent, wenn die Vereinigung von Basen von E_i eine Basis von V bildet. (Übung!) Nach Proposition 7.14 ist das genau dann der Fall, wenn T diagonalisierbar ist. $\square$

Beispiel 7.18.

(1) Beispiel 7.16 und Korollar 7.17 zeigen erneut, dass die Matrix

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix} \in M_3(\mathbb{Q})$$

diagonalisierbar ist.

(2) Die reelle Matrix

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in M_3(\mathbb{R})$$

hat die Eigenwerte 1 und 2, und die Eigenräume

$$\operatorname{Eig}(A, 1) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle, \quad \operatorname{Eig}(A, 2) = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle.$$

Da $\dim \operatorname{Eig}(A, 1) = \dim \operatorname{Eig}(A, 2) = 1$ und somit

$$\dim \operatorname{Eig}(A, 1) + \dim \operatorname{Eig}(A, 2) = 2 < 3 = \dim \mathbb{R}^3,$$

so ist A nach Korollar 7.17 nicht diagonalisierbar.

Definition 7.19. Sei V ein endlich-dimensionaler K-Vektorraum und sei $\lambda \in K$ ein Eigenwert eines Endomorphismus $T\colon V \to V$. Die *algebraische Vielfachheit* von λ ist

$$\mu(\chi_T, \lambda);$$

also die Vielfachheit von λ als Nullstelle des charakteristischen Polynoms χ_T.

Proposition 7.20. *Sei V ein endlich-dimensionaler K-Vektorraum, sei $\lambda \in K$ ein Eigenwert eines Endomorphismus $T\colon V \to V$, seien a bzw. g die algebraische bzw. geometrische Vielfachheit von λ. Dann gilt*

$$g \leq a.$$

Beweis. Per Definition ist

$$g = \dim \operatorname{Eig}(T, \lambda).$$

Sei nun $\{v_1, \ldots, v_g\}$ eine Basis vom Eigenraum $\operatorname{Eig}(T, \lambda)$. Wir ergänzen diese Basis zu

$$\text{einer Basis } \mathcal{B} := \{v_1, \ldots, v_n\} \text{ von } V.$$

Die Darstellungsmatrix von T bezüglich $\mathcal{B}$ hat dann die folgende Form:

$$A = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & B \\ 0 & & \lambda & \\ \hline & 0 & & C \end{array}\right),$$

wobei $B \in M_{g,n-g}(K)$ und $C \in M_{n-g}(K)$. Dann gilt:

$$\begin{aligned} \chi_A(t) = \det(A - tI_n) &= \det \left(\begin{array}{ccc|c} \lambda - t & & 0 & \\ & \ddots & & B \\ 0 & & \lambda - t & \\ \hline & 0 & & C - tI_{n-g} \end{array}\right) \\ &= \det \begin{pmatrix} \lambda - t & & 0 \\ & \ddots & \\ 0 & & \lambda - t \end{pmatrix} \cdot \det(C - tI_{n-g}) = (-1)^g (t - \lambda)^g \chi_C(t). \end{aligned}$$

Aus dieser Formel folgt sofort, dass $a = \mu(\chi_T, \lambda) \geq g$. □

Korollar 7.21. *Sei V ein endlich-dimensionaler K-Vektorraum und seien $\lambda_1, \ldots, \lambda_n$ verschiedene Eigenwerte einer linearen Abbildung $T\colon V \to V$. Dann ist T genau dann diagonalisierbar, wenn:*

(a) das charakteristische Polynom χ_T zerfällt in lineare Faktoren, und

(b) für jedes i gilt $\mu(\chi_T, \lambda_i) = \dim \operatorname{Eig}(T, \lambda_i)$.

Beweis. Setze $E_i := \operatorname{Eig}(T, \lambda_i)$ für jedes i. Nach Korollar 7.17 ist T diagonalisierbar genau dann, wenn

$$\sum_{i=1}^{n} \dim E_i = \dim V. \tag{36}$$

Andererseits: Proposition 7.20 und Lemma 3.43(a) ergeben

$$\sum_{i=1}^{n} \dim E_i \leq \sum_{i=1}^{n} \mu(\chi_T, \lambda_i) \leq \deg \chi_T = \dim V.$$

Dann ist (36) äquivalent zur Aussage, dass diese beide Ungleichungen tatsächlich Gleichungen sind. Die Gleichung

$$\sum_{i=1}^{n} \dim E_i = \sum_{i=1}^{n} \mu(\chi_T, \lambda_i)$$

ist äquivalent zu (b), und die Gleichung

$$\sum_{i=1}^{n} \mu(\chi_T, \lambda_i) = \deg \chi_T$$

ist äquivalent zu (a) nach Lemma 3.43(b). □

Beispiel 7.22. Wir betrachten wieder die folgende Matrix aus Beispielen 7.8 und 7.16:

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}.$$

Dort haben wir berechnet, dass

$$\chi_A(x) = -(x+1)^2(x-8),$$

und damit ist

$$\mu(\chi_A, -1) = 2 \quad \text{und} \quad \mu(\chi_A, 8) = 1.$$

Wir haben auch berechnet, dass

$$\dim \operatorname{Eig}(A, -1) = 2 \quad \text{und} \quad \dim \operatorname{Eig}(A, 8) = 1.$$

Es folgt aus Korollar 7.21, dass A diagonalisierbar ist, da das Polynom χ_A in lineare Faktoren zerfällt, und die geometrische und algebraische Vielfachheit aller Eigenwerte übereinstimmen.

8 Skalarprodukte

Definition 8.1. Sei V ein reeller bzw. komplexer Vektorraum. Ein *Skalarprodukt* ist eine Abbildung

$$\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}, \quad (x, y) \mapsto \langle x, y \rangle$$

bzw.

$$\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}, \quad (x, y) \mapsto \langle x, y \rangle$$

mit den folgenden Eigenschaften:

(S1) $\langle \cdot\,, \cdot \rangle$ ist linear im ersten Argument, d.h.

$$\langle \alpha_1 x_1 + \alpha_2 x_2, y \rangle = \alpha_1 \langle x_1, y \rangle + \alpha_2 \langle x_2, y \rangle;$$

(S2) für alle $x, y \in V$ gilt

$$\langle x, y \rangle = \overline{\langle y, x \rangle};$$

insbesondere ist im reellen Fall:

$$\langle x, y \rangle = \langle y, x \rangle;$$

(S3) für alle $x \in V$ gilt $\langle x, x \rangle \in \mathbb{R}$ und

$$\langle x, x \rangle \geq 0;$$

ferner gilt:

$$\langle x, x \rangle = 0 \quad \Longleftrightarrow \quad x = 0.$$

Ein Vektorraum, der mit einem solchen reellen oder komplexen Skalarprodukt versehen ist, heißt *Skalarproduktraum* und wird oft mit

$$\big(V, \langle \cdot\,, \cdot \rangle\big)$$

bezeichnet. Ist V reell, so ist V ein *euklidischer Raum*; ist V komplex, so ist V ein *unitärer Raum*.

Bemerkung 8.2. Sei $\big(V, \langle \cdot\,, \cdot \rangle\big)$ ein Skalarproduktraum.

(1) Ist V ein $\mathbb{R}$-Vektorraum, so ist $\langle \cdot\,, \cdot \rangle$ linear auch im zweiten Argument:

$$\begin{aligned} \langle y, \alpha_1 x_1 + \alpha_2 x_2 \rangle &\overset{(S2)}{=} \langle \alpha_1 x_1 + \alpha_2 x_2, y \rangle \\ &\overset{(S1)}{=} \alpha_1 \langle x_1, y \rangle + \alpha_2 \langle x_2, y \rangle \overset{(S2)}{=} \alpha_1 \langle y, x_1 \rangle + \alpha_2 \langle y, x_2 \rangle. \end{aligned}$$

In diesem Fall sagen wir, dass das Skalarprodukt *bilinear* ist.

