Lineare Algebra und Analytische Geometrie

Manuel Kauers

Institut für Algebra
Johannes Kepler Universität

Version: 24. Juni 2016

Inhalt

Teil I Algebraische Strukturen **5**

1 Mengen . . . 6
2 Relationen . . . 10
3 Funktionen . . . 16
4 Gruppen . . . 23
5 Ringe . . . 32
6 Körper . . . 35

Teil II Vektoren und Matrizen **39**

7 Vektoren . . . 40
8 Matrizen . . . 43
9 Gleichungssysteme . . . 51
10 Lineare Unabhängigkeit und Rang . . . 61
11 Inhomogene Systeme . . . 66
12 Determinanten . . . 73

Teil III Vektorräume und Lineare Abbildungen **84**

13 Vektorräume . . . 85
14 Basis und Dimension . . . 89
15 Konstruktionen . . . 96
16 Lineare Abbildungen und Isomorphie . . . 104
17 Koordinatendarstellungen . . . 113
18 Der Dualraum . . . 116

Teil IV Anwendungen **121**

19 Affine und Projektive Geometrie . . . 122
20 Farbräume . . . 127
21 Graphentheorie . . . 129
22 C-finite Folgen . . . 136
23 Kodierungstheorie . . . 142
24 Lineare Algebra in Maple, Mathematica und Sage . . . 145

Teil V Eigenwerte **153**
25 Polynome . . . 154
26 Diagonalisierbarkeit . . . 162
27 Annihilierende Polynome . . . 170
28 Der Satz von Cayley-Hamilton . . . 174
29 Invariante Unterräume . . . 179
30 Die Jordan-Normalform . . . 184

Teil VI Skalarprodukte **191**
31 Metrische, normierte und euklidische Räume . . . 192
32 Positivdefinitheit . . . 202
33 Orthogonalsysteme . . . 206
34 Adjungierte Abbildungen . . . 216
35 Projektionen und Isometrien . . . 220
36 Die Singulärwertzerlegung . . . 226

Teil VII Modultheorie **232**
37 Grundbegriffe . . . 233
38 Konstruktionen . . . 238
39 Freiheit und Torsion . . . 243
40 Lineare Gleichungssysteme über Ringen . . . 247
41 Die Smith-Normalform . . . 253
42 Kurze Vektoren in Gittern . . . 258

Teil VIII Algorithmische Lineare Algebra **267**
43 Komplexität . . . 268
44 Dünn besetzte Matrizen . . . 271
45 Strukturierte Matrizen . . . 276
46 Schnelle Multiplikation . . . 285
47 Numerische Algorithmen . . . 290
48 Symbolische Algorithmen . . . 296

Teil IX Anwendungen **304**
49 Eliminationstheorie . . . 305
50 Hauptachsentransformation . . . 309
51 PageRank . . . 314
52 Rationale erzeugende Funktionen . . . 316
53 Positive Polynome . . . 321
54 Signalverarbeitung . . . 321

Teil I

Algebraische Strukturen

1 Mengen

Die Aufgabe der Mathematik ist es, „Zusammenhänge“ zwischen „Eigenschaften“ von (abstrakten) „Objekten“ zu erkennen und zu beschreiben. Objekte sind z. B. Zahlen, Funktionen, Punkte usw. Eigenschaften sind z. B. gerade/ungerade, prim, negativ, stetig, usw. Ein Zusammenhang ist z. B. $x > y \wedge y > z \Rightarrow x > z$ (wenn x größer ist als y und y seinerseits größer als z, dann folgt daraus, dass x auch größer als z ist).

Objekte werden zu „Mengen“ (engl. *set*) zusammengefasst. („Sei A die Menge aller Objekte mit der Eigenschaft ...“.) Man schreibt $x \in A$, falls x ein Objekt ist, das die Eigenschaft hat und $x \notin A$, falls nicht. Im Fall $x \in A$ sagt man, x ist ein *Element* der Menge A.

Mengen sind selbst auch abstrakte Objekte, können also in anderen Mengen enthalten sein. Zum Beispiel gilt $1 \in \{1\}$ und $\{1\} \in \{\{1\}\}$, jedoch $1 \notin \{\{1\}\}$ und $\{1\} \notin \{1\}$.

Bei der Konstruktion von Mengen muss man aufpassen, dass man sich nicht in Widersprüche verwickelt.

Beispiel. Sei A die Menge aller Mengen M mit der Eigenschaft $M \notin M$. Für die Menge A gilt dann

$$A \in A \iff A \notin A$$

Also: A ist genau dann selbst in A enthalten, wenn A nicht in A enthalten ist.

Das kann nicht sein. Haben wir hier etwas verbotenes getan? Für Juristen gilt der Grundsatz *Was nicht explizit verboten ist, ist erlaubt.* In der Mathematik gilt dagegen: *Was nicht explizit erlaubt ist, ist verboten!* Wir müssen also, wenn wir mit Mengen hantieren wollen, vorher festlegen, welche Gesetze für die Theorie der Mengen gelten sollen. Die folgenden haben sich als zweckmäßig erwiesen:

Axiom.

1. Sind $a_1, \dots, a_n$ endlich viele Objekte, dann existiert eine Menge A mit der Eigenschaft

   $$x \in A \iff x = a_1 \vee x = a_2 \vee \cdots \vee x = a_n.$$

   Notation: $\{a_1, \dots, a_n\} := A$. Im Fall $n = 0$ schreibt man $\emptyset := \{\}$ und nennt diese Menge die *leere Menge.*

2. Die Menge $\mathbb{N} := \{0, 1, 2, \dots\}$ der natürlichen Zahlen existiert.

3. Ist A eine Menge von Mengen (d. h. jedes Element von A ist eine Menge), so existieren die Mengen $\bigcup_{a \in A} a$ und $\bigcap_{a \in A} a$ mit

   $$x \in \bigcup_{a \in A} a \iff \exists\, a \in A : x \in a,$$
   $$x \in \bigcap_{a \in A} a \iff \forall\, a \in A : x \in a.$$

   Die Menge $\bigcup_{a \in A} a$ heißt die *Vereinigung* (engl. *union*) der Mengen in A, die Menge $\bigcap_{a \in A} a$ heißt der *Schnitt* (engl. *intersection*) der Mengen in A. Ist $A = \{a_1, \dots, a_n\}$ eine endliche Menge von Mengen, so schreibt man statt $\bigcup_{a \in A} a$ und $\bigcap_{a \in A} a$ auch $\bigcup_{i=1}^{n} a_i$ oder $a_1 \cup a_2 \cup \cdots \cup a_n$ bzw. $\bigcap_{i=1}^{n} a_i$ oder $a_1 \cap \cdots \cap a_n$.

4. Seien A, B Mengen, $f\colon A \to B$ eine Funktion (siehe Abschnitt 3), und $p\colon A \to \{\mathrm{T}, \mathrm{F}\}$. Dann existiert eine Menge C mit

$$y \in C \iff \exists\, x \in A : p(x) = \mathrm{T} \wedge f(x) = y.$$

   Notation: $\{\, f(x) : x \in A \wedge p(x) \,\} := C$; im Fall $f = \mathrm{id}$ auch $\{\, x \in A : p(x) \,\} := C$.

5. Seien $A_1, \ldots, A_n$ Mengen. Dann existiert auch die Menge C mit

$$z \in C \iff \exists\, x_1 \in A_1 \ \cdots\ \exists\, x_n \in A_n : z = (x_1, \ldots, x_n)$$

   Dabei ist $(x_1, \ldots, x_n)$ das *Tupel* bestehend aus den *Komponenten* $x_1, \ldots, x_n$. Notation: $A_1 \times A_2 \cdots \times A_n := C$. Man nennt die Menge C das *kartesische Produkt* der Mengen $A_1, \ldots, A_n$. Statt $A \times A \times \cdots \times A$ schreibt man auch A^n.

Beispiel.

1. $\{1, 1, 2\} = \{1, 2\} = \{2, 1\} = \{1, 2, 2, 2, 2, 1\}$
2. $\{1, 3, 4, 5\} \cup \{3, 5, 7, 8\} = \{1, 3, 4, 5, 7, 8\}$
3. $\{1, 3, 4, 5\} \cap \{3, 5, 7, 8\} = \{3, 5\}$
4. $\{\, x^2 : x \in \mathbb{N} \wedge x \text{ prim} \,\} = \{4, 9, 25, 49, 121, \ldots\}$
5. $\{\, \text{Geburtsdatum}(S) : S \text{ ist Vorlesungsteilnehmer} \,\}$
6. $\{1, 2\} \times \{a, b\} = \{(1, a), (2, a), (1, b), (2, b)\}$
7. $\{a, b\} \times \{1, 2\} = \{(a, 1), (a, 2), (b, 1), (b, 2)\}$

Definition 1. Seien A, B Mengen.

1. Die Mengen A, B heißen *(zueinander) gleich* (engl. *equal*), geschrieben $A = B$, falls gilt

$$\forall\, x : (x \in A \iff x \in B).$$

2. A heißt *Teilmenge* (engl. *subset*) von B, notiert $A \subseteq B$, falls $\forall\, x : x \in A \Rightarrow x \in B$.
3. A heißt *Obermenge* (engl. *superset*) von B, notiert $A \supseteq B$, falls $\forall\, x : x \in B \Rightarrow x \in A$.
4. A und B heißen *(zueinander) disjunkt* (engl. *disjoint*), falls $A \cap B = \emptyset$.
5. $A \setminus B := \{\, a \in A : a \notin B \,\}$ heißt das *Komplement* von B in A.

Satz 1. Seien A, B, C Mengen.

1. Wenn $A \subseteq B$ und $B \subseteq C$, dann $A \subseteq C$
2. $A = B \iff A \subseteq B \wedge B \subseteq A$
3. Ist $B \neq \emptyset$ eine Menge von Mengen, so gilt $A \cap \bigcup_{b \in B} b = \bigcup_{b \in B}(A \cap b)$
4. Ist $B \neq \emptyset$ eine Menge von Mengen, so gilt $A \cup \bigcap_{b \in B} b = \bigcap_{b \in B}(A \cup b)$
5. Ist $B \neq \emptyset$ eine Menge von Mengen, so gilt $A \setminus \bigcap_{b \in B} b = \bigcup_{b \in B}(A \setminus b)$ und $A \setminus \bigcup_{b \in B} b = \bigcap_{b \in B}(A \setminus b)$.
6. $A \cap B \subseteq A \subseteq A \cup B$
7. $A = B \iff B = A$, $A \cap B = B \cap A$, $A \cup B = B \cup A$

Beweis.

1. Annahmen: $A \subseteq B$ und $B \subseteq C$

   zu zeigen: $A \subseteq C$.

   Sei $x \in A$ beliebig. Nach Annahme $A \subseteq B$ folgt per Definition von „$\subseteq$“, dass $x \in B$. Daraus folgt nach Annahme $B \subseteq C$ per Definition von „$\subseteq$“, dass $x \in C$.

   Da x ein beliebiges Objekt von A war, ist also gezeigt: $\forall\, x : (x \in A \Rightarrow x \in C)$. Also gilt $A \subseteq C$, was zu zeigen war.

2. Aussagen der Form $A \iff B$ zeigt man, indem man zunächst $A \Rightarrow B$ und dann $B \Rightarrow A$ zeigt.

   „$\Rightarrow$“ Annahme: $A = B$

   zu zeigen: $A \subseteq B \wedge B \subseteq A$.

   Aus Symmetriegründen genügt es, $A \subseteq B$ zu zeigen. Das Argument für $B \subseteq A$ geht dann genauso.

   Sei $x \in A$ beliebig. Nach Annahme $A = B$ und per Definition von „$=$“ gilt $x \in B$. Da x beliebig war, folgt $x \in B$.

   „$\Leftarrow$“ Annahmen: $A \subseteq B$ und $B \subseteq A$.

   zu zeigen: $A = B$.

   Nach Definition von „$=$“ ist zu zeigen $\forall\, x : (x \in A \iff x \in B)$.

   Sei x ein beliebiges Objekt. Aus der Annahme $A \subseteq B$ folgt $x \in A \Rightarrow x \in B$ und aus der Annahme $B \subseteq A$ folgt $x \in B \Rightarrow x \in A$. Beides zusammen gibt $x \in A \iff x \in B$.

3. Wegen Punkt 2 kann man die Gleichheit zweier Mengen zeigen, indem man zeigt, dass jede eine Teilmenge der anderen ist.

   „$\subseteq$“ Sei $x \in A \cap \bigcup_{b \in B} b$ beliebig. Wir zeigen $x \in \bigcup_{b \in B}(A \cap b)$.

   Dazu ist zu zeigen $\exists\, b \in B : x \in A \cap b$. Nach Annahme gilt $x \in A$ und $x \in \bigcup_{b \in B} b$, d. h. $\exists\, b \in B : x \in b$.

Wähle ein $b \in B$ mit $x \in b$. Dann ist wegen $x \in A$ auch $x \in A \cap b$. Es folgt $\exists\, b \in B : x \in A \cap b$, und damit die Behauptung.

„$\supseteq$“ Sei $x \in \bigcup_{b\in B}(A \cap b)$ beliebig. Wir zeigen $x \in A \cap \bigcup_{b\in B} b$.

Dazu ist zu zeigen $x \in A$ und $x \in \bigcup_{b\in B} b$, d. h. $\exists\, b \in B : x \in b$.

Wähle $b \in B$ mit $x \in A \cap b$. Dann gilt $x \in A$ und $x \in b$. Also $\exists\, b \in B : x \in b$, also $x \in \bigcup_{b\in B} b$, also $x \in A \cap \bigcup_{b\in B} b$.

4. „$\subseteq$“ Sei $x \in A \cup \bigcap_{b\in B} b$ beliebig. Wir zeigen $x \in \bigcap_{b\in B}(A \cup b)$.

   Dazu ist zu zeigen $\forall\, b \in B : x \in A \cup b$. Nach Annahme gilt $x \in A$ oder $x \in \bigcap_{b\in B} b$, d. h. $\forall\, b \in B : x \in b$.

   Für beliebiges $b \in B$ gilt also $x \in A$ oder $x \in b$, also $x \in A \cup b$. Da $b \in B$ beliebig war, folgt $\forall\, b \in B : x \in A \cup b$, und damit die Behauptung.

   „$\supseteq$“ Sei $x \in \bigcap_{b\in B}(A \cup b)$ beliebig. Wir zeigen $x \in A \cup \bigcap_{b\in B} b$.

   Dazu ist zu zeigen $x \in A$ oder $x \in \bigcap_{b\in B} b$, d. h. $x \in A$ oder $\forall\, b \in B : x \in b$. Wir zeigen: Wenn $x \notin A$ ist, dann muss $\forall\, b \in B : x \in b$ gelten.

   Sei $b \in B$ beliebig. Dann gilt nach Voraussetzung $x \in A \cup b$, und daraus folgt $x \in A$ oder $x \in b$. Nach Annahme gilt $x \notin A$, also $x \in b$. Da b beliebig war, folgt also $\forall\, b \in B : x \in b$, und damit die Behauptung.

5., 6., 7. Übung. ∎

Beispiel. $A = \{x, y, z\}$, $B = \{1, 2, 3, 4\}$

$$A \left\{ \begin{array}{c|cccc} x & (x,1) & (x,2) & (x,3) & (x,4) \\ y & (y,1) & (y,2) & (y,3) & (y,4) \\ z & (z,1) & (z,2) & (z,3) & (z,4) \end{array} \right\} A \times B$$

$$\begin{array}{c|c} & \underbrace{1 \qquad 2 \qquad 3 \qquad 4}_{B} \end{array}$$

Satz 2. Seien A, B, C Mengen.

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

Beweis.

1. „$\subseteq$“ Sei $x \in (A \cap B) \times C$ beliebig. Zu zeigen: $x \in (A \times C) \cap (B \times C)$.

   Nach Definition von $\times$ gilt $x = (x_1, x_2)$ für gewisse x_1, x_2 mit $x_1 \in A \cap B$ und $x_2 \in C$. Aus $x_1 \in A \cap B$ folgt $x_1 \in A$ und $x_1 \in B$. Demnach gilt $(x_1, x_2) \in A \times C$ und $(x_1, x_2) \in B \times C$, also $(x_1, x_2) \in (A \times C) \cap (B \times C)$.

   „$\supseteq$“ Sei $x \in (A \cap C) \times (B \times C)$ beliebig. Zu zeigen: $x \in (A \cap B) \times C$.

Nach Definition gilt $x \in A \times C$ und $x \in B \times C$.

$$\begin{array}{ccc} \Downarrow & & \Downarrow \\ x = (x_1, x_2) \text{ mit} & & x = (x_1, x_2) \text{ mit} \\ x_1 \in A, x_2 \in C & & x_1 \in B, x_2 \in C \end{array}$$

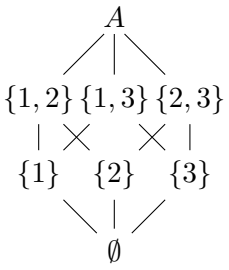
$$\Rightarrow \quad \underbrace{x_1 \in A \wedge x_1 \in B}_{\Downarrow} \wedge x_2 \in C$$

$$\Rightarrow \quad \underbrace{x_1 \in A \cap B}_{x = (x_1, x_2) \in (A \cap B) \times C}$$

2. Übung. ∎

Axiom. Sei A eine Menge. Dann existiert eine Menge B mit $\forall\, x : x \in B \iff x \subseteq A$. Man schreibt $\mathcal{P}(A) := B$ und nennt diese Menge die *Potenzmenge* (engl. *power set*) von A.

Beispiel. Für $A = \{1, 2, 3\}$ ist $\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$.



Ein weiteres Axiom zur Mengenlehre kommt später in Abschnitt 14. Darüber hinaus gibt es einige weitere, die wir nicht brauchen werden, zum Beispiel eines, das besagt, dass jede Kette $a \in b$, $b \in c$, $c \in d$, $d \in e$, ... nach endlich vielen Schritten abbrechen muss, d. h. die „Schachtelungstiefe“ bei Mengen von Mengen von Mengen von... darf nicht unendlich werden.

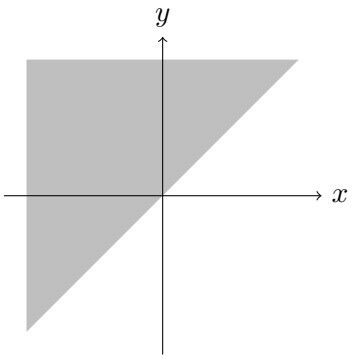
2 Relationen

Definition 2. Sei A eine Menge. Eine Teilmenge $R \subseteq A \times A$ heißt *Relation* auf A. Statt $(x, y) \in R$ schreibt man xRy.

Beispiel.

05-06

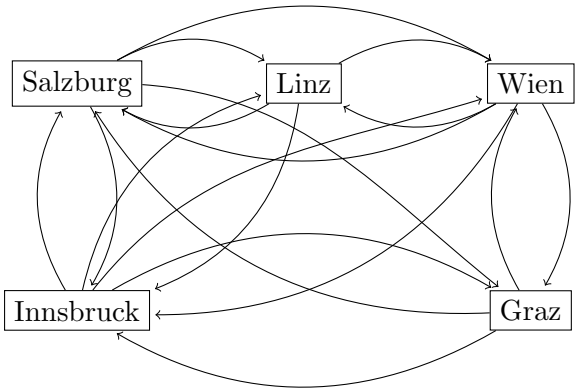
1. $\leq = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x \text{ ist kleiner oder gleich } y\}$ ist eine Relation. Die Menge $\mathbb{R}^2$ kann man sich als die Punkte einer Ebene vorstellen. Eine Teilmenge von $\mathbb{R}^2$ kann man dann dadurch darstellen, dass man die entsprechenden Punkte der Ebene markiert. In der folgenden Abbildung sind alle Punkte (x, y) markiert, die zur Relation $\leq$ gehören.



2. A die Menge aller Städte in Österreich mit mehr als 100000 Einwohnern (2015).

   R die Menge aller $(x, y) \in A \times A$ für die gilt: es gibt eine direkte Zugverbindung von x nach y.

   In der Informatik nennt man so etwas einen *Graphen.* Ein Graph ist also ein Paar $G = (V, E)$, wobei V eine Menge und E eine Relation auf V ist. Elemente von V nennt man *Knoten* (engl. *vertex*) und Elemente von E *Kanten* (engl. *edge*).



3. Ist A die Menge aller Teilnehmer der Vorlesungsklausur, und ist R die Menge aller $(x, y) \in A \times A$, für die gilt, dass x und y die gleiche Note bekommen, so ist R eine Relation auf A. 05-06

4. *Teilbarkeit* (engl. *divisibility*): $A = \mathbb{Z}$, $R = \{(x, y) \in A \times A : \exists\ z \in \mathbb{Z} : x \cdot z = y\}$. Statt xRy schreibt man typischerweise $x \mid y$. Zum Beispiel gilt $3 \mid 15$ und $4 \nmid 15$.

5. $\subseteq$ ist eine Relation auf $\mathcal{P}(A)$: $\subseteq = \{(U, V) \in \mathcal{P}(A) \times \mathcal{P}(A) : U \subseteq V\}$.

6. Ist M die Menge aller Menschen, so ist $\heartsuit := \{(x, y) \in M \times M : x \text{ mag } y\}$ eine Relation auf M. Zum Beispiel könnte gelten Heinz $\heartsuit$ Erika. 05-06

Definition 3. Sei A eine Menge und R eine Relation auf A.

1. R heißt *reflexiv,* falls gilt $\forall\, x \in A : xRx$ und *irreflexiv,* falls gilt $\forall\, x \in A : \neg(xRx)$.
2. R heißt *symmetrisch,* falls gilt:

   $$\forall\, x, y \in A : xRy \Rightarrow yRx$$

   und *antisymmetrisch,* falls

   $$\forall\, x, y \in A : (xRy \wedge yRx) \Rightarrow x = y$$

3. R heißt *transitiv,* falls gilt

   $$\forall\, x, y, z \in A : (xRy \wedge yRz) \Rightarrow xRz.$$

4. R heißt *total,* falls gilt

   $$\forall\, x, y \in A : xRy \vee yRx$$

Beispiel. Die Relationen aus dem vorangegangenen Beispiel haben folgende Eigenschaften:

Bsp	1	2	3	4	5	6
reflexiv	ja	nein	ja	ja	ja	unklar
irreflexiv	nein	ja	nein	nein	nein	nein
symmetrisch	nein	ja	ja	nein	nein	nein
antisymmetrisch	ja	nein	nein	ja	ja	nein
transitiv	ja	nein	ja	ja	ja	nein
total	ja	nein	nein	nein	nein	nein

Definition 4. Sei A eine Menge und $R \subseteq A \times A$ eine Relation auf A. Wenn R reflexiv, transitiv und antisymmetrisch ist, dann heißt R eine *Halbordnung* auf A. Ist R außerdem total, so heißt R *(Total-)Ordnung* (engl. *order(ing)*) auf A.

Beispiel.

1. $\subseteq$, $\leq$ und $|$ sind Halbordnungen auf $\mathcal{P}(A), \mathbb{R}, \mathbb{N}$. $\leq$ ist sogar eine Totalordnung, aber $|$ und $\subseteq$ sind es nicht.
2. Auf $M = \mathbb{Z} \times \mathbb{Z}$ wird eine Halbordnung $\leq$ definiert durch

   $$(a_1, a_2) \leq (b_1, b_2) :\Longleftrightarrow a_1 \leq b_1 \wedge a_2 \leq b_2,$$

   wobei mit $\leq$ auf der rechten Seite die übliche Ordnung auf $\mathbb{Z}$ gemeint ist.

   Es gilt dann zum Beispiel $(3, 5) \leq (7, 8)$. Bei dieser Halbordnung handelt es sich nicht um eine Totalordnung, weil zum Beispiel die Elemente $(3, 7)$ und $(5, 3)$ nicht miteinander vergleichbar sind, d.h. es gilt weder $(3, 7) \leq (5, 3)$ noch $(5, 3) \leq (3, 7)$.

Definition 5. Sei A eine Menge und $R \subseteq A \times A$ eine Relation auf A. Wenn R reflexiv, symmetrisch und transitiv ist, dann heißt R eine *Äquivalenzrelation.*

Beispiel.

1. Für jede Menge A ist die Gleichheitsrelation $=$ eine Äquivalenzrelation, denn für alle Objekte x, y, z gilt $x = x$ (Reflexivität), $x = y \iff y = x$ (Symmetrie) und $x = y \wedge y = z \Rightarrow x = z$ (Transitivität).

   Im allgemeinen darf man sich eine Äquivalenzrelation vorstellen als eine abgeschwächte Variante der Gleichheitsrelation, bei der bestimmte irrelevante Eigenschaften ignoriert werden.

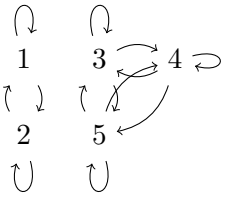
2. Sei $A = \{\square, \blacksquare, \square, \square, \blacksquare, \circ, \bullet, \bullet, \bullet, \triangle, \blacktriangle, \blacktriangle, \triangle\}$. Man kann sich für diese Menge verschiedene Äquivalenzrelationen vorstellen. Hier sind ein paar Möglichkeiten:

   - $x \sim y$, falls x und y die gleiche Form haben – dann gilt z. B. $\square \sim \blacksquare$ und $\square \not\sim \circ$.
   - $x \sim y$, falls x und y die gleiche Farbe haben – dann gilt z. B. $\square \sim \circ$ und $\square \not\sim \blacksquare$.
   - $x \sim y$, falls x und y die gleiche Höhe haben – dann gilt z. B. $\square \sim \blacktriangle$ und $\square \not\sim \triangle$.
   - $x \sim y$, falls x und y sich höchstens in der Farbe unterscheiden – dann gilt z. B. $\square \sim \blacksquare$ und $\square \not\sim \bullet$.

3. Sei $m \in \mathbb{Z}$ und $\equiv_m := \{\, (x, y) \in \mathbb{Z}^2 : m \mid x - y \,\}$. Dann ist $\equiv_m$ eine Äquivalenzrelation auf $\mathbb{Z}$. Es gilt zum Beispiel:

$$\begin{aligned} 0 \equiv_3 3 \equiv_3 6 \equiv_3 \ 9 \equiv_3 -3 \equiv_3 -6 \equiv_3 15 \equiv_3 \cdots \\ 1 \equiv_3 4 \equiv_3 7 \equiv_3 10 \equiv_3 -2 \equiv_3 -5 \equiv_3 16 \equiv_3 \cdots \\ 2 \equiv_3 5 \equiv_3 8 \equiv_3 11 \equiv_3 -1 \equiv_3 -4 \equiv_3 17 \equiv_3 \cdots \end{aligned}$$

4. Die Kantenmenge V des Graphen $G = (E, V)$, der durch folgendes Diagramm gegeben ist, ist eine Äquivalenzrelation auf $E = \{1, 2, 3, 4, 5\}$.



5. Ist A die Menge aller Schüler einer Volksschule und

$$R = \{(x, y) \in A \times A : x \text{ und } y \text{ gehen in dieselbe Klasse}\},$$

   dann ist R eine Äquivalenzrelation auf A. Klassenbildung ist symptomatisch für Äquivalenzrelationen.

Definition 6. Ist $\sim$ eine Äquivalenzrelation auf A und ist $x \in A$, so heißt

$$[x]_\sim := \{y \in A : x \sim y\}$$

die *Äquivalenzklasse* von x (bezüglich $\sim$). Man schreibt $A/\sim := \{\, [x]_\sim : x \in A \,\}$ für die Menge aller Äquivalenzklassen von Elementen von A.

Beispiel. Was sind bei den Äquivalenzrelationen aus dem vorherigen Beispiel die Äquivalenzklassen?

1. Bezüglich $=$ ist die Äquivalenzklasse eines Elementes $x \in A$ genau die Menge $\{x\}$, die nur dieses Element enthält.

2. Die genannten Äquivalenzrelationen auf $A = \{\square, \blacksquare, \square, \square, \blacksquare, \circ, \bullet, \bullet, \bullet, \triangle, \blacktriangle, \blacktriangle, \triangle\}$ teilen A in folgende Äquivalenzklassen auf:

   - gleiche Form: $\{\square, \blacksquare, \square, \square, \blacksquare\}$, $\{\circ, \bullet, \bullet, \bullet\}$, $\{\triangle, \blacktriangle, \blacktriangle, \triangle\}$
   - gleiche Farbe: $\{\square, \square, \circ, \triangle, \triangle\}$, $\{\blacksquare, \blacksquare, \bullet, \bullet, \blacktriangle\}$, $\{\square, \bullet, \blacktriangle\}$
   - gleiche Höhe: $\{\square, \blacksquare, \square, \circ, \triangle, \blacktriangle, \blacktriangle\}$, $\{\square, \blacksquare, \bullet, \bullet, \triangle\}$
   - gleich bis auf Farbe: $\{\square, \blacksquare, \square\}$, $\{\square, \blacksquare\}$, $\{\circ, \bullet\}$, $\{\bullet, \bullet\}$, $\{\triangle, \blacktriangle, \blacktriangle\}$, $\{\triangle\}$

3. Die beiden Zusammenhangskomponenten $\{1, 2\}$ und $\{3, 4, 5\}$ sind die Äquivalenzklassen.

4. Wir haben

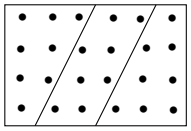
$$\begin{aligned}
[0]_{\equiv_3} &= \{\ldots, -3, 0, 3, 6, 9, \ldots\} \\
[1]_{\equiv_3} &= \{\ldots, -2, 1, 4, 7, 10, \ldots\} \\
[2]_{\equiv_3} &= \{\ldots, -1, 2, 5, 8, 11, \ldots\} \\
[3]_{\equiv_3} &= [0]_{\equiv_3} \\
[4]_{\equiv_3} &= [1]_{\equiv_3}
\end{aligned}$$

   und so weiter.

5. In diesem Fall sind die Äquivalenzklassen genau die Schulklassen.

Satz 3. Es seien A eine Menge, $\sim$ eine Äquivalenzrelation auf A, und $x, y \in A$. Dann gilt:

1. $x \sim y \iff [x]_\sim = [y]_\sim$
2. $x \not\sim y \iff [x]_\sim \cap [y]_\sim = \emptyset$
3. $A = \bigcup_{C \in A/\sim} C$



Beweis.

1. „⇒“ Annahme $x \sim y$, zu zeigen: $[x]_\sim = [y]_\sim$.

   Aus Symmetriegründen genügt es, $[x]_\sim \subseteq [y]_\sim$ zu zeigen.

   Sei also $z \in [x]_\sim$. Dann gilt:

$$\begin{aligned} z \in [x]_\sim &\overset{\text{Def.}}{\Longrightarrow} x \sim z \\ &\overset{\text{Sym.}}{\Longrightarrow} z \sim x \\ &\overset{\text{Ann. + Trans.}}{\Longrightarrow} z \sim y \\ &\overset{\text{Sym.}}{\Longrightarrow} y \sim z \\ &\overset{\text{Def.}}{\Longrightarrow} z \in [y]_\sim. \end{aligned}$$

   „⇐“ Annahme: $[x]_\sim = [y]_\sim$, zu zeigen: $x \sim y$.

   Wegen Reflexivität gilt jedenfalls $x \in [x]_\sim$. Wegen $[x]_\sim = [y]_\sim$ also auch $y \in [x]_\sim$, und damit nach Definition auch $x \sim y$.

2. „⇒“ Annahme: $x \not\sim y$, zu zeigen: $[x]_\sim \cap [y]_\sim = \emptyset$.

   Widerspruchsbeweis: wäre $[x]_\sim \cap [y]_\sim \neq \emptyset$, so gäbe es ein $z \in [x]_\sim \cap [y]_\sim$. Für dieses z gilt dann $z \sim x$ und $z \sim y$, also wegen Symmetrie und Transitivität auch $x \sim y$, im Widerspruch zur Annahme $x \not\sim y$.