(2) Ist V ein $\mathbb{C}$-Vektorraum, so gilt

$$\begin{aligned}\langle y, \alpha_1 x_1 + \alpha_2 x_2\rangle &\overset{(S2)}{=} \overline{\langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle} \\ &\overset{(S1)}{=} \overline{\alpha_1\langle x_1, y\rangle + \alpha_2\langle x_2, y\rangle} \\ &= \overline{\alpha_1}\overline{\langle x_1, y\rangle} + \overline{\alpha_2}\overline{\langle x_2, y\rangle} \overset{(S2)}{=} \overline{\alpha_1}\langle y, x_1\rangle + \overline{\alpha_2}\langle y, x_2\rangle.\end{aligned}$$

In diesem Fall sagen wir, dass das Skalarprodukt *sesquilinear* ist.

(3) Für jedes $x \in V$ gilt:

$$\langle 0, x\rangle = \langle 0 + 0, x\rangle = \langle 0, x\rangle + \langle 0, x\rangle,$$

und damit $\langle 0, x\rangle = 0$.

(4) Jeder Unterraum eines Skalarproduktraums ist auch ein Skalarproduktraum.

Beispiel 8.3.

(1) Betrachten wir den $\mathbb{R}$-Vektorraum $\mathbb{R}^n$ und seien $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ und $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Die Formel

$$\langle x, y\rangle := \sum_{i=1}^{n} x_i y_i$$

definiert das *kanonische Skalarprodukt* auf $\mathbb{R}^n$.

(2) Betrachten wir den $\mathbb{C}$-Vektorraum $\mathbb{C}^n$ und seien $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$ und $y = (y_1, \ldots, y_n) \in \mathbb{C}^n$. Die Formel

$$\langle x, y\rangle := \sum_{i=1}^{n} x_i \overline{y_i}$$

definiert das *kanonische Skalarprodukt* auf $\mathbb{C}^n$.

Definition 8.4. Sei V ein $\mathbb{R}$- oder $\mathbb{C}$-Vektorraum. Eine *Norm auf* V ist eine Abbildung

$$\|\cdot\| \colon V \to \mathbb{R}, \quad x \mapsto \|x\|$$

mit den folgenden Eigenschaften:

(N1) für alle $x \in V$ gilt $\|x\| \geq 0$ und

$$\|x\| = 0 \quad \Longleftrightarrow \quad x = 0;$$

(N2) $\|\alpha x\| = |\alpha| \cdot \|x\|$ für jeden Skalar α und jedes $x \in V$;

(N3) es gilt die *Dreiecksungleichung*:

$$\|x + y\| \leq \|x\| + \|y\| \quad \text{für alle } x, y \in V.$$

Ein mit einer Norm versehener Vektorraum V heißt *normiert* und wird oft mit

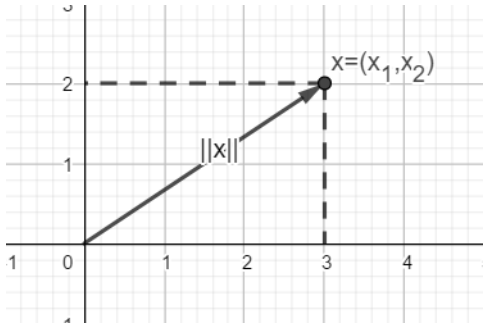
$$(V, \|\cdot\|)$$

bezeichnet.

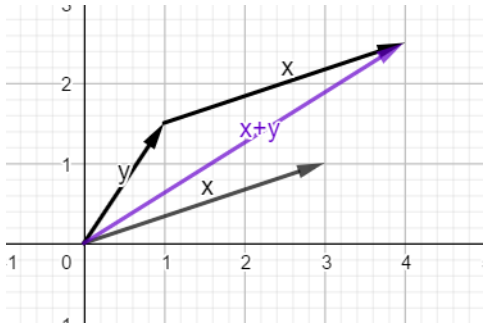
Beispiel 8.5. Die Abbildung

$$\|\cdot\| \colon \mathbb{R}^n \to \mathbb{R}, \quad x = (x_1, \ldots, x_n) \mapsto \sqrt{x_1^2 + \cdots + x_n^2} = \sqrt{\langle x, x \rangle}$$

ist die *kanonische Norm* auf $\mathbb{R}^n$. Die Norm $\|\cdot\|$ misst die Länge eines Vektors.



Die Dreieckungleichung lässt sich im Dreieck visualisieren:



Die Cauchy-Schwarz-Ungleichung

Wir werden zeigen, dass die Formel

$$\|x\| := \sqrt{\langle x, x\rangle}$$

eine Norm auf einem Skalarproduktraum $\big(V, \langle\cdot\,,\cdot\rangle\big)$ definiert. Das Einzige, was hierbei schwierig zu zeigen ist, ist die Dreiecksungleichung. Das folgt aus dem nächsten wichtigen Satz.

Satz 8.6 (Cauchy-Schwarz-Ungleichung)**.** *Sei* $\big(V, \langle\cdot\,,\cdot\rangle\big)$ *ein Skalarproduktraum. Für jedes* $x \in V$ *setze*

$$\|x\| := \sqrt{\langle x, x\rangle}.$$

Dann gilt für alle $x, y \in V$*:*

$$|\langle x, y\rangle| \leq \|x\|\|y\|,$$

mit der Gleichung genau dann, wenn x *und* y *linear abhängig sind. (Wir sagen auch, dass* x *und* y proportional *sind).*

Beweis. Wenn $y = 0$, dann ist die Aussage trivial. Wir dürfen also annehmen, dass $y \neq 0$. Setze

$$\lambda := \frac{\langle x, y\rangle}{\|y\|^2}.$$

Dann gilt:

$$\begin{aligned}
0 &\leq \langle x - \lambda y, x - \lambda y\rangle = \langle x, x\rangle - \langle x, \lambda y\rangle - \langle \lambda y, x\rangle + \langle \lambda y, \lambda y\rangle \\
&= \langle x, x\rangle - \overline{\lambda}\langle x, y\rangle - \lambda\langle x, y\rangle + \lambda\overline{\lambda}\langle y, y\rangle \\
&= \|x\|^2 - \frac{\overline{\langle x, y\rangle}}{\|y\|^2}\langle x, y\rangle - \frac{\langle x, y\rangle}{\|y\|^2}\overline{\langle x, y\rangle} + \frac{\langle x, y\rangle\overline{\langle x, y\rangle}}{\|y\|^2\|y\|^2}\|y\|^2 \\
&= \|x\|^2 - \frac{\langle x, y\rangle\overline{\langle x, y\rangle}}{\|y\|^2} = \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2},
\end{aligned}$$

und damit

$$|\langle x, y\rangle|^2 \leq \|x\|^2\|y\|^2.$$

Die Gleichheit gilt genau dann, wenn

$$\langle x - \lambda y, x - \lambda y\rangle = 0,$$

oder äquivalent:

$$x - \lambda y = 0.$$

Also, die Gleichung impliziert, dass x und y linear abhängig sind. Die Umgekehrte Richtung ist einfach zu zeigen. $\square$

Korollar 8.7. *Sei* $\big(V, \langle \cdot\,, \cdot \rangle\big)$ *ein Skalarproduktraum. Für jedes* $x \in V$ *setze*

$$\|x\| := \sqrt{\langle x, x \rangle}.$$

Dann definiert $\|\cdot\| \colon V \to \mathbb{R}$ *eine Norm auf* V.

Beweis. Die Eigenschaften (N1) und (N2) der Definition der Norm folgen sofort. Wir müssen (N3) zeigen: also, dass

$$\|x + y\| \leq \|x\| + \|y\| \quad \text{für alle } x, y \in V.$$

Diese Ungleichung ist äquivalent zu

$$\|x + y\|^2 \leq (\|x\| + \|y\|)^2,$$

und wenn wir die beiden Seiten entwickeln, bekommen wir, dass die letzte Ungleichung äquivalent ist zu:

$$\|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2,$$

oder äquivalent:

$$\langle x, y \rangle + \langle y, x \rangle \leq 2\|x\|\|y\|.$$

Aber dies folgt aus der folgender Rechnung:

$$\langle x, y \rangle + \langle y, x \rangle = \langle x, y \rangle + \overline{\langle x, y \rangle} = 2\,\mathrm{Re}\langle x, y \rangle \leq 2|\langle x, y \rangle| \leq 2\|x\|\|y\|,$$

wobei die letzte Ungleichung die Cauchy-Schwarz-Ungleichung ist. □

Bemerkung 8.8. Falls die Norm $\|\cdot\|$ eines normierten Vektorraums $\big(V, \|\cdot\|\big)$ von einem Skalarprodukt $\langle \cdot\,, \cdot \rangle$ induziert wird, d.h.

$$\|x\| = \sqrt{\langle x, x \rangle} \quad \text{für alle } x \in V,$$

dann erfüllt $\|\cdot\|$ die *Parallelogrammidentität*

$$\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right),$$

denn die folgenden Gleichungen gelten:

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2,$$
$$\|x - y\|^2 = \langle x - y, x - y \rangle = \|x\|^2 - \langle x, y \rangle - \langle y, x \rangle + \|y\|^2.$$

Beispiel 8.9. Die *Maximumsnorm* auf $\mathbb{R}^n$ ist wie folgt definiert:

$$\text{für alle } x = (x_1, \dots, x_n) \in \mathbb{R}^n \quad \text{setze} \quad \|x\|_\infty = \max_{i=1}^{n} \{|x_i|\}.$$

Sei nun $n \geq 2$, und seien $e_1 = (1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1)$. Dann gilt:

$$\begin{aligned} &\|e_1\|_\infty = 1, \qquad \|e_n\|_\infty = 1, \\ &\|e_1 + e_n\|_\infty = \|(1, 0, \dots, 0, 1)\|_\infty = 1, \\ &\|e_1 - e_n\|_\infty = \|(1, 0, \dots, 0, -1)\|_\infty = 1. \end{aligned}$$

Damit ist die Parallelogrammidentität nicht erfüllt, und deswegen ist, nach vorigem Beispiel, diese Norm nicht von einem Skalarprodukt induziert.

Lemma 8.10. *Eine Norm auf einem normierten Raum $\big(V, \|\cdot\|\big)$ ist von einem Skalarprodukt $\langle \cdot\,, \cdot \rangle$ genau dann induziert, wenn die Parallelogrammidentität erfüllt ist.*

Beweis.
$\Longrightarrow$: Folgt aus Bemerkung 8.8.

$\Longleftarrow$: Falls V ein reeller Vektorraum ist, setze

$$\langle x, y \rangle := \frac{1}{4} \left(\|x + y\|^2 - \|x - y\|^2 \right);$$

falls V ein komplexer Vektorraum ist, setze

$$\langle x, y \rangle := \frac{1}{4} \left(\|x + y\|^2 - \|x - y\|^2 \right) + \frac{i}{4} \left(\|x + iy\|^2 - \|x - iy\|^2 \right).$$

Die Behauptung folgt durch Nachrechnen. $\square$

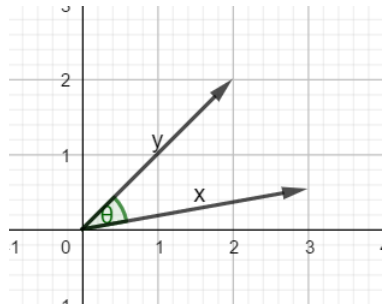
Orthogonalität und Orthonormalität

Bemerkung 8.11. Sei $\big(V, \langle \cdot\,, \cdot \rangle\big)$ ein Skalarproduktraum. Für alle $x, y \in V \setminus \{0\}$ folgt aus der Cauchy-Schwarz-Ungleichung:

$$-1 \leq \frac{\langle x, y \rangle}{\|x\| \|y\|} \leq 1.$$

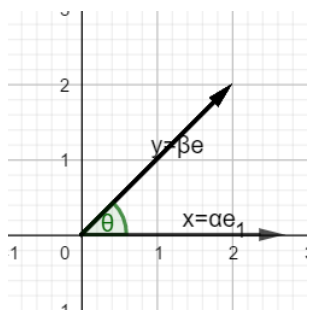
Wir definieren den *Winkel* θ zwischen x und y als

$$\theta = \arccos \frac{\langle x, y \rangle}{\|x\| \|y\|}.$$



Die Begründung: Angenommen, $V = \mathbb{R}^2$. Nach Rotation und Spiegelung können wir annehmen, dass

$$x = \alpha e_1 \quad \text{und} \quad y = \beta e, \text{ wobei } \|e\| = 1 \text{ und } \alpha, \beta \in \mathbb{R}.$$



Dann sehen wir, dass

$$\frac{\langle x, y\rangle}{\|x\|\|y\|} = \frac{\alpha\beta\langle e_1, e\rangle}{\alpha \underbrace{\|e_1\|}_{=1} \beta \underbrace{\|e\|}_{=1}} = \langle e_1, e\rangle = \langle (1,0), (\cos\theta, \sin\theta)\rangle = \cos\theta.$$

Definition 8.12. Sei $\big(V, \langle\cdot\,,\cdot\rangle\big)$ ein Skalarproduktraum.

(1) Zwei Vektoren $x, y \in V$ heißen *orthogonal*, und wir schreiben $x \perp y$, falls

$$\langle x, y\rangle = 0.$$

(2) Eine Teilmenge $U \subseteq V$ heißt *orthogonal*, falls

$$\langle u, v\rangle = 0 \quad \text{für alle } u, v \in U,\ u \neq v.$$

(3) Eine orthogonale Teilmenge $U \subseteq V$ heißt *orthonormal*, falls

$$\|u\| = 1 \quad \text{für alle } u \in U.$$

Bemerkung 8.13. Sei $\big(V, \langle\cdot\,,\cdot\rangle\big)$ ein Skalarproduktraum.

(1) Der Vektor $0 \in V$ ist orthogonal zu jedem anderen Vektor in V.

(2) Seien $x, y \in V \setminus \{0\}$, und sei θ der Winkel zwischen x und y. Aus der Formel
$$\theta = \arccos \frac{\langle x, y\rangle}{\|x\|\|y\|}$$
folgt, dass x und y genau dann orthogonal sind, wenn
$$\theta = \frac{\pi}{2}.$$

(3) Ist U eine orthogonale Teilmenge von V mit $0 \notin U$, so ist die Menge
$$\left\{ \frac{u}{\|u\|} \;\middle|\; u \in U \right\}$$
orthonormal.

(4) In einer orthonormalen Teilmenge $U = \{u_i \mid i \in I\}$ von V gilt
$$\langle u_i, u_j\rangle = \delta_{ij} \quad \text{für alle } u_i, u_j \in U.$$

Proposition 8.14. *Sei $\big(V, \langle\cdot\,,\cdot\rangle\big)$ ein Skalarproduktraum und sei U eine orthogonale Teilmenge von V mit $0 \notin U$. Dann ist U linear unabhängig.*

Beweis. Seien $u_1, \ldots, u_n \in U$ und Skalare $\alpha_1, \ldots, \alpha_n$ mit
$$\sum_{i=1}^{n} \alpha_i u_i = 0.$$
Daraus folgt für jedes j:
$$0 = \langle 0, u_j\rangle = \left\langle \sum_{i=1}^{n} \alpha_i u_i, u_j \right\rangle = \sum_{i=1}^{n} \alpha_i \langle u_i, u_j\rangle \underset{\text{wenn } i\neq j}{\overset{u_i \perp u_j}{=}} \alpha_j \langle u_j, u_j\rangle = \alpha_j \|u_j\|^2,$$
und damit $\alpha_j = 0$ für alle j. Deswegen sind $u_1, \ldots, u_n$ linear unabhängig. $\square$

Definition 8.15. Sei $\mathcal{B}$ eine Basis eines endlich-dimensionalen Skalarproduktraums $\big(V, \langle\cdot\,,\cdot\rangle\big)$. Ist $\mathcal{B}$ eine orthonormale Menge, so heißt $\mathcal{B}$ *Orthonormalbasis von V*.

Beispiel 8.16. Die kanonische Basis $\{e_1, \ldots, e_n\}$ von $\mathbb{R}^n$ ist eine Orthonormalbasis bezüglich des kanonischen Skalarprodukts.