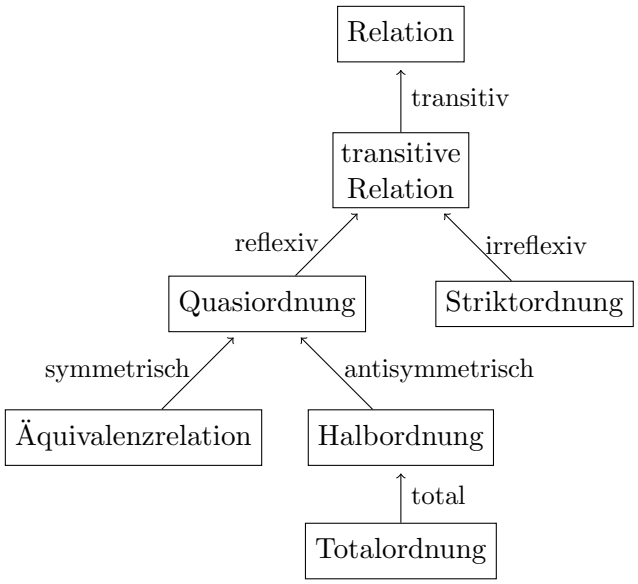
   „⇐“ Annahme: $[x]_\sim \cap [y]_\sim = \emptyset$, zu zeigen: $x \not\sim y$.

   Wäre $x \sim y$, dann wäre $[x]_\sim = [y]_\sim$ nach Teil 1, also $[x]_\sim \cap [y]_\sim = [x]_\sim \neq \emptyset$, da zumindest $x \in [x]_\sim$ (wegen Reflexivität). Damit ist $x \sim y$ ausgeschlossen und es bleibt nur $x \not\sim y$.

3. „⊆“ Sei $y \in A$. Zu zeigen: $y \in \bigcup_{[x]_\sim \in A/\sim} [x]_\sim$. Das folgt direkt aus $y \in [y]_\sim \in A/\sim$.

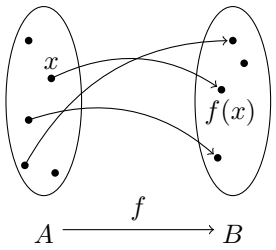
   „⊇“ Sei $y \in \bigcup_{[x]_\sim \in A/\sim} [x]_\sim$. Dann gibt es ein $[x]_\sim \in A/\sim$ mit $y \in [x]_\sim$. Wegen $[x]_\sim \subseteq A$ folgt $y \in A$. ∎

Wir haben in Definition 3 verschiedene Eigenschaften eingeführt, die eine Relation haben kann oder auch nicht. In den Definitionen 4 und 5 haben wir Relationen betrachtet, die mehrere dieser Eigenschaften zugleich erfüllen. Eine Übersicht über weitere Arten von Relationen und ihre Beziehungen zueinander gibt das folgende Diagramm. Ein Pfeil $A \xrightarrow{e} B$ zwischen zwei Relationsarten bedeutet, dass jede Relation vom Typ A eine Relation vom Typ B mit der Eigenschaft e ist.



3 Funktionen

Idee: Objekte anderen Objekten zuordnen.



Definition 7. Seien A, B Mengen und $R \subseteq A \times B$.

1. R heißt *linkseindeutig,* falls gilt:

$$\forall\ (x_1, y_1), (x_2, y_2) \in R : y_1 = y_2 \Rightarrow x_1 = x_2$$

2. R heißt *rechtseindeutig,* falls gilt:

$$\forall\ (x_1, y_1), (x_2, y_2) \in R : x_1 = x_2 \Rightarrow y_1 = y_2$$

3. R heißt *linkstotal,* falls gilt:

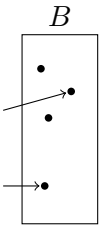
$$\forall\ x \in A\ \exists\ y \in B : (x, y) \in R$$

4. R heißt *rechtstotal,* falls gilt:

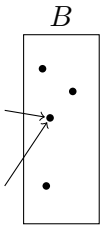
$$\forall\ y \in B\ \exists\ x \in A : (x, y) \in R$$

Beispiel.

1. Linkseindeutig:

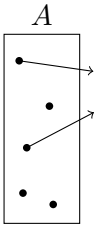


erlaubt

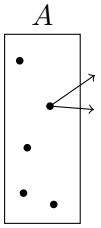


nicht erlaubt

2. Rechtseindeutig:

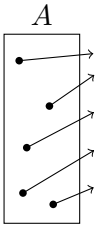


erlaubt

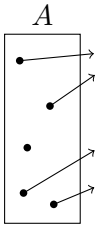


nicht erlaubt

3. Linkstotal:

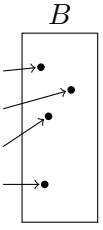


erlaubt

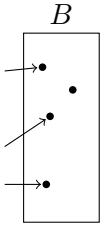


nicht erlaubt

4. Rechtstotal:



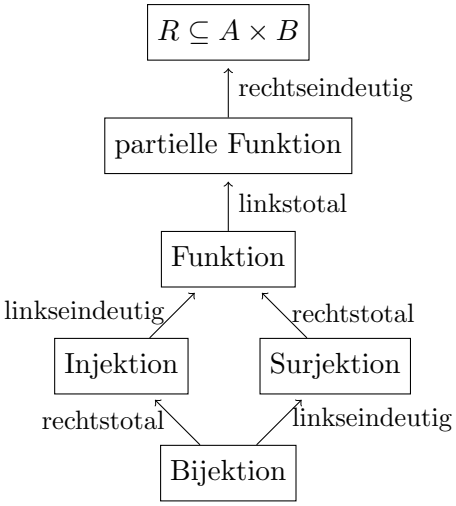
erlaubt



nicht erlaubt

Definition 8. Seien A, B Mengen, $R \subseteq A \times B$.

1. Ist R rechtseindeutig, so heißt R eine *partielle Funktion* und statt $(x, y) \in R$ schreibt man $y = R(x)$.
2. Ist R außerdem linkstotal, so heißt R *Funktion* oder *Abbildung* (engl. *map*) und statt $R \subseteq A \times B$ schreibt man $R\colon A \to B$. Für die Menge aller Funktionen $R\colon A \to B$ schreibt man B^A.
3. Eine linkseindeutige Funktion heißt *injektiv*.
4. Eine rechtstotale Funktion heißt *surjektiv*.
5. Eine Funktion, die sowohl injektiv als auch surjektiv ist, heißt *bijektiv*.



Beispiel.

1. Die *Identitätsfunktion* $\mathrm{id}_A\colon A \to A$ mit $\mathrm{id}_A(x) = x$ für alle $x \in A$ ist eine Funktion.
2. $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{1}{x}$ ist eine partielle Funktion, aber keine Funktion. Man sollte deshalb besser schreiben $f = \{(x, y) \in \mathbb{R}^2 : y = \frac{1}{x}\}$.
3. $f\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = \frac{1}{x}$ ist eine Funktion.

   Diese Funktion ist injektiv. Zum Beweis betrachtet man $x_1, x_2 \in \mathbb{R} \setminus \{0\}$ mit $f(x_1) = f(x_2)$ und zeigt, dass dann $x_1 = x_2$ sein muss. In der Tat bedeutet $f(x_1) = f(x_2)$, dass $\frac{1}{x_1} = \frac{1}{x_2}$, und also $x_1 = x_2$, wie gefordert.
4. $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ ist eine Funktion.

   Diese Funktion ist nicht injektiv, weil z. B. $y = 4$ zwei verschiedene Urbilder hat (nämlich $x = 2$ und $x = -2$).

   Sie ist auch nicht surjektiv, weil z. B. $y = -1$ gar kein Urbild hat.
5. $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = \mathrm{e}^x$ ist injektiv, aber nicht surjektiv.

6. $f\colon \mathbb{R} \to \{\, x \in \mathbb{R} : x > 0 \,\}$, $f(x) = \mathrm{e}^x$ ist injektiv und surjektiv, also bijektiv.

7. $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x + \frac{1}{1+x^2}$ ist surjektiv, aber nicht injektiv.

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

Für eine Funktion ist es nicht wesentlich, dass man sie durch einen Funktionsausdruck beschreiben kann. Tatsächlich lassen sich fast alle Funktionen nicht durch eine endlich große Formel beschreiben. Auch (und vor allem) solche Funktionen sind immer mitgemeint, wenn es heißt „Sei f eine (beliebige) Funktion“.

Eine bijektive Funktion von A nach B existiert offenbar genau dann, wenn A und B gleich viele Elemente haben. Man spricht bei einer Bijektion deshalb auch von einer 1:1-Zuordnung. Für eine Menge A kann man an dieser Stelle die *Mächtigkeit* (engl. *cardinality*) $|A|$ dadurch definieren, dass man sagt $|A| := n$, falls es eine bijektive Abbildung $f\colon A \to \{1, 2, \ldots, n\}$ gibt.

Gibt es für ein bestimmtes $n \in \mathbb{N}$ eine solche Abbildung, so sagt man, A ist *endlich* (engl. *finite*). Anderenfalls sagt man, A ist *unendlich* (engl. *infinite*) und schreibt $|A| = \infty$. Wenn A und B beide endlich sind und $|A| = |B|$ gilt, dann gibt es immer eine Bijektion $f\colon A \to B$. Aber Vorsicht: aus $|A| = |B| = \infty$ folgt im allgemeinen **nicht,** dass es eine Bijektion $f\colon A \to B$ gibt. Man kann zum Beispiel zeigen, dass es keine Bijektion von $\mathbb{Q}$ nach $\mathbb{R}$ gibt, obwohl beide Mengen unendlich sind.

Definition 9. Sind $f\colon A \to B$ und $g\colon B \to C$ Funktionen, so heißt $g \circ f\colon A \to C$ mit $(g \circ f)(x) := g(f(x))$ die *Verkettung* (oder *Komposition*) von f und g.

Satz 4.

1. Die Verkettung injektiver Funktionen ist injektiv.

2. Die Verkettung surjektiver Funktionen ist surjektiv.

3. Die Verkettung bijektiver Funktionen ist bijektiv.

Beweis.

1. Seien $f\colon A \to B$ und $g\colon B \to C$ injektive Funktionen und $h = g \circ f$.

   Zu zeigen: h ist injektiv, also $\forall\, x_1, x_2 \in A : h(x_1) = h(x_2) \Rightarrow x_1 = x_2$.

   Seien also $x_1, x_2 \in A$ mit $h(x_1) = h(x_2)$, d. h. $g(f(x_1)) = g(f(x_2))$. Da g nach Annahme injektiv ist, folgt zunächst $f(x_1) = f(x_2)$. Und daraus folgt, da nach Annahme auch f injektiv ist, $x_1 = x_2$, was zu zeigen war.

2., 3. Übung. ∎

Satz 5. $f\colon A \to B$ ist genau dann bijektiv, wenn es eine Funktion $g\colon B \to A$ gibt mit $g \circ f = \mathrm{id}_A$ und $f \circ g = \mathrm{id}_B$.

Diese Funktion g ist dann eindeutig bestimmt und ihrerseits bijektiv.

Beweis. „⇒“ Annahme: $f\colon A \to B$ ist bijektiv, zu zeigen: Es gibt $g\colon B \to A$ mit $g \circ f = \mathrm{id}_A$.
Betrachte $g = \{\, (y, x) \in B \times A : (x, y) \in f \subseteq A \times B \,\}$.

1. g ist rechtseindeutig, weil f nach Annahme injektiv und damit linkseindeutig ist.
2. g ist linkstotal, weil f nach Annahme surjektiv und damit rechtstotal ist.

Damit ist g eine Funktion.
Für jedes $x \in A$ gilt $(x, f(x)) \in f$ und $(f(x), x) \in g$, also $g(f(x)) = x$. Damit ist gezeigt $g \circ f = \mathrm{id}_A$.
Für jedes $y \in B$ gibt es ein $x \in A$ mit $f(x) = y$, weil f surjektiv ist. Nach Definition von g gilt dann $g(y) = x$ und damit $f(g(y)) = y$. Damit ist gezeigt $f \circ g = \mathrm{id}_B$.

„⇐“ Annahme: Es gibt ein $g\colon B \to A$ mit $g \circ f = \mathrm{id}_A$ und $f \circ g = \mathrm{id}_B$. zu zeigen: f ist bijektiv, d. h. injektiv und surjektiv.
Injektiv: Seien $x_1, x_2 \in A$ mit $f(x_1) = f(x_2)$. Dann gilt $\underbrace{g(f(x_1))}_{=x_1} = \underbrace{g(f(x_2))}_{=x_2}$.
Surjektiv: Sei $y \in B$. Nach Annahme gilt $f(g(y)) = y$, also existiert ein $x \in A$, nämlich $x = g(y)$ mit $f(x) = y$.

Eindeutigkeit: Sind $g, \tilde{g}\colon B \to A$ zwei verschiedene Funktionen mit $g \circ f = \tilde{g} \circ f = \mathrm{id}_A$, dann gibt es zumindest ein $y \in B$ mit $g(y) \neq \tilde{g}(y)$.
Wähle so ein $y \in B$ und setze $x_1 := g(y)$ und $x_2 := \tilde{g}(y)$.
Da f surjektiv ist, existiert $x \in A$ mit $f(x) = y$. Wegen $g \circ f = \tilde{g} \circ f = \mathrm{id}_A$ gilt

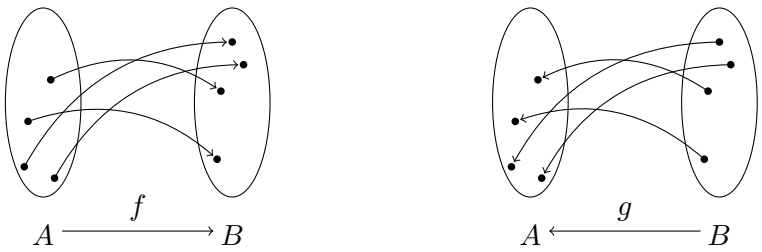
$$\underbrace{g(\underbrace{f(x)}_{=y})}_{=x_1} = x \qquad \text{und} \qquad \underbrace{\tilde{g}(\underbrace{f(x)}_{=y})}_{=x_2} = x,$$

also $x_1 = x_2$. Das ist ein Widerspruch zu der Annahme, dass g und $\tilde{g}$ verschiedene Funktionen sind.

Bijektivität: Zu zeigen ist, dass g injektiv und surjektiv ist.

1. g ist linkseindeutig (also injektiv) weil f rechtseindeutig (da Funktion) ist.
2. g ist rechtstotal (also surjektiv) weil f linkstotal (da Funktion) ist. ∎

Beispiel.



Das Schaubild für g ergibt sich aus dem Schaubild für f, indem man alle Pfeile umkehrt. Nach dem Satz ist f genau dann bijektiv, wenn dabei wieder eine Funktion herauskommt.

Definition 10. Sei $f\colon A \to B$ eine Funktion.

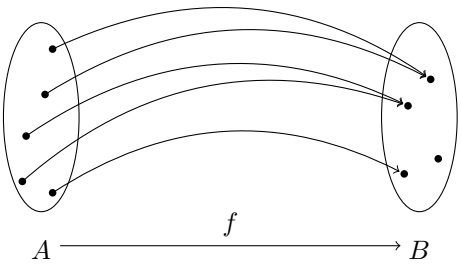
1. Für $U \subseteq A$ definiert man
$$f(U) := \{\, y \in B \mid \exists\, x \in U : f(x) = y \,\} \subseteq B$$
und nennt diese Menge das *Bild* (engl. *image*) von U unter f.

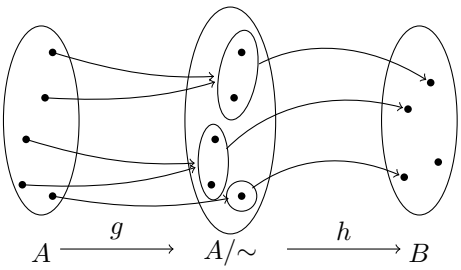
2. Für $V \subseteq B$ definiert man
$$f^{-1}(V) := \{\, x \in A \mid \exists\, y \in V : f(x) = y \,\} \subseteq A$$
und nennt diese Menge das *Urbild* (engl. *preimage*) von V unter f.

3. Ist f bijektiv und $g\colon B \to A$ wie im vorherigen Satz, dann heißt $f^{-1} := g$ die *Umkehrfunktion* (oder *Inverse*) von f.

Satz 6. (Homomorphiesatz für Mengen) Sei $f\colon A \to B$ eine Funktion.

1. Durch $x \sim y \iff f(x) = f(y)$ wird eine Äquivalenzrelation auf A definiert.
2. Die Funktion $g\colon A \to A/\!\sim$, $g(x) = [x]_\sim$ ist surjektiv.
3. Es gibt eine injektive Funktion $h\colon A/\!\sim \to B$ so dass $f = h \circ g$.
4. Diese Funktion h ist eindeutig bestimmt.
5. f ist genau dann surjektiv, wenn h bijektiv ist.





Beweis.

1., 2. Übung

3. Betrachte die Funktion $g\colon A \to A/\!\sim$ mit $g(x) = [x]_\sim$ für alle $x \in A$ und „definiere“ die Funktion $h\colon A/\!\sim \to B$ durch $h([x]_\sim) = f(x)$. Das ist möglich, weil $[x]_\sim = [y]_\sim \iff x \sim y \iff f(x) = f(y)$. (Man sagt, die Definition ist „repräsentantenunabhängig“, oder „h ist wohldefiniert“.)

   h ist injektiv, denn wenn $[x]_\sim, [y]_\sim \in A/\!\sim$ so sind, dass $h([x]_\sim) = h([y]_\sim)$ gilt, dann $f(x) = f(y)$, und dann $x \sim y$, und dann $[x]_\sim = [y]_\sim$.

   Es gilt $f = h \circ g$, denn für alle $x \in A$ gilt $h(g(x)) = h([x]_\sim) = f(x)$.

4. Wenn $\bar{h}\colon A/\!\sim \to B$ eine andere Funktion mit $f = \bar{h} \circ g$ ist, müsste es ein $[x]_\sim \in A/\!\sim$ geben mit $h([x]_\sim) \neq \bar{h}([x]_\sim)$, obwohl doch $h([x]_\sim) = h(g(x)) = f(x) = \bar{h}(g(x)) = \bar{h}([x]_\sim)$ gelten soll.

5. Übung ∎

Beispiel.

1. Seien $A = \{1,2,3,4,5,6\}$, $B = \{a,b,c\}$ und

$$f\colon \begin{array}{c|c|c|c|c|c} 1 & 2 & 3 & 4 & 5 & 6 \\ \hline a & b & a & c & c & a \end{array}.$$

   Dann ist

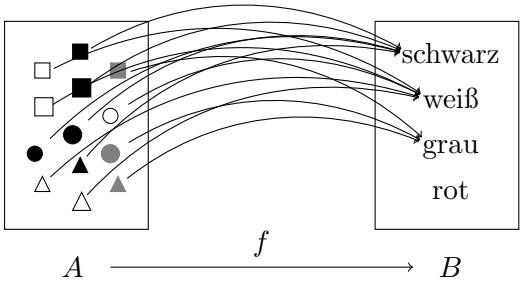
$$A/\!\sim = \big\{\{1,3,6\},\{2\},\{4,5\}\big\}$$

   Die Funktionen g und h aus dem Satz sind gegeben durch

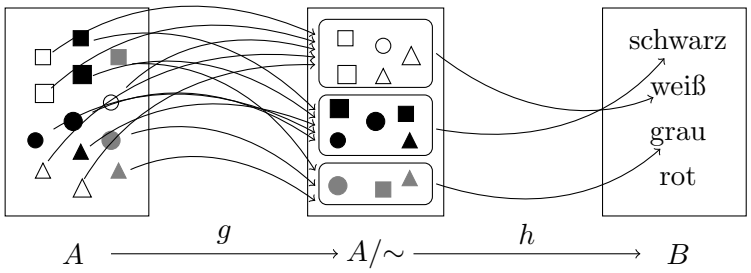
$$g\colon \begin{array}{c|c|c|c|c|c} 1 & 2 & 3 & 4 & 5 & 6 \\ \hline \{1,3,6\} & \{2\} & \{1,3,6\} & \{4,5\} & \{4,5\} & \{1,3,6\} \end{array}$$

$$h\colon \begin{array}{c|c|c} \{1,3,6\} & \{2\} & \{4,5\} \\ \hline a & b & c \end{array}$$

2. Seien $A = \{\square, \blacksquare, \blacksquare, \square, \blacksquare, \circ, \bullet, \bullet, \bullet, \triangle, \blacktriangle, \blacktriangle, \triangle\}$, $B = \{\text{schwarz}, \text{weiß}, \text{grau}, \text{rot}\}$, und sei $f\colon A \to B$ die Funktion, die jedem Element aus A dessen Farbe zuordnet:



Die Zerlegung von f in $f = h \circ g$ sieht dann wie folgt aus:



3. Sei A die Menge der Teilnehmer der Vorlesungsklausur, $B = \{1, 2, 3, 4, 5\}$ die Menge der erreichbaren Noten, $f\colon A \to B$ die Funktion, die jedem Teilnehmer seine Beurteilung zuordnet.

   In diesem Fall gruppiert $\sim$ die Teilnehmer entsprechend ihrer Note. Wenn jede Note mindestens einmal vergeben wird, wenn also f surjektiv ist, dann ist

   $$A/\sim = \{f^{-1}(\{1\}), f^{-1}(\{2\}), f^{-1}(\{3\}), f^{-1}(\{4\}), f^{-1}(\{5\})\}.$$

   Wenn dagegen z. B. nur die Noten 1 und 5 vergeben werden, dann ist $A/\sim = \{f^{-1}(\{1\}), f^{-1}(\{5\})\}$, weil dann $f^{-1}(\{2\}) = f^{-1}(\{3\}) = f^{-1}(\{4\}) = \emptyset$ keine Äquivalenzklassen sind. (Äquivalenzklassen sind niemals leer.)

   Jedenfalls bildet die Funktion g aus dem Satz jeden Klausurteilnehmer $t \in A$ auf die Menge $[t]_\sim$ aller Teilnehmer ab, die die gleiche Note wie t bekommen. Die Funktion h bildet dann jede dieser Mengen auf die Note ab, die die Teilnehmer dieser Menge bekommen.

4 Gruppen

Definition 11. Sei A eine Menge. Eine Funktion $\circ\colon A \times A \to A$ heißt *Verknüpfung.* Statt $\circ(x, y)$ schreibt man $x \circ y$.
Sei $\circ\colon A \times A \to A$ eine Verknüpfung.

1. $\circ$ heißt *assoziativ,* falls $\forall\ x, y, z \in A : (x \circ y) \circ z = x \circ (y \circ z)$
2. $\circ$ heißt *kommutativ,* falls $\forall\ x, y \in A : x \circ y = y \circ x$
3. $e \in A$ heißt *Neutralelement* (bezüglich $\circ$), falls gilt: $\forall\ x \in A : x \circ e = e \circ x = x$.
4. Ist $e \in A$ ein Neutralelement, so heißt $x \in A$ *invertierbar,* falls

   $$\exists\ y \in A : x \circ y = y \circ x = e.$$

   Ein solches Element y heißt dann ein *Inverses* von x.

Beispiel.

1. Sei $A = \{1, 2, 3\}$ und die Verknüpfung $\circ\colon A \times A \to A$ gegeben durch

$\circ$	1	2	3
1	2	3	1
2	1	3	2
3	3	2	1

   .

   Dann gilt $(1 \circ 1) \circ 2 = 3$ und $1 \circ (1 \circ 2) = 1$. Diese Verknüpfung ist also nicht assoziativ.

2. $+$ und $\cdot$ sind Verknüpfungen auf $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, etc.

   Diese Verknüpfungen sind assoziativ und kommutativ.

   0 ist ein Neutralelement bezüglich $+$ und 1 ist ein Neutralelement bezüglich $\cdot$.

3. Sei A eine Menge und A^A die Menge aller Funktionen $f\colon A \to A$. Dann ist die Komposition eine Verknüpfung auf A^A, die assoziativ aber im allgemeinen nicht kommutativ ist. Die Identitätsfunktion $\mathrm{id}_A \in A^A$ ist ein Neutralelement, und eine Funktion $f \in A^A$ ist genau dann invertierbar, wenn sie bijektiv ist.

4. $\cap$ und $\cup$ sind Verknüpfungen auf $\mathcal{P}(A)$.

Satz 7. Sei $\circ\colon A \times A \to A$ eine assoziative Verknüpfung.

1. Sind e_1, e_2 Neutralelemente von $\circ$, so gilt $e_1 = e_2$.

2. Ist e ein Neutralelement von $\circ$, $x \in A$ invertierbar, und sind y_1, y_2 Inverse von x, so gilt $y_1 = y_2$.

   Notation: $x^{-1} := y_1 = y_2$.

3. Ist $x \in A$ invertierbar, so ist auch x^{-1} invertierbar und es gilt $(x^{-1})^{-1} = x$.

4. Sind $x, y \in A$ invertierbar, so ist auch $x \circ y$ invertierbar und es gilt $(x \circ y)^{-1} = y^{-1} \circ x^{-1}$.

Beweis.

1. Nach Definition gilt $\forall\, x \in A : x \circ e_1 = e_1 \circ x = x$ und $\forall\, x \in A : x \circ e_2 = e_2 \circ x = x$.

   Aus dem ersten folgt mit $x = e_2$, dass $e_2 \circ e_1 = e_1 \circ e_2 = e_2$ ist, und aus dem zweiten folgt mit $x = e_1$, dass $e_1 \circ e_2 = e_2 \circ e_1 = e_1$ ist.

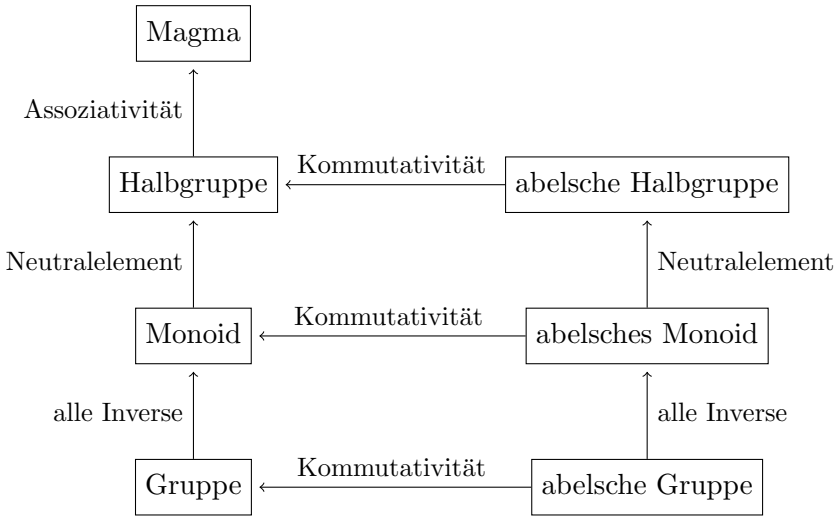
   Aus beidem zusammen folgt $e_1 = e_1 \circ e_2 = e_2$, wie behauptet.

2. Es gilt $x \circ y_1 = e$, also $y_2 \circ (x \circ y_1) = y_2 \circ e$, also $(y_2 \circ x) \circ y_1 = y_2$, also $e \circ y_1 = y_2$, also $y_1 = y_2$.

3., 4. Übung.

Definition 12. Sei A eine Menge und $\circ\colon A \times A \to A$ eine Verknüpfung.

1. Das Paar $(A, \circ)$ heißt *Magma.*
2. Ist $\circ$ assoziativ, so heißt $(A, \circ)$ eine *Halbgruppe* (engl. *semi group*).
3. Eine Halbgruppe mit Neutralelement heißt *Monoid.*
4. Ein Monoid, in dem jedes Element invertierbar ist, heißt *Gruppe* (engl. *group*).
5. Ein[e] Halbgruppe/Monoid/Gruppe heißt *abelsch,* wenn $\circ$ kommutativ ist.



Typischerweise verwendet man $\circ$ und $*$ als Symbole für Verknüpfungen, wenn man allgemein über Gruppen spricht. Bei konkreten Beispielen für Gruppen ist es üblich, das Symbol $+$ für die Verknüpfung zu wählen, wenn es sich um eine abelsche Gruppe handelt. In diesem Fall schreibt man $-x$ statt x^{-1} für das Inverse von x, und man kann auch z. B. $5x$ statt $x + x + x + x + x$ schreiben. Für das Neutralelement verwendet man dann typischerweise das Symbol 0 (Null).

Handelt es sich nicht um eine abelsche Gruppe, und sind $\circ$ und $*$ zu unhandlich, dann kann man auch den Multiplikationspunkt $\cdot$ als Verknüpfungssymbol verwenden. Statt $x \cdot y$ schreibt man dann auch einfach xy, und statt $xxxxx$ schreibt man dann x^5. Für das Neutralelement bietet sich in diesem Fall das Symbol 1 (Eins) an.

Ist $(G, \circ)$ eine Gruppe, dann nennt man G die *Trägermenge* der Gruppe. Wenn die Verknüpfung aus dem Zusammenhang klar ist, sagt man auch einfach, „G ist eine Gruppe“.

Beispiel.

1. Die erste Verknüpfung im vorherigen Beispiel ist nur ein Magma.
2. $(\{3, 4, 5, \dots\}, +)$ ist eine Halbgruppe, aber kein Monoid.
3. $(\mathbb{N}, +)$, $(\mathbb{N} \setminus \{0\}, \cdot)$, $(\mathbb{Z} \setminus \{0\}, \cdot)$ sind Monoide. 05-06
4. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{Q} \setminus \{0\}, \cdot)$, $(\mathbb{R}, +)$, $(\mathbb{R} \setminus \{0\}, \cdot)$, etc. sind abelsche Gruppen.

5. Die Menge $\{ x \in \mathbb{Q} : x > 0 \}$ aller positiven rationalen Zahlen bildet zusammen mit der Multiplikation eine Gruppe. Ebenso die Menge aller positiven reellen Zahlen.

6. Ist A eine Menge, so sind $(\mathcal{P}(A), \cup)$, $(\mathcal{P}(A), \cap)$ Monoide, aber keine Gruppen. Definiert man auf $\mathcal{P}(A)$ die Verknüpfung

$$\oplus\colon \mathcal{P}(A) \times \mathcal{P}(A) \to \mathcal{P}(A), \quad U \oplus V := (U \cup V) \setminus (U \cap V),$$

so ist $(\mathcal{P}(A), \oplus)$ eine abelsche Gruppe. Das Neutralelement ist dann die leere Menge $\emptyset$ und jedes Element $U \in \mathcal{P}(A)$ ist zu sich selbst invers.

7. Ist A eine Menge, so ist $(A^A, \circ)$ ein Monoid, aber keine Gruppe. Aber die Menge $S(A) := \{ f \in A^A : f \text{ ist bijektiv} \}$ bildet zusammen mit der Verkettung als Verknüpfung eine Gruppe. Diese Gruppe ist nicht kommutativ. Man nennt sie Gruppe die *symmetrische Gruppe.* Statt $S(\{1, 2, \ldots, n\})$ schreibt man auch S_n und nennt dann die Elemente *Permutationen.* Bsp.:

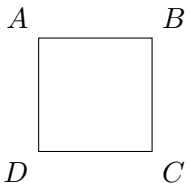
$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 5 & 2 \end{pmatrix}, \quad \pi^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 3 & 1 & 4 \end{pmatrix}$$

und

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$
$$\neq \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Mehr über Permutationen in Abschnitt 12.

8. Betrachte ein Quadrat mit den Eckpunkten A, B, C, D.



Eine Symmetrie lässt sich auffassen als eine Funktion $\{A, B, C, D\} \to \{A, B, C, D\}$, die das Quadrat als Ganzes fest lässt.

Das Quadrat hat folgende Symmetrien:

- $\sigma = \begin{pmatrix} A & B & C & D \\ D & C & B & A \end{pmatrix}$ – das ist die Spiegelung an der horizontalen Achse
- $\rho = \begin{pmatrix} A & B & C & D \\ B & C & D & A \end{pmatrix}$ – das ist die Rotation um 90° entgegen dem Uhrzeigersinn

sowie einige weitere, die sich aus diesen durch Komposition bilden lassen. Die komplette Liste lautet:

$$G = \{\mathrm{id}, \rho, \rho^2, \rho^3, \sigma, \sigma\rho, \sigma\rho^2, \sigma\rho^3\}.$$

Diese Menge G bildet zusammen mit der Komposition eine Gruppe, die sogenannte *Symmetriegruppe* des Quadrats. Die Gruppe ist nicht abelsch; es gilt aber zum Beispiel $\sigma\rho = \rho^3\sigma$.

9. Sei $m \in \mathbb{N} \setminus \{0\}$ und definiere $\equiv_m$ durch $a \equiv_m b \iff m \mid a - b$ für alle $a, b \in \mathbb{Z}$. Dies ist eine Äquivalenzrelation, die $\mathbb{Z}$ in m Äquivalenzklassen aufteilt:

$$\mathbb{Z}_m := \mathbb{Z}/\equiv_m = \{\, [0]_{\equiv_m}, [1]_{\equiv_m}, \ldots, [m-1]_{\equiv_m} \}.$$

Definiere $+, \cdot \colon \mathbb{Z}_m \times \mathbb{Z}_m \to \mathbb{Z}_m$ durch

$$[x]_{\equiv_m} + [y]_{\equiv_m} := [x + y]_{\equiv_m}, \qquad [x]_{\equiv_m} \cdot [y]_{\equiv_m} := [x \cdot y]_{\equiv_m}.$$

Diese Definitionen sind repräsentantenunabhängig, d. h. für x, x' und y, y' mit $x \equiv_m x'$ und $y \equiv_m y'$ gilt stets $(x + y) \equiv_m (x' + y')$ und $(x \cdot y) \equiv_m (x' \cdot y')$. (Beweis: Übung.)

$(\mathbb{Z}_m, +)$ ist eine Gruppe.

$(\mathbb{Z}_m \setminus \{[0]_{\equiv_m}\}, \cdot)$ ist ein Monoid. Man kann zeigen, dass dieses Monoid genau dann eine Gruppe ist, wenn m eine Primzahl ist.

10. Betrachte die sechs Funktionen $f_i \colon \mathbb{R} \setminus \{0, 1\} \to \mathbb{R} \setminus \{0, 1\}$ definiert durch

$$f_1(x) = x, \qquad f_2(x) = \frac{1}{1-x}, \quad f_3(x) = \frac{1}{x},$$
$$f_4(x) = \frac{x-1}{x}, \quad f_5(x) = 1 - x, \quad f_6(x) = \frac{x}{x-1}.$$

Die Menge $G = \{f_1, \ldots, f_6\}$ bildet zusammen mit der Komposition eine Gruppe.

11. Sei $G = \{\, x \in \mathbb{R} : -1 < x < 1 \,\}$ und definiere

$$x \oplus y := \frac{x + y}{1 + xy},$$

wobei die Symbole auf der rechten Seite die übliche Bedeutung haben. Dann ist $(G, \oplus)$ eine abelsche Gruppe. Man sieht sofort, dass $\oplus$ kommutativ ist, dass 0 ein neutrales Element ist, und dass $-x$ das Inverse von x ist. Assoziativität lässt sich leicht nachrechnen:

$$(x \oplus y) \oplus z = \frac{\frac{x+y}{1+xy} + z}{1 + \frac{x+y}{1+xy} z} = \frac{x + y + z + xyz}{1 + xy + xz + yz} = \frac{x + \frac{y+z}{1+yz}}{1 + x\frac{y+z}{1+yz}} = x \oplus (y \oplus z).$$

Jetzt muss man sich aber auch noch davon überzeugen, dass $\oplus$ tatsächlich eine Verknüpfung ist. Dazu ist zu zeigen, dass $1 + xy \neq 0$ für alle $x, y \in G$ (sonst wäre für manche $x, y \in G$ der Ausdruck auf der rechten Seite nicht definiert), und dass $|\frac{x+y}{1+xy}| < 1$ für jede Wahl von $x, y \in G$ (sonst würde uns die Verknüpfung von manchen $x, y \in G$ aus der Gruppe herauswerfen).

Zunächst ist klar, dass mit $|x| < 1$ und $|y| < 1$ auch $|xy| < 1$ ist, und damit $xy > -1$, und damit $1 + xy > 0$, also insbesondere $1 + xy \neq 0$.

Als nächstes gilt mit $|x| < 1$ und $|y| < 1$ insbesondere $1 - x > 0$ und $1 - y > 0$, und also auch $(1-x)(1-y) > 0$. Nun ist $(1-x)(1-y) = 1 - x - y + xy$. Also $1 + xy > x + y$, also $1 > \frac{x+y}{1+xy}$ (denn wir haben ja vorher schon gezeigt, dass $1 + xy > 0$ ist). Auf ähnliche Weise zeigt man $\frac{x+y}{1+xy} > -1$.

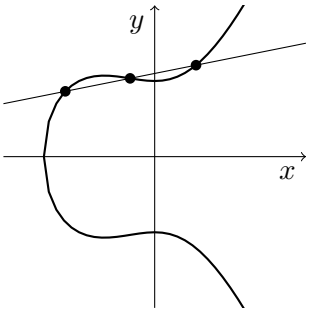
Die Gruppe $(G, \oplus)$ erklärt die Addition von Geschwindigkeiten in der speziellen Relativitätstheorie.

12. Betrachte die Menge

$$E = \{\, (x, y) \in \mathbb{R}^2 : y^2 = x^3 + x^2 + 1 \,\} \cup \{\mathcal{O}\},$$

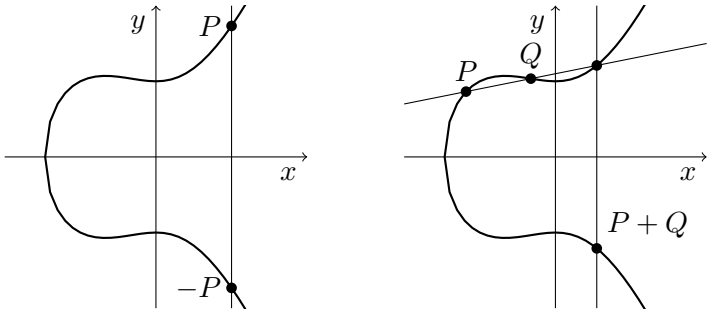
die aus den Punkten einer Kurve in der Ebene sowie dem zusätzlichen Symbol $\mathcal{O}$ besteht. Jede Gerade durch zwei Punkte dieser Kurve schneidet die Kurve in einem dritten Punkt, wenn man

- tangentiale Berührungen doppelt zählt, und
- bei senkrechten Geraden das spezielle Symbol $\mathcal{O}$ als den dritten Punkt ansieht.



Auf E lässt sich eine Verknüpfung $+$ definieren, indem man fordert, dass gelte $P + Q + R = \mathcal{O}$ für je drei Punkte $P, Q, R \in E$, die auf einer Geraden liegen.

Mit dieser Verknüpfung wird E zu einer abelschen Gruppe. Das Neutralelement ist $\mathcal{O}$. Inverse bekommt man durch Spiegelung an der horizontalen Achse. Die Verknüpfung führt man durch, indem man zu gegebenen P und Q zunächst den zugehörigen dritten Punkt bestimmt, und diesen dann noch an der horizontalen Achse spiegelt.



Die Gruppe E ist ein Beispiel für eine sogenannte *elliptische Kurve.* Solche Gruppen werden in der Kryptographie eingesetzt.

13. Seien $A = \{a_1, \ldots, a_n\}$, $\bar{A} = \{\bar{a}_1, \ldots, \bar{a}_n\}$ zwei disjunkte (d. h. $A \cap \bar{A} = \emptyset$) Mengen mit $|A| = |\bar{A}| = n$.

Betrachte die Elemente von $A \cup \bar{A}$ als Buchstaben und Tupel von Elementen als Wörter, z. B. $(a_1, a_3, \bar{a}_1, a_2)$, $(a_3, \bar{a}_1, \bar{a}_2)$, etc.

Es sei $W(A)$ die Menge all solcher Tupel (von beliebiger aber stets endlicher Länge), in denen nie a_i und $\bar{a}_i$ (mit dem selben Index) unmittelbar nebeneinander stehen.

Definiere $\circ\colon W(A) \times W(A) \to W(A)$ so, dass $w_1 \circ w_2$ durch Aneinanderhängen von w_1 und w_2 und anschliessendes Löschen aller Vorkommen $(a_i, \bar{a}_i)$ und $(\bar{a}_i, a_i)$ entsteht, zum Beispiel:

$$(a_1, \bar{a}_2, \bar{a}_3) \circ (a_3, \bar{a}_1, a_2) = (a_1, \bar{a}_2, \bar{a}_1, a_2).$$

Dann ist $(W(A), \circ)$ eine Gruppe, die sogenannte *freie Gruppe* über A.

Das Neutralelement ist das „leere Wort“ ().

Inverse ergeben sich durch Rückwärtslesen des Tupels und Vertauschen aller a_i mit den zugehörigen $\bar{a}_i$, zum Beispiel:

$$(a_1, \bar{a}_2, a_3)^{-1} = (\bar{a}_3, a_2, \bar{a}_1).$$

14. Seien $(A, \circ), (B, *)$ zwei Gruppen [zwei Halbgruppen, zwei Monoide] und $G = A \times B$. Auf G wird durch
$$(a_1, b_1) \odot (a_2, b_2) := (a_1 \circ a_2, b_1 * b_2)$$
eine Verknüpfung $\odot\colon G \times G \to G$ definiert, mit der G zu einer Gruppe [einer Halbgruppe, einem Monoid] wird.

Definition 13. Sei $(G, \circ)$ eine Gruppe. $U \subseteq G$ heißt eine Untergruppe, falls gilt $U \neq \emptyset$ und $\forall\, u, v \in U : u \circ v \in U \wedge u^{-1} \in U$.

Beispiel.

1. $(\mathbb{Z}, +)$ ist eine Untergruppe von $(\mathbb{Q}, +)$; $(\mathbb{Q}, +)$ ist eine Untergruppe von $(\mathbb{R}, +)$.
2. $(\{\, x \in \mathbb{Q} : x > 0 \,\}, \cdot)$ ist eine Untergruppe von $(\mathbb{Q} \setminus \{0\}, \cdot)$; $(\mathbb{Q} \setminus \{0\}, \cdot)$ und $(\{\, x \in \mathbb{R} : x > 0 \,\}, \cdot)$ sind Untergruppen von $(\mathbb{R} \setminus \{0\}, \cdot)$.
3. Die Symmetriegruppe des Rechtecks ist eine Untergruppe der Symmetriegruppe des Quadrats.
4. S_3 lässt sich auffassen als Untergruppe von S_5, wenn man sich jede Funktion
$$f\colon \{1, 2, 3\} \to \{1, 2, 3\}$$
aus S_3 zu einer Funktion $f\colon \{1, 2, 3, 4, 5\} \to \{1, 2, 3, 4, 5\}$ mit $f(4) = 4$ und $f(5) = 5$ fortgesetzt denkt.

Definition 14. Seien $(G_1, \circ)$ und $(G_2, *)$ Gruppen. Eine Funktion $h\colon G_1 \to G_2$ heißt *Homomorphismus,* falls gilt:

$$\forall\, x, y \in G_1 : h(x \circ y) = h(x) * h(y).$$

Sei e_2 das Neutralelement von G_2. Die Menge

$$\ker h := h^{-1}(\{e_2\}) = \{\, x \in G_1 : h(x) = e_2 \,\} \subseteq G_1$$

heißt der *Kern* (engl. *kernel*) und

$$\operatorname{im} h := h(G_1) = \{\, h(x) : x \in G_1 \,\} \subseteq G_2$$

heißt das *Bild* (engl. *image*) von h.

Ein bijektiver Homomorphismus heißt *Isomorphismus.* Wenn es einen Isomorphismus von G_1 nach G_2 gibt, sagt man, G_1 und G_2 sind (zueinander) *isomorph.* Notation in diesem Fall: $G_1 \cong G_2$.

Beispiel.

1. Die Abbildung $h\colon S_3 \to S_5$, die im vorigen Beispiel beschrieben wurde, ist ein Homomorphismus. Statt zu sagen, S_3 ist eine Untergruppe von S_5, wäre es sauberer zu sagen $h(S_3)$ ist eine Untergruppe von S_5.

2. Die Abbildung $f\colon \mathbb{Z} \to \mathbb{Z}_m$, $x \mapsto [x]_{\equiv_m}$ ist ein Homomorphismus zwischen den Gruppen $(\mathbb{Z}, +)$ und $(\mathbb{Z}_m, +)$.

   Es gilt $\ker f = [0]_{\equiv_m} = m\mathbb{Z} = \{0, m, -m, 2m, -2m, \dots\}$.

   Der Homomorphismus ist auch mit der Multiplikation verträglich, d. h. es gilt $[xy]_{\equiv_m} = [x]_{\equiv_m}[y]_{\equiv_m}$ für alle $x, y \in \mathbb{Z}$.

3. Die Abbildung $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto \exp(x)$ ist ein Homomorphismus zwischen $(\mathbb{R}, +)$ und $(\mathbb{R} \setminus \{0\}, \cdot)$.

4. Sind $(A, \circ), (B, *)$ zwei Gruppen und $G = A \times B$ zusammen mit

   $$(a_1, b_1) \odot (a_2, b_2) := (a_1 \circ a_2, b_1 * b_2).$$

   Dann ist $h\colon G \to A$, $(a, b) \mapsto a$ ein Homomorphismus.

5. Die Abbildung $f\colon \mathbb{Z} \to \mathbb{Z}_7 \setminus \{[0]_{\equiv_7}\}$,

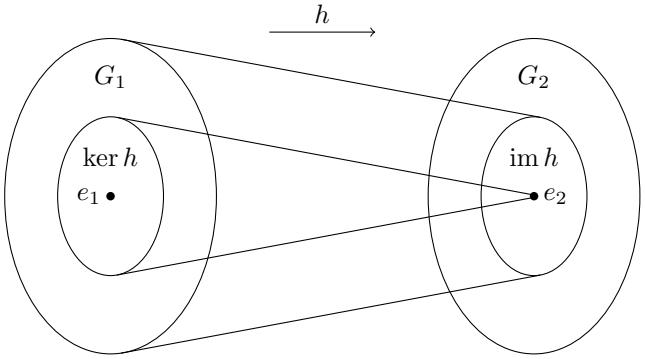
   $$x \mapsto [2]^x_{\equiv_7} := \begin{cases} [2]^x_{\equiv_7} & \text{falls } x > 0 \\ [1]_{\equiv_7} & \text{falls } x = 0 \\ ([2]^{-1}_{\equiv_7})^{|x|} & \text{falls } x < 0 \end{cases}$$

   ist ein Homomorphismus von $(\mathbb{Z}, +)$ nach $(\mathbb{Z}_7 \setminus \{[0]_{\equiv_7}\}, \cdot)$.

   Es gilt $\operatorname{im} f = \{[1]_{\equiv_7}, [2]_{\equiv_7}, [4]_{\equiv_7}\} \subseteq \mathbb{Z}_7 \setminus \{[0]_{\equiv_7}\}$ und $\ker f = \{0, 3, -3, 6, -6, \dots\} \subseteq \mathbb{Z}$.

Satz 8. Sei $h\colon G_1 \to G_2$ ein Homomorphismus. Dann gilt:

1. Ist e_1 das Neutralelement von G_1 und e_2 das Neutralelement von G_2, so gilt $h(e_1) = e_2$.
2. $\ker h$ ist eine Untergruppe von G_1.
3. $\operatorname{im} h$ ist eine Untergruppe von G_2.



Beweis.

1. Es gilt $h(e_1) = h(e_1 \circ e_1) = h(e_1) * h(e_1)$. Multipliziert man diese Gleichung mit dem Inversen von $h(e_1)$ in G_2, so erhält man $e_2 = h(e_1)$.

2. Nach Teil 1 gilt zunächst $e_1 \in \ker h$ und damit insbesondere $\ker h \neq \emptyset$. Darüber hinaus bleibt zu zeigen: für alle $u, v \in \ker h$ gilt $u \circ v \in \ker h$ und $u^{-1} \in \ker h$.

   Seien $u, v \in \ker h$ beliebig. Es gilt $h(u) = h(v) = e_2$, weil $u, v \in \ker h$. Folglich gilt:

   $$h(u \circ v) = h(u) * h(v) = e_2 * e_2 = e_2,$$

   und also $u \circ v \in \ker h$.

   Es gilt $e_2 = h(e_1) = h(u \circ u^{-1}) = h(u) * h(u^{-1}) = e_2 * h(u^{-1}) = h(u^{-1})$.

   (Daraus folgt übrigens auch $h(u)^{-1} = h(u^{-1})$.)

3. Nach Teil 1 gilt zunächst $e_2 \in \operatorname{im} h$. Darüber hinaus bleibt zu zeigen: für alle $u, v \in \operatorname{im} h$ gilt $u * v \in \operatorname{im} h$ und $u^{-1} \in \operatorname{im} h$.

   Seien $u, v \in \operatorname{im} h$ beliebig. Dann gibt es $a, b \in G_1$ mit $u = h(a)$ und $v = h(b)$.

   Es gilt $u * v = h(a) * h(b) = h(a \circ b) \in \operatorname{im} h$.

   Außerdem $u^{-1} = h(a)^{-1} = h(a^{-1}) \in \operatorname{im} h$. ∎

Satz 9. Sei $(G, +)$ eine abelsche Gruppe und H eine Untergruppe von G, und sei $\sim$ definiert durch $a \sim b \iff a + (-b) \in H$.

Dann ist $\sim$ eine Äquivalenzrelation auf G und $G/H := G/\sim$ zusammen mit $*\colon G/H \times G/H \to G/H$, $[x]_\sim * [y]_\sim := [x + y]_\sim$ ist wieder eine abelsche Gruppe.

Beweis. Dass $\sim$ eine Äquivalenzrelation ist, kann man sich zur Übung selbst überlegen.

Wir zeigen, dass $*$ wohldefiniert ist. Zu zeigen ist, dass für $x, x', y, y' \in G$ mit $x \sim x'$ und $y \sim y'$ gilt $x + y \sim x' + y'$.

Aus $x \sim x'$ und $y \sim y'$ folgt $x + (-x') \in H$ und $y + (-y') \in H$. Da H eine Untergruppe von G ist, folgt $x + (-x') + y + (-y') \in H$. Da G abelsch ist, ist $x + (-x') + y + (-y') = x + y + (-x') + (-y') = (x + y) + (-(x' + y'))$, also gilt $x + y \sim x' + y'$, wie behauptet.

Um schließlich zu zeigen, dass G/H eine Gruppe ist, überzeugt man sich, dass die nötigen Gesetze erfüllt sind. Assoziativität folgt zum Beispiel aus

$$[x] * ([y] * [z]) = [x] * [y + z] = [x + (y + z)] = [(x + y) + z] = [x + y] * [z] = ([x] * [y]) * [z].$$

Die Rechnungen für die anderen Gesetze sind ähnlich und bleiben zur Übung überlassen. ∎

Beispiel. Sei $m \in \mathbb{N} \setminus \{0\}$. Dann ist $m\mathbb{Z} = \{\, mx : x \in \mathbb{Z} \,\} = \{\dots, -m, 0, m, 2m, \dots\}$ eine Untergruppe von $(\mathbb{Z}, +)$. Es gilt $\mathbb{Z}/m\mathbb{Z} = \mathbb{Z}_m$.

Satz 10. (Homomorphiesatz für abelsche Gruppen) Es seien $(G_1, +), (G_2, \oplus)$ abelsche Gruppen und $f\colon G_1 \to G_2$ ein Homomorphismus.

1. Die Funktion
$$g\colon G_1 \to G_1/\ker f, \quad g(x) := [x]_\sim$$
ist ein surjektiver Homomorphismus.

2. Es existiert genau ein injektiver Homomorphismus $h\colon G_1/\ker f \to G_2$ mit $f = h \circ g$.

3. Dieser Homomorphismus h ist genau dann bijektiv, wenn f surjektiv ist.

   Insbesondere gilt immer $G_1/\ker f \cong \operatorname{im} f$.

Beweis. Nach Satz 6 ist bloß noch zu zeigen, dass die Funktionen g und h Homomorphismen sind. (Beachte: $x \sim y \iff x + (-y) \in \ker f \iff f(x + (-y)) = e \iff f(x) \oplus f(-y) = e \iff f(x) \oplus f(y)^{-1} = e \iff f(x) = f(y)$, d. h. die hier verwendete Äquivalenzrelation ist ein Spezialfall der Relation aus Satz 6.)

Seien $x, y \in G$.

Wir zeigen zuerst: $g(x + y) = g(x) * g(y)$. In der Tat folgt $g(x + y) = [x + y]_\sim = [x]_\sim * [y]_\sim$ direkt aus der Definition von $*$ aus Satz 9. Damit ist g ein Homomorphismus.

Wir zeigen nun: $h(x*y) = h(x) \oplus h(y)$. In der Tat gilt: $h([x]_\sim * [y]_\sim) = h([x+y]_\sim) = f(x+y) = f(x) \oplus f(y) = h([x]_\sim) \oplus h([y]_\sim)$. Damit ist h ein Homomorphismus. ∎

Beispiel. Sei $f\colon \mathbb{Z} \to \mathbb{Z}_7$, $f(x) = [3]^x_{\equiv_7}$. Es gilt $\ker f = 6\mathbb{Z}$ und $\operatorname{im} f = \mathbb{Z}_7 \setminus \{[0]_{\equiv_7}\}$. Die Gruppe $(\mathbb{Z}_6, +)$ ist also isomorph zur Gruppe $(\mathbb{Z}_7 \setminus \{[0]_{\equiv_7}\}, \cdot)$ und der Isomorphismus h aus dem Satz erlaubt es, die Multiplikation in $\mathbb{Z}_7$ auf die Addition in $\mathbb{Z}_6$ zurückzuführen.

Wählt man $f\colon \mathbb{Z} \to \mathbb{Z}_7$, $f(x) = [2]^x_{\equiv_7}$, so ist $\ker f = 3\mathbb{Z}$ und $\operatorname{im} f = (\{[1]_{\equiv_7}, [2]_{\equiv_7}, [4]_{\equiv_7}\}, \cdot)$. In diesem Fall ist f nicht surjektiv, und der Satz liefert nur einen Isomorphismus zwischen $(\mathbb{Z}_3, +)$ und der Untergruppe $(\{[1]_{\equiv_7}, [2]_{\equiv_7}, [4]_{\equiv_7}\}, \cdot)$ von $(\mathbb{Z}_7 \setminus \{[0]_{\equiv_7}\}, \cdot)$.

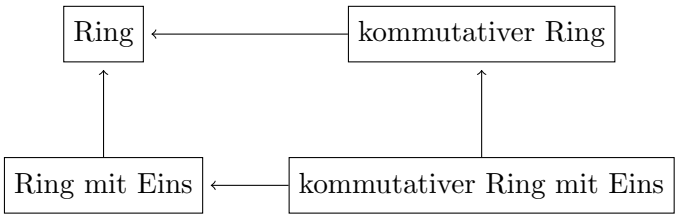
5 Ringe

Definition 15. Sei R eine Menge, $+\colon R \times R \to R$ und $\cdot\colon R \times R \to R$ seien zwei Verknüpfungen, so dass $(R, +)$ eine abelsche Gruppe ist. Ihr Neutralelement nennen wir *Null* (Symbol: 0), und das Inverse von $x \in R$ bezüglich $+$ schreiben wir $-x$. Statt $x + (-y)$ schreibt man $x - y$. Statt $x \cdot y$ schreibt man auch xy.

1. $(R, +, \cdot)$ heißt *Ring,* falls $(R, \cdot)$ eine Halbgruppe ist und gilt
$$\forall\, x, y, z \in R : (x + y) \cdot z = x \cdot z + y \cdot z$$
$$\forall\, x, y, z \in R : x \cdot (y + z) = x \cdot y + x \cdot z.$$

2. Wenn außerdem $(R\setminus\{0\}, \cdot)$ ein Monoid ist, nennt man das Neutralelement *Eins* (Symbol: 1) und $(R, +, \cdot)$ einen *Ring mit Eins.*

3. Ein Ring (mit oder ohne Eins) heißt *kommutativ,* falls $\cdot$ kommutativ ist.



Wenn man in einer Definition einen Begriff einführt, dann ist man normalerweise bemüht, für den Begriff ein Wort aus der Umgangssprache zu verwenden, das den Sachverhalt der Definition treffend beschreibt. Zum Beispiel haben wir vorher definiert, dass eine Relation R „symmetrisch“ zu nennen ist, wenn gilt $xRy \iff yRx$, und „reflexiv“, wenn gilt xRx. Theoretisch hätten wir auch das Wort „reflexiv“ für die Eigenschaft $xRy \iff yRx$ und das Wort „symmetrisch“ für die Eigenschaft xRx vergeben können, es wäre bloß sehr verwirrend, wenn die gemeinte Eigenschaft nicht mit der üblichen Bedeutung des Wortes zusammenpasst.

Der Begriff „Ring“, der oben eingeführt wurde, ist von zentraler Bedeutung in der Algebra. Es besteht allerdings kein offensichtlicher Zusammenhang mit der umgangssprachlichen Bedeutung des Wortes. Man hätte genauso gut „Haus“ oder „Baum“ als Wort verwenden können.

Satz 11. Sei $(R, +, \cdot)$ ein Ring mit Eins, und sei $x \in R$. Dann gilt $x0 = 0 = 0x$ und $-x = (-1)x$.

Beweis. Sei $x \in R$ beliebig.

Es gilt $x0 = x(0 + 0) = x0 + x0$. Addiert man auf beiden Seiten $-(x0)$, so bekommt man $x0 + (-x0) = x0 + x0 + (-x0)$, also $0 = x0 + 0 = x0$, wie behauptet. Der Beweis von $0x = 0$ geht analog.

Als nächstes gilt $0 = 0x = (1+(-1))x = 1x+(-1)x = x+(-1)x$. Damit ist $(-1)x$ ein Inverses von x bezüglich $+$, und wegen der Eindeutigkeit der Inversen muss gelten $(-1)x = -x$. ∎

Beispiel.

1. $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ sind kommutative Ringe mit Eins. Aber $(\mathbb{N}, +, \cdot)$ ist kein Ring, weil $(\mathbb{N}, +)$ keine Gruppe ist.

2. Ist A eine Menge, so ist $(\mathcal{P}(A), \oplus, \cap)$ ein Ring mit Eins. Dabei ist wie zuvor $U \oplus V := (U \cup V) \setminus (U \cap V)$ definiert. Die Null ist $\emptyset$ und die Eins ist A.

3. Für jedes $m \in \mathbb{N} \setminus \{0\}$ ist $(\mathbb{Z}_m, +, \cdot)$ ein kommutativer Ring mit Eins, der sogenannte *Restklassenring* (engl. *residue class ring*) modulo m.

4. Sei $m \in \mathbb{N} \setminus \{0, 1\}$ und $m\mathbb{Z} = \{\, mx : x \in \mathbb{Z} \,\} = \{\dots, -m, 0, m, 2m, \dots\}$. Dann ist $(m\mathbb{Z}, +, \cdot)$ ein kommutativer Ring, aber ohne Eins.

5. Sei $p \in \mathbb{Z} \setminus \{0\}$ eine Primzahl und $\mathbb{Z}_{(p)}$ die Menge aller rationalen Zahlen $u/v \in \mathbb{Q}$ mit $p \nmid v$. Dann ist $\mathbb{Z}_{(p)}$ ein Ring, der sogenannte Ring der *p-adischen ganzen Zahlen.*

6. Sei $(R, +, \cdot)$ ein Ring. Die Menge $\mathbb{R}^{\mathbb{N}}$ aller Funktionen $f\colon \mathbb{N} \to R$ (d. h. aller Folgen in R) bildet einen Ring, wenn man definiert

$$\begin{aligned} (f + g)\colon \mathbb{N} \to R, &\quad (f + g)(n) := f(n) + g(n) \\ (f \cdot g)\colon \mathbb{N} \to R, &\quad (f \cdot g)(n) := f(n)g(n). \end{aligned}$$

   Allgemeiner kann man statt $\mathbb{N}$ hier auch irgendeine andere Menge nehmen.

7. Die Menge $\mathbb{R}^{\mathbb{N}}$ bildet auch einen Ring, wenn man $+$ wie vorher definiert, und $\cdot$ durch

$$(a_0, a_1, a_2, \dots) \cdot (b_0, b_1, b_2, \dots) := (c_0, c_1, c_2, \dots)$$

mit $c_n := \sum_{k=0}^{n} a_k b_{n-k}$. (Beweis durch Nachrechnen der nötigen Gesetze.)

Zum Beispiel gilt

$$(1, 2, 3, \dots) \cdot (1, 2, 3, \dots) := (c_0, c_1, c_2, \dots)$$

mit

$$\begin{aligned} c_0 &= \sum_{k=0}^{0} a_k b_{n-k} = a_0 b_0 = 1 \\ c_1 &= \sum_{k=0}^{1} a_k b_{n-k} = a_0 b_1 + a_1 b_0 = 4 \\ c_2 &= \sum_{k=0}^{2} a_k b_{n-k} = a_0 b_2 + a_1 b_1 + a_2 b_1 = 10. \end{aligned}$$

Definiert man $X := (0, 1, 0, 0, \dots)$, so bietet es sich an, eine Folge $(a_n)_{n=0}^{\infty} \in R^{\mathbb{N}}$ in der Form $\sum_{n=0}^{\infty} a_n X^n$ zu schreiben. Man beachte, dass

$$\begin{aligned} X^0 &= (1, 0, 0, \dots) = 1 \\ X^1 &= (0, 1, 0, 0, \dots) \\ X^2 &= (0, 0, 1, 0, 0, \dots) \\ &\vdots \\ X^n &= (0, \dots, 0, \underset{\substack{\uparrow \\ \text{Index } n}}{1}, 0, 0, \dots). \end{aligned}$$

Die oben definierte Multiplikation entspricht dann genau dem üblichen Multiplikationsgesetz für Potenzreihen, freilich ohne dass dabei irgendwo von Konvergenz die Rede ist.

Statt $R^{\mathbb{N}}$ schreibt man auch $R[[X]]$ und nennt die Elemente *formale Potenzreihen* (engl. *formal power series*).

8. Die Teilmenge

$$R[X] := \left\{ \sum_{n=0}^{\infty} a_n X^n \in R[[X]] : \exists\, N \in \mathbb{N}\ \forall\, n \geq N : a_n = 0 \right\}$$

ist abgeschlossen unter $+$ und $\cdot$ und bildet deshalb selbst auch einen Ring. Man sagt, $R[X]$ ist ein *Unterring* von $R[[X]]$.

Die Elemente von $R[X]$ heißen *Polynome* über R.

Beispiel:

$$\begin{aligned}&(a_0 + a_1X + a_2X^2)(b_0 + b_1X)\\ &= a_0b_0 + (a_0b_1 + a_1b_0)X + (a_2b_0 + a_1b_1)X^2 + a_2b_1X^3\end{aligned}$$

Das Nullpolynom ist das Polynom $p = \sum_{n=0}^{\infty} a_nX^n \in R[X]$ mit $a_n = 0$ für alle $n \in \mathbb{N}$. Für alle anderen Polynome gibt es genau einen Index $n \in \mathbb{N}$, so dass $a_n \neq 0$ und $a_k = 0$ für alle $k > n$ ist. Diesen Index nennt man den *Grad* des Polynoms (engl. *degree*), geschrieben $\deg p := n$. Für das Nullpolynom definiert man $\deg 0 := -\infty$.

Mehr über Polynome in Abschnitt 25.

9. Sei R ein Ring und G ein Monoid.

   Betrachte die Menge $R[G]$ aller Funktionen $c\colon G \to R$ mit der Eigenschaft $c(g) \neq 0$ für höchstens endlich viele $g \in G$.

   Für $c_1, c_2 \in R[G]$ definiere

   $$(c_1 + c_2)\colon G \to R, \quad g \mapsto c_1(g) + c_2(g)$$

   und

   $$(c_1 \cdot c_2)\colon G \to R, \quad g \mapsto \sum_{h_1,h_2 \in G : h_1 \cdot h_2 = g} c_1(h_1) \cdot c_2(h_2).$$

   Dann ist $(R[G], +, \cdot)$ ein Ring.

   Schreibt man die Funktionen $c \in R[G]$ in der Form

   $$c = \alpha_1 \cdot g_1 + \alpha_2 \cdot g_2 + \cdots + \alpha_n \cdot g_n,$$

   wenn $c(g_1) = \alpha_1$, $c(g_2) = \alpha_2$, …, $c(g_n) = \alpha_n$ und $c(g) = 0$ für alle $g \in G \setminus \{g_1, \ldots, g_n\}$, dann entsprechen die obigen Definitionen von $+$ und $\cdot$ gerade den gewohnten Rechnenregeln. (Aber Vorsicht: wenn G nicht abelsch ist, ist $\cdot$ nicht kommutativ!)

   Beispiel. Sei $R = \mathbb{Z}$ und $G = \{a, b, c\}$ mit der Verknüpfung $\circ$, die durch folgende Tabelle definiert ist:

$\circ$	a	b	c
a	a	b	c
b	b	c	a
c	c	a	b

   Dann gilt zum Beispiel

   $$\begin{aligned}(2a + 3b)(3c - 5b) &= 6(a \circ c) - 10(a \circ b) + 9(b \circ c) - 15(b \circ b)\\ &= 6c - 10b + 9a - 15c\\ &= 9a - 10b - 9c.\end{aligned}$$

6 Körper

Definition 16. Ein kommutativer Ring $(K, +, \cdot)$ heißt *Körper* (engl. *field*), falls $(K \setminus \{0\}, \cdot)$ eine abelsche Gruppe ist.