Proposition 8.17. *Sei $\{v_1, \dots, v_n\}$ eine Orthonormalbasis eines endlich-dimensionalen Skalarproduktraums $\big(V, \langle \cdot\,, \cdot \rangle\big)$. Dann sind die Koeffizienten in der Darstellung*

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

eines Vektors $v \in V$ durch die Formel

$$\alpha_i = \langle v, v_i \rangle \quad \textit{für jedes } i$$

gegeben.

Beweis. Wie der Beweis von Proposition 8.14. □

Satz 8.18 (Gram-Schmidt-Verfahren)**.** *Jeder endlich-dimensionale Skalarproduktraum $\big(V, \langle \cdot\,, \cdot \rangle\big)$ hat eine Orthonormalbasis.*

Beweis. Fixiere eine beliebige Basis $\{v_1, \dots, v_n\}$ von V. Setze

$$f_1 := v_1 \quad \text{und} \quad e_1 := \frac{f_1}{\|f_1\|},$$

und definiere induktiv für alle $i \geq 2$:

$$\begin{aligned} f_i &:= v_i - \langle v_i, e_1 \rangle e_1 - \langle v_i, e_2 \rangle e_2 - \cdots - \langle v_i, e_{i-1} \rangle e_{i-1}, \\ e_i &:= \frac{f_i}{\|f_i\|}. \end{aligned}$$

Wir behaupten, dass $\{e_1, \dots, e_n\}$ eine Orthonormalbasis von V ist.

Wir zeigen dies, indem wir per Induktion nach $k = 1, \dots, n$ die folgenden drei Aussagen beweisen:

(i) $f_k \neq 0$,

(ii) $\langle e_1, \dots, e_k \rangle = \langle v_1, \dots, v_k \rangle$,

(iii) $\langle e_i, e_j \rangle = \delta_{ij}$ für alle $1 \leq i, j \leq k$.

Für $k = n$ bekommen wie die Behauptung oben.

Der Fall $k = 1$ ist trivial, und damit dürfen wir annehmen, dass $k \geq 2$.

Wir zeigen zuerst (i). Offensichtlich ist f_{k+1} eine lineare Kombination von $e_1, \dots, e_{k-1}, v_k$, und v_k ist eine lineare Kombination von $e_1, \dots, e_{k-1}, f_k$. Damit ist

$$\langle e_1, \dots, e_{k-1}, f_k \rangle = \langle e_1, \dots, e_{k-1}, v_k \rangle. \tag{37}$$

Da

$$\langle e_1, \dots, e_{k-1} \rangle = \langle v_1, \dots, v_{k-1} \rangle$$

per Induktionsschritt für (ii), so gilt

$$\langle e_1, \dots, e_{k-1}, f_k \rangle = \langle v_1, \dots, v_{k-1}, f_k \rangle$$

und

$$\langle e_1, \dots, e_{k-1}, v_k \rangle = \langle v_1, \dots, v_{k-1}, v_k \rangle. \tag{38}$$

Zusammen mit (37) ergibt sich:

$$\langle v_1, \dots, v_{k-1}, f_k \rangle = \langle v_1, \dots, v_{k-1}, v_k \rangle.$$

Aber $\{v_1, \dots, v_k\}$ ist eine minimale erzeugende Menge von $\langle v_1, \dots, v_k \rangle$, und damit $f_k \neq 0$, welches (i) zeigt.

Für (ii): Per Definition von e_k gilt

$$\langle e_1, \dots, e_{k-1}, e_k \rangle = \langle e_1, \dots, e_{k-1}, f_k \rangle,$$

und dies ist gleich $\langle v_1, \dots, v_k \rangle$ nach (37) und (38).

Für (iii): Da $\langle e_i, e_j \rangle = \delta_{ij}$ für alle $1 \leq i, j \leq k-1$ per Induktionsschritt für (iii), so müssen wir zeigen, dass

$$\langle e_i, e_k \rangle = 0 \text{ für } i \leq k-1, \quad \text{und} \quad \langle e_k, e_k \rangle = 1.$$

Die zweite Gleichung gilt nach der Definition von e_k. Für die erste Gleichung: Es gilt

$$\begin{aligned}
\langle e_i, e_k \rangle &= \left\langle e_i, \frac{f_k}{\|f_k\|} \right\rangle = \frac{1}{\|f_k\|} \langle e_i, f_k \rangle = \frac{1}{\|f_k\|} \left\langle e_i, v_k - \sum_{j=1}^{k-1} \langle v_k, e_j \rangle e_j \right\rangle \\
&= \frac{1}{\|f_k\|} \left(\langle e_i, v_k \rangle - \sum_{j=1}^{k-1} \overline{\langle v_k, e_j \rangle} \underbrace{\langle e_i, e_j \rangle}_{=\delta_{ij}} \right) \\
&= \frac{1}{\|f_k\|} \left(\langle e_i, v_k \rangle - \overline{\langle v_k, e_i \rangle} \right) = 0,
\end{aligned}$$

und der Beweis ist komplett. □

Bemerkung 8.19. Ist im obigen Beweis $\{v_1, \dots, v_n\}$ bereits eine Orthonormalbasis von V, dann ergibt das Gram-Schmidt-Verfahren, dass

$$e_i = v_i \quad \text{für alle } i = 1, \dots, n.$$

Korollar 8.20 (Basisergänzung)**.** *Sei $\big(V, \langle \cdot\,, \cdot \rangle\big)$ ein n-dimensionaler Skalarproduktraum. Sei U ein Unterraum von V und $\{u_1, \ldots, u_m\}$ eine Orthonormalbasis von U. Dann gibt es $u_{m+1}, \ldots, u_n \in V$, sodass $\{u_1, \ldots, u_n\}$ eine Orthonormalbasis von V ist.*

Beweis. Wir ergänzen $\{u_1, \ldots, u_m\}$ zu einer beliebigen Basis

$$\{u_1, \ldots, u_m, v_{m+1}, \ldots, v_n\} \text{ von } V.$$

Dann wenden wir das Gram-Schmidt-Verfahren auf diese Basis an und nutzen Bemerkung 8.19. □

Beispiel 8.21. Sei $U = \langle v_1, v_2, v_3 \rangle \subseteq \mathbb{R}^4$, wobei

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 3 \\ 1 \\ 1 \\ 3 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 2 \\ -2 \\ 4 \\ 0 \end{pmatrix}.$$

Wir werden eine Orthonormalbasis für U finden und diese zu einer Orthonormalbasis von $\mathbb{R}^4$ ergänzen.

Man prüft zuerst, dass v_1, v_2, v_3 linear unabhängig sind. Dann wenden wir das Gram-Schmidt-Verfahren an: es gilt

$$f_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{und} \quad e_1 = \frac{f_1}{\|f_1\|} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Es folgt:

$$\begin{aligned} f_2 &= v_2 - \langle v_2, e_1 \rangle e_1 \\ &= \begin{pmatrix} 3 \\ 1 \\ 1 \\ 3 \end{pmatrix} - \left\langle \begin{pmatrix} 3 \\ 1 \\ 1 \\ 3 \end{pmatrix}, \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle \cdot \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \\ 3 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \end{aligned}$$

und damit

$$e_2 = \frac{f_2}{\|f_2\|} = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}.$$

Es ist weiterhin:

$$
\begin{aligned}
f_3 &= v_3 - \langle v_3, e_1\rangle e_1 - \langle v_3, e_2\rangle e_2 \\
&= \begin{pmatrix} 2 \\ -2 \\ 4 \\ 0 \end{pmatrix} - \left\langle \begin{pmatrix} 2 \\ -2 \\ 4 \\ 0 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle \cdot \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \\
&- \left\langle \begin{pmatrix} 2 \\ -2 \\ 4 \\ 0 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \right\rangle \cdot \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 4 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - 0 = \begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix}
\end{aligned}
$$

und somit

$$
e_3 = \frac{f_3}{\|f_3\|} = \frac{1}{2\sqrt{5}}\begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix}.
$$