Genau wie im Fall der Ringe ist nicht unmittelbar klar, warum sich der Begriff „Körper“ für diese Struktur eingebürgert hat. Es besteht jedenfalls kein offensichtlicher Zusammenhang zu den Körpern der Geometrie (Würfel, Kugeln, usw.). Der englische Begriff *field* ist auch nicht besonders gut motiviert. Der Hintergrund ist wahrscheinlich einfach, dass es in der Umgangssprache kein Wort gibt, das den Sachverhalt treffend beschreibt.

Satz 12. Sei $(K, +, \cdot)$ ein Körper und $x, y \in K$. Dann gilt $xy = 0 \Rightarrow x = 0 \vee y = 0$.

Beweis. Seien $x, y \in K$ mit $xy = 0$. Zu zeigen, $x = 0$ oder $y = 0$. Wir nehmen an, dass $x \neq 0$ ist, und zeigen, dass dann $y = 0$ sein muss. Da K ein Körper ist, gibt es in $K \setminus \{0\}$ zu jedem Element ein Inverses bezüglich $\cdot$. Da x nach Annahme nicht Null ist, existiert also x^{-1}. Aus $xy = 0$ folgt dann $x^{-1}xy = x^{-1}0$, also $1y = 0$, also $y = 0$, wie behauptet. ∎

Beispiel.

1. $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ sind Körper, aber $(\mathbb{Z}, +, \cdot)$ nicht, da z. B. $2 \in \mathbb{Z}$ nicht bezüglich $\cdot$ invertierbar ist. ($2 \in \mathbb{Q}$ natürlich schon).

2. Man kann zeigen, dass $(\mathbb{Z}_m, +, \cdot)$ genau dann ein Körper ist, wenn m eine Primzahl ist.

3. Sei $\mathbb{Q}(\sqrt{2}) := \{\, a + b\sqrt{2} : a, b \in \mathbb{Q} \,\} \subseteq \mathbb{R}$. Zusammen mit der üblichen Addition und Multiplikation aus $\mathbb{R}$ bildet $\mathbb{Q}(\sqrt{2})$ einen Ring, denn

$$\underbrace{(a + b\sqrt{2})}_{\in\mathbb{Q}(\sqrt{2})} + \underbrace{(c + d\sqrt{2})}_{\in\mathbb{Q}(\sqrt{2})} = \underbrace{(a + c) + (b + d)\sqrt{2}}_{\in\mathbb{Q}(\sqrt{2})}$$

   und

$$\underbrace{(a + b\sqrt{2})}_{\in\mathbb{Q}(\sqrt{2})}\underbrace{(c + d\sqrt{2})}_{\in\mathbb{Q}(\sqrt{2})} = \underbrace{(ac + 2bd) + (bc + ad)\sqrt{2}}_{\in\mathbb{Q}(\sqrt{2})},$$

   d. h. $\mathbb{Q}(\sqrt{2}) \subseteq \mathbb{R}$ ist abgeschlossen unter $+$ und $\cdot$ und damit ein Unterring von $\mathbb{R}$.

   Es ist außerdem ein Körper, denn

$$\begin{aligned}\frac{1}{a + b\sqrt{2}} &= \frac{a - b\sqrt{2}}{(a + b\sqrt{2})(a - b\sqrt{2})} \\ &= \underbrace{\underbrace{\frac{a}{a^2 - 2b^2}}_{\in\mathbb{Q}} - \underbrace{\frac{b}{a^2 - 2b^2}}_{\in\mathbb{Q}}\sqrt{2}}_{\in\mathbb{Q}(\sqrt{2})}.\end{aligned}$$

   Beachte dabei, dass für $a, b \neq 0$ immer gilt $a^2 - 2b^2 \neq 0$, weil $\sqrt{2} \notin \mathbb{Q}$.

4. Sei $\mathbb{C} := \mathbb{R} \times \mathbb{R}$ mit

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

Dann ist $\mathbb{C}$ ein Körper. Man hat dort $\underset{\in\mathbb{C}}{1} = (\underset{\in\mathbb{R}}{1}, 0)$ und definiert $\mathrm{i} := (0,1)$, so dass

$$\mathbb{C} = \{\, a + b\mathrm{i} : a, b \in \mathbb{R} \,\}.$$

Beachte: $\mathrm{i}^2 = (0,1) \cdot (0,1) = (-1,0) = -1$, also „$\mathrm{i} = \sqrt{-1}$“.

Die Elemente von $\mathbb{C}$ nennt man *komplexe Zahlen.* Das Element $\mathrm{i} \in \mathbb{C}$ nennt man die *imaginäre Einheit.* Für eine komplexe Zahl $c = a + b\mathrm{i} \in \mathbb{C}$ mit $a, b \in \mathbb{R}$ nennt man $\mathrm{Re}(c) := a$ den *Realteil* und $\mathrm{Im}(c) := b$ den *Imaginärteil.*

5. Sei K ein Körper und $R = K[X]$ der Ring der Polynome über K. Dieser Ring ist kein Körper, aber man kann sich einen Körper überlegen, der $K[X]$ enthält. Und zwar so:

   Auf $R \times (R \setminus \{0\})$ wird durch

   $$(p_1, p_2) \sim (q_1, q_2) \iff \exists\, u, v \in R \setminus \{0\} : (up_1, up_2) = (vq_1, vq_2)$$

   eine Äquivalenzrelation definiert. (Beweis: Übung.) Die Elemente von $K(X) := (R \times (R \setminus \{0\}))/\!\sim$ heißen *rationale Funktionen.*

   Statt $[(p_1, p_2)]_\sim$ schreibt man $\frac{p_1}{p_2}$ und statt $\frac{p_1}{1}$ einfach p_1.

   Rationale Funktionen sind also Brüche von Polynomen, und die Äquivalenzrelation gewährleistet das Kürzen und Erweitern von Brüchen.

   Beachte: Rationale Funktionen sind **nicht** Funktionen im Sinn von Abschnitt 3, sondern heißen bloß so. In Wahrheit handelt es sich um rein algebraische Gebilde. Es ist deshalb auch egal, ob der Nenner für eine bestimmte „Belegung“ von X mit einem Element von K Null wird. Das Symbol X steht hier nicht für eine Variable, sondern für das Element $(0, 1, 0, 0, \dots) \in K[X]$ (vgl. Bsp. 8 nach Def. 15).

   Jedenfalls gilt: $K(X)$ ist ein Körper, wenn man Addition und Multiplikation definiert durch

   $$\begin{aligned} \frac{p_1}{p_2} + \frac{q_1}{q_2} &= \frac{p_1 q_2 + p_2 q_1}{q_1 q_2}, \\ \frac{p_1}{p_2} \cdot \frac{q_1}{q_2} &= \frac{p_1 p_2}{q_1 q_2}. \end{aligned}$$

   In ganz ähnlicher Weise wie $K(X)$ aus $K[X]$ erzeugt wird, erhält man den Körper $\mathbb{Q}$ aus dem Ring $\mathbb{Z}$.

6. Sei K ein Körper und $R = K[[X]]$ der Ring der formalen Potenzreihen über K. Dieser Ring ist kein Körper, da nicht jedes $a \in R \setminus \{0\}$ ein multiplikatives Inverses in $K[[X]]$ hat. Zum Beispiel ist $a = X$ nicht invertierbar. Man kann aber zeigen, dass eine formale Potenzreihe $\sum_{n=0}^{\infty} a_n X^n$ genau dann invertierbar ist, wenn $a_0 \neq 0$ ist. Das motiviert folgende Konstruktion:

   Betrachte die Menge

   $$K((X)) := \{0\} \cup \Big\{ \sum_{n=0}^{\infty} a_n X^n \in K[[X]] : a_0 \neq 0 \Big\} \times \mathbb{Z}$$

Wir schreiben $X^e \sum_{n=0}^{\infty} a_n X^n$ oder $\sum_{n=e}^{\infty} a_{n-e} X^n$ statt $\left(\sum_{n=0}^{\infty} a_n X^n, e\right)$ und definieren Addition und Multiplikation in der Weise, die durch diese Notation suggeriert wird, also insbesondere $(X^{e_1} a_1) \cdot (X^{e_2} a_2) := X^{e_1+e_2}(a_1 a_2)$, wobei $a_1 a_2$ das Produkt in $K[[X]]$ bezeichnet. (Beachte: der Koeffizient von X^0 in $a_1 a_2$ ist genau das Produkt der Koeffizienten von X^0 in a_1 und a_2, also von Null verschieden.)

Von 0 verschiedene Elemente $X^e a \in K((X))$ kann man dann immer in $K((X))$ invertieren: $(X^e a)^{-1} = X^{-e} a^{-1}$, wobei a^{-1} das multiplikative Inverse von a in $K[[X]]$ ist.

Die Elemente von $K((X))$ heißen *formale Laurent-Reihen*.

$K((X))$ ist ein Körper.

Allgemein bedeutet „K ist ein Körper“, dass man mit den Elementen von K in gewohnter Weise rechnen kann, d. h. dass es in K eine Addition, eine Subtraktion, eine Multiplikation und eine Division gibt, die den gewohnten Rechenregeln gehorchen. Das allein ist für die Theorie der Linearen Algebra entscheidend. Es ist unbedeutend, ob wir in $\mathbb{Q}$ oder in $\mathbb{R}$ oder in irgendeinem anderen Körper rechnen. Wir werden Definitionen und Sätze deshalb für einen beliebigen Körper $\mathbb{K}$ formulieren, anstatt mehrmals für verschiedene konkrete Körper. In Beispielen betrachten wir meist $\mathbb{K} = \mathbb{Q}$ oder $\mathbb{K} = \mathbb{R}$.

Teil II

Vektoren und Matrizen

7 Vektoren

Ab jetzt sei $\mathbb{K} = (\mathbb{K}, +, \cdot)$ immer ein (beliebiger) Körper.

Definition 17. Ein Element von $\mathbb{K}^n = \mathbb{K} \times \cdots \times \mathbb{K}$ heißt *Vektor* (im engeren Sinne, vgl. Def. 29). Die *Addition* von Vektoren ist definiert durch

$$+ \colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}^n, \quad (a_1, \ldots, a_n) \underset{\uparrow}{+} (b_1, \ldots, b_n) := (a_1 \underset{\uparrow}{+} b_1, a_2 + b_2, \ldots, a_n + b_n).$$

Vektor-Addition — Körper-Addition

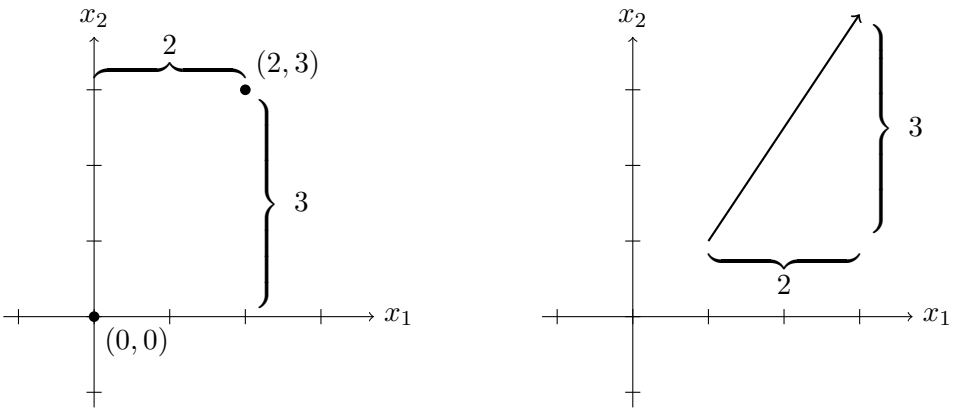
Die *Skalarmultiplikation* ist definiert durch

$$\cdot \colon \mathbb{K} \times \mathbb{K}^n \to \mathbb{K}^n, \quad \alpha \underset{\uparrow}{\cdot} (a_1, \ldots, a_n) := (\alpha \underset{\uparrow}{\cdot} a_1, \alpha\, a_2, \ldots, \alpha\, a_n).$$

Skalarmultipliation — Körper-Multiplikation

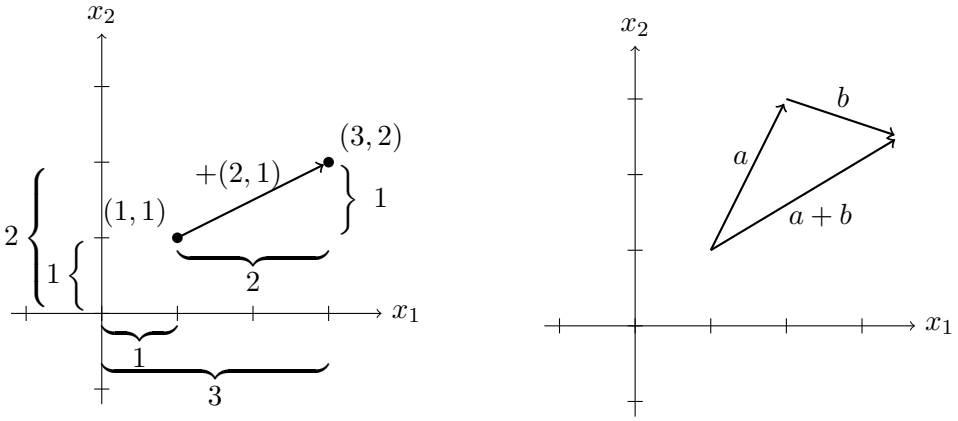
Statt $(a_1, \ldots, a_n)$ schreibt man auch $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$.

Beispiel. $\mathbb{K} = \mathbb{R}$, $n = 2$. Vektoren repräsentieren Punkte in der Ebene oder „Richtungen“.

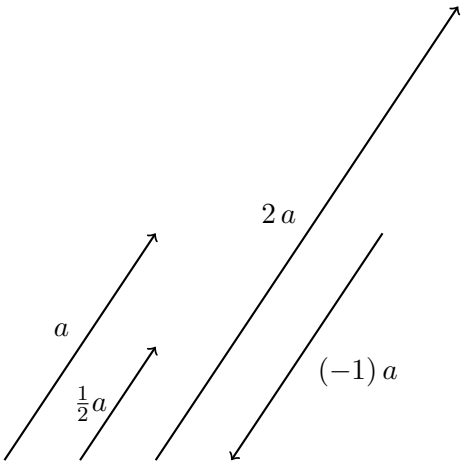


Bei der Interpretation von Richtungen als Pfeilen ist der Startpunkt des Pfeils ohne Bedeutung. Pfeile, die gleich lang sind und in die gleiche Richtung zeigen, veranschaulichen den gleichen Vektor, selbst wenn ihre Startpunkte verschieden sind.

Addition mit einem Vektor entspricht der Verschiebung eines Punktes bzw. der Kombination zweier Richtungen.



Skalarmultiplikation entspricht der Streckung oder Stauchung eines Vektors.



Für $n > 2$ ist die geometrische Anschauung entsprechend, wenn auch weniger angenehm zu zeichnen.

Satz 13. $(\mathbb{K}^n, +)$ ist eine abelsche Gruppe. Ihr Neutralelement ist $0 := (0, \ldots, 0)$.

Beweis. Folgt unmittelbar aus der Definition von $+$ und der Tatsache, dass $(\mathbb{K}, +)$ eine abelsche Gruppe mit Neutralelement 0 ist. ∎

Satz 14. Für alle $\alpha, \beta \in \mathbb{K}$ und alle $v, w \in \mathbb{K}^n$ gilt:

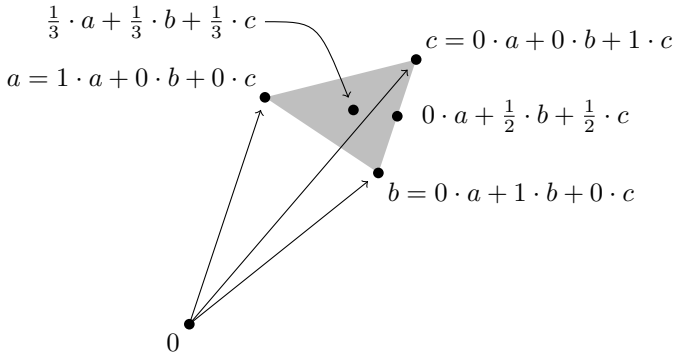
1. $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$
2. $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$
3. $\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$
4. $1 \cdot v = v$, $(-1) \cdot v = -v$
5. $\alpha v = 0 \Rightarrow \alpha = 0 \vee v = 0$

Beweis. Übung. ∎

Mit Vektoren kann man geometrische Objekte beschreiben. Zum Beispiel ist

$$\Delta(a,b,c) := \{\, \alpha a + \beta b + \gamma c : \alpha,\beta,\gamma \in [0,1], \alpha+\beta+\gamma = 1 \,\}$$

das Dreieck mit den Eckpunkten $a, b, c \in \mathbb{R}^n$:

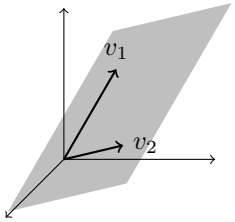


Ein Ausdruck der Form

$$\alpha_1 v_1 + \cdots + \alpha_m v_m$$

mit $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ und $v_1, \ldots, v_m \in \mathbb{K}^n$ heißt *Linearkombination* von $v_1, \ldots, v_m$. Von besonderem Interesse in der linearen Algebra sind die geometrischen Objekte, die aus allen Linearkombinationen von bestimmten gegebenen Vektoren bestehen.

Sind zum Beispiel v_1, v_2 zwei Vektoren im $\mathbb{R}^3$, so ist $\{\, \alpha v_1 + \beta v_2 : \alpha, \beta \in \mathbb{R} \,\}$ eine Ebene durch die Punkte $(0,0)$, v_1 und v_2, ...



... *es sei denn,* dass für ein bestimmtes Paar $(\alpha, \beta) \in \mathbb{R}^2 \setminus \{(0,0)\}$ gilt $\alpha v_1 + \beta v_2 = 0$. In diesem Fall wäre $\{\, \alpha v_1 + \beta v_2 : \alpha, \beta \in \mathbb{R} \,\} = \{\alpha v_1 : \alpha \in \mathbb{R}\}$ bloß eine Gerade durch die Punkte $(0,0)$ und v_1 (und v_2 ein Punkt irgendwo auf dieser Gerade), *es sei denn,* dass sogar $v_1 = v_2 = 0$ ist. In diesem Fall wäre $\{\, \alpha v_1 + \beta v_2 : \alpha, \beta \in \mathbb{R} \,\} = \{(0,0)\}$ bloß ein isolierter Punkt.

Definition 18. $v_1, \ldots, v_m \in \mathbb{K}^n$ heißen *linear abhängig,* falls es $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ gibt, von denen mindestens eins von 0 verschieden ist, und für die gilt $\alpha_1 v_1 + \cdots + \alpha_m v_m = 0$. Anderenfalls heißen $v_1, \ldots, v_m$ *linear unabhängig.*

Beispiel. Die Vektoren

$$\begin{pmatrix}1\\2\\3\end{pmatrix}, \quad \begin{pmatrix}4\\5\\6\end{pmatrix}, \quad \begin{pmatrix}7\\8\\9\end{pmatrix}$$

sind linear abhängig, weil

$$1 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + (-2) \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + 1 \cdot \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Achtung: Aus der paarweisen linearen Unabhängigkeit von v_i und v_j für alle i, j folgt im allgemeinen **nicht,** dass $v_1, \ldots, v_m$ insgesamt linear unabhängig sind. Im vorliegenden Beispiel sind $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und $\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ linear unabhängig, und $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und $\begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}$ sind linear unabhängig, und $\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ und $\begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}$ sind linear unabhängig, aber alle drei Vektoren miteinander sind abhängig.

Für Teilmengen $A \subseteq B \subseteq \mathbb{K}^n$ gilt nur: wenn A linear abhängig ist, dann auch B, und wenn B linear unabhängig ist, dann auch A. Aber aus der linearen Unabhängigkeit von A folgt nichts über die lineare Unabhängigkeit von B, und aus der linearen Abhängigkeit von B folgt nichts über die lineare Abhängigkeit von A (vgl. Satz 38 in Abschnitt 14).

Anschauung: Wenn $v_1, v_2, v_3 \in \mathbb{R}^3$ linear abhängig sind, dann liegt der Punkt v_3 in der Ebene durch die Punkte $(0, 0, 0)$, v_1 und v_2.

Satz 15. Seien $v_1, \ldots, v_m \in \mathbb{K}^n$ und

$$U := \{ \alpha_1 v_1 + \cdots + \alpha_m v_m : \alpha_1, \ldots, \alpha_m \in \mathbb{K} \} \subseteq \mathbb{K}^n$$

die Menge aller Linearkombinationen von $v_1, \ldots, v_m$. Dann ist $(U, +)$ eine Untergruppe von $(\mathbb{K}^n, +)$.

Beweis. Zu zeigen: (a) $U \neq \emptyset$, (b) $\forall\, a, b \in U : a + b \in U$, (c) $\forall\, a \in U : -a \in U$.

zu (a): $0 = 0v_1 + \cdots + 0v_m \in U$.

zu (b): Seien $a, b \in U$. Nach Definition von U gibt es dann $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ und $\beta_1, \ldots, \beta_m \in \mathbb{K}$ mit

$$a = \alpha_1 v_1 + \cdots + \alpha_m v_m$$
$$b = \beta_1 v_1 + \cdots + \beta_m v_m.$$

Daraus folgt durch Addition der beiden Gleichungen

$$a + b = \underbrace{(\alpha_1 + \beta_1)}_{\in \mathbb{K}} v_1 + \cdots + \underbrace{(\alpha_m + \beta_m)}_{\in \mathbb{K}} v_m \in U.$$

zu (c): Sei $a \in U$, etwa $a = \alpha_1 v_1 + \cdots + \alpha_m v_m$ für gewisse $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$. Dann gilt auch $-a = -(\alpha_1 v_1 + \cdots + \alpha_m v_m) = (-\alpha_1)v_1 + \cdots + (-\alpha_m)v_m \in U$. ∎

8 Matrizen

Definition 19. Seien $n, m, k \in \mathbb{N}$. Eine *Matrix* ist ein Element von $\mathbb{K}^{n\times m} := (\mathbb{K}^m)^n$. Schreibweise:

$$A = ((a_{i,j}))_{i=1,j=1}^{n,m} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix} \in \mathbb{K}^{n\times m}.$$

Der Vektor $\begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix} \in \mathbb{K}^n$ heißt die *j-te Spalte* von A, der Vektor $(a_{i,1}, \ldots, a_{i,m}) \in \mathbb{K}^m$ heißt die *i-te Zeile* von A.

Konvention: Bei einem Doppelindex (i, j) bezieht sich immer die erste Komponente auf die Zeile und die zweite auf die Spalte, in der der Matrixeintrag $a_{i,j}$ steht. Insbesondere hat eine $(n \times m)$-Matrix stets n Zeilen und m Spalten.

Für

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix} \in \mathbb{K}^{n\times m} \quad \text{und} \quad B = \begin{pmatrix} b_{1,1} & \cdots & b_{1,k} \\ \vdots & \ddots & \vdots \\ b_{m,1} & \cdots & b_{m,k} \end{pmatrix} \in \mathbb{K}^{m\times k}$$

wird das *Matrixprodukt*

$$A \cdot B := \begin{pmatrix} c_{1,1} & \cdots & c_{1,k} \\ \vdots & \ddots & \vdots \\ c_{n,1} & \cdots & c_{n,k} \end{pmatrix} \in \mathbb{K}^{n\times k}$$

definiert durch $c_{i,j} := \sum_{\ell=1}^{m} a_{i,\ell} b_{\ell,j}$ $(i = 1, \ldots, n,\ j = 1, \ldots, k)$.

Beispiel.

1.
$$\begin{pmatrix} \boxed{1 & 2 & 3} \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} \boxed{1} & -1 \\ \boxed{3} & 0 \\ \boxed{2} & 4 \end{pmatrix} = \begin{pmatrix} \boxed{1\cdot 1 + 2\cdot 3 + 3\cdot 2} & 1\cdot(-1) + 2\cdot 0 + 3\cdot 4 \\ 4\cdot 1 + 5\cdot 3 + 6\cdot 2 & 4\cdot(-1) + 5\cdot 0 + 6\cdot 4 \end{pmatrix}$$
$$= \begin{pmatrix} 13 & 11 \\ 31 & 20 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$

2.
$$\begin{pmatrix} 1 & -1 \\ \boxed{3 & 0} \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & \boxed{3} \\ 4 & 5 & \boxed{6} \end{pmatrix} = \begin{pmatrix} 1\cdot 1 + (-1)\cdot 4 & 1\cdot 2 + (-1)\cdot 5 & 1\cdot 3 + (-1)\cdot 6 \\ 3\cdot 1 + 0\cdot 4 & 3\cdot 2 + 0\cdot 5 & \boxed{3\cdot 3 + 0\cdot 6} \\ 2\cdot 1 + 4\cdot 4 & 2\cdot 2 + 4\cdot 5 & 2\cdot 3 + 4\cdot 6 \end{pmatrix}$$
$$= \begin{pmatrix} -3 & -3 & -3 \\ 3 & 6 & 9 \\ 18 & 24 & 30 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

3. Das Matrixprodukt
$$\begin{pmatrix} 1 & 3 & 0 \\ -1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$
ist nicht definiert. Die Formate passen nicht. In einem Matrixprodukt $A \cdot B$ muss A genau so viele Spalten haben wie B Zeilen hat.

4. Vektoren (d. h. Elemente von $\mathbb{K}^n$) lassen sich als spezielle Matrizen auffassen, je nach Bedarf als Elemente von $\mathbb{K}^{1\times n}$ (dann spricht man von Zeilenvektoren) oder von $\mathbb{K}^{n\times 1}$ (dann spricht man von Spaltenvektoren). Insbesondere kann man Matrizen mit Vektoren multiplizieren, wenn die Formate passen. Zum Beispiel:

$$\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{pmatrix}\begin{pmatrix}1\\ -1\\ 2\end{pmatrix} = \begin{pmatrix}1\cdot 1+2\cdot(-1)+3\cdot 2\\ 4\cdot 1+5\cdot(-1)+6\cdot 2\\ 7\cdot 1+8\cdot(-1)+9\cdot 2\end{pmatrix} = \begin{pmatrix}5\\ 11\\ 17\end{pmatrix}$$

$$(1,-1,2)\cdot\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{pmatrix} = (11,13,15).$$

   Für Vektoren in $\mathbb{K}^n$ wollen wir uns nicht festlegen, ob sie nun Zeilen- oder Spaltenvektoren sein sollen. Es sind einfach Vektoren. Wenn es auf eine bestimmte Interpretation ankommt, wird immer aus dem Zusammenhang klar sein, welche gemeint ist.

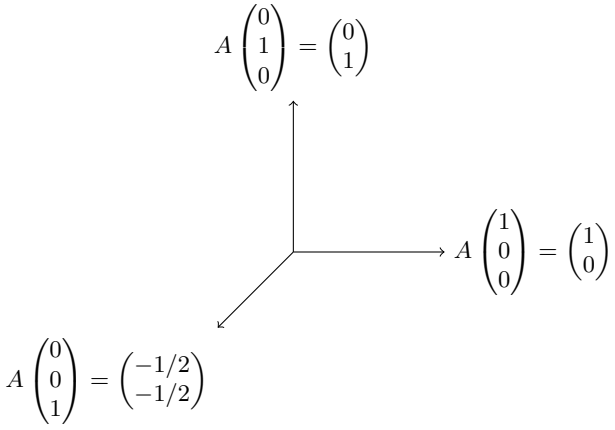
   Insbesondere wollen wir folgende notationelle Konvention machen: Sind $v_1,\dots,v_m \in \mathbb{K}^n$ Vektoren, so bezeichnet $(v_1, v_2, \dots, v_m) \in \mathbb{K}^{n\times m}$ die Matrix, deren Spalten die $v_1,\dots,v_m$ sind. (Dabei werden die v_j also als Spaltenvektoren interpretiert.) Dagegen soll mit der Notation $\begin{pmatrix}v_1\\ \vdots\\ v_m\end{pmatrix} \in \mathbb{K}^{m\times n}$ die Matrix gemeint sein, deren Zeilen die $v_1,\dots,v_m$ sind. (Dabei werden die v_i also als Zeilenvektoren interpretiert.)

Eine Matrix $A \in \mathbb{K}^{n\times m}$ beschreibt eine Funktion $\phi\colon \mathbb{K}^m \to \mathbb{K}^n$, $\phi(v) := A\cdot v$. Diese Funktion „transformiert“ den Raum $\mathbb{K}^m$ auf eine bestimmte Weise in (eine Teilmenge von) $\mathbb{K}^n$. Eine klassische Anwendung ist das Zeichnen von dreidimensionalen Objekten auf zweidimensionalem Papier. Hierbei ist $m = 3$, $n = 2$, und eine mögliche Matrix ist

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

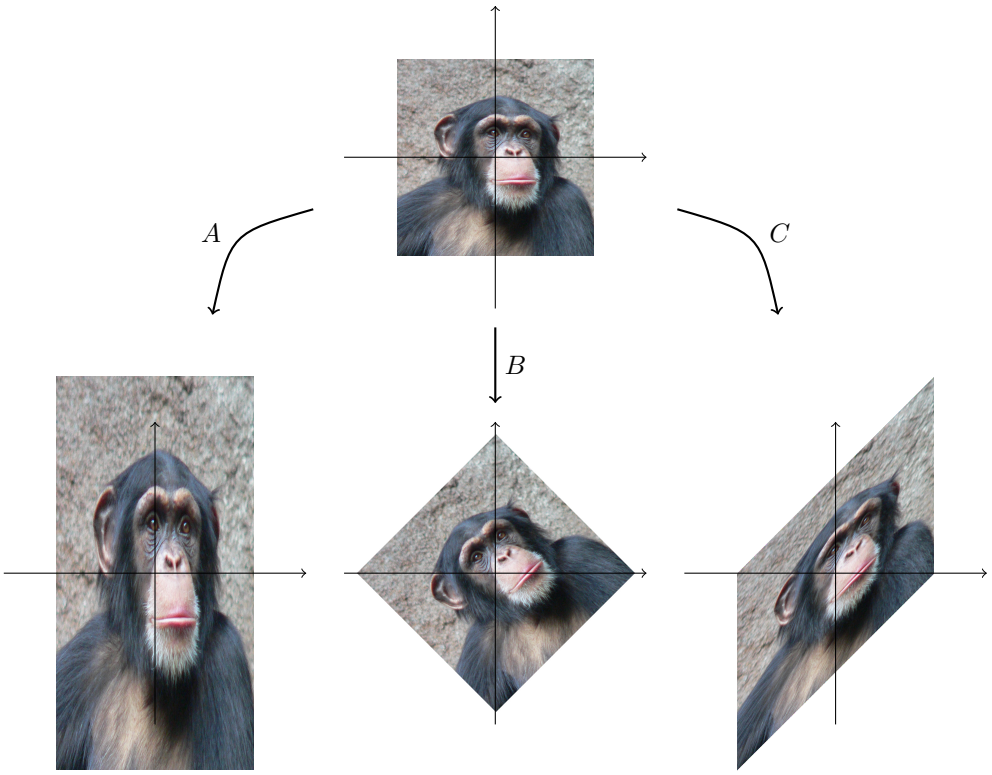
Ein Punkt $(x, y, z) \in \mathbb{R}^3$ wird von A auf den Punkt $(x - \frac{1}{2}z, y - \frac{1}{2}z) \in \mathbb{R}^2$ abgebildet. Beachte, dass mehrere Punkte der Ebene dasselbe Urbild im dreidimensionalen Raum haben koennen. Zum Beispiel ist $(1, 1, 0) \neq (2, 2, 2)$, aber $\phi(1, 1, 0) = \phi(2, 2, 2) = (1, 1)$.

Ist $A \in \mathbb{K}^{n\times m}$ und ist $e_i = (0,\dots,0,1,0,\dots,0)$ der i-te *Einheitsvektor,* also der Vektor, an dessen i-ter Komponente eine 1 steht und dessen andere Komponenten alle 0 sind, so ist $A \cdot e_j$ genau die j-te Spalte von A und $e_i \cdot A$ die i-te Zeile. Die Spalten von A zeigen also an, auf welche Punkte die Einheitsvektoren abgebildet werden.



Weitere Beispiele im Fall $n = m = 2$:

- $A = \begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$ dehnt die Ebene in vertikaler Richtung.
- $B = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}$ dreht die Ebene um 45 Grad entgegen dem Uhrzeigersinn.
- $C = \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$ bewirkt eine sogenannte *Scherung* der Ebene.



Satz 16. Für alle $A \in \mathbb{K}^{n\times m}$, $B \in \mathbb{K}^{m\times k}$, $C \in \mathbb{K}^{k\times \ell}$ gilt $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

Beweis. Seien $A = ((a_{i,j}))_{i=1,j=1}^{n,m}$, $B = ((b_{i,j}))_{i=1,j=1}^{m,k}$, $C = ((c_{i,j}))_{i=1,j=1}^{k,\ell}$. Der (i,j)-te Eintrag von $(A \cdot B) \cdot C$ lautet dann

$$\sum_{u=1}^{k}\left(\sum_{v=1}^{m} a_{i,v}b_{v,u}\right)c_{u,j} = \sum_{u=1}^{k}\sum_{v=1}^{m} a_{i,v}b_{v,u}c_{u,j} = \sum_{v=1}^{m}\sum_{u=1}^{k} a_{i,v}b_{v,u}c_{u,j} = \sum_{v=1}^{m}\left(a_{i,v}\sum_{u=1}^{k} b_{v,u}c_{u,j}\right),$$

und letzteres ist gerade der (i,j)-te Eintrag von $A \cdot (B \cdot C)$. Da alle Einträge von $(A \cdot B) \cdot C$ und $A \cdot (B \cdot C)$ übereinstimmen, folgt die Behauptung. ∎

Daraus folgt insbesondere, dass Matrixmultiplikation der Verkettung der entsprechenden Funktionen entspricht. Außerdem folgt, dass $(\mathbb{K}^{n\times n}, \cdot)$ ein Monoid ist, mit der sogenannten *Einheitsmatrix*

$$I_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

als Neutralelement. Ist es auch eine Gruppe?

Beispiel.

1. $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ ist invertierbar: für $B = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$ gilt
$$A \cdot B = B \cdot A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$

2. $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ ist nicht invertierbar, denn gäbe es eine Inverse $B = \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix}$, dann müsste gelten:
$$\begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} b_{1,1}+b_{1,2} & 0 \\ b_{2,1}+b_{2,2} & 0 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
also
$$\begin{aligned} b_{1,1}+b_{1,2} &= 1 \qquad & 0 &= 0 \\ b_{2,1}+b_{2,2} &= 0 \qquad & 0 &= 1. \end{aligned}$$
Die vierte Gleichung ist offensichtlich nicht zu erfüllen, egal wie wir die $b_{i,j}$ wählen.

Definition 20. $\mathrm{GL}(n,\mathbb{K}) := \{\, A \in \mathbb{K}^{n\times n} : \exists\, B \in \mathbb{K}^{n\times n} : AB = BA = I_n \,\}$ heißt die *lineare Gruppe* der Größe n über dem Körper $\mathbb{K}$.
Die Matrix B mit $AB = BA = I_n$ heißt *Inverse* von A. Notation: $A^{-1} = B$.

Wegen Satz 16 ist klar, dass $\mathrm{GL}(n,\mathbb{K})$ zusammen mit der Matrixmultiplikation eine Gruppe bildet. Diese Gruppe ist nicht kommutativ, z. B. gilt

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix},$$
$$\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}.$$

In diesem Zusammenhang sei an die Rechenregel $(AB)^{-1} = B^{-1}A^{-1}$ erinnert.

Satz 17.

1. $\mathbb{K}^{n\times n}$ bildet zusammen mit der Addition

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix} + \begin{pmatrix} b_{1,1} & \cdots & b_{1,m} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \cdots & b_{n,m} \end{pmatrix} := \begin{pmatrix} a_{1,1}+b_{1,1} & \cdots & a_{1,m}+b_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1}+b_{n,1} & \cdots & a_{n,m}+b_{n,m} \end{pmatrix}$$

und der Matrixmultiplikation einen (nicht-kommutativen!) Ring mit Eins.

2. Für die Skalarmultiplikation

$$\cdot\colon \mathbb{K}\times\mathbb{K}^{n\times m}\to\mathbb{K}^{n\times m},\quad \alpha\begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix} := \begin{pmatrix} \alpha a_{1,1} & \cdots & \alpha a_{1,m} \\ \vdots & \ddots & \vdots \\ \alpha a_{n,1} & \cdots & \alpha a_{n,m} \end{pmatrix}$$

und die Addition gelten die Gesetze $\alpha(A+B)=\alpha A+\alpha B$, $(\alpha+\beta)A=\alpha A+\beta A$, $(\alpha\beta)A=\alpha(\beta A)$, $1A=A$, und $(\alpha A)C=\alpha(AC)$, jeweils für alle $\alpha,\beta\in\mathbb{K}$, $A,B\in\mathbb{K}^{n\times m}$, $C\in\mathbb{K}^{m\times k}$.

Beweis. Übung. ∎

Definition 21. Die *Transposition* von Matrizen ist definiert durch

$$\cdot^\top\colon \mathbb{K}^{n\times m}\to\mathbb{K}^{m\times n},\quad \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix}^\top = \begin{pmatrix} a_{1,1} & \cdots & a_{n,1} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{n,m} \end{pmatrix}.$$

Beispiel. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}^\top = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$.

Satz 18.

1. Für alle $A\in\mathbb{K}^{n\times m}$ und $B\in\mathbb{K}^{m\times k}$ gilt $(AB)^\top=B^\top A^\top$.
2. Für alle $A\in\mathrm{GL}(n,\mathbb{K})$ gilt $(A^\top)^{-1}=(A^{-1})^\top$.

Beweis.

1. Es seien $a_{i,j}$ und $b_{i,j}$ die Einträge von A bzw. B in der i-ten Zeile und j-ten Spalte. Der (i,j)-te Eintrag von $(AB)^\top$ ist nach Definition der Transposition gleich dem (j,i)-ten Eintrag von AB, und dieser ist nach Definition der Matrixmultiplikation gleich

$$\sum_{\ell=1}^{m} a_{j,\ell}b_{\ell,i}.$$

Der (i,j)-te Eintrag von $B^\top A^\top$ ist

$$\sum_{\ell=1}^{m} \underbrace{b_{\ell,i}}_{\substack{\uparrow \\ (i,\ell)\text{-Eintrag von } B^\top}} \underbrace{a_{j,\ell}}_{\substack{\uparrow \\ (\ell,j)\text{-Eintrag von } A^\top}} = \sum_{\ell=1}^{m} a_{j,\ell} b_{\ell,i}.$$

Die Einträge sind also gleich.

2. Es gilt $AA^{-1} = I_n$. Mit Teil 1 des Satzes folgt

$$\underbrace{(AA^{-1})^\top}_{=(A^{-1})^\top A^\top} = I_n^\top = I_n,$$

also $(A^{-1})^\top A^\top = I_n$. ∎

Definition 22. Sei $\pi \in S_n$ eine Permutation. Dann heißt die Matrix $A = ((a_{i,j}))_{i,j=1}^n \in \mathbb{K}^{n\times n}$ mit

$$a_{i,j} = \begin{cases} 1 & \text{falls } \pi(i) = j \\ 0 & \text{sonst} \end{cases}$$

die zu π gehörige *Permutationsmatrix.*

Beispiel. Die zu $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} \in S_4$ gehörige Permutationsmatrix lautet

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Zum Beispiel steht in Zeile 2 eine Eins in der vierten Spalte, weil $\pi(2) = 4$ ist.

Multiplikation einer Permutationsmatrix mit einem Vektor permutiert die Einträge des Vektors:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0a+1b+0c+0d \\ 0a+0b+0c+1d \\ 0a+0b+1c+0d \\ 1a+0b+0c+0d \end{pmatrix} = \begin{pmatrix} b \\ d \\ c \\ a \end{pmatrix}.$$

A ist genau dann eine Permutationsmatrix, wenn in jeder Zeile und in jeder Spalte genau eine Eins und sonst nur Nullen stehen.

Satz 19. Es sei P_π die Permutationsmatrix zu einer Permutation $\pi \in S_n$. Dann gilt:

1. $\forall\ \pi, \sigma \in S_n : P_{\sigma\pi} = P_\pi P_\sigma$,
2. $\forall\ \pi \in S_n : P_{\pi^{-1}} = P_\pi^{-1} = P_\pi^\top$.
3. Für jede Matrix $A = ((a_{i,j}))_{i,j=1}^n$ und jede Permutation $\pi \in S_n$ gilt

$$((a_{\pi(i),\pi(j)}))_{i,j=1}^n = P_\pi A P_\pi^{-1}.$$

Beweis.

1. Seien $\pi, \sigma \in S_n$ beliebig. Nach Definition ist

$$P_\pi = ((a_{i,j}))_{i,j=1}^n \quad \text{mit } a_{i,j} = \begin{cases} 1 & \text{falls } \pi(i) = j \\ 0 & \text{sonst} \end{cases}$$
$$P_\sigma = ((b_{i,j}))_{i,j=1}^n \quad \text{mit } b_{i,j} = \begin{cases} 1 & \text{falls } \sigma(i) = j \\ 0 & \text{sonst} \end{cases}$$

   Sei $P_\pi P_\sigma = ((c_{i,j}))_{i,j=1}^n$. Dann gilt:

$$\begin{aligned} c_{i,j} &= \sum_{k=1}^{n} \underbrace{a_{i,k} b_{k,j}}_{= \begin{cases} 1 & \text{falls } \pi(i) = k \text{ und } \sigma(k) = j \\ 0 & \text{sonst} \end{cases}} \\ &= \begin{cases} 1 & \text{falls } \sigma(\pi(i)) = j \\ 0 & \text{sonst} \end{cases} \\ &= \begin{cases} 1 & \text{falls } (\sigma\pi)(i) = j \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

2. Aus der Definition folgt unmittelbar $P_{\mathrm{id}} = I_n$. Unter Verwendung von Teil 1 erhält man daraus

$$I_n = P_{\mathrm{id}} = P_{\pi\pi^{-1}} = P_{\pi^{-1}} P_\pi,$$

   und damit $P_\pi^{-1} = P_{\pi^{-1}}$.

   Die Behauptung $P_{\pi^{-1}} = P_\pi^\top$ folgt aus $\pi(i) = j \iff \pi^{-1}(j) = i$.

3. Wir verwenden die Notation

$$\delta_{u,v} := \begin{cases} 1 & \text{falls } u = v \\ 0 & \text{sonst} \end{cases}$$

   für $u, v \in \mathbb{N}$. Dann ist $P_\pi = ((\delta_{\pi(i),j}))_{i,j=1}^n$ und nach Teil 2 $P_\pi^{-1} = ((\delta_{i,\pi(j)}))_{i,j=1}^n$. Für den Eintrag an Position (i,j) von $P_\pi A P_\pi^{-1}$ ergibt sich deshalb

$$\underbrace{\sum_{k=1}^{n} \delta_{\pi(i),k} \underbrace{\sum_{l=1}^{n} a_{k,l} \delta_{l,\pi(j)}}_{a_{k,\pi(j)}}}_{= a_{\pi(i),\pi(j)}},$$

   wie behauptet. ∎

Die Abbildung $h\colon S_n \to \mathrm{GL}(n, \mathbb{K})$, $\pi \mapsto P_{\pi^{-1}}$ ist ein injektiver Gruppenhomomorphismus.

9 Gleichungssysteme

Viele Probleme in der linearen Algebra lassen sich zurückführen auf ein Problem von folgendem Typ:

> Gegeben $A \in \mathbb{K}^{n\times m}$, finde alle $x \in \mathbb{K}^m$, so dass $Ax = 0 \in \mathbb{K}^n$.

Man nennt dieses Problem ein *lineares Gleichungssystem* und $L = \{\, x \in \mathbb{K}^m : Ax = 0 \,\}$ dessen Lösungsmenge. Jedes Element $x \in L$ heißt *Lösung* des Gleichungssystems. Man sagt auch, L ist der *Kern* der Matrix A, Notation: $\ker A := L$.

- Wie findet man Lösungen eines linearen Gleichungssystems?
- Was lässt sich über die Struktur der Lösungsmenge sagen?

Schreibe

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix}$$

für die bekannten Daten und

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

für die unbekannten Daten. Dann lautet das Problem also, alle $x_1, \ldots, x_m \in \mathbb{K}$ zu finden mit

$$\begin{aligned} & a_{1,1}x_1 + \cdots + a_{1,m}x_m = 0 \\ \wedge \quad & a_{2,1}x_1 + \cdots + a_{2,m}x_m = 0 \\ \wedge \quad & \cdots \\ \wedge \quad & a_{n,1}x_1 + \cdots + a_{n,m}x_m = 0. \end{aligned}$$

Das System besteht aus n Gleichungen und m Variablen.
Beispiel:

$$\begin{aligned} 2x_1 + 7x_2 &= 0 \\ 3x_1 - 2x_2 &= 0. \end{aligned}$$

Beobachtung: Die Lösungsmenge ändert sich nicht, wenn man

- Gleichungen mit einer von 0 verschiedenen Konstanten multipliziert,
- Das c-fache einer Gleichung zu einer anderen Gleichung dazuaddiert, für ein beliebiges $c \in \mathbb{K}$,
- Gleichungen vertauscht.

All diese Operationen lassen sich durch eine Operation vom gleichen Typ wieder rückgängig machen.

Idee: Verwende diese Operationen, um ein gegebenes Gleichungssystem systematisch in eine Form zu bringen, aus der sich die Lösungen leicht ablesen lassen.

Beispiel.

1.

$$\begin{array}{r|l} \text{I} & 2x_1 + 7x_2 = 0 \\ \text{II} & 3x_1 - 2x_2 = 0 \end{array}$$

$$\overset{\text{II}\to\text{II}-\frac{3}{2}\text{I}}{\iff} \quad \begin{array}{r|l} \text{I} & 2x_1 + 7x_2 = 0 \\ \text{II} & 0x_1 - \frac{25}{2}x_2 = 0 \end{array}$$

$$\overset{\text{II}\to-\frac{2}{25}\text{II}}{\iff} \quad \begin{array}{r|l} \text{I} & 2x_1 + 7x_2 = 0 \\ \text{II} & 0x_1 + 1x_2 = 0 \end{array}$$

$$\overset{\text{I}\to\text{I}-7\text{II}}{\iff} \quad \begin{array}{r|l} \text{I} & 2x_1 + 0x_2 = 0 \\ \text{II} & 0x_1 + 1x_2 = 0 \end{array}$$

$$\overset{\text{I}\to\frac{1}{2}\text{I}}{\iff} \quad \begin{array}{r|l} \text{I} & 1x_1 + 0x_2 = 0 \\ \text{II} & 0x_1 + 1x_2 = 0 \end{array}$$

$$\iff \quad (x_1, x_2) = (0, 0)$$

Also ist in diesem Fall $L = \{\binom{0}{0}\}$ die Lösungsmenge des Gleichungssystems.

Handlichere Schreibweise für die gleiche Rechnung:

$$\begin{pmatrix} 2 & 7 \\ 3 & -2 \end{pmatrix} \begin{array}{l} \urcorner -3/2 \\ \leftarrow\!\lrcorner + \end{array} \quad \leftrightarrow \quad \begin{pmatrix} 2 & 7 \\ 0 & -\frac{25}{2} \end{pmatrix} \begin{array}{l} \\ \mid \cdot -\frac{2}{25} \end{array} \quad \leftrightarrow$$

$$\begin{pmatrix} 2 & 7 \\ 0 & 1 \end{pmatrix} \begin{array}{l} \leftarrow\!\lrcorner + \\ \lrcorner -7 \end{array} \quad \leftrightarrow \quad \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{array}{l} \mid \cdot \frac{1}{2} \\ \\ \end{array} \quad \leftrightarrow \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

2.

$$\begin{array}{r|l} \text{I} & 2x_1 - 3x_2 = 0 \\ \text{II} & -4x_1 + 6x_2 = 0 \end{array}$$

$$\overset{\text{II}\to\text{II}+2\text{I}}{\iff} \quad \begin{array}{r|l} \text{I} & 2x_1 - 3x_2 = 0 \\ \text{II} & 0x_1 + 0x_2 = 0 \end{array}$$

$$\iff \quad 2x_1 - 3x_2 = 0$$

In diesem Fall kann x_2 beliebig gewählt werden, z. B. $x_2 = \alpha \in \mathbb{Q}$, und für jede Wahl gibt es genau eine passende Wahl von x_1, nämlich $x_1 = \frac{3}{2}\alpha$. Die Lösungsmenge hat also die Gestalt

$$L = \{\, \alpha \begin{pmatrix} 3/2 \\ 1 \end{pmatrix} : \alpha \in \mathbb{Q} \,\}.$$

3. Ein Beispiel mit drei Gleichungen und drei Variablen:

$$\begin{array}{r|l} \text{I} & 0x_1 + 4x_2 - x_3 = 0 \\ \text{II} & 1x_1 + 2x_2 + 0x_3 = 0 \\ \text{III} & 1x_1 - x_2 + 2x_3 = 0 \end{array}$$

Wir verwenden gleich die handlichere Schreibweise:

$$\begin{pmatrix} 0 & 4 & -1 \\ 1 & 2 & 0 \\ 1 & -1 & 2 \end{pmatrix} \begin{matrix} \leftarrow \\ \leftarrow \\ \\ \end{matrix} \leftrightarrow \begin{pmatrix} 1 & 2 & 0 \\ 0 & 4 & -1 \\ 1 & -1 & 2 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow + \end{matrix}^{-1} \leftrightarrow$$

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 4 & -1 \\ 0 & -3 & 2 \end{pmatrix} \begin{matrix} \\ \leftarrow + \\ \\ \end{matrix}^{\frac{3}{4}} \leftrightarrow \begin{pmatrix} 1 & 2 & 0 \\ 0 & 4 & -1 \\ 0 & 0 & \frac{5}{4} \end{pmatrix} \; | \cdot \tfrac{4}{5} \leftrightarrow \begin{pmatrix} 1 & 2 & 0 \\ 0 & 4 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \leftarrow + \\ \\ \\ \end{matrix} \leftrightarrow$$

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \leftarrow + \\ | \cdot \frac{1}{4} \\ \\ \end{matrix}^{-2} \leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Die Lösungsmenge ist also offensichtlich $L = \{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \}$.

Definition 23.

1. Eine Matrix A heißt in *Treppenform* (TF) (engl. *echelon form*), falls gilt:

   (a) $A = (1, *, *, \ldots, *) \in \mathbb{K}^{1 \times m}$, oder

   (b) $A = 0$, oder

   (c) $A = \begin{pmatrix} 1 & * & \cdots & * \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$ für eine Matrix $B \in \mathbb{K}^{(n-1)\times(m-1)}$ in Treppenform, oder

   (d) $A = \begin{pmatrix} 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$ für eine Matrix $B \in \mathbb{K}^{n\times(m-1)}$ in Treppenform.

   Dabei stehen die Symbole $*$ für beliebige (nicht notwendigerweise identische) Elemente von $\mathbb{K}$.

2. Ist A in Treppenform, so heißen die Stellen (i, j) mit $a_{i,1} = a_{i,2} = \cdots = a_{i,j-1} = 0$ und $a_{i,j} = 1$ die *Treppenstufen* von A.

3. A heißt *Treppennormalform* (TNF) (engl. *reduced echelon form*), falls A in Treppenform ist und zusätzlich für alle ihre Treppenstufen (i, j) gilt $a_{1,j} = a_{2,j} = \cdots = a_{i-1,j} = 0$.

Beispiel. Eine Matrix in Treppenform:

$$\begin{pmatrix} 1 & * & * & * & * & * & * & * & * & * & * & * & * & * & * \\ 0 & 0 & 1 & * & * & * & * & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Eine Matrix in Treppennormalform:

$$\begin{pmatrix} 1 & * & 0 & * & * & * & 0 & 0 & 0 & * & * & 0 & * & * & * \\ 0 & 0 & 1 & * & * & * & 0 & 0 & 0 & * & * & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & * & * & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & * & * & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

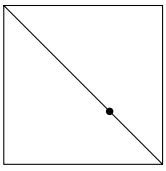
Definition 24. Zwei Matrizen $A, B \in \mathbb{K}^{n\times m}$ heißen *äquivalent,* geschrieben $A \leftrightarrow B$, falls es Matrizen $E_1, \ldots, E_k \in \mathbb{K}^{n\times n}$ gibt, so dass $A = E_k E_{k-1} \cdots E_1 B$, wobei jedes E_i eine der folgenden drei möglichen Formen hat:

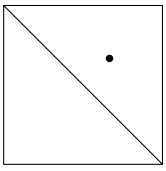
1. Alle Einträge jenseits der Diagonalen sind 0, die Einträge auf der Diagonale sind nicht 0, und mindestens $n-1$ dieser Einträge sind 1.

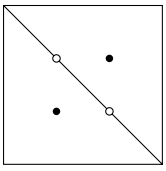
2. Alle Einträge auf der Diagonalen sind 1, die Einträge jenseits der Diagonale sind bis auf höchstens eine Ausnahme 0.

3. Eine Permutationsmatrix für eine Permutation, die $n-2$ Punkte fest lässt und die beiden anderen miteinander vertauscht.

Matrizen dieser Form heißen *Elementarmatrizen.*

1. 

2. 

3. 

Beachte: Die Multiplikation einer Matrix von links mit einer dieser Matrizen entspricht genau der Anwendung der vorher genannten Zeilenoperationen. Multiplikation dieser Matrizen von rechts bewirkt die entsprechende Operation auf den Spalten der Matrix.

Satz 20.

1. $\leftrightarrow$ ist eine Äquivalenzrelation auf $\mathbb{K}^{n\times m}$.
2. Jede Äquivalenzklasse enthält genau eine Matrix in Treppenormalform.
3. Zwei Matrizen haben genau dann den gleichen Kern, wenn sie äquivalent sind.

Beweis. (Skizze)

1. Übung.
2. Dass es mindestens eine TNF gibt, ergibt sich aus dem im folgenden beschriebenen Gauß-Algorithmus, der eine solche Form berechnet.

   Für die Eindeutigkeit argumentiert man per Induktion über die Anzahl der Spalten der Matrix und verwendet die rekursive Struktur der Definition.
3. Man konstruiert eine bijektive Abbildung zwischen den Matrizen in Treppennormalform und den Mengen, die als Lösungsmenge eines Gleichungssystems auftreten können. ∎

Diese Argumente sind zugegebenermaßen nicht das, was wir normalerweise unter einem Beweis verstehen. Das ist unbefriedigend, zumal der Satz durchaus eine zentrale Bedeutung für den Aufbau der linearen Algebra hat und man tunlichts vermeiden sollte, Folgerungen aus Behauptungen zu ziehen, die nicht lückenlos bewiesen wurden. Selbstverständlich kann man den Beweis des Satzes auch nach allen Regeln der Kunst formal sauber aufschreiben. Nur wird er dann recht technisch und länglich, und man sieht nicht wirklich, was vor sich geht. Wir wollen deshalb ausnahmsweise darauf verzichten, und uns bloß anhand von Beispielen von der Richtigkeit der Aussagen überzeugen. Wer dem nicht traut, sollte keine Schwierigkeiten haben, in der einschlägigen Literatur einen ausformulierten formalen Beweis zu finden.

Um die Lösungsmenge eines Gleichungssystems zu bestimmen, bringt man die Matrix zuerst in Treppennormalform. Daraus kann man dann eine explizite Darstellung der Lösungsmenge ablesen. Wir formulieren dieses Vorgehen als drei separate Algorithmen. Der erste bringt eine gegebene Matrix in Treppenform, der zweite eine gegebene Treppenform in Treppennormalform, und der dritte bestimmt aus einer gegebenen Treppennormalform eine Beschreibung des Kerns. Alle drei Algorithmen bezeichnet man als *Gauß-Algorithmus* oder *Gauß-Elimination.*

Algorithmus 1. Eingabe: $A = ((a_{i,j}))_{i=1,j=1}^{n,m} \in \mathbb{K}^{n\times m}$
Ausgabe: Eine zu A äquivalente Treppenform.

```
1   r := 1
2   für c = 1, ..., m:
3     wenn ∃ p ∈ {r, ..., n} : A[p, c] ≠ 0, dann:
4       wähle so ein p
5       wenn p ≠ r, dann vertausche die Zeilen p und r
6       multipliziere die Zeile r mit 1/A[r, c]
7       für i = r + 1, ..., n:
8         addiere das (−A[i, c])-fache der Zeile r zur Zeile i
9       r := r + 1
```

10 gib A als Ergebnis zurück

Der Algorithmus ist so beschrieben, dass er die Einträge der Matrix A im Laufe der Rechnung verändert. Mit der Notation $A[i,j]$ sind die Körperelemente gemeint, die zum aktuellen Zeitpunkt gerade an Position (i,j) in der Matrix stehen. Insbesondere gilt also $A[r,c] \neq 0$ in Schritt 6, denn durch die Vertauschung in Schritt 5 steht jetzt in Schritt r, was vorher in Zeile p stand, und diese Zeile war in Schritt 4 gerade so gewählt, dass $A[p,c] \neq 0$ ist.

Algorithmus 2. Eingabe: $A \in \mathbb{K}^{n\times m}$ in Treppenform
Ausgabe: Eine zu A äquivalente Treppennormalform.

1 für $r = n, \ldots, 1$:

2 wenn $\exists\, j \in \{1,\ldots,m\} : A[r,j] \neq 0$, dann:

3 wähle das kleinste solche j

4 für $i = 1, \ldots, r-1$:

5 addiere das $(-A[i,j])$-fache von Zeile r zur Zeile i

6 gib A als Ergebnis zurück

Beispiel. Wie kommt man von einer TNF zur entsprechenden Lösungsmenge? Betrachte

$$\begin{pmatrix} 1 & 0 & 3 & 0 & 6 & 0 \\ & 1 & 2 & 0 & 5 & 0 \\ & & & 1 & 4 & 0 \\ & & & & & 1 \end{pmatrix}.$$

Als Gleichungssystem geschrieben:

$$\begin{array}{c|cccccccc} \text{I} & x_1 & & +3x_3 & & +6x_5 & & = 0 \\ \text{II} & & x_2 & +2x_3 & & +5x_5 & & = 0 \\ \text{III} & & & & x_4 & +4x_5 & & = 0 \\ \text{IV} & & & & & & x_6 & = 0 \end{array}$$

Wir stellen jede Gleichung nach der ersten vorkommenden Variablen frei und ergänzen zur Verdeutlichung der Situation triviale Gleichungen für die Variablen, die keine eigene Gleichung haben:

$$\begin{aligned} x_1 &= -3x_3 - 6x_5 \\ x_2 &= -2x_3 - 5x_5 \\ x_3 &= x_3 \\ x_4 &= -4x_5 \\ x_5 &= x_5 \\ x_6 &= 0 \end{aligned}$$

In dieser Darstellung sieht man, dass die Lösungsmenge von zwei frei wählbaren Parametern abhängt, die x_3 und x_5 entsprechen. Für jede (beliebige) Wahl von x_3 und x_5 sind die Werte aller anderen Variablen dann eindeutig bestimmt. Die Lösungsmenge lässt sich also schreiben

als

$$L = \Big\{ \alpha \begin{pmatrix} -3 \\ -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -6 \\ -5 \\ 0 \\ -4 \\ 1 \\ 0 \end{pmatrix} : \alpha, \beta \in \mathbb{Q} \Big\}.$$

Trick: Die Vektoren, die in der Beschreibung der Lösungsmenge stehen, kann man wie folgt bekommen. Ergänze die TNF mit zusätzlichen Zeilen der Form $(0, \ldots, 0, -1, 0, \ldots, 0)$, und zwar so, dass eine Matrix entsteht, die unterhalb der Diagonalen lauter Nullen hat, und auf deren Diagonalen nur $+1$ und -1 stehen. Die Lösungsmenge L besteht dann genau aus den Linearkombinationen aller Spalten, bei denen -1 auf der Diagonalen steht.

Im vorliegenden Beispiel ist in der TNF eine solche Zeile zwischen der dritten und der vierten sowie zwischen der vorletzten und der letzten Zeile einzufügen. Man erhält dann

$$\begin{array}{l} \\ \\ \text{neu} \rightarrow \\ \\ \text{neu} \rightarrow \\ \\ \\ \\ \end{array} \begin{pmatrix} 1 & 0 & \boxed{3} & 0 & \boxed{6} & 0 \\ 0 & 1 & 2 & 0 & 5 & 0 \\ \boxed{0} & \boxed{0} & \boxed{-1} & \boxed{0} & \boxed{0} & \boxed{0} \\ 0 & 0 & 0 & 1 & 4 & 0 \\ \boxed{0} & \boxed{0} & \boxed{0} & \boxed{0} & \boxed{-1} & \boxed{0} \\ 0 & 0 & \boxed{0} & 0 & \boxed{0} & 1 \\ & & \downarrow & & \downarrow & \\ & & b_1 & & b_2 & \end{pmatrix}$$

Algorithmus 3. Eingabe: $A \in \mathbb{K}^{n \times m}$ in TNF
Ausgabe: Eine Menge $\{b_1, \ldots, b_k\}$, so dass

$$L = \{ \alpha_1 b_1 + \cdots + \alpha_k b_k : \alpha_1, \ldots, \alpha_k \in \mathbb{K} \}$$

die Lösungsmenge des Gleichungssystems $Ax = 0$ ist.

1 für $r = 1, \ldots, m$:

2 wenn $A[r, r] = 0$, dann:

3 füge $-e_r = (0, \ldots, 0, -1, 0 \ldots, 0) \in \mathbb{K}^m$ als zusätzliche Zeile zwischen der $r$ten und der $(r-1)$ten ein.

4 $B = \emptyset$

5 für $c = 1, \ldots, m$:

6 wenn $A[c, c] = -1$, dann:

7 $$B = B \cup \{ \begin{pmatrix} A[1, c] \\ \vdots \\ A[n, c] \end{pmatrix} \}$$

8 gib B als Ergebnis zurück

Algorithmus 4. (Gauß-Algorithmus)
Eingabe: $A \in \mathbb{K}^{n \times m}$
Ausgabe: Eine Menge $\{b_1, \ldots, b_k\}$, so dass

$$L = \{ \alpha_1 b_1 + \cdots + \alpha_k b_k : \alpha_1, \ldots, \alpha_k \in \mathbb{K} \}$$

die Lösungsmenge des Gleichungssystems $Ax = 0$ ist.

1 Berechne eine Treppenform B von A mit Algorithmus 1

2 Berechne eine Treppennormalform C von B mit Algorithmus 2

3 Berechne $\{b_1, \ldots, b_k\}$ aus C mit Algorithmus 3

4 gib $\{b_1, \ldots, b_k\}$ als Ergebnis zurück.

Beispiel. Es sei die Matrix

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 2 & 2 & 1 & 0 \\ 0 & 1 & 1 & 2 & 1 \\ 2 & 2 & 2 & 1 & 1 \end{pmatrix}$$

gegeben. Die Einträge seien als Elemente von $\mathbb{Z}_3$ zu verstehen, d.h. es gilt z.B. $2 \cdot 2 = 1$ und $1 + 2 = 0$. Gesucht ist der Kern von A.