Also lautet unsere gefundene Orthonormalbasis für U:

$$
\mathcal{B}_U = \left\{ \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}, \frac{1}{2\sqrt{5}}\begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix} \right\}.
$$

Wähle nun zum Beispiel

$$
v_4 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \notin U.
$$

Damit ist $\langle v_1, v_2, v_3, v_4\rangle = \mathbb{R}^4$, und Gram-Schmidt ergibt:

$$
\begin{aligned}
f_4 &= v_4 - \langle v_4, e_1\rangle e_1 - \langle v_4, e_2\rangle e_2 - \langle v_4, e_3\rangle e_3 \\
&= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle \cdot \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \right\rangle \cdot \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \\
&- \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{2\sqrt{5}}\begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix} \right\rangle \cdot \frac{1}{2\sqrt{5}}\begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix} = \frac{1}{20}\begin{pmatrix} 9 \\ 3 \\ -3 \\ -9 \end{pmatrix}
\end{aligned}
$$

und

$$e_4 = \frac{f_4}{\|f_4\|} = \frac{1}{6\sqrt{5}} \begin{pmatrix} 9 \\ 3 \\ -3 \\ -9 \end{pmatrix}.$$

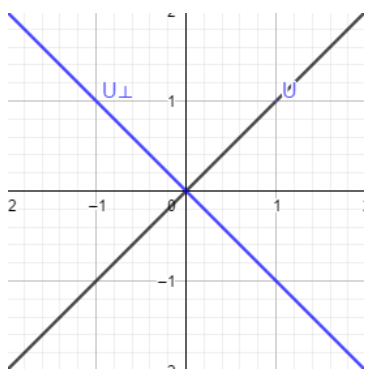
Also lautet unsere gefundene Orthonormalbasis für $\mathbb{R}^4$

$$\mathcal{B}_{\mathbb{R}^4} = \left\{ \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}, \frac{1}{2\sqrt{5}} \begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix}, \frac{1}{6\sqrt{5}} \begin{pmatrix} 9 \\ 3 \\ -3 \\ -9 \end{pmatrix} \right\}.$$

Definition 8.22. Sei U eine Teilmenge eines Skalarproduktraums $\big(V, \langle \cdot\,, \cdot \rangle\big)$. Das *orthogonale Komplement von* U ist die Menge

$$U^\perp := \{v \in V \mid \langle v, u \rangle = 0 \text{ für alle } u \in U\} = \{v \in V \mid v \perp u \text{ für alle } u \in U\}.$$

Beispiel 8.23. Das orthogonale Komplement einer Gerade in $\mathbb{R}^2$:



Bemerkung 8.24. Sei $\big(V, \langle \cdot\,, \cdot \rangle\big)$ ein Skalarproduktraum und sei U eine Teilmenge von V. Dann gilt (Übung!):

(1) $U^\perp$ ist ein Unterraum von V,

(2) wenn U ein Unterraum von V ist, dann $(U^\perp)^\perp = U$,

(3) $V^\perp = \{0\}$ und $\{0\}^\perp = V$.

Bemerkung 8.25. Sei $\big(V, \langle \cdot\,, \cdot \rangle\big)$ ein Skalarproduktraum und sei U ein endlich-dimensionaler Unterraum von V. Sei $\{u_1, \ldots, u_n\}$ eine Orthonormalbasis von U. Dann gehört ein Vektor $v \in V$ dem orthogonalen Komplement $U^\perp$ genau dann, wenn $v \perp u_i$ für alle $i = 1, \ldots, n$. Die Implikation $\Longrightarrow$ ist

trivial; wir zeigen nun die Implikation $\Longleftarrow$. Und zwar, sei $u \in U$. Dann gibt es Skalare $\alpha_1, \ldots, \alpha_n$, sodass

$$u = \sum_{i=1}^{n} \alpha_i v_i,$$

und damit:

$$\langle v, u \rangle = \left\langle v, \sum_{i=1}^{n} \alpha_i v_i \right\rangle = \sum_{i=1}^{n} \overline{\alpha_i} \langle v, v_i \rangle = 0,$$

also $v \in U^\perp$.

Proposition 8.26. *Sei U ein endlich-dimensionaler Unterraum eines Skalarproduktraums $\big(V, \langle \cdot\,, \cdot \rangle\big)$. Dann gilt*

$$V = U \oplus U^\perp.$$

Beweis. Wir zeigen zuerst, dass

$$V = U + U^\perp. \tag{39}$$

Sei $\{u_1, \ldots, u_n\}$ eine Orthonormalbasis von U. Für jedes $v \in V$ definiere

$$u_v := \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_n \rangle u_n \in U, \quad w_v := v - u_v.$$

Um (39) zu zeigen, reicht es zu zeigen, dass $w_v \in U^\perp$. Nach Bemerkung 8.25 reicht es dann zu zeigen, dass $w_v \perp u_j$ für alle j. Und zwar:

$$\begin{aligned} \langle w_v, u_j \rangle &= \langle v - \langle v, u_1 \rangle u_1 - \cdots - \langle v, u_n \rangle u_n, u_j \rangle \\ &= \langle v, u_j \rangle - \sum_{i=1}^{n} \langle v, u_i \rangle \underbrace{\langle u_i, u_j \rangle}_{=\delta_{ij}} = \langle v, u_j \rangle - \langle v, u_j \rangle = 0. \end{aligned}$$

Dies zeigt (39).

Es bleibt zu zeigen, dass die Summe direkt ist. Nach Lemma 4.42 reicht es zu zeigen, dass $U \cap U^\perp = \{0\}$. Und zwar, sei $v \in U \cap U^\perp$. Dann ist $v \perp u$ für alle $u \in U$. Insbesondere, $v \perp v$. Äquivalent:

$$\langle v, v \rangle = 0,$$

und damit $v = 0$, welches zu zeigen war. $\square$

Definition 8.27. Sei U ein Unterraum eines Skalarproduktraums $\big(V, \langle \cdot\,, \cdot \rangle\big)$ mit der Eigenschaft, dass

$$V = U \oplus U^{\perp}.$$

Die eindeutige Abbildung

$$P\colon V \to V \text{ mit } \ker P = U^{\perp} \text{ und } \operatorname{Im} P = U$$

heißt *orthogonale Projektion* auf U. Diese Abbildung ist mit der obigen Notation

$$P(v) = u_v.$$

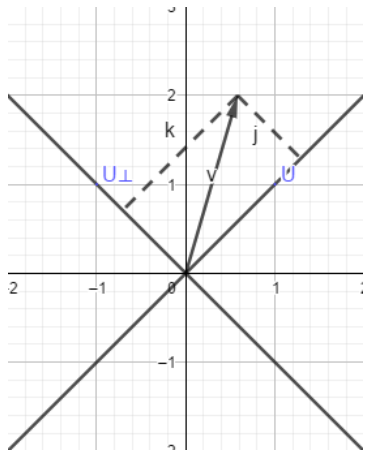


Abbildung 1: orthogonale Projektion auf eine Geraden in $\mathbb{R}^2$

Isometrien

In diesem Abschnitt schreiben wir alle Vektoren als Spaltenvektoren.

Bemerkung 8.28.

(i) Betrachte den Skalarproduktraum $\big(\mathbb{R}^n, \langle \cdot\,, \cdot \rangle\big)$ bzw. $\big(\mathbb{C}^n, \langle \cdot\,, \cdot \rangle\big)$ mit dem kanonischen Skalarprodukt. Dann gilt, für alle $x, y \in \mathbb{R}^n$ bzw. $x, y \in \mathbb{C}^n$:

$$\langle x, y \rangle = x^T y \quad \text{bzw.} \quad \langle x, y \rangle = x^T \overline{y}.$$

(ii) Seien $A, B \in M_n(\mathbb{C})$ zwei Matrizen mit der Eigenschaft, dass

$$x^T A \overline{y} = x^T B \overline{y} \quad \text{für alle } x, y \in \mathbb{C}^n.$$

Dann gilt $A = B$. (Ersetze $x = e_i$ und $y = e_j$ für alle $1 \leq i, j \leq n$.)