Schritt 1: Berechnung einer Treppenform von A. Der Algorithmus arbeitet sich dazu von links nach rechts durch die Spalten der Matrix. In jedem Schritt ist nur die jeweils eingerahmte Teilmatrix relevant. Dieser eingerahmte Teil wird in jedem Schritt um eine Spalte kürzer.

$$\begin{pmatrix} \boxed{1} & \boxed{0} & \boxed{1} & \boxed{0} & \boxed{1} \\ \boxed{1} & \boxed{2} & \boxed{2} & \boxed{1} & \boxed{0} \\ \boxed{0} & \boxed{1} & \boxed{1} & \boxed{2} & \boxed{1} \\ \boxed{2} & \boxed{2} & \boxed{2} & \boxed{1} & \boxed{1} \end{pmatrix} \begin{matrix} \\ \overset{2}{\leftarrow}+ \\ \\ \leftarrow+ \end{matrix} \quad \longleftrightarrow \quad \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & \boxed{2} & \boxed{1} & \boxed{1} & \boxed{2} \\ 0 & \boxed{1} & \boxed{1} & \boxed{2} & \boxed{1} \\ 0 & \boxed{2} & \boxed{0} & \boxed{1} & \boxed{2} \end{pmatrix} \quad | \cdot 2$$

$$\longleftrightarrow \quad \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & \boxed{1} & \boxed{2} & \boxed{2} & \boxed{1} \\ 0 & \boxed{1} & \boxed{1} & \boxed{2} & \boxed{1} \\ 0 & \boxed{2} & \boxed{0} & \boxed{1} & \boxed{2} \end{pmatrix} \begin{matrix} \\ \overset{2}{\leftarrow}+ \\ \\ \leftarrow+ \end{matrix} \quad \longleftrightarrow \quad \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 2 & 2 & 1 \\ 0 & 0 & \boxed{2} & \boxed{0} & \boxed{0} \\ 0 & 0 & \boxed{2} & \boxed{0} & \boxed{0} \end{pmatrix} \quad | \cdot 2$$

$$\longleftrightarrow \quad \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 2 & 2 & 1 \\ 0 & 0 & \boxed{1} & \boxed{0} & \boxed{0} \\ 0 & 0 & \boxed{2} & \boxed{0} & \boxed{0} \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow+ \end{matrix} \quad \longleftrightarrow \quad \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 2 & 2 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \boxed{0} & \boxed{0} \end{pmatrix}$$

Schritt 2: Transformation der Treppenform in eine Treppennormalform. Hier geht man von rechts nach links durch die Matrix. Von links nach rechts vorzugehen würde auch zum richtigen Ergebnis führen, ist aber meist mit etwas mehr Rechenaufwand verbunden. Im vorliegenden Beispiel ist es egal, weil es nur eine Treppenstufe gibt (nämlich die in der dritten Zeile), über der Einträge stehen, die von Null verschieden sind. Dieser Rechenschritt könnte theoretisch die Einträge in der eingerahmten Teilmatrix ändern. Im vorliegenden Fall ändern sich nur die Einträge oberhalb der dritten Treppenstufe, weil die Einträge rechts dieser Treppenstufe

zufällig Null sind.

$$\begin{pmatrix} 1 & 0 & \boxed{1} & 0 & \boxed{1} \\ 0 & 1 & \boxed{2} & \boxed{2} & \boxed{1} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \longleftarrow + \\ \longleftarrow + \\ \longleftarrow 2 \end{matrix} \quad \longleftrightarrow \quad \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Schritt 3: Bestimmung der Lösungsvektoren aus der Treppennormalform. Wir müssen die TNF durch hinzufügen geeigneter negativer Einheitsvektoren zu einer Matrix ergänzen, und zwar so, dass

- die Diagonale am Ende so viele Einträge hat wie die Matrix Spalten hat (also hier fünf),
- auf der Diagonalen keine Nullen mehr stehen,
- unterhalb der Diagonalen nur Nullen stehen.

Nullzeilen kann man, wenn man will, streichen. Im vorliegenden Beispiel erhält man also, da wir in $\mathbb{Z}_3$ rechnen und dort $-1 = 2$ gilt, die Matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Diese Matrix ist **nicht** äquivalent zur TNF, sondern nur eine Hilfsmatrix, die es erleichtert, die Lösungsvektoren abzulesen. Dieses sind nämlich genau jene Spalten, bei denen auf der Diagonale -1 steht. Die Lösungsmenge lautet also

$$\ker A = \{\alpha \begin{pmatrix} 0 \\ 2 \\ 0 \\ 2 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 2 \end{pmatrix} : \alpha, \beta \in \mathbb{Z}_3\}.$$

Satz 21. Die Menge $\{b_1, \ldots, b_k\}$, die von Algorithmus 4 berechnet wird, ist linear unabhängig.

Beweis. Wir dürfen annehmen, dass die b_i in der Reihenfolge indiziert sind, in der der Algorithmus sie findet.
Zu zeigen: für alle $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ gilt:

$$\alpha_1 b_1 + \cdots + \alpha_k b_k = \underset{\in \mathbb{K}^m}{0} \quad \Rightarrow \quad \alpha_1 = \cdots = \alpha_k = \underset{\in \mathbb{K}}{0}.$$

Seien also $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ so dass $\alpha_1 b_1 + \cdots + \alpha_k b_k = 0$. Angenommen, nicht alle α_i sind Null. Dann gibt es ein $i \in \{1, \ldots, k\}$, so dass $\alpha_i \neq 0$ und $\alpha_{i+1} = \cdots = \alpha_k = 0$. Wenn $j \in \{1, \ldots, m\}$ der maximale Index ist, so dass die j-te Komponente von b_i von Null verschieden

ist, dann ist diese Komponente -1 und die j-te Komponente der Vektoren $b_1, \dots, b_{i-1}$ ist Null. Das folgt aus Zeile 3 von Algorithmus 3. Die j-te Komponente der Linearkombination $\alpha_1 b_1 + \cdots + \alpha_k b_k$ ist dann $\alpha_1 0 + \cdots + \alpha_{i-1} 0 - \alpha_i + 0(b_{i+1})_j + \cdots + 0(b_k)_j = -\alpha_i \neq 0$. Damit kann die Linearkombination nicht der Nullvektor sein. ∎

Später werden wir sagen, $\{b_1, \dots, b_k\}$ ist eine „Basis“ des Lösungsraums, und $k = |\{b_1, \dots, b_k\}|$ ist dessen „Dimension“.

Algorithmus 4 berechnet also nicht die komplette Lösungsmenge, sondern nur eine endliche Menge von Vektoren, durch die sich alle (evtl. unendlich vielen) Lösungen darstellen lassen. Wozu dann der ganze Aufwand? Die Matrix A selbst ist schließlich auch eine endliche Menge von Vektoren, durch die sich alle Lösungen darstellen lassen: $L = \{\, x \in \mathbb{K}^m : Ax = 0 \,\}$. Warum ist eine Basis besser?

In der Tat kann man nicht pauschal sagen, dass eine Darstellung besser ist als die andere. Es kommt darauf an, was man machen will. Wenn man z. B. einen Vektor $x \in \mathbb{K}^m$ gegeben hat und wissen will, ob er in L liegt, dann ist eine Basis zunächst nicht sehr hilfreich. Einfacher ist es, Ax auszurechnen und zu schauen, ob 0 rauskommt. Umgekehrt, wenn man einen konkreten Lösungsvektor sucht, dann ist A nicht sehr hilfreich, aber eine Basis schon (jedes Basiselement ist ja insbesondere eine Lösung; wähle $\alpha_i = 1$ und alle $\alpha_j = 0$ für $j \neq i$).

Man sagt, „$L = \{\, x \in \mathbb{K}^m : Ax = 0 \,\}$“ ist eine *implizite* Darstellung von L, und man nennt „$L = \{\alpha_1 b_1 + \cdots + \alpha_k b_k : \alpha_1, \dots, \alpha_k \in \mathbb{K}\} = \{\, (b_1, \dots, b_k)x : x \in \mathbb{K}^k \,\}$“ eine *explizite* Darstellung.

Algorithmus 4 ist also ein Algorithmus, der eine implizite Darstellung in eine explizite Darstellung umwandelt. Geht es auch umgekehrt? Klar!

Algorithmus 5. Eingabe: $\{b_1, \dots, b_k\} \subseteq \mathbb{K}^m$
Ausgabe: Eine Matrix $A \in \mathbb{K}^{n \times m}$, so dass $\ker A = \{\, \alpha_1 b_1 + \cdots + \alpha_k b_k : \alpha_1, \dots, \alpha_k \in \mathbb{K} \,\}$.

1 Sei $B = \begin{pmatrix} b_1 \\ \vdots \\ b_k \end{pmatrix} \in \mathbb{K}^{k \times m}$.

2 Berechne $a_1, \dots, a_n \in \mathbb{K}^m$, so dass

$$\{\, x \in \mathbb{K}^m : Bx = 0 \,\} = \{\, \alpha_1 a_1 + \cdots + \alpha_n a_n : \alpha_1, \dots, \alpha_n \in \mathbb{K} \,\}.$$

3 gib $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{K}^{n \times m}$ als Ergebnis zurück.

Um die Korrektheit dieses Algorithmus zu zeigen, muss man beweisen

$$\begin{aligned} & \{\, x \in \mathbb{K}^m : Bx = 0 \,\} = \{\, \alpha_1 a_1 + \cdots + \alpha_n a_n : \alpha_1, \dots, \alpha_n \in \mathbb{K} \,\} \\ \Rightarrow \quad & \{\, x \in \mathbb{K}^m : Ax = 0 \,\} = \{\, \alpha_1 b_1 + \cdots + \alpha_k b_k : \alpha_1, \dots, \alpha_n \in \mathbb{K} \,\}. \end{aligned}$$

„$\supseteq$“ Nach Annahme gilt

$$\underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_k \end{pmatrix}}_{\in \mathbb{K}^{k \times m}} \underbrace{(a_1, \dots, a_n)}_{\in \mathbb{K}^{m \times n}} = 0 \in \mathbb{K}^{k \times n} \quad \Rightarrow \quad \underbrace{\left(\begin{pmatrix} b_1 \\ \vdots \\ b_k \end{pmatrix} (a_1, \dots, a_n) \right)^\top}_{= \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} (b_1, \dots, b_k)} = 0 \in \mathbb{K}^{n \times k},$$

$$= A$$

also gilt $Ab_i = 0$ für alle i, und damit auch

$$\begin{aligned} &A \cdot (\alpha_1 b_1 + \cdots + \alpha_k b_k) \\ &= \alpha_1 A b_1 + \cdots + \alpha_k A b_k \\ &= \alpha_1 0 + \cdots + \alpha_k 0 = 0 \end{aligned}$$

für jede Wahl von $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$.

„$\subseteq$“ Diese Richtung ist nicht ganz so offensichtlich. Wir werden in Abschnitt 16 darauf zurückkommen.

10 Lineare Unabhängigkeit und Rang

Satz 22. Seien $b_1, \ldots, b_m \in \mathbb{K}^n$ und sei $T := \begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix} \in \mathbb{K}^{m \times n}$ eine Treppenform von $B := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^{m \times n}$. Dann gilt: $b_1, \ldots, b_m$ sind genau dann linear abhängig, wenn $t_m = (0, \ldots, 0)$ ist.

Beweis. „$\Leftarrow$“ $t_m = (0, \ldots, 0)$. Es gibt Elementarmatrizen $E_1, \ldots, E_k$, so dass

$$T = E_k \cdots E_1 B.$$

Wenn $U = E_k \cdots E_1$ ist und $(u_{m,1}, \ldots, u_{m,m})$ der m-te Zeilenvektor von U, dann gilt also $0 = t_m = u_{m,1} b_1 + \cdots + u_{m,m} b_m$.

Da die E_i invertierbare Matrizen sind, ist auch U eine invertierbare Matrix. Als solche kann U keine Nullzeile enthalten, denn es muss ja $U^{-1}U = I_m$ gelten, und wäre z.B. die m-te Zeile von U komplett Null, so wäre auch die m-te Zeile von I_m komplett Null, was nicht der Fall ist.

Es gilt also, dass $(u_{m,1}, \ldots, u_{m,m})$ nicht der Nullvektor ist, und also $b_1, \ldots, b_m$ linear abhängig sind.

„$\Rightarrow$“ $b_1, \ldots, b_m$ sind linear abhängig, etwa

$$\alpha_1 b_1 + \cdots + \alpha_m b_m = 0$$

für gewisse $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$, von denen nicht alle Null sind.

Es gilt also

$$\begin{aligned} (\alpha_1, \ldots, \alpha_m) \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}}_{= U \begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix}} &= (0, \ldots, 0) \end{aligned}$$

für ein invertierbares $U \in \mathbb{K}^{m\times m}$. Mit $(\beta_1, \dots, \beta_m) := (\alpha_1, \dots, \alpha_m)U$ gilt deshalb

$$\beta_1 t_1 + \cdots + \beta_m t_m = (0, \dots, 0).$$

Da U invertierbar ist und nicht alle α_i Null sind, sind auch nicht alle β_i Null, denn wären alle β_i Null, dann wären wegen $(\alpha_1, \dots, \alpha_m) = (\beta_1, \dots, \beta_m)U^{-1}$ auch alle α_i Null.

Also sind $t_1, \dots, t_m$ linear abhängig.

Sei i minimal, so dass $\beta_i \neq 0$. Wir können dann o.B.d.A. annehmen, dass $\beta_i = -1$ ist. Dann gilt

$$t_i = \beta_{i+1} t_{i+1} + \cdots + \beta_m t_m. \qquad (*)$$

Wir zeigen, dass $t_i = 0$ ist. Wegen der Treppenform folgt dann $t_{i+1} = \cdots = t_m = 0$.

Wäre $t_i \neq 0$, dann wäre es wegen der Treppenform von der Form $(0, \dots, 0, \underset{\uparrow}{1}, *, \dots, *)$ mit der k

1 an einem bestimmten Index k, und $t_{i+1}, \dots, t_m$ wären von der Form $(0, \dots, 0, 0, \underset{\uparrow}{*}, \dots, *)$,
$k+1$

mit einer 0 am Index k. Für die k-te Komponente der Gleichung $(*)$ würde dann gelten

$$1 = \beta_{i+1} 0 + \cdots + \beta_m 0 = 0.$$

Widerspruch. ∎

Der Beweis zeigt übrigens auch, dass die Menge aller Zeilen einer Treppenform, die von 0 verschieden sind, stets linear unabhängig ist. Wenn also $b_1, \dots, b_m \in \mathbb{K}^n$ linear abhängig sind, dann übersetzen sich die Linearkombinationen $b_1, \dots, b_m$, die 0 ergeben, in Linearkombinationen von der Form

$$0t_1 + \cdots + 0t_k + \alpha_{k+1} 0 + \cdots + \alpha_m 0 = 0$$

für beliebige $\alpha_{k+1}, \dots, \alpha_m \in \mathbb{K}$.

Man kann also sagen, dass eine Treppenform den linear unabhängigen und den linear abhängigen Anteil von $b_1, \dots, b_m$ voneinander trennt. Es ist deshalb von Interesse, wie viele Nullzeilen eine Treppenfrom enthält.

Definition 25. Sei $A \in \mathbb{K}^{n\times m} \setminus \{0\}$. Sei $T = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}$ eine Treppenform von A und $k \in \{1, \dots, n\}$ maximal mit $t_k \neq 0$. Dann heißt $\operatorname{Rang} A := k$ der *Rang* (engl. *rank*) von A.

Für die Nullmatrix definiert man $\operatorname{Rang} 0 := 0$.

Der vorherige Satz sagt also, dass die Zeilen von $A \in \mathbb{K}^{n\times m}$ genau dann linear abhängig sind, wenn $\operatorname{Rang} A < n$ ist.

Wenn T und T' zwei verschiedene Treppenformen von A sind, so müssen doch beide den gleichen Rang haben. Die Definition hängt also nicht von der Wahl der Treppenform ab und ist deshalb zulässig.

Allgemein gilt $\operatorname{Rang} A \leq n$. Übrigens gilt auch $\operatorname{Rang} A \leq m$. Es folgt also aus $m < n$ direkt, dass die Zeilen von A linear abhängig sind.

Beachte außerdem: Wenn E eine Elementarmatrix ist, dann gilt $\operatorname{Rang} A = \operatorname{Rang} EA$.

Beispiel.

1. Rang $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = 3$, da $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ und diese TF drei Stufen hat.
2. Rang $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 2$, da $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \leftrightarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$ und diese TF zwei Stufen hat.
3. Rang $\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix} = 1$, da $\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix} \leftrightarrow \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ und diese TF eine Stufe hat.

Satz 23. Sei $A \in \mathbb{K}^{n \times m}$. Dann gilt:

1. Ist irgendeine Wahl von k Zeilen von A linear unabhängig, so ist $\operatorname{Rang} A \geq k$.
2. Ist jede Wahl von k Zeilen von A linear abhängig, so ist $\operatorname{Rang} A < k$.

Mit anderen Worten: $\operatorname{Rang} A = k$ genau dann, wenn k die maximale Anzahl von linear unabhängigen Zeilen von A ist.

Beweis. Seien $a_1, \ldots, a_n \in \mathbb{K}^m$ die Zeilenvektoren von A.

1. Wir zeigen: wenn $\operatorname{Rang} A < k$, dann sind $a_1, \ldots, a_k$ linear abhängig. Nehmen wir also an, es gilt $\operatorname{Rang} A < k$. Dann hat jede Treppenform von A die Form

$$T = \begin{pmatrix} t_1 \\ \vdots \\ t_{k-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

   Es gibt eine invertierbare Matrix $U \in \mathbb{K}^{m \times m}$ mit $A = UT$. Damit lässt sich jedes a_i als Linearkombination von $t_1, \ldots, t_n$ darstellen, und damit auch als Linearkombination von $t_1, \ldots, t_{k-1}$, etwa

$$\begin{aligned} a_1 &= c_{1,1}t_1 + \cdots + c_{1,k-1}t_{k-1} + 0 + \cdots + 0 \\ a_2 &= c_{2,1}t_1 + \cdots + c_{2,k-1}t_{k-1} \\ &\vdots \\ a_k &= c_{k,1}t_1 + \cdots + c_{k,k-1}t_{k-1}. \end{aligned}$$

   Die Matrix

$$C = \begin{pmatrix} c_{1,1} & \cdots & c_{1,k-1} \\ \vdots & \ddots & \vdots \\ c_{k,1} & \cdots & c_{k,k-1} \end{pmatrix}$$

hat mehr Zeilen als Spalten, deshalb müssen wegen $\operatorname{Rang} C \leq k-1$ nach Satz 22 ihre Zeilen linear abhängig sein, etwa

$$\begin{aligned}&\alpha_1(c_{1,1},\ldots,c_{1,k-1})\\&+\alpha_2(c_{2,1},\ldots,c_{2,k-1})\\&+\cdots\\&+\alpha_k(c_{k,1},\ldots,c_{k,k-1})=0\end{aligned}$$

für gewisse $\alpha_1,\ldots,\alpha_k \in \mathbb{K}$, von denen nicht alle Null sind.

Dann gilt auch

$$\begin{aligned}0 &= \underbrace{\sum_{i=1}^{k} \alpha_i(c_{i,1},\ldots,c_{i,k-1})}_{=0} \begin{pmatrix} t_1 \\ \vdots \\ t_{k-1} \end{pmatrix} \\ &= \alpha_1 a_1 + \cdots + \alpha_k a_k,\end{aligned}$$

d. h. $\{a_1,\ldots,a_k\}$ ist linear abhängig.

2. Wir zeigen: wenn $\operatorname{Rang} A \geq k$ ist, dann gibt es eine Wahl von k Zeilen von A, die linear unabhängig sind. Sei $M = \{a_{i_1},\ldots,a_{i_\ell}\}$ eine Menge von Zeilen von A, die linear unabhängig ist und in dem Sinn maximal, dass $M \cup \{a\}$ linear abhängig ist für jede Zeile a von A, die nicht schon in M ist. Dann gilt

$$A \leftrightarrow \begin{pmatrix} a_{i_1} \\ \vdots \\ a_{i_\ell} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad (*)$$

denn für jede Zeile $a \in \{a_1,\ldots,a_n\} \setminus M$ gibt es nach Wahl von M eine lineare Abhängigkeit $\alpha_0 a + \alpha_{i_1} a_{i_1} + \cdots + \alpha_{i_\ell} a_{i_\ell} = 0$, und in dieser muss $\alpha_0 \neq 0$ sein, weil sonst schon $a_{i_1},\ldots,a_{i_\ell}$ linear abhängig wären. Wenn aber $\alpha_0 \neq 0$ ist, kann man O.B.d.A. annehmen, dass $\alpha_0 = -1$ ist, dass es also eine Darstellung $a = \alpha_{i_1} a_{i_1} + \cdots + \alpha_{i_\ell} a_{i_\ell}$ gibt. Daher lässt sich die Zeile a durch geeignete Zeilenoperationen mithilfe der Zeilen in M in eine Nullzeile überführen.

Aus der Form $(*)$ folgt $\operatorname{Rang} A \leq \ell$. Zusammen mit der Annahme $\operatorname{Rang} A \geq k$ folgt $\ell \geq k$. ∎

Satz 24. Für alle $A \in \mathbb{K}^{n\times m}$ gilt $\operatorname{Rang} A = \operatorname{Rang} A^\top$.

Beweis. Wegen $(A^\top)^\top = A$ genügt es zu zeigen, dass $\operatorname{Rang} A \geq \operatorname{Rang} A^\top$.

Ist $\operatorname{Rang} A = k$, so ist $A = UT$ für eine Matrix U, die das Produkt von Elementarmatrizen ist, und eine Treppenform T, bei der die ersten k Zeilen von 0 verschieden sind, und die letzten $n-k$ Zeilen Null sind. Dann ist $A^\top = T^\top U^\top$. Die Matrix $T^\top$ hat die Form

$$\underbrace{\begin{pmatrix} * & \cdots & * & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ * & \cdots & * & 0 & \cdots & 0 \end{pmatrix}}_{k \qquad\qquad n-k}.$$

Für die Treppenformen dieser Matrix sind nur die ersten k Spalten relevant. Der Rang der Matrix $T^\top$ ist daher höchstens k.

Nach Satz 23 sind deshalb je $k+1$ Zeilen von $T^\top$ linear abhängig. Sei $\tilde{T} \in \mathbb{K}^{n\times(k+1)}$ eine beliebige Wahl von $k+1$ Spalten von T, so dass $\tilde{T}^\top \in \mathbb{K}^{(k+1)\times n}$ eine beliebige Wahl von $k+1$ Zeilen von $T^\top$ ist. Dann gibt es also $(\alpha_1, \ldots, \alpha_{k+1}) \neq 0$ mit

$$(\alpha_1, \ldots, \alpha_{k+1})\tilde{T}^\top = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{K}^n.$$

Und dann gilt auch

$$(\alpha_1, \ldots, \alpha_{k+1})\tilde{T}^\top U^\top = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Damit ist gezeigt, dass jede Wahl von $k+1$ Zeilen von $A^\top = T^\top U^\top$ linear abhängig ist. Nach Satz 23 folgt $\operatorname{Rang} A^\top < k+1$, d. h. $\operatorname{Rang} A^\top \leq k$, wie behauptet. ∎

Wegen Satz 24 gilt Satz 23 auch für die Spalten von A. Für die Berechnung des Rangs einer Matrix bedeutet das, dass man sowohl Zeilen- als auch Spaltenoperationen anwenden darf, um A in eine Form zu bringen, aus der man den Rang ablesen kann. Spaltenoperationen sind ja Zeilenoperationen auf der Transponierten, und solche Operationen ändern den Rang nicht.

Durch Zeilen- und Spaltenoperationen lässt sich jede Matrix $A \in \mathbb{K}^{n\times m}$ mit $\operatorname{Rang} A = k$ auf die Form

$$\underbrace{\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}}_{k \qquad\qquad m-k} \begin{matrix} \left.\vphantom{\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}}\right\} k \\ \left.\vphantom{\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}}\right\} n-k \end{matrix}$$

bringen.

Beispiel.

$$\begin{array}{c}\begin{array}{ccc} -2 & + & \end{array}\\ \begin{pmatrix}1&2&0\\0&3&1\\2&1&3\end{pmatrix}\end{array} \rightsquigarrow \begin{pmatrix}1&0&0\\0&3&1\\2&-3&3\end{pmatrix}\begin{array}{l} -2\\ \\ +\end{array} \rightsquigarrow \begin{pmatrix}1&0&0\\0&3&-1\\0&-3&3\end{pmatrix}\begin{array}{l} \\ \\ +\end{array}$$

$$\rightsquigarrow \begin{array}{c}\begin{array}{ccc} & :3 & \end{array}\\ \begin{pmatrix}1&0&0\\0&3&-1\\0&0&2\end{pmatrix}\end{array} \rightsquigarrow \begin{pmatrix}1&0&0\\0&1&-1\\0&0&2\end{pmatrix}\begin{array}{l} \\ \\ \mid :2\end{array} \rightsquigarrow \begin{array}{c}\begin{array}{ccc} & + & \end{array}\\ \begin{pmatrix}1&0&0\\0&1&-1\\0&0&1\end{pmatrix}\end{array}$$

$$\rightsquigarrow \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

Also ist Rang $\begin{pmatrix}1&2&0\\0&3&1\\2&1&3\end{pmatrix} = 3$.

Man braucht sich aber gar nicht die Mühe zu machen, eine Matrix bis zu dieser Form zu bringen. Es genügt schon, durch Zeilen- und Spaltenoperationen alle Einträge unterhalb der Diagonale zu Null zu machen. Danach ist der Rang genau die Anzahl der Elemente auf der Diagonale, die von Null verschieden sind. Im vorliegenden Beispiel hätte man also schon bei der ersten Matrix in der zweiten Zeile aufhören können.

11 Inhomogene Systeme

Gegeben seien $A \in \mathbb{K}^{n\times m}$, $b \in \mathbb{K}^n$, und gesucht seien alle $x \in \mathbb{K}^m$ mit $Ax = b$.

Mit $A = ((a_{i,j}))_{i=1,j=1}^{n,m}$, $b = (b_1, \dots, b_n)$ und $x = (x_1, \dots, x_m)$ ist also folgendes System von Gleichungen zu lösen:

$$\begin{aligned} a_{1,1}x_1 + \cdots + a_{1,m}x_m &= b_1 \\ &\vdots \\ a_{n,1}x_1 + \cdots + a_{n,m}x_m &= b_n. \end{aligned}$$

Der Fall $b = (0, \dots, 0)$ wurde schon in Abschnitt 9 behandelt. In diesem Fall spricht man von einem *homogenen* Gleichungssystem. Den Fall $b \neq 0$ bezeichnet man als *inhomogenes* Gleichungssystem. (engl. *homogeneous*/*inhomogeneous*)

Wie sieht die Lösungsmenge eines inhomogenen Gleichungssystems aus, und wie bestimmt man sie?

Beachte:

- x ist genau dann eine Lösung von $Ax = b$ wenn für jedes x_h mit $Ax_h = 0$ auch $x + x_h$ eine Lösung ist.

  Es genügt also, eine einzige Lösung x des Gleichungssystems $Ax = b$ zu finden. Alle weiteren unterscheiden sich von dieser dann nur noch um Lösungen des homogenen Systems $Ax = 0$.

- $x = (x_1, \ldots, x_n) \in \mathbb{K}^m$ ist genau dann eine Lösung von $Ax = b$, wenn $\tilde{x} = (x_1, \ldots, x_n, -1) \in \mathbb{K}^{m+1}$ eine Lösung von $(A|b)\tilde{x} = 0$ ist.

  Dabei ist $(A|b) \in \mathbb{K}^{n\times(m+1)}$ die Matrix, die aus A entsteht, wenn man b als zusätzliche Spalte rechts anfügt.

Idee: Berechne zunächst die Lösungsmenge des homogenen Systems $(A|b)\tilde{x} = 0$ und bestimme dann die Vektoren der Lösungsmenge, für die die $(m+1)$-te Koordinate -1 ist.

Betrachten wir dazu eine Treppenform $T = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \in \mathbb{K}^{n\times(m+1)}$ von $(A|b)$. Es gibt zwei Fälle zu unterscheiden:

1. Die letzte von 0 verschiedene Zeile hat mehr als einen Eintrag:

$$T = \begin{pmatrix} * & \cdots & \cdots & \cdots & \cdots & * \\ \vdots & & & & & \vdots \\ * & \cdots & \cdots & \cdots & \cdots & * \\ & & 1 & * & \cdots & * \\ & & & & & \end{pmatrix}.$$

   In diesem Fall können wir die $(m+1)$te Koordinate des Lösungsvektors (des homogenen Systems) frei wählen. Insbesondere ist -1 eine mögliche Wahl. Die weiteren Koordinaten $x_1, \ldots, x_m$ ergeben sich dann wie üblich. Im allgemeinen werden manche von ihnen durch Gleichungen bestimmt sein und andere frei wählbar.

2. Die letzte von 0 verschiedene Zeile hat genau einen Eintrag:

$$T = \begin{pmatrix} * & \cdots & \cdots & \cdots & \cdots & * \\ \vdots & & & & & \vdots \\ * & \cdots & \cdots & \cdots & \cdots & * \\ & & & & & 1 \\ & & & & & \end{pmatrix}.$$

   In diesem Fall entspricht die letzte Zeile der Treppenform der Gleichung $1 \cdot x_{m+1} = 0$, d. h. alle Lösungen $(x_1, \ldots, x_{m+1})$ des homogenen Systems haben 0 als letzte Koordinate. Insbesondere gibt es keine Lösung mit -1 als letzter Koordinate. Das inhomogene System hat in diesem Fall also *keine Lösung.*

Beispiel.

1. $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $b = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$

$$\left(\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 6 \end{array}\right) \begin{array}{l} \\ \leftarrow\!\!\rfloor{\scriptstyle +} \end{array}\!\!\!

2. $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, $b = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$

$$\left(\begin{array}{cc|c} 1 & 2 & 3 \\ 2 & 4 & 1 \end{array}\right) \begin{array}{l} -2 \\ + \end{array} \leftrightarrow \left(\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & 0 & -5 \end{array}\right)$$

Die Lösungsmenge lautet $L = \emptyset$.

3. $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 2 \\ 7 & 8 & 9 & 3 \end{array}\right) \begin{array}{l} -4 \quad -7 \\ + \\ \quad\quad + \end{array} \leftrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -2 \\ 0 & -6 & -12 & -4 \end{array}\right) \begin{array}{l} \\ -2 \quad | : (-3) \\ + \end{array}$$

$$\leftrightarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 2/3 \\ 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} + \\ -2 \\ \\ \end{array}$$

$$\leftrightarrow \left(\begin{array}{ccc|c} 1 & 0 & -1 & -1/3 \\ 0 & 1 & 2 & 2/3 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Die Lösungsmenge lautet

$$L = \left\{ \begin{pmatrix} -1/3 \\ 2/3 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} : \alpha \in \mathbb{R} \right\}.$$

Beachte: das homogene Gleichungssystem $Ax = 0$ hat die Lösungsmenge

$$L_h = \left\{ \alpha \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} : \alpha \in \mathbb{R} \right\}.$$

Die Ergebnisse der vorangegangenen Diskussion lassen sich wie folgt als Satz zusammenfassen:

Satz 25. Sei $A \in \mathbb{K}^{n \times m}$, $b \in \mathbb{K}^n$ und sei $L = \{ x \in \mathbb{K}^m : Ax = b \}$. Dann gilt: $L = \emptyset$ oder es gibt ein $x_0 \in \mathbb{K}^m$ so dass

$$L = \{ x_0 + x_h : x_h \in \ker A \} \subseteq \mathbb{K}^m.$$

(Zur Erinnerung: $\ker A = \{ x \in \mathbb{K}^m : Ax = 0 \}$.)
Insbesondere gilt: wenn $\ker A = \{0\}$ und $n = m$, dann ist $|L| = 1$. Wenn $\ker A \neq \{0\}$, dann ist $|L| = 0$ oder $|L| > 1$. Im Fall $|L| > 1$ hat L mindestens so viele Elemente wie der Körper $\mathbb{K}$.

Bei der Berechnung der Lösungsmenge entfällt der größte Teil der Rechenarbeit auf A und nur ein vergleichsweise kleiner Teil auf b. Wenn man also mehrere inhomogene Gleichungssysteme mit dem selben A zu lösen hat, sollte man die Rechenschritte, die A betreffen, nicht mehrmals durchführen. Stattdessen bietet es sich an, alle Systeme gleichzeitig zu lösen.

Beispiel. $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}$, $b_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $b_2 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$.

$$
\begin{aligned}
(A \mid b_1\, b_2) &= \left(\begin{array}{ccc|cc} 1 & 0 & 2 & 1 & 3 \\ 0 & 1 & -1 & 2 & 2 \\ 1 & -1 & 0 & 3 & 1 \end{array}\right) \begin{array}{l} \\ \\ \leftarrow + \end{array} \begin{array}{l} \\ \\ -1 \end{array} \\
&\leftrightarrow \left(\begin{array}{ccc|cc} 1 & 0 & 2 & 1 & 3 \\ 0 & 1 & -1 & 2 & 2 \\ 0 & -1 & -2 & 2 & -2 \end{array}\right) \begin{array}{l} \\ \\ \leftarrow + \end{array} \\
&\leftrightarrow \left(\begin{array}{ccc|cc} 1 & 0 & 2 & 1 & 3 \\ 0 & 1 & -1 & 2 & 2 \\ 0 & 0 & -3 & 4 & 0 \end{array}\right) \begin{array}{l} \\ \leftarrow + \\ \mid : (-3) \end{array} \begin{array}{l} + \\ + \\ -2 \end{array} \\
&\leftrightarrow \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 11/3 & 3 \\ 0 & 1 & 0 & 2/3 & 2 \\ 0 & 0 & 1 & -4/3 & 0 \end{array}\right)
\end{aligned}
$$

Daraus folgt: Die Lösungsmenge von $Ax = b_1$ ist $\{\begin{pmatrix} 11/3 \\ 2/3 \\ -4/3 \end{pmatrix}\}$ und die Lösungsmenge von $Ax = b_2$ ist $\{\begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}\}$.