(iii) Seien $A, B \in M_n(\mathbb{R})$ zwei Matrizen mit der Eigenschaft, dass

$$x^T A y = x^T B y \quad \text{für alle } x, y \in \mathbb{R}^n.$$

Dann gilt $A = B$, ähnlich wie in (ii).

Definition 8.29. Seien $\big(V, \langle \cdot\,, \cdot \rangle\big)$ und $\big(W, \langle \cdot\,, \cdot \rangle\big)$ zwei Skalarprodukträume. Ein Isomorphismus $F\colon V \to W$ heißt *Isometrie*, falls

$$\underbrace{\langle F(v), F(w) \rangle}_{\text{Skalaprodukt in } W} = \underbrace{\langle v, w \rangle}_{\text{Skalaprodukt in } V}.$$

Wenn V und W euklidisch sind, man nennt F auch *orthogonale Abbildung*; und wenn V und W unitär sind, man nennt F auch eine *unitäre Abbildung*.

Bemerkung 8.30. Sei $F\colon V \to W$ eine Isometrie. Dann gilt:

(i) $\|F(v)\| = \|v\|$ für alle $v \in V$,

(ii) für alle $v, w \in V$ gilt:

$$v \perp w \quad \Longleftrightarrow \quad F(v) \perp F(w),$$

(iii) F^{-1} ist auch eine Isometrie.

Satz 8.31. *Sei $\big(V, \langle \cdot\,, \cdot \rangle\big)$ ein n-dimensionaler $\mathbb{R}$- bzw. $\mathbb{C}$-Skalarproduktraum. Dann ist V* isometrisch *zu $\big(\mathbb{R}^n, \langle \cdot\,, \cdot \rangle\big)$ bzw. $\big(\mathbb{C}^n, \langle \cdot\,, \cdot \rangle\big)$, wobei $\langle \cdot\,, \cdot \rangle$ das kanonische Skalarprodukt ist.*

Beweis. Wir wählen eine orthonormale Basis $\{v_1, \ldots, v_n\}$ von V, und sei $\{e_1, \ldots, e_n\}$ die kanonische Basis von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ ist. Definiere $F(v_i) := e_i$ für jedes i. Dies ist die gewünschte Isometrie. $\square$

Definition 8.32.

(1) Eine Matrix $A \in M_n(\mathbb{R})$ heißt *orthogonal*, falls die assoziierte Abbildung $T\colon \mathbb{R}^n \to \mathbb{R}^n,\ x \mapsto Ax$ eine Isometrie ist.

(2) Eine Matrix $A \in M_n(\mathbb{C})$ heißt *unitär*, falls die assoziierte Abbildung $T\colon \mathbb{C}^n \to \mathbb{C}^n,\ x \mapsto Ax$ eine Isometrie ist.

Sei $A \in \mathrm{GL}(n, \mathbb{C})$ unitär. Dann gilt

$$\langle Az, Aw \rangle = \langle z, w \rangle \quad \text{für alle } z, w \in \mathbb{C}^n. \tag{40}$$

Da

$$\langle Az, Aw \rangle = (Az)^T \overline{(Aw)} = z^T A^T \overline{A} \overline{w}$$

und

$$\langle z, w \rangle = z^T \overline{w} = z^T I_n \overline{w},$$

so folgt aus (40):

$$z^T (A^T \overline{A}) \overline{w} = z^T I_n \overline{w} \quad \text{für alle } z, w \in \mathbb{C}^n.$$

Daraus folgt, nach Bemerkung 8.28:

$$A^T \overline{A} = I_n.$$

Wenn wir diese Gleichung konjugieren, bekommen wir

$$\overline{A}^T A = I_n,$$

und damit

$$A^{-1} = \overline{A}^T.$$

Wir haben also gezeigt:

Lemma 8.33.

(1) Eine Matrix $A \in M_n(\mathbb{R})$ ist orthogonal genau dann, wenn $A^{-1} = A^T$.

(2) Eine Matrix $A \in M_n(\mathbb{C})$ ist unitär genau dann, wenn $A^{-1} = \overline{A}^T$.

Lemma/Definition 8.34. *Die Menge*

$$O(n) := \{A \in \mathrm{GL}(n, \mathbb{R}) \mid A^{-1} = A^T\}$$

ist eine Gruppe und heißt orthogonale Gruppe *(Übung!).*

Die Menge

$$U(n) := \{A \in \mathrm{GL}(n, \mathbb{C}) \mid A^{-1} = \overline{A}^T\}$$

ist eine Gruppe und heißt unitäre Gruppe *(Übung!).*

Proposition 8.35.

(1) Sei $A \in \mathrm{GL}(n, \mathbb{C})$ bzw. $A \in \mathrm{GL}(n, \mathbb{R})$. Die folgenden Aussagen sind äquivalent.

(a) *A ist unitär bzw. orthogonal.*

(b) *Die Spaltenvektoren von A bilden eine Orthonormalbasis von $\mathbb{C}$ bzw. $\mathbb{R}$.*

(c) *Die Zeilenvektoren von A bilden eine Orthonormalbasis von $\mathbb{C}$ bzw. $\mathbb{R}$.*

(2) *Sei $\mathcal{B}$ eine Orthonormalbasis von $\mathbb{C}$ bzw. $\mathbb{R}$. Sei $T \in \mathrm{End}(\mathbb{C})$ bzw. $T \in \mathrm{End}(\mathbb{R})$. Dann ist T unitär bzw. orthogonal genau dann, wenn die Darstellungsmatrix von T bezüglich $\mathcal{B}$ unitär bzw. orthogonal ist.*

(3) *Sei $A \in \mathrm{GL}(n, \mathbb{C})$ unitär bzw. sei $A \in \mathrm{GL}(n, \mathbb{R})$ orthogonal. Sei λ ein Eigenwert von A. Dann gilt*

$$|\lambda| = 1.$$

Beweis. (1) Sei $A = \begin{pmatrix} a^1 & \cdots & a^n \end{pmatrix}$. Aus der Gleichung $\overline{A}^T A = I_n$ folgt

$$\overline{a^i}^T a^j = \delta_{ij} \quad \text{für alle } i, j.$$

Äquivalent:

$$\langle a^j, a^i \rangle = \delta_{ij} \quad \text{für alle } i, j,$$

und das bedeutet genau, dass die Spalten von A eine Orthonormalbasis bilden. Auf ähnliche Weise zeigt man die Aussage für Zeilen.

(2) Übung!

(3) Sei v ein Eigenvektor zu λ. Dann gilt

$$0 \neq \|v\| = \|Av\| = \|\lambda v\| = |\lambda| \cdot \|v\|,$$

und damit ist $|\lambda| = 1$. □

Satz 8.36 (QR-Zerlegung)**.** *Sei $A \in \mathrm{GL}(n, \mathbb{R})$. Dann existiert eine Matrix $Q \in O(n)$ und eine obere Dreiecksmatrix R, sodass*

$$A = QR.$$

Beweis. Sei $A = \begin{pmatrix} a^1 & \cdots & a^n \end{pmatrix}$. Sei $\{v^1, \ldots, v^n\}$ die Orthonormalbasis von $\mathbb{R}^n$, die aus der Basis $\{a^1, \ldots, a^n\}$ durch das Gram-Schmidt-Verfahren gewonnen wird. Dann wissen wir (siehe den Beweis vom Gram-Schmidt-Verfahren), dass für jedes $i = 1, \ldots, n$ gilt

$$\langle v^1, \ldots, v^i \rangle = \langle a^1, \ldots, a^i \rangle.$$

Insbesondere ist v^{i+1} orthogonal zu jedem $a^1, \dots, a^i$; insbesondere gilt

$$\langle v^j, a^i \rangle = 0 \quad \text{für alle } 1 \leq i < j \leq n. \tag{41}$$

Definiere

$$Q := \begin{pmatrix} v^1 & \cdots & v^n \end{pmatrix} \in O(n) \quad \text{und} \quad R := Q^T A.$$

Dann folgt aus (41):

$$R = \begin{pmatrix} (v^1)^T \\ \vdots \\ (v^n)^T \end{pmatrix} \begin{pmatrix} a^1 & \cdots & a^n \end{pmatrix} = \begin{pmatrix} 0 & & * \\ \vdots & \ddots & \\ 0 & \cdots & 0 \end{pmatrix}.$$

Andererseits, da Q nach Proposition 8.35 eine orthogonale Matrix ist, so gilt

$$R = Q^T A = Q^{-1} A,$$

und damit $A = QR$. □

Spektralsätze

Notation 8.37. Sei V ein K-Vektorraum und sei $F \in \mathrm{End}(V)$. Ein Unterraum $U \subseteq V$ ist *F-invariant*, falls

$$F(U) \subseteq U.$$

Insbesondere ist dann $F|_U \in \mathrm{End}(U)$.