Und was, wenn uns jetzt noch jemand nach der Lösung x von $Ax = 3b_1 - 7b_2$ fragt? Antwort: Wir können entweder noch einmal von vorne losrechnen. Oder wir kombinieren die Lösung einfach aus den schon bekannten Lösungen. Tatsächlich lautet die Lösung

$$x = 3\begin{pmatrix} 11/3 \\ 2/3 \\ -4/3 \end{pmatrix} - 7\begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} \quad [= \begin{pmatrix} -20 \\ -12 \\ -4 \end{pmatrix}].$$

Begründung: Wenn $Ax_1 = b_1$ und $Ax_2 = b_2$, dann ist $A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A x_1 + \alpha_2 A x_2 = \alpha_1 b_1 + \alpha_2 b_2$ für alle $\alpha_1, \alpha_2 \in \mathbb{K}$.

Das gleichzeitige Lösen mehrerer inhomogener Gleichungssysteme lässt sich auch auffassen als das Lösen von Matrixgleichungen. Gegeben: $A \in \mathbb{K}^{n\times m}$, $B \in \mathbb{K}^{n\times k}$, gesucht: alle $X \in \mathbb{K}^{m\times k}$ mit $AX = B$. Um so eine Gleichung zu lösen, bringt man einfach die erweiterte Matrix $(A|B) \in \mathbb{K}^{n\times(m+k)}$ in Treppen(normal)form und liest daraus die Lösungsmenge $L \subseteq \mathbb{K}^{m\times k}$ ab.

Der wichtigste Spezialfall ist, wenn $A = \mathbb{K}^{n\times n}$ und $B = I_n \in \mathbb{K}^{n\times n}$. Die Gleichung $AX = I_n$ ist genau dann lösbar, wenn A eine invertierbare Matrix ist. Die Lösung ist dann $X = A^{-1}$.

Beispiel.

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

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \\ \leftarrow +\,\rfloor^{-3} \\ \\ \end{array} \leftrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & -9 & -3 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \\ \leftarrow +\,\rfloor^{-1} \\ \\ \end{array}$$

$$\leftrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & -9 & -3 & 1 & 0 \\ 0 & 0 & 11 & 3 & -1 & 1 \end{array}\right) \begin{array}{l} \\ \\ |:11\ \rfloor_{9} \end{array} \begin{array}{l} \leftarrow\!\!\!

Beweis. Die Äquivalenzen (3)⇔(4)⇔(5) folgen aus den Sätzen 23 und 24.

(1)⇒(2). Sei $x \in \mathbb{K}^n$ so, dass $Ax = 0$ ist. Dann ist $x = A^{-1}Ax = A^{-1}0 = 0$.

(2)⇒(5). Wären die Spalten von $A = (a_1, \ldots, a_n) \in \mathbb{K}^{n\times n}$ linear abhängig, dann gäbe es ein $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n \setminus \{0\}$ mit

$$\alpha_1 a_1 + \cdots + \alpha_n a_n = 0,$$

also $A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = 0$, also $(\alpha_1, \ldots, \alpha_n) \in \ker A$, also $\ker A \neq \{0\}$.

(3)⇒(6). Wenn $\operatorname{Rang} A = n$ ist, ist die TNF von A die Einheitsmatrix I_n. Es gibt also Elementarmatrizen $E_1, \ldots, E_m \in \mathbb{K}^{n\times n}$ mit $I_n = E_m \cdots E_1 A$. Da jede Elementarmatrix invertierbar ist und ihr Inverses wieder eine Elementarmatrix ist, ist $A = E_1^{-1} \cdots E_m^{-1}$ die gewünschte Darstellung.

(6)⇒(1). Jede Elementarmatrix ist invertierbar und das Produkt invertierbarer Matrizen ist invertierbar. ∎

Das Verfahren zum Invertieren von Matrizen kann man allgemein auch dazu verwenden, zu gegebenem $A \in \mathbb{K}^{n\times m}$ eine invertierbare Matrix $U \in \mathbb{K}^{n\times n}$ zu finden, so dass UA in Treppen(normal)form ist: die Treppen(normal)form von $(A|I_n)$ ist $(T|U)$, wobei T die Treppen(normal)form von A ist und U die gesuchte Transformationsmatrix.

Beispiel.

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 7 & 8 & 9 & 0 & 0 & 1 \end{array}\right) \leftrightarrow \cdots \leftrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & -5/3 & 2/3 & 0 \\ 0 & 1 & 2 & 4/3 & -1/3 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right).$$

Für $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ ist also $U = \frac{1}{3}\begin{pmatrix} -5 & 2 & 0 \\ 4 & -1 & 0 \\ 3 & -6 & 3 \end{pmatrix}$ eine Matrix, so dass

$$UA = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

die Treppennormalform von A ist.

Noch allgemeiner: Wenn wir sowohl Zeilen- als auch Spaltenoperationen anwenden wollen, zum Beispiel um den Rang einer Matrix A zu bestimmen, dann können wir Matrizen $U \in \mathbb{K}^{n\times n}$ und $V \in \mathbb{K}^{m\times m}$ finden, so dass

$$UAV = D := \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Dazu betrachtet man die Matrix $\begin{pmatrix} A & I_n \\ I_m & 0 \end{pmatrix}$ und wendet darauf nur Zeilen- und Spaltenoperationen an, die die oberen n Zeilen oder die linken m Spalten betreffen. Das Ergebnis ist eine Matrix $\begin{pmatrix} D & U \\ V & 0 \end{pmatrix}$.

Beispiel. $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$

$$\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ 4 & 5 & 6 & 0 & 1 \\ \hline 1 & 0 & 0 & & \\ 0 & 1 & 0 & & \\ 0 & 0 & 1 & & \end{array}\right) \begin{array}{l} \\ \\ \\ \\ \\ \end{array} \begin{array}{l} \\ \\ -4 \\ + \\ \\ \end{array} \rightsquigarrow \begin{array}{c} \overset{-3 \qquad +}{\ulcorner\!\!\!

12 Determinanten

In diesem Abschnitt sind alle Matrizen quadratisch, d. h. wir betrachten hier nur Matrizen, die gleich viele Zeilen wie Spalten haben.

Das Ziel ist, eine Funktion $\det\colon \mathbb{K}^{n\times n} \to \mathbb{K}$ zu konstruieren, so dass der Wert $\det(A) \in \mathbb{K}$ etwas darüber aussagt, ob $A \in \mathbb{K}^{n\times n}$ einen nichtleeren Kern hat.

Dazu ist etwas Vorbereitung nötig. Man erinnere sich, dass S_n die Gruppe der bijektiven Funktionen $\{1,\dots,n\} \to \{1,\dots,n\}$ ist, und dass ihre Elemente auch als Permutationen bezeichnet werden.

Definition 26.

1. Eine Permutation $\pi \in S_n$ heißt *Zyklus* (engl. *cycle*), falls es paarweise verschiedene $k_1,\dots,k_m \in \{1,\dots,n\}$ gibt, so dass

$$\pi(k_1) = k_2, \quad \pi(k_2) = k_3, \quad \dots, \quad \pi(k_{m-1}) = k_m, \quad \pi(k_m) = k_1$$

   sowie $\pi(k) = k$ für alle $k \in \{1,\dots,n\} \setminus \{k_1,\dots,k_m\}$ gilt.

   Schreibweise: $\pi = (k_1\ k_2\ \dots\ k_m)$.

   Man nennt m die *Länge* des Zyklus.

2. Zwei Permutationen π_1, π_2 heißen (zueinander) *disjunkt,* falls gilt

$$\forall\ k \in \{1,\dots,n\} : \pi_1(k) = k \vee \pi_2(k) = k.$$

3. Ein Zyklus der Länge zwei heißt *Transposition.*

4. Ein $k \in \{1,\dots,n\}$ mit $\pi(k) = k$ heißt *Fixpunkt* von $\pi \in S_n$.

Beispiel.

1. $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} = (1\ 2\ 4\ 3) = (3\ 1\ 2\ 4)$ ist ein Zyklus. Beachte: Ein Zyklus lässt sich auf verschiedene Weise schreiben.

2. $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = (1\ 3)$ ist eine Transposition.

3. $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$ ist kein Zyklus, lässt sich aber als Produkt (Komposition) der beiden disjunkten Zyklen (1 3) und (2 4) schreiben.

Satz 27.

1. Sind $\pi_1, \pi_2 \in S_n$ disjunkt, so gilt $\pi_1\pi_2 = \pi_2\pi_1$. (Zur Erinnerung: Im allgemeinen ist die Verknüpfung von Permutationen nicht kommutativ.)

2. Jedes $\pi \in S_n$ lässt sich als endliches Produkt von disjunkten und von id verschiedenen Zyklen schreiben. Diese Darstellung ist bis auf die Reihenfolge der Faktoren eindeutig.

3. Sind $k_1, \ldots, k_m \in \{1, \ldots, m\}$ paarweise verschieden, so gilt

$$(k_1\ k_2\ \ldots\ k_m) = (k_1\ k_2)(k_2\ k_3)\cdots(k_{m-1}\ k_m),$$
$$(k_1\ k_2\ \ldots\ k_m)^{-1} = (k_m\ k_{m-1}\ \ldots\ k_1).$$

4. Sind $\tau_1, \ldots, \tau_m \in S_n$ Transpositionen mit $\tau_1 \cdots \tau_m = \mathrm{id}$, so ist m gerade.

5. $|S_n| = n! := 1 \cdot 2 \cdots (n-1) \cdot n$.

Beweis.

1. Sei $k \in \{1, \ldots, n\}$ beliebig. Falls $\pi_1(k) = \pi_2(k) = k$ ist, gilt $(\pi_1\pi_2)(k) = (\pi_2\pi_1)(k) = k$.

   Wenn $\pi_1(k) \neq k$ ist, dann ist $\pi_2(k) = k$, weil π_1, π_2 disjunkt sind. Da π_1 bijektiv ist, gilt dann auch $\pi_1(\pi_2(k)) = \pi_1(k)$.

   Aus $\pi_1(k) \neq k$ und der Bijektivität von π_1 folgt auch, dass $\pi_1(\pi_1(k)) \neq \pi_1(k)$. Aus der Disjunktheit von π_1, π_2 folgt deshalb $\pi_2(\pi_1(k)) = \pi_1(k)$.

   Damit ist gezeigt $\pi_1(\pi_2(k)) = \pi_2(\pi_1(k))$.

   Der Fall $\pi_2(k) \neq k$ geht genauso. Insgesamt ist also gezeigt, dass für alle $k \in \{1, \ldots, n\}$ gilt $(\pi_1\pi_2)(k) = (\pi_2\pi_1)(k)$, wie behauptet.

2. Die *Existenz* folgt aus folgendem konstruktiven Argument.

   1 Setze $B := \{1, \ldots, n\}$.
   2 Solange $B \neq \emptyset$
   3 Wähle ein beliebiges $k \in B$.
   4 Falls $\pi(k) = k$, dann
   5 Setze $B := B \setminus \{k\}$
   6 ansonsten
   7 Bestimme $m \in \{1, \ldots, n\}$ mit $\pi^m(k) = k$ und $\pi^i(k) \neq k$ für $i = 1, \ldots, m-1$. So ein m existiert.
   8 Notiere den Zyklus $(k\ \pi(k)\ \ldots\ \pi^{m-1}(k))$.
   9 Setze $B := B \setminus \{k, \ldots, \pi^{m-1}(k)\}$.

   Da in jeder Iteration die Menge B um wenigstens ein Element kleiner wird, wird dieses Verfahren nach endlich vielen Schritten fertig. Offensichtlich ist π das Produkt aller notierten Zyklen und diese sind $\neq \mathrm{id}$ und paarweise zueinander disjunkt.

Eindeutigkeit: Hätte $\pi \in S_n$ zwei verschiedene Darstellungen als Produkt disjunkter Zyklen, etwa

$$\pi = \sigma_1 \cdots \sigma_m = \tilde{\sigma}_1 \cdots \tilde{\sigma}_{\tilde{m}},$$

dann müsste $\sigma_i \notin \{\tilde{\sigma}_1, \ldots, \tilde{\sigma}_{\tilde{m}}\}$ für mindestens ein $i \in \{1, \ldots, m\}$ gelten. (O.B.d.A. können wir annehmen $m \geq \tilde{m}$.) Wähle so ein σ_i.

Da $\sigma_i \neq \mathrm{id}$ ist, gibt es ein $k \in \{1, \ldots, n\}$ mit $\sigma_i(k) \neq k$. Wegen Disjunktheit ist dann auch $\pi(k) = \sigma_i(k) \neq k$. Damit muss es ein $\tilde{\imath} \in \{1, \ldots, \tilde{m}\}$ geben mit $\pi(k) = \tilde{\sigma}_{\tilde{\imath}}(k) \neq k$. Induktiv zeigt man $\forall\ \ell \in \mathbb{N} : \pi^\ell(k) = \sigma_i^\ell(k) = \tilde{\sigma}_{\tilde{\imath}}^\ell(k)$. Aber dann ist $\sigma_i = \tilde{\sigma}_{\tilde{\imath}}$, im Widerspruch zur Wahl von σ_i.

3. Übung.

4. Betrachte die Funktion

$$F\colon S_n \to \mathbb{N}, \quad \pi \mapsto |\{\, (i,j) \in \{1, \ldots, n\}^2 : i < j \wedge \pi(i) > \pi(j) \,\}|.$$

(Beispiel: $F(\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 5 & 3 \end{pmatrix}) = |\{(1,3),(2,3),(2,5),(4,5)\}| = 4.$)

Sei $\pi \in S_n$ beliebig und $\tau = (i\ j)$ mit $i < j$ eine Transposition.

1. Fall: $\pi(i) > \pi(j)$. Dann gilt $(\pi\tau)(i) < (\pi\tau)(j)$ und $(\pi\tau)(k) = \pi(k)$ für alle $k \in \{1, \ldots, n\} \setminus \{i, j\}$. Es gilt also $F(\pi\tau) = F(\pi) - 1$.
2. Fall: $\pi(i) < \pi(j)$. Dann ist $(\pi\tau)(i) > (\pi\tau)(j)$ und $(\pi\tau)(k) = \pi(k)$ für alle $k \in \{1, \ldots, n\} \setminus \{i, j\}$. Es gilt also $F(\pi\tau) = F(\pi) + 1$.
3. Fall: $\pi(i) = \pi(j)$. Dieser Fall kann nicht auftreten, weil $i < j$, also $i \neq j$ und π bijektiv, also injektiv.

Damit ist gezeigt, dass für jede Permutation $\pi \in S_n$, die ein Produkt einer ungeraden Anzahl von Transpositionen ist, auch $F(\pi)$ ungerade ist. Da $F(\mathrm{id}) = 0$ gerade ist, folgt die Behauptung.

5. Induktion nach n. Für $n = 1$ gilt $S_1 = \{\mathrm{id}\}$, also $|S_1| = 1$. Sei nun $n \in \mathbb{N}$ so, dass $|S_n| = n!$ gilt. Wir zeigen $|S_{n+1}| = (n+1)!$.

Zunächst ist klar, dass die Permutationen von $\{1, \ldots, n\}$ genauso zahlreich sind wie die Permutationen von $\{1, \ldots, n+1\}$, die $n+1$ als Fixpunkt haben. Nach Induktionsannahme gibt es $n!$ viele Permutationen von $\{1, \ldots, n\}$. Für jede solche Permutation π und jede Wahl von $k \in \{1, \ldots, n+1\}$ ist $(n+1\ k) \circ \pi$ eine Permutation von $\{1, \ldots, n+1\}$. Da all diese Permutationen paarweise verschieden sind, folgt $|S_{n+1}| \geq (n+1)|S_n|$.

Umgekehrt gilt: ist $\sigma \in S_{n+1}$ beliebig, so ist $\pi := (n+1\ \sigma(n+1)) \circ \sigma$ eine Permutation, die $n+1$ als Fixpunkt hat. Da Transpositionen selbstinvers sind, gilt auch $\sigma = (n+1\ \pi(n+1)) \circ \pi$, so dass sich also jedes Element von S_{n+1} als Produkt einer Permutation mit Fixpunkt $n+1$ und einer Transposition $(n+1\ k)$ schreiben lässt. Daher werden mit der vorher beschriebenen Konstruktion alle Elemente von S_{n+1} erreicht und es folgt $|S_{n+1}| = (n+1)|S_n|$. ∎

Wegen Teil 2 und 3 lässt sich jedes $\pi \in S_n$ als Produkt von (nicht notwendigerweise disjunkten) Transpositionen schreiben. Diese Darstellung ist zwar nicht eindeutig, aber wegen Teil 4 gilt: Entweder haben alle Darstellungen von $\pi \in S_n$ eine gerade Anzahl von Transpositionen, oder alle Darstellungen haben eine ungerade Anzahl. Deshalb ist folgende Definition erlaubt:

Definition 27. Sei $\pi \in S_n$. Dann heißt

$$\operatorname{sgn}(\pi) := \begin{cases} 1 & \text{falls } \pi \text{ Produkt einer geraden Anzahl von Transpositionen ist} \\ -1 & \text{falls } \pi \text{ Produkt einer ungeraden Anzahl von Transpositionen ist} \end{cases}$$

das *Vorzeichen* (engl. *sign*) von π.

Ist F wie im Beweis von Satz 27, so gilt $\operatorname{sgn}(\pi) = (-1)^{F(\pi)}$.
Außerdem ist $\operatorname{sgn}\colon S_n \to \{-1, 1\}$ ein Gruppenhomomorphismus, d. h. es gilt $\operatorname{sgn}(\pi\sigma) = \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma)$.

Beispiel.

$$\operatorname{sgn}\left(\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}\right) = 1, \qquad \operatorname{sgn}\left(\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}\right) = -1$$

Definition 28. Sei $A = ((a_{i,j}))_{i,j=1}^n \in \mathbb{K}^{n\times n}$. Dann heißt

$$\det(A) := \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i,\pi(i)} \in \mathbb{K}$$

die *Determinante* von A. Statt $\det(A)$ schreibt man auch

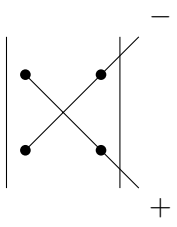
$$\begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}.$$

Beispiel.

1. $n = 2$, $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$.

$$S_2 = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\} = \{ \underset{\substack{\uparrow \\ \text{sgn} = 1}}{\text{id}}, \overset{\substack{\text{sgn} = -1 \\ \downarrow}}{(1\ 2)} \}.$$

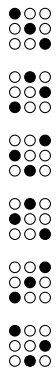
$\det A = 1 \cdot a_{1,1}a_{2,2} + (-1)a_{1,2}a_{2,1} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$.



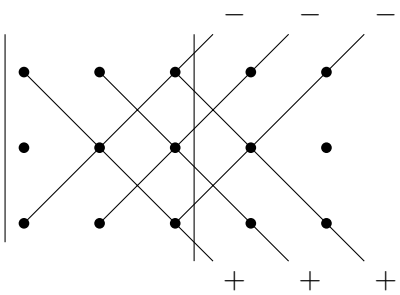
2. $n = 3$, $A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$.

$$S_3 = \{ \underset{+}{\text{id}}, \underset{-}{(1\ 2)}, \underset{-}{(1\ 3)}, \underset{-}{(2\ 3)}, \underset{+}{(1\ 3\ 2)}, \underset{+}{(1\ 2\ 3)} \}.$$

$$\begin{aligned}
\det(A) &= a_{1,1}a_{2,2}a_{3,3} \\
&+ a_{1,3}a_{2,1}a_{3,2} \\
&+ a_{1,2}a_{2,3}a_{3,1} \\
&- a_{1,2}a_{2,1}a_{3,3} \\
&- a_{1,3}a_{2,2}a_{3,1} \\
&- a_{1,1}a_{2,3}a_{3,2}
\end{aligned}$$



Für die Handrechnung ist es nützlich, die 3×5-Matrix zu betrachten, die man aus A erhält, wenn man die erste Spalte in die vierte und die zweite in die fünfte kopiert. Dann lässt sich die Determinante berechnen, indem man für die drei absteigenden und die drei aufsteigenden Diagonalen jeweils das Produkt der Elemente berechnet, bei den aufsteigenden Diagonalen das Vorzeichen ändert, und die Ergebnisse aufaddiert.



Ab $n = 4$ wird die Berechnung von $\det(A)$ mit der Definition unangenehm. (Beachte: $|S_4| = 24$, $|S_5| = 120$.) Wir werden aber in Kürze sehen, dass die Berechnung von $\det(A)$ mit dem gleichen Aufwand möglich ist wie die Lösung eines Gleichungssystems.

Satz 28. Für alle $A \in \mathbb{K}^{n\times n}$ gilt $\det(A) = \det(A^\top)$.

Beweis. Zunächst gilt

$$\prod_{i=1}^{n} a_{i,\pi(i)} = \prod_{i=1}^{n} a_{\sigma(i),\pi(\sigma(i))}$$

für jedes beliebige $\sigma \in S_n$, weil $(\mathbb{K} \setminus \{0\}, \cdot)$ kommutativ ist (σ vertauscht bloß die Reihenfolge der Faktoren im Produkt).

Für $\sigma = \pi^{-1}$ folgt daraus insbesondere

$$\prod_{i=1}^{n} a_{i,\pi(i)} = \prod_{i=1}^{n} a_{\pi^{-1}(i),i}.$$

Zweitens gilt $\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi^{-1})$ wegen der Teile 3 und 4 von Satz 27.

Drittens gilt $\{\, \pi : \pi \in S_n \,\} = \{\, \pi^{-1} : \pi \in S_n \,\}$, weil S_n eine Gruppe ist. Deshalb gilt $\sum_{\pi\in S_n} f(\pi) = \sum_{\pi\in S_n} f(\pi^{-1})$ für jede Funktion $f\colon S_n \to \mathbb{K}$ (es ändert sich nur die Reihenfolge der Summanden, aber nicht der Wert der Summe).

Aus allem zusammen folgt

$$\begin{aligned}\det(A) &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i,\pi(i)} \\ &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi^{-1}) \prod_{i=1}^{n} a_{\pi^{-1}(i),i} \\ &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{\pi(i),i} = \det(A^\top).\end{aligned}$$

■

Satz 29. Sei $A \in \mathbb{K}^{n\times n}$.

1. Entsteht $B \in \mathbb{K}^{n\times n}$ aus A durch Multiplikation einer Zeile mit $\lambda \in \mathbb{K}$, so gilt $\det(B) = \lambda \det(A)$.
2. Entsteht $B \in \mathbb{K}^{n\times n}$ aus A durch Vertauschen zweier Zeilen, so gilt $\det(B) = -\det(A)$.
3. Entsteht $B \in \mathbb{K}^{n\times n}$ aus A dadurch, dass man das λ-fache einer Zeile von A zu einer anderen Zeile dazuaddiert, so gilt $\det(B) = \det(A)$.

Beweis. Schreibe jeweils $A = ((a_{i,j}))_{i,j=1}^n$, $B = ((b_{i,j}))_{i,j=1}^n$.

1. $\det(B) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \underbrace{\prod_{i=1}^{n} b_{i,\pi(i)}}_{= \lambda \prod_{i=1}^n a_{i,\pi(i)}} = \lambda \det(A)$.
2. Sei $\tau \in S_n$ die entsprechende Transposition. gilt:

$$\begin{aligned}\det(B) &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{\tau(i),\pi(i)} \\ &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i,(\pi\tau)(i)} \qquad (\text{weil } \tau^2 = \mathrm{id}) \\ &= \operatorname{sgn}(\tau) \sum_{\pi\in S_n} \operatorname{sgn}(\pi\tau) \prod_{i=1}^{n} a_{i,(\pi\tau)(i)} \\ &= -\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \\ &= -\det(A).\end{aligned}$$

3. Wegen Teil 2 können wir o.B.d.A. annehmen, dass das λ-fache der zweiten Zeile zur ersten addiert wird, also

$$b_{1,j} = a_{1,j} + \lambda a_{2,j}$$

$$b_{i,j} = a_{i,j} \qquad (i > 1).$$

Dann gilt:

$$\begin{aligned}\det(B) &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \underbrace{\prod_{i=1}^{n} b_{i,\pi(i)}}_{= (a_{1,\pi(1)} + \lambda a_{2,\pi(1)}) \prod_{i=2}^{n} a_{i,\pi(i)}} \\ &= \det(A) + \lambda \underbrace{\sum_{\pi\in S_n} \operatorname{sgn}(\pi) a_{2,\pi(1)} \prod_{i=2}^{n} a_{i,\pi(i)}}_{= \begin{vmatrix} a_{2,1} & \cdots & a_{2,n} \\ a_{2,1} & \cdots & a_{2,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix} =: \Delta}.\end{aligned}$$

Wir sind fertig, wenn wir zeigen können, dass $\Delta = 0$ ist.

Dazu nutzen wir aus, dass Δ die Determinante einer Matrix mit zwei identischen Zeilen ist. Vertauscht man diese beiden Zeilen, so bleibt Δ gleich, und aus Teil 2 folgt deshalb $\Delta = -\Delta$, also $2\Delta = 0$. Daraus folgt $\Delta = 0$, jedoch nur, wenn $\mathbb{K}$ ein Körper ist, in dem $2 \neq 0$ gilt! Das dürfen wir nicht ohne weiteres annehmen.

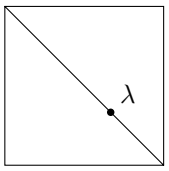
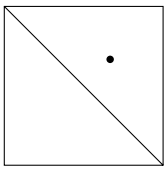
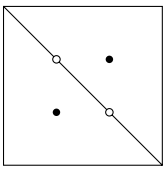
Es geht auch ohne diese Annahme: Durch $\sigma \sim \tau \iff \sigma = \tau \vee \sigma = \tau \circ (1\ 2)$ wird auf S_n eine Äquivalenzrelation erklärt. Jede Äquivalenzklasse hat genau zwei Elemente, und für beide Elemente π einer Äquivalenzklasse hat $a_{2,\pi(1)} a_{1,\pi(2)} \prod_{i=3}^{n} a_{i,\pi(i)}$ denselben Wert und $\operatorname{sgn}(\pi)$ unterschiedliches Vorzeichen. Daher gilt

$$\Delta = \sum_{[\pi]\in S_n/\sim} \underbrace{(1 + (-1))}_{=0} a_{2,\pi(1)} a_{2,\pi(2)} \prod_{i=3}^{n} a_{i,\pi(i)} = 0.$$

■

Wegen Satz 28 gilt Satz 29 auch für Spaltenoperationen statt Zeilenoperationen.

Für die Identitätsmatrix I_n folgt leicht aus der Definition, dass $\det(I_n) = 1$. Mit Satz 29 folgt daraus für die Determinate von Elementarmatrizen:

$\det\Big($  $\Big) = \lambda,$ $\det\Big($  $\Big) = 1,$ $\det\Big($  $\Big) = -1.$

Damit lassen sich Determinanten durch Zeilen- und Spaltenoperationen ausrechnen, indem man sie auf eine Form bringt, aus der man den Wert direkt ablesen kann. Insbesondere kann man den Wert einer Determinante direkt ablesen bei Matrizen, bei denen auf einer Seite der Diagonalen lauter Nullen stehen:

$$\begin{vmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{vmatrix} = \lambda_1 \cdots \lambda_n.$$

Beispiel.

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{vmatrix} \begin{matrix} -1 \quad -1 \\ + \\ + \end{matrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 2 & 8 \end{vmatrix} \begin{matrix} \\ \\ \mid :8 \end{matrix} = 8 \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & \frac{1}{4} & 1 \end{vmatrix} \begin{matrix} + \; -1/4 \\ \\ \\ \end{matrix} = 8 \begin{vmatrix} 1 & \frac{3}{4} & 1 \\ 0 & \frac{1}{4} & 3 \\ 0 & 0 & 1 \end{vmatrix} = 2.$$

Satz 30. (Ergänzung zum Satz 26) $A \in \mathbb{K}^{n\times n}$ ist genau dann invertierbar, wenn $\det(A) \neq 0$ gilt.

Beweis. „$\Rightarrow$“ Ist A invertierbar, so lässt sich A durch Zeilenumformungen auf die Treppennormalform I_n bringen. Nach Satz 29 ändert sich bei jeder Zeilenumformung der Wert der Determinante höchstes durch Multiplikation mit einem $\lambda \in \mathbb{K} \setminus \{0\}$. Daraus folgt $\det(A) \neq 0$.

„$\Leftarrow$“ Ist A nicht invertierbar, so gilt $\operatorname{Rang} A < n$ und jede Treppenform T von A enthält eine Nullzeile. Für solche Treppenformen muss gelten $\det(T) = 0$. Da sich A durch Zeilenumformungen in T überführen lässt und jede Zeilenumformung den Wert der Determinaten höchstes mit einer Konstanten multipliziert, folgt $\det(A) = 0$. ∎

Satz 31. Für alle $A, B \in \mathbb{K}^{n\times n}$ gilt: $\det(AB) = \det(A)\det(B)$.

Beweis. Falls $\operatorname{Rang} A < n$ ist, ist auch $\operatorname{Rang} AB < n$, und es gilt sowohl $\det(AB) = 0$ als auch $\det(A)\det(B) = 0$.

Falls $\operatorname{Rang} A = n$ ist, ist A nach Satz 26 ein Produkt endlich vieler Elementarmatrizen, etwa $A = E_1 \cdots E_m$. Nach Satz 29 bewirkt die Multiplikation einer Matrix mit einer Elementarmatrix E die Multiplikation ihrer Determinante mit $\det(E)$. Also gilt

$$\det(AB) = \underbrace{\det(E_1)\cdots\det(E_m)}_{= \det(\underbrace{E_1\cdots E_m I_n}_{= A})} \det(B).$$

∎

Aus Satz 31 und $\det(I_n) = 1$ folgt direkt $\det(A^{-1}) = \frac{1}{\det(A)}$ für invertierbare Matrizen $A \in \mathbb{K}^{n\times n}$. Darüber hinaus folgt, dass

$$\det\colon \operatorname{GL}(n, \mathbb{K}) \to (\mathbb{K} \setminus \{0\}, \cdot)$$

ein Gruppenhomomorphismus ist. Sein Kern

$$\ker\det := \{\, A \in \operatorname{GL}(n, \mathbb{K}) : \det(A) = 1 \,\}$$

heißt die *spezielle lineare Gruppe* und wird $\operatorname{SL}(n, \mathbb{K})$ geschrieben. (Beachte: der Kern eines Gruppenhomomorphismus ist nach Teil 2 von Satz 2 eine Gruppe.)

Determinanten sind nützlich, um Vektoren auf lineare Unabhängigkeit zu untersuchen. Vor allem dann, wenn es nicht um konkrete Vektoren geht, so dass das Gauß-Verfahren sich nicht ohne weiteres anwenden lässt.

Beispiel.

1. Seien $\phi_1, \dots, \phi_n \in \mathbb{K}$ paarweise verschieden und $v_1, \dots, v_n \in \mathbb{K}^n$ definiert durch

$$v_i = (1, \phi_i, \phi_i^2, \dots, \phi_i^{n-1}) \in \mathbb{K}^n$$

04-03 für $i = 1, \dots, n$. Dann ist $\{v_1, \dots, v_n\}$ linear unabhängig. Zum Beweis zeigt man, dass

$$\begin{vmatrix} 1 & \phi_1 & \cdots & \phi_1^{n-1} \\ 1 & \phi_2 & \cdots & \phi_2^{n-1} \\ \vdots & & & \vdots \\ 1 & \phi_n & \cdots & \phi_n^{n-1} \end{vmatrix} = \underbrace{\prod_{i=2}^{n} \prod_{j=1}^{i-1} \underbrace{(\phi_i - \phi_j)}_{\neq 0}}_{\neq 0}$$

(*Vandermonde Determinante*). 04-03

2. Für welche Werte von $\alpha \in \mathbb{R}$ sind die Vektoren

$$\begin{pmatrix} 1 \\ \alpha \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 1-\alpha \\ 0 \\ 4 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ -2\alpha \end{pmatrix}$$

linear abhängig? Um das zu beantworten, berechnen wir die Determinante

$$\begin{vmatrix} 1 & 1-\alpha & 1 \\ \alpha & 0 & 1 \\ 2 & 4 & -2\alpha \end{vmatrix} = -2\alpha^3 + 2\alpha^2 + 2\alpha - 2 = -2(\alpha - 1)^2(\alpha + 1).$$

Die gesuchten Werte für α sind genau jene, für die die Determinante Null wird. Das ist offensichtlich genau dann der Fall, wenn $\alpha = 1$ oder $\alpha = -1$ ist.