Satz 8.38 (Spektralsatz für unitäre Isomorphismen)**.** *Sei $\big(V, \langle \cdot\,, \cdot \rangle\big)$ ein n-dimensionaler unitärer Vektorraum. Sei $F \in \mathrm{Iso}(V)$ ein unitärer Isomorphismus. Dann existiert eine Orthonormalbasis von V aus Eigenvektoren von F. Insbesondere ist F diagonalisierbar.*

Beweis. Der Beweis ist per Induktion nach n. Die Aussage ist einfach, wenn $n = 1$. Wir nehmen also an, dass $n \geq 2$.

Sei $\lambda \in \mathbb{C}$ ein Eigenwert von F. Nach Proposition 8.35 gilt $|\lambda| = 1$, und insbesondere $\lambda \neq 0$. Sei $v \neq 0$ ein Eigenvektor zu λ. Nach Proposition 8.26 gilt

$$V = \langle v \rangle \oplus \langle v \rangle^\perp.$$

Wir behaupten, dass der Unterraum $\langle v \rangle^\perp$ F-invariant ist. Und zwar, sei $w \in \langle v \rangle^\perp$, oder äquivalent: $\langle v, w \rangle = 0$. Da F unitär ist, so folgt, dass

$$0 = \langle \underbrace{F(v)}_{=\lambda v}, F(w) \rangle = \lambda \langle v, F(w) \rangle.$$

Es folgt, dass $F(w) \in \langle v \rangle^\perp$, welches die Behauptung zeigt.

Es gilt also $F|_{\langle v \rangle^\perp} \in \operatorname{End}\left(\langle v \rangle^\perp\right)$, und man sieht sofort, dass auch $F|_{\langle v \rangle^\perp}$ unitär ist. Per Induktion gibt es eine Orthonormalbasis $\{e_2, \ldots, e_n\}$ von $\langle v \rangle^\perp$, die aus Eigenvektoren von $F|_{\langle v \rangle^\perp}$ besteht; diese sind dann auch Eigenvektoren von F. Die gewünschte Basis von V ist nun

$$\left\{ \frac{v}{\|v\|}, e_2, \ldots, e_n \right\}.$$

Letztlich ist F diagonalisierbar nach Proposition 7.14. $\square$

Man kann auch zeigen:

Satz 8.39 (Spektralsatz für orthogonale Isomorphismen)**.** *Sei $\left(V, \langle \cdot, \cdot \rangle\right)$ ein n-dimensionaler euklidischer Vektorraum. Sei $F \in \operatorname{Iso}(V)$ ein orthogonaler Isomorphismus. Dann gibt es eine Orthonormalbasis $\mathcal{B}$ von $\mathbb{R}^n$, sodass die Darstellungsmatrix von F bezüglich $\mathcal{B}$ die Gestalt*

$$\mathcal{A} = \begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & 0 & \\ & & 1 & & & & & & & \\ & & & -1 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & -1 & & & & \\ & & & & & & \mathcal{A}_1 & & & \\ & 0 & & & & & & \ddots & & \\ & & & & & & & & & \mathcal{A}_m \end{pmatrix}$$

hat, wobei

$$\mathcal{A}_i = \begin{pmatrix} \cos\alpha_i & -\sin\alpha_i \\ \sin\alpha_i & \cos\alpha_i \end{pmatrix} \quad \text{für } \alpha_i \in (0, 2\pi)\backslash\{\pi\}.$$

Insbesondere ist F ist damit eine Komposition aus den Spiegelungen an Hyperebenen und Drehungen in einer Ebene.

Definition 8.40. Eine Matrix $A \in M_n(\mathbb{C})$ ist *symmetrisch*, falls

$$A = A^T,$$

und sie ist *hermitesch*, falls

$$A = \overline{A}^T.$$

Beispiel 8.41. Eine hermitesche Matrix $A \in M_2(\mathbb{C})$ hat die Form

$$\begin{pmatrix} a_1 + ib_1 & a_2 + ib_2 \\ a_2 - ib_2 & a_3 + ib_3 \end{pmatrix}, \quad \text{wobei } a_i, b_i \in \mathbb{R}.$$

Satz 8.42. *Sei $\langle \cdot\,, \cdot \rangle$ das kanonische Skalarprodukt auf $\mathbb{C}^n$ bzw. auf $\mathbb{R}^n$ und sei $A \in M_n(\mathbb{C})$ bzw. $A \in M_n(\mathbb{R})$. Dann ist A hermitesch bzw. symmetrisch genau dann, wenn*

$$\langle Az, w \rangle = \langle z, Aw \rangle \quad \textit{für alle Vektoren } w, z. \tag{42}$$

Beweis. Wir zeigen die Aussage für hermitesche Matrizen; die Aussage für symmetrische Matrizen zeigt man analog. Da

$$\langle Az, w \rangle = (Az)^T \overline{w} = z^T A^T \overline{w}$$

und

$$\langle z, Aw \rangle = z^T \overline{(Aw)} = z^T \overline{A} \overline{w}.$$

Damit ist (42) äquivalent zu

$$z^T A^T \overline{w} = z^T \overline{A} \overline{w} \quad \text{für alle } w, z \in \mathbb{C}^n.$$

Dies ist äquivalent, nach Bemerkung 8.28, zu $A^T = \overline{A}$, welches zu $\overline{A}^T = A$ äquivalent ist. □

Satz 8.43. *Sei $A \in M_n(\mathbb{R})$ eine symmetrische bzw. sei $A \in M_n(\mathbb{C})$ eine hermitesche Matrix. Dann sind alle Eigenwerte von A reell.*

Beweis. Wir zeigen die Aussage für hermitesche Matrizen; derselbe Beweis funktioniert für symmetrische Matrizen. Sei $\lambda \in \mathbb{C}$ ein Eigenwert von A und sei $v \in \mathrm{Eig}(A, \lambda) \backslash \{0\}$. Dann gilt

$$\langle v, Av \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \langle v, v \rangle = \overline{\lambda} \|v\|^2$$

und

$$\langle Av, v \rangle = \langle \lambda v, v \rangle = \lambda \langle v, v \rangle = \lambda \|v\|^2.$$

Da A hermitesch ist, so folgt

$$\overline{\lambda} \|v\|^2 = \lambda \|v\|^2,$$

und damit $\lambda = \overline{\lambda}$. Also, $\lambda \in \mathbb{R}$. □

Satz 8.44 (Hauptachsentransformation)**.** *Sei $A \in M_n(\mathbb{C})$ eine hermitesche Matrix. Dann existiert eine Orthonormalbasis $\{v_1, \dots, v_n\} \subseteq \mathbb{C}^n$ aus Eigenvektoren von A, und A ist diagonalisierbar.*

Präziser: Es gibt eine unitäre Matrix $Q \in U(n)$ such that

$$Q^{-1}AQ = \overline{Q}^T AQ = \operatorname{diag}(\lambda_1, \dots, \lambda_n),$$

wobei $\lambda_i \in \mathbb{R}$ die Eigenwerte von A sind.

Beweis. Der Beweis ist per Induktion nach n. Die Aussage ist einfach, wenn $n = 1$. Wir nehmen also an, dass $n \geq 2$. Sei

$$F \colon \mathbb{C}^n \to \mathbb{C}^n, \quad x \mapsto Ax$$

die assoziierte lineare Abbildung.

Sei λ_1 ein Eigenwert von A. Nach Satz 8.43 gilt $\lambda_1 \in \mathbb{R}$. Sei v_1 ein Eigenvektor von λ_1 mit $\|v_1\| = 1$. Nach Proposition 8.26 gilt

$$V = \langle v_1 \rangle \oplus \langle v_1 \rangle^\perp.$$

Wir behaupten, dass der Unterraum $\langle v_1 \rangle^\perp$ F-invariant ist. Und zwar, sei $w \in \langle v_1 \rangle^\perp$, oder äquivalent: $\langle w, v_1 \rangle = 0$. Damit ist nach Satz 8.42:

$$\langle F(w), v_1 \rangle = \langle Aw, v_1 \rangle = \langle w, Av_1 \rangle = \langle w, \lambda_1 v_1 \rangle = \overline{\lambda_1} \langle w, v_1 \rangle = 0.$$

Es folgt, dass $F(w) \in \langle v_1 \rangle^\perp$, welches die Behauptung zeigt.

Es gilt also $F|_{\langle v_1 \rangle^\perp} \in \operatorname{End}\left(\langle v_1 \rangle^\perp\right)$, und man sieht sofort, dass die Darstellungsmatrix B von $F|_{\langle v_1 \rangle^\perp}$ auch hermitesch ist. Per Induktion gibt es eine Orthonormalbasis $\{v_2, \dots, v_n\}$ von $\langle v_1 \rangle^\perp$, die aus Eigenvektoren von B besteht; diese sind dann auch Eigenvektoren von A. Die gewünschte Basis von V ist nun

$$\{v_1, v_2, \dots, v_n\}.$$

Letztlich ist A diagonalisierbar nach Proposition 7.14.

Nun zeigen wir die zweite Aussage. Betrachte die Vektoren v_i als Spalten und setze

$$Q_1 := \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \overset{8.35}{\in} U(n) \quad \text{und} \quad A_1 := \overline{Q_1}^T A Q_1.$$

Man berechnet einfach, dass

$$A_1 = \overline{Q_1}^T A Q_1 = \left(\begin{array}{c|c} \lambda_1 & 0 \\ \hline 0 & A_2 \end{array}\right),$$

wobei $A_2 \in M_{n-1}(\mathbb{C})$. Es ist einfach zu sehen, dass A_1 hermitesch ist, und daraus folgt, dass A_2 hermitesch ist. Per Induktionsschritt gibt es dann $Q_2 \in U(n-1)$, sodass

$$\overline{Q_2}^T A_2 Q_2 = \operatorname{diag}(\lambda_2, \ldots, \lambda_n),$$

wobei $\lambda_i \in \mathbb{R}$ die Eigenwerte von A_2 sind, damit auch die Eigenwerte von A_1 sind, und damit auch von A. Definiere

$$Q := Q_1 \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_2 \end{array}\right) \in U(n).$$

Nun gilt:

$$\begin{aligned}
\overline{Q}^T A Q &= \overline{\left(Q_1 \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_2 \end{array}\right)\right)}^T \cdot A \cdot \left(Q_1 \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_2 \end{array}\right)\right) \\
&= \overline{\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_2 \end{array}\right)}^T \cdot \overline{Q_1}^T \cdot A \cdot Q_1 \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_2 \end{array}\right) \\
&= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \overline{Q_2}^T \end{array}\right) \cdot \left(\begin{array}{c|c} \lambda_1 & 0 \\ \hline 0 & A_2 \end{array}\right) \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_2 \end{array}\right) \\
&= \left(\begin{array}{c|c} \lambda_1 & 0 \\ \hline 0 & \overline{Q_2}^T A_2 Q_2 \end{array}\right) = \left(\begin{array}{c|ccc} \lambda_1 & & 0 & \\ \hline & \lambda_2 & & \\ 0 & & \ddots & \\ & & & \lambda_n \end{array}\right).
\end{aligned}$$

Das beweist den Satz. □

Im reellen Fall gilt:

Satz 8.45 (Hauptachsentransformation über $\mathbb{R}$)**.** *Sei $A \in M_n(\mathbb{R})$ eine symmetrische Matrix. Dann existiert eine Orthonormalbasis $\{v_1, \ldots, v_n\} \subseteq \mathbb{R}^n$ aus Eigenvektoren von A, und A ist diagonalisierbar.*

Präziser: Es gibt eine unitäre Matrix $Q \in O(n)$ such that

$$Q^{-1}AQ = Q^TAQ = \operatorname{diag}(\lambda_1, \dots, \lambda_n),$$

wobei $\lambda_i \in \mathbb{R}$ die Eigenwerte von A sind.

Korollar 8.46. *Sei $A \in M_n(\mathbb{R})$ eine symmetrische bzw. sei $A \in M_n(\mathbb{C})$ eine hermitesche Matrix. Dann hat das charakteristische Polynom von A die Zerlegung*

$$\chi_A(t) = \prod_{i=1}^{n}(t - \lambda_i),$$

wobei $\lambda_i \in \mathbb{R}$ die (nicht notwendigerweise verschiedenen) Eigenwerte von A sind.

Beweis. Folgt aus der Hauptachsentransformation und Korollar 7.21. □

Beispiel 8.47. Sei

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \in M_3(\mathbb{R}).$$

Wir werden dem vorigen Beweis folgen, um eine Matrix $Q \in O(3)$ und eine Diagonalmatrix $D \in M_3(\mathbb{R})$ zu finden, sodass $Q^TAQ = D$.

Man sieht zuerst sofort, dass A symmetrisch ist und

$$\chi_A(t) = \det \begin{pmatrix} 2-t & 1 & 0 \\ 1 & 2-t & 1 \\ 0 & 1 & 2-t \end{pmatrix} = (2-t)\left((2-t)^2 - 2\right).$$

Damit sind die Eigenwerte von A:

$$\lambda_1 = 2, \qquad \lambda_2 = 2 + \sqrt{2}, \qquad \lambda_3 = 2 - \sqrt{2}.$$

Ferner ist

$$\operatorname{Eig}(A, \lambda_1) = \ker \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\rangle,$$

und $v_1 := \dfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ ist ein Vektor in $\operatorname{Eig}(A, \lambda_1)$ der Norm 1.

Setze

$$H := \langle v_1 \rangle^\perp = \left\{ x \in \mathbb{R}^3 \;\middle|\; \left\langle x, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\rangle = 0 \right\}.$$

Wir benutzen nun Gram-Schmidt-Verfahren, um eine orthonormale Basis von H zu finden. Und zwar, man berechnet, dass $H = \langle v_2, v_3 \rangle$, wobei

$$v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad v_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

und $\{v_1, v_2, v_3\}$ ist eine Orthonormalbasis von $\mathbb{R}^3$. Setze

$$Q_1 := \begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{pmatrix}$$

und

$$A_1 := Q_1^T A Q_1 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & \sqrt{2} \\ 0 & \sqrt{2} & 2 \end{pmatrix};$$

folgend, setze

$$A_2 := \begin{pmatrix} 2 & \sqrt{2} \\ \sqrt{2} & 2 \end{pmatrix}.$$

Die Eigenwerte von A_2 sind $2 \pm \sqrt{2}$, und

$$\mathrm{Eig}(A_2, \lambda_2) = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle.$$

Damit ist $w_1 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ein Vektor in $\mathrm{Eig}(A_2, \lambda_2)$ der Norm 1. Es gilt

$$H' := \langle w_1 \rangle^\perp = \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle,$$

und letztlich ist $w_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ein Vektor in H' der Norm 1. Definiere

$$Q_2 := \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}.$$

Dann berechnen wir einfach, dass

$$Q_2^T A_2 Q_2 = \begin{pmatrix} 2+\sqrt{2} & 0 \\ 0 & 2-\sqrt{2} \end{pmatrix},$$

und endlich ist die gesuchte Matrix Q, die A zu der Diagonalmatrix

$$D := \operatorname{diag}(2, 2+\sqrt{2}, 2-\sqrt{2})$$

transformiert:

$$Q = Q_1 \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$