Satz 32. Sind $v, w, v_2, \dots, v_n \in \mathbb{K}^n$, so gilt

$$\det\big(v, v_2, \cdots, v_n\big) + \det\big(w, v_2, \cdots, v_n\big) = \det\big(v + w, v_2, \cdots, v_n\big).$$

Beweis. Übung. ∎

Satz 33. (Laplace-Entwicklung) Sei $A = ((a_{i,j}))_{i,j=1}^n \in \mathbb{K}^{n \times n}$. Es sei $A^{(i,j)} \in \mathbb{K}^{(n-1)\times(n-1)}$ die Matrix, die aus A entsteht, wenn man die i-te Zeile und die j-te Spalte löscht. Dann gilt:

$$\det(A) = a_{1,1} \det(A^{(1,1)}) - a_{1,2} \det(A^{(1,2)}) + a_{1,3} \det(A^{(1,3)}) \pm \cdots + (-1)^n a_{1,n} \det(A^{(1,n)}).$$

Beweis. Wegen Satz 29 und Satz 32 gilt zunächst

$$\begin{aligned}\det(A) = a_{1,1}&\begin{vmatrix} 1 & 0 & \cdots & 0 \\ a_{2,1} & \cdots & \cdots & a_{2,n} \\ \vdots & & & \vdots \\ a_{n,1} & \cdots & \cdots & a_{n,n} \end{vmatrix} \\ + a_{1,2}&\begin{vmatrix} 0 & 1 & 0 & \cdots & 0 \\ a_{2,1} & \cdots & \cdots & \cdots & a_{2,n} \\ \vdots & & & & \vdots \\ a_{n,1} & \cdots & \cdots & \cdots & a_{n,n} \end{vmatrix} \\ + \cdots& \\ + a_{1,n}&\begin{vmatrix} 0 & \cdots & 0 & 1 \\ a_{2,1} & \cdots & \cdots & a_{2,n} \\ \vdots & & & \vdots \\ a_{n,1} & \cdots & \cdots & a_{n,n} \end{vmatrix}.\end{aligned}$$

Daher, und wegen Teil 2 von Satz 29, genügt es, die Behauptung für den Fall

$$A = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ a_{2,1} & \cdots & \cdots & a_{2,n} \\ \vdots & & & \vdots \\ a_{n,1} & \cdots & \cdots & a_{n,n} \end{vmatrix}$$

zu zeigen. Nach Definition gilt für diesen Fall

$$\begin{aligned}\det(A) &= \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \underbrace{\prod_{i=1}^{n} a_{i,\pi(i)}}_{= 0, \text{ außer wenn } \pi(1) = 1} \\ &= \sum_{\pi \in S_n : \pi(1) = 1} \operatorname{sgn}(\pi) a_{1,1} \prod_{i=2}^{n} a_{i,\pi(i)} \\ &= \begin{vmatrix} a_{2,2} & \cdots & a_{2,n} \\ \vdots & & \vdots \\ a_{n,2} & \cdots & a_{n,n} \end{vmatrix},\end{aligned}$$

denn die Permutationen von $\{1, \ldots, n\}$, die 1 fest lassen, sind offenbar genau die Permutationen von $\{2, \ldots, n\}$. ∎

Beispiel.

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{vmatrix} = 1 \begin{vmatrix} 6 & 7 & 8 \\ 10 & 11 & 12 \\ 14 & 15 & 16 \end{vmatrix} - 2 \begin{vmatrix} 5 & 7 & 8 \\ 9 & 11 & 12 \\ 13 & 15 & 16 \end{vmatrix} + 3 \begin{vmatrix} 5 & 6 & 8 \\ 9 & 10 & 12 \\ 13 & 14 & 16 \end{vmatrix} - 4 \begin{vmatrix} 5 & 6 & 7 \\ 9 & 10 & 11 \\ 13 & 14 & 15 \end{vmatrix}.$$

Sprechweise: „Die Determinante wird nach der ersten Zeile entwickelt.“ Wegen der Sätze 28 und 29 gilt Satz 33 natürlich analog für andere Zeilen oder Spalten, zum Beispiel können wir auch nach der zweiten Spalte entwickeln:

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{vmatrix} = -2 \begin{vmatrix} 5 & 7 & 8 \\ 9 & 11 & 12 \\ 13 & 15 & 16 \end{vmatrix} + 6 \begin{vmatrix} 1 & 3 & 4 \\ 9 & 11 & 12 \\ 13 & 15 & 16 \end{vmatrix} - 10 \begin{vmatrix} 1 & 3 & 4 \\ 5 & 7 & 8 \\ 13 & 15 & 16 \end{vmatrix} + 14 \begin{vmatrix} 1 & 3 & 4 \\ 5 & 7 & 8 \\ 9 & 11 & 12 \end{vmatrix}.$$

Zweckmäßig ist es, für die Entwicklung eine Zeile oder Spalte auszuwählen, in der viele Nullen stehen.

Man beachte bei der Entwicklung das Vorzeichenmuster für die Koeffizienten:

$$\begin{matrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{matrix}$$

Satz 34. (Cramersche Regel) Sei $A \in \mathbb{K}^{n\times n}$ invertierbar, $b \in \mathbb{K}^n$ und $x \in \mathbb{K}^n$ so, dass $Ax = b$ gilt. Dann ist

$$x = \Big(\frac{\det A^{(1)}}{\det A}, \ldots, \frac{\det A^{(n)}}{\det A}\Big),$$

wobei $A^{(i)}$ die Matrix ist, die aus A entsteht, wenn man die i-te Spalte durch b ersetzt.

Beweis. Schreibe $A = (a_1, \ldots, a_n)$ mit $a_1, \ldots, a_n \in \mathbb{K}^n$ und $x = (x_1, \ldots, x_n)$ mit $x_1, \ldots, x_n \in \mathbb{K}$. Dann gilt

$$\begin{aligned} x_i \det(A) &= \det(a_1, \ldots, a_{i-1}, x_i a_i, a_{i+1}, \ldots, a_n) \\ &= \det(a_1, \ldots, a_{i-1}, \underbrace{\sum_{j=1}^{n} x_j a_j}_{=b}, a_{i+1}, \ldots, a_n) = \det A^{(i)} \end{aligned}$$

für jedes $i = 1, \ldots, n$. ∎

Teil III

Vektorräume und Lineare Abbildungen

13 Vektorräume

Definition 29. Sei $(V, +)$ eine abelsche Gruppe und $\cdot\colon \mathbb{K} \times V \to V$ so, dass

1. $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$
2. $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$
3. $\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$
4. $1 \cdot v = v$

für alle $\alpha, \beta \in \mathbb{K}$ und alle $v, w \in V$. Dann heißt $(V, +, \cdot)$ ein *Vektorraum* (engl. *vector space*) über $\mathbb{K}$, oder einfach: ein $\mathbb{K}$-Vektorraum. Die Elemente von V heißen *Vektoren* (im weiteren Sinn; vgl. Def. 17). Das Neutralelement von V heißt *Nullvektor* und wird mit dem Symbol 0 bezeichnet. Die Operation $\cdot$ heißt *Skalarmultiplikation.* Statt $\alpha \cdot v$ schreibt man auch αv.

Beispiel.

1. $\mathbb{K}$ ist ein Vektorraum über sich selbst.
2. $\mathbb{K}^n$ ist ein Vektorraum über $\mathbb{K}$.
3. $\mathbb{K}^{n\times m}$ ist ein Vektorraum über $\mathbb{K}$.
4. Seien $v_1, \ldots, v_k \in \mathbb{K}^n$ und

   $$V = \{\, \alpha_1 v_1 + \cdots + \alpha_k v_k : \alpha_1, \ldots, \alpha_k \in \mathbb{K} \,\} \subseteq \mathbb{K}^n$$

   die Menge aller Linearkombinationen von $v_1, \ldots, v_k$. Dann ist V ein Vektorraum. (Beachte: $v, w \in V, \alpha, \beta \in \mathbb{K} \Rightarrow \alpha v + \beta w \in V$.) Man sagt, V „wird von $v_1, \ldots, v_k$ aufgespannt“ oder „erzeugt“.

   Insbesondere lassen sich jeder Matrix $A \in \mathbb{K}^{n\times m}$ in natürlicher Weise vier Vektorräume zuordnen:

   - der *Spaltenraum* $\operatorname{im} A$ (engl. *column space*) – das ist die Teilmenge von $\mathbb{K}^n$, die von den Spaltenvektoren von A aufgespannt wird
   - der *Zeilenraum* $\operatorname{coim} A$ (engl. *row space*) – das ist die Teilmenge von $\mathbb{K}^m$, die von den Zeilenvektoren von A aufgespannt wird
   - der *Kern* $\ker A$ – das ist die Menge aller $x \in \mathbb{K}^m$ mit $Ax = 0$
   - der *Ko-Kern* $\operatorname{coker} A$ – das ist die Menge aller $x \in \mathbb{K}^n$ mit $xA = 0$.

   Dass Spaltenraum und Zeilenraum Vektorräume sind, ist klar. Dass der Kern ein Vektorraum ist, ergibt sich aus den Ergebnissen des Abschnitts über Gleichungssysteme. Dass auch der Ko-Kern ein Vektorraum ist, folgt dann unmittelbar aus $\operatorname{coker} A = \ker A^\top$.
5. $\mathbb{Q}(\sqrt{2})$ ist ein Vektorraum über $\mathbb{Q}$ (vgl. Bsp. 3 nach Def. 16).
6. $\mathbb{C}$ ist ein Vektorraum über $\mathbb{R}$ (vgl. Bsp. 4 nach Def. 16).

7. $\mathbb{R}$ ist ein Vektorraum über $\mathbb{Q}$, und auch über $\mathbb{Q}(\sqrt{2})$.

8. $\mathbb{K}[X]$ ist ein Vektorraum über $\mathbb{K}$. Ebenso die Obermenge $\mathbb{K}[[X]]$ und die Untermenge

$$\mathbb{K}[X]_{\leq 3} := \{\, \alpha_0 + \alpha_1 X + \alpha_2 X^2 + \alpha_3 X^3 : \alpha_0, \ldots, \alpha_3 \in \mathbb{K} \,\} \subseteq \mathbb{K}[X]$$

   aller Polynome vom Grad höchstens drei.

9. Die Menge $\mathbb{K}^{\mathbb{N}}$ aller Folgen in $\mathbb{K}$ bildet einen Vektorraum über $\mathbb{K}$, wenn man definiert

$$\begin{aligned}(a_0, a_1, \ldots) + (b_0, b_1, \ldots) &:= (a_0 + b_0, a_1 + b_1, \ldots)\\ \alpha(a_0, a_1, \ldots) &:= (\alpha a_0, \alpha a_1, \ldots).\end{aligned}$$

   Allgemeiner: Die Menge $\mathbb{K}^A$ aller Funktionen $f\colon A \to \mathbb{K}$ bildet in natürlicher Weise einen Vektorraum über $\mathbb{K}$.

10. Im Fall $\mathbb{K} = \mathbb{R}$ bildet die Menge aller konvergenten Folgen einen Vektorraum, denn es gilt ja: sind $(a_n)_{n=0}^{\infty}$, $(b_n)_{n=0}^{\infty}$ konvergent, so ist auch $(\alpha a_n + \beta b_n)_{n=0}^{\infty}$ konvergent, für jede Wahl von Konstanten $\alpha, \beta \in \mathbb{R}$.

    Auch die Menge $N \subseteq \mathbb{R}^{\mathbb{N}}$ aller Nullfolgen bildet einen Vektorraum über $\mathbb{R}$, denn mit $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 0$ gilt auch $\lim_{n\to\infty}(\alpha a_n + \beta b_n) = 0$ für jede Wahl von Konstanten $\alpha, \beta \in \mathbb{R}$.

11. Die Menge $C(\mathbb{R}, \mathbb{R})$ aller stetigen Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ ist ein Vektorraum über $\mathbb{R}$, denn wenn f, g stetig sind, so ist auch $\alpha f + \beta g$ stetig, für jede Wahl von Konstanten $\alpha, \beta \in \mathbb{R}$.

    Ebenso die Menge $C^1(\mathbb{R}, \mathbb{R})$ aller differenzierbaren Funktionen $f\colon \mathbb{R} \to \mathbb{R}$, deren Ableitung f' stetig ist, sowie die Menge $C^\infty(\mathbb{R}, \mathbb{R})$ aller beliebig oft differenzierbaren Funktionen.

12. Seien $a, b, c\colon [0,1] \to \mathbb{R}$ stetige Funktionen und sei V die Menge aller Funktionen $f\colon [0,1] \to \mathbb{R}$, die mindestens zwei mal differenzierbar sind und für die gilt

$$a(x)f(x) + b(x)f'(x) + c(x)f''(x) = 0$$

    für alle $x \in [0,1]$. Eine solche Gleichung nennt man eine (lineare) *Differentialgleichung* (zweiter Ordnung), und die Funktionen f heißen *Lösungen* (engl. *solution*) der Differentialgleichung.

    Die Menge $V \subseteq C^2([0,1], \mathbb{R})$ bildet einen Vektorraum über $\mathbb{R}$, denn wenn f, g Lösungen sind und $\alpha, \beta \in \mathbb{R}$ Konstanten, dann folgt aus

$$\begin{aligned} a\,f + b\,f' + c\,f'' &= 0 \qquad \mid \cdot\,\alpha\\ a\,g + b\,g' + c\,g'' &= 0 \qquad \mid \cdot\,\beta\end{aligned}$$

    durch Addition, dass

$$a\,(\alpha f + \beta g) + b\,(\alpha f + \beta g)' + c\,(\alpha f + \beta g)'' = 0,$$

    also $\alpha f + \beta g \in V$.

13. Seien $a, b, c\colon \mathbb{Z} \to \mathbb{K}$ Funktionen und sei V die Menge aller Funktionen $f\colon \mathbb{Z} \to \mathbb{K}$, für die gilt
$$a(n)f(n) + b(n)f(n+1) + c(n)f(n+2) = 0$$
für alle $n \in \mathbb{Z}$. Eine solche Gleichung heißt (lineare) *Rekurrenz* (zweiter Ordnung), und die Funktionen f heißen *Lösungen* der Rekurrenz.

    Die Menge $V \subseteq \mathbb{K}^{\mathbb{Z}}$ bildet einen Vektorraum über $\mathbb{K}$, denn wenn f, g Lösungen sind und $\alpha, \beta \in \mathbb{K}$ Konstanten, dann folgt aus
$$\begin{aligned} a(n)f(n) + b(n)f(n+1) + c(n)f(n+2) &= 0 \qquad &&| \cdot \alpha \\ a(n)g(n) + b(n)g(n+1) + c(n)g(n+2) &= 0 \qquad &&| \cdot \beta \end{aligned}$$
durch Addition, dass auch die Funktion $h\colon \mathbb{Z} \to \mathbb{K}$, $h(n) := \alpha f(n) + \beta g(n)$ die Rekurrenz erfüllt.

Sie werden im Verlauf Ihres Studiums noch viele weitere Vektorräume kennenlernen. Es lohnt sich deshalb, im folgenden beim Begriff „Vektor“ nicht nur an Pfeile zu denken, mit denen eine bestimmte geometrische Anschauung verbunden ist, sondern eine allgemeinere Vorstellung des Begriffs zu entwickeln, die die Beispiele oben miteinschließt. Ein Vektor ist ab jetzt einfach ein Element eines Vektorraums, und ein Vektorraum ist alles, was die Bedingungen aus Definition 29 erfüllt, egal ob man sich darunter räumlich etwas vorstellen kann oder nicht.

Satz 35. Sei V ein Vektorraum über $\mathbb{K}$. Dann gilt:

1. $\forall\, v \in V : 0 \cdot v = 0$
2. $\forall\, v \in V : (-1) \cdot v = -v$
3. $\forall\, \alpha \in \mathbb{K} : \alpha \cdot 0 = 0$
4. $\forall\, \alpha \in \mathbb{K}\ \forall\, v \in V : \alpha v = 0 \Rightarrow \alpha = 0 \vee v = 0$

Beweis. Übung ∎

Definition 30. Sei V ein $\mathbb{K}$-Vektorraum. Eine Teilmenge $\emptyset \neq U \subseteq V$ heißt *Untervektorraum* (oder einfach: *Unterraum*, engl: *subspace*) von V, falls gilt:
$$\forall\, u, v \in U\ \forall\, \alpha, \beta \in \mathbb{K} : \alpha u + \beta v \in U.$$

Beispiel.

1. Sind $u_1, \ldots, u_k \in \mathbb{K}^n$, so ist die Menge U aller Linearkombinationen von $u_1, \ldots, u_k$ ein Unterraum von $\mathbb{K}^n$. Man sagt dann, U ist die *lineare Hülle* (engl. *span*) von $u_1, \ldots, u_k$ in $\mathbb{K}^n$. Schreibweise: $V = \langle u_1, \ldots, u_k \rangle$ oder $V = \operatorname{span}(u_1, \ldots, u_k)$.

    Allgemeiner kann man statt $\mathbb{K}^n$ irgendeinen Vektorraum V betrachten: für jede Wahl von $u_1, \ldots, u_k \in V$ ist die Menge $\langle u_1, \ldots, u_k \rangle$ aller Linearkombinationen von $u_1, \ldots, u_k$ ein Unterraum von V.
2. $\mathbb{Q}(\sqrt{2})$ ist als $\mathbb{Q}$-Vektorraum ein Unterraum von $\mathbb{R}$.

3. $\mathbb{K}[X]_{\leq 3}$ ist ein Unterraum von $\mathbb{K}[X]$, und $\mathbb{K}[X]$ ist ein Unterraum von $\mathbb{K}[[X]]$.

4. Die Nullfolgen in $\mathbb{R}$ bilden einen Unterraum des Raums der konvergenten Folgen, und diese bilden einen Unterraum des Raums aller Folgen in $\mathbb{R}$.

5. Die differenzierbaren Funktionen bilden einen Unterraum des Raums aller stetigen Funktionen, und diese einen Unterraum des Raums aller reellen Funktionen.

6. Die Menge aller Lösungen $f\colon [0,1] \to \mathbb{R}$ der linearen Differentialgleichung

$$a(x)f(x) + b(x)f'(x) + c(x)f''(x) = 0$$

bildet einen Unterraum des Vektorraums $C^2([0,1],\mathbb{R})$ aller zweimal stetig differenzierbaren Funktionen.

Satz 36. Sei V ein $\mathbb{K}$-Vektorraum und $U \subseteq V$ ein Unterraum von V. Dann ist U auch ein $\mathbb{K}$-Vektorraum.

Beweis. Es ist zu zeigen, dass $(U,+)$ eine abelsche Gruppe ist und dass die Skalarmultiplikation die Gesetze aus Definition 29 erfüllt. Dass die Gesetze an sich erfüllt sind, folgt schon daraus, dass sie für V erfüllt sind und $U \subseteq V$ ist. Es könnte höchstens sein, dass U nicht unter allen Operationen abgeschlossen ist. Dass das nicht so ist, garantiert die Bedingung aus Definition 30, z. B. mit $\alpha = \beta = 1$ gilt $\forall\ u,v \in U : u+v \in U$ und mit $\alpha = -1, \beta = 0$ gilt $\forall\ u \in U : -u \in U$. Damit ist $(U,+)$ eine Untergruppe von $(V,+)$, und also eine Gruppe. Ähnlich argumentiert man für die Skalarmultiplikation. ∎

Satz 37. Sei V ein $\mathbb{K}$-Vektorraum und $U_1, U_2 \subseteq V$ seien Unterräume von V. Dann gilt:

1. Der Schnitt $U_1 \cap U_2$ ist ein Unterraum von V.

2. $U_1 + U_2 := \{\, u_1 + u_2 : u_1 \in U_1, u_2 \in U_2 \,\}$ ist ein Unterraum von V.

Beweis.

1. Zu zeigen: $\forall\ u,v \in U_1 \cap U_2\ \forall\ \alpha,\beta \in \mathbb{K} : \alpha u + \beta v \in U_1 \cap U_2$.

   Seien $u,v \in U_1 \cap U_2$ und $\alpha,\beta \in \mathbb{K}$ beliebig. Dann gilt $u,v \in U_1$, und weil U_1 Unterraum ist, folgt $\alpha u + \beta v \in U_1$. Ebenso gilt $\alpha u + \beta v \in U_2$, weil $u,v \in U_2$ und U_2 Unterraum ist.

   Also ist $\alpha u + \beta v \in U_1 \cap U_2$, was zu zeigen war.

2. Zu zeigen: $\forall\ u,v \in U_1 + U_2\ \forall\ \alpha,\beta \in \mathbb{K} : \alpha u + \beta v \in U_1 + U_2$.

   Seien $u,v \in U_1 + U_2$ und $\alpha,\beta \in \mathbb{K}$ beliebig. Dann gibt es $u_1, v_1 \in U_1$ und $u_2, v_2 \in U_2$ mit $u = u_1 + u_2$ und $v = v_1 + v_2$. Da U_1 und U_2 Unterräume sind, gilt

$$\alpha u_1 + \beta v_1 \in U_1 \quad \text{und} \quad \alpha u_2 + \beta v_2 \in U_2,$$

   und folglich

$$\underbrace{(\alpha u_1 + \beta v_1) + (\alpha u_2 + \beta v_2)}_{=\alpha(u_1+u_2)+\beta(v_1+v_2)} \in U_1 + U_2.$$

∎

Beispiel. Sind $v_1, \ldots, v_k \in \mathbb{K}^n$ und schreibt man $\mathbb{K}v_i := \{\, \alpha v_i : \alpha \in \mathbb{K} \,\}$ $(i = 1, \ldots, k)$, so ist $\mathbb{K}v_1 + \cdots + \mathbb{K}v_k = \langle v_1, \ldots, v_k \rangle$ der Unterraum aller Linearkombinationen von $v_1, \ldots, v_k$.

Die Vereinigung $U_1 \cup U_2$ zweier Unterräume ist genau dann wieder ein Unterraum, wenn $U_1 \subseteq U_2$ oder $U_2 \subseteq U_1$ ist. Allgemein gilt: $U_1 + U_2$ ist der kleinste Vektorraum, der $U_1 \cup U_2$ enthält.

14 Basis und Dimension

Definition 31. Sei V ein $\mathbb{K}$-Vektorraum und $B \subseteq V$.

1. B heißt *linear abhängig,* falls es für eine endliche Wahl paarweise verschiedener Vektoren $v_1, \ldots, v_k \in B$ Körperelemente $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ gibt, von denen mindestens eines nicht Null ist, so dass gilt
$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0.$$
Anderenfalls heißt B *linear unabhängig.*

2. B heißt *Erzeugendensystem* (engl. *generating set*) von V, falls gilt: für alle $v \in V$ existieren $v_1, \ldots, v_k \in B$ und $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$, so dass
$$v = \alpha_1 v_1 + \cdots + \alpha_k v_k.$$
Schreibweise in diesem Fall: $V = \langle B \rangle$ oder $V = \operatorname{span}(B)$. Sprechweise: „Die Elemente von B spannen V auf“ oder „erzeugen V“.

3. Ein linear unabhängiges Erzeugendensystem von V heißt *Basis* von V.

Beispiel. Sei $V = \mathbb{Q}^4$.

1. $\left\{ \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \right\}$ ist eine Basis von V.
Allgemeiner: ist $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ der i-te Einheitsvektor, so ist $\{e_1, \ldots, e_n\}$ eine Basis von $\mathbb{K}^n$, die sogenannte *Standardbasis.*

2. $\left\{ \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\end{pmatrix} \right\}$ ist auch eine Basis von V.

3. $\left\{ \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}4\\3\\2\\1\end{pmatrix} \right\}$ ist ein Erzeugendensystem, aber keine Basis von V, da die Menge nicht linear unabhängig ist. Es gilt nämlich:
$$1\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + 1\begin{pmatrix}1\\1\\0\\0\end{pmatrix} + 1\begin{pmatrix}1\\1\\1\\0\end{pmatrix} + 1\begin{pmatrix}1\\1\\1\\1\end{pmatrix} + (-1)\begin{pmatrix}4\\3\\2\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix}.$$

4. $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ ist zwar linear unabhängig, aber kein Erzeugendensystem von V, da z. B. der Vektor $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ nicht als Linearkombination von $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ dargestellt werden kann.

Für jeden Vektorraum V gilt: V selbst ist ein Erzeugendensystem. Allerdings ist V niemals linear unabhängig, da immer $0 \in V$ gilt und $1 \cdot 0 = 0$ zeigt, dass $\{0\} \subseteq V$ linear abhängig ist.

Satz 38. Sei V ein $\mathbb{K}$-Vektorraum, und seien $B_1, B_2 \subseteq V$ mit $B_1 \subseteq B_2$.

1. Ist B_1 linear abhängig, so ist auch B_2 linear abhängig.
2. Ist B_2 linear unabhängig, so ist auch B_1 linear unabhängig.
3. Ist B_1 ein Erzeugendensystem von V, so ist auch B_2 ein Erzeugendensystem von V.
4. Ist B_2 kein Erzeugendensystem von V, so ist auch B_1 kein Erzeugendensystem von V.

Beweis.

1. Zu zeigen: es gibt $v_1, \ldots, v_k \in B_2$ und $(\alpha_1, \ldots, \alpha_k) \in \mathbb{K}^k \setminus \{0\}$ mit $\alpha_1 v_1 + \cdots + \alpha_k v_k = 0$. Nach Annahme gilt dies für B_1 anstelle von B_2, und damit wegen $B_1 \subseteq B_2$ erstrecht auch für B_2.
2. Folgt direkt aus Teil 1.
3. Zu zeigen: für alle $v \in V$ existieren $v_1, \ldots, v_k \in B_2$ und $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ so dass $v = \alpha_1 v_1 + \cdots + \alpha_k v_k$. Nach Annahme gilt dies für B_1, und wegen $B_1 \subseteq B_2$ erstrecht auch für B_1.
4. Folgt direkt aus Teil 3. ∎

Satz 39. Sei V ein $\mathbb{K}$-Vektorraum und $B \subseteq V$. Dann sind folgende Aussagen äquivalent:

1. B ist eine Basis von V.
2. B ist ein Erzeugendensystem von V und für jedes $b \in B$ gilt, dass $B \setminus \{b\}$ kein Erzeugendensystem von V ist.
3. B ist linear unabhängig und für jedes $v \in V \setminus B$ gilt, dass $B \cup \{v\}$ linear abhängig ist.
4. Jedes $v \in V$ lässt sich in eindeutiger Weise als Linearkombination von Elementen aus B schreiben.

Beweis.

(1)⇒(2) Gäbe es ein $b \in B$, so dass $B \setminus \{b\}$ auch ein Erzeugendensystem von V ist, dann gäbe es für dieses b eine Darstellung

$$b = \alpha_1 b_1 + \cdots + \alpha_k b_k$$

mit $b_1, \ldots, b_k \in B \setminus \{b\}$ und $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$. Dann ist aber $\alpha_1 b_1 + \cdots + \alpha_k b_k + (-1)b = 0$, d. h. B wäre linear abhängig. Das steht im Widerspruch zur Annahme, dass B eine Basis ist.

(2)⇒(3) zu zeigen: (a) B ist linear unabhängig und (b) $B \cup \{v\}$ ist linear abhängig für jedes $v \in V \setminus B$.

(a) Wäre B linear abhängig, so gäbe es eine Abhängigkeit

$$\alpha_1 b_1 + \cdots + \alpha_k b_k = 0$$

für gewisse $b_1, \ldots, b_k \in B$ und $(\alpha_1, \ldots, \alpha_k) \in \mathbb{K}^k \setminus \{0\}$. O.B.d.A. können wir annehmen, dass $\alpha_1 \neq 0$. Dann aber ist

$$b_1 = (-\frac{\alpha_2}{\alpha_1})b_2 + \cdots + (-\frac{\alpha_k}{\alpha_1})b_k.$$

Damit kann jede Darstellung eines Vektors v, in der b_1 vorkommt, in eine andere übersetzt werden, in der b_1 nicht vorkommt. Also ist auch $B \setminus \{b_1\}$ ein Erzeugendensystem, im Widerspruch zur Annahme.

(b) Sei $v \in V \setminus B$ beliebig. Da B ein Erzeugendensystem ist, gibt es $b_1, \ldots, b_k \in B$ und $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ mit $v = \alpha_1 b_1 + \cdots + \alpha_k b_k$. Aber dann gilt

$$\alpha_1 b_1 + \cdots + \alpha_k b_k + (-1)v = 0,$$

d. h. $\{b_1, \ldots, b_k, v\}$ ist linear abhängig, und also auch $B \cup \{v\}$.

(3)⇒(4) (a) Existenz: zu zeigen ist, dass sich jedes $v \in V$ als Linearkombination von Elementen aus B schreiben lässt. Für $v \in B$ ist das offensichtlich. Nehmen wir also an, $v \notin B$. Nach Voraussetzung $B \cup \{v\}$ linear abhängig, d. h. es gibt $b_1, \ldots, b_k \in B$ und $(\alpha_1, \ldots, \alpha_k) \in \mathbb{K}^k \setminus \{0\}$ mit

$$\alpha_1 b_1 + \cdots + \alpha_k b_k + \alpha_{k+1} v = 0.$$

Es kann nicht $\alpha_{k+1} = 0$ sein, sonst wäre $\{b_1, \ldots, b_k\}$ linear abhängig, und damit auch B. Wenn aber $\alpha_{k+1} \neq 0$ ist, haben wir

$$v = (-\frac{\alpha_1}{\alpha_{k+1}})b_1 + \cdots + (-\frac{\alpha_k}{\alpha_{k+1}})b_k,$$

d. h. v lässt sich als Linearkombination von Elementen aus B schreiben.

(b) Eindeutigkeit: Sind

$$v = \alpha_1 b_1 + \cdots + \alpha_k b_k$$

und

$$v = \tilde{\alpha}_1 b_1 + \cdots + \tilde{\alpha}_k b_k$$

zwei Darstellungen eines Vektors $v \in V$ durch Elemente von B, so folgt

$$0 = (\alpha_1 - \tilde{\alpha}_1)b_1 + \cdots + (\alpha_k - \tilde{\alpha}_k)b_k,$$

und da B linear unabhängig ist, folgt $\alpha_1 = \tilde{\alpha}_1, \ldots, \alpha_k = \tilde{\alpha}_k$.

(4)⇒(1) Es ist klar, dass B ein Erzeugendensystem ist. Wäre B nicht linear unabhängig, so gäbe es $b_1, \ldots, b_k \in B$ und $(\alpha_1, \ldots, \alpha_k) \in \mathbb{K}^k \setminus \{0\}$ mit

$$\alpha_1 b_1 + \cdots + \alpha_k b_k = 0.$$

Eine andere Darstellung von $0 \in V$ durch Elemente von B ist aber $0b_1 + \cdots + 0b_k = 0$, im Widerspruch zur vorausgesetzten Eindeutigkeit solcher Darstellungen. ∎

Satz 40. Sei V ein $\mathbb{K}$-Vektorraum, $M_1 \subseteq V$ sei ein Erzeugendensystem von V und $M_2 \subseteq V$ sei linear unabhängig. Dann gilt $|M_1| \geq |M_2|$.

Beweis. Angenommen nicht. Dann ist $k := |M_1| < |M_2|$ insbesondere endlich und es gibt paarweise verschiedene Vektoren $v_1, \ldots, v_{k+1} \in M_2$.

Nach Satz 38 ist $\{v_1, \ldots, v_{k+1}\}$ linear unabhängig, weil M_2 linear unabhängig ist. Wir zeigen, dass $\{v_1, \ldots, v_{k+1}\}$ linear abhängig ist und kommen so zu einem Widerspruch zur Annahme $|M_1| < |M_2|$.

Schreibe $M_1 = \{b_1, \ldots, b_k\}$. Da M_1 nach Voraussetzung ein Erzeugendensystem ist, lässt sich jedes $v_i \in M_2 \subseteq V$ als Linearkombination von $b_1, \ldots, b_k$ schreiben, etwa

$$\begin{aligned} v_1 &= \alpha_{1,1}b_1 + \alpha_{1,2}b_2 + \cdots + \alpha_{1,k}b_k \\ &\vdots \\ v_{k+1} &= \alpha_{k+1,1}b_1 + \alpha_{k+1,2}b_2 + \cdots + \alpha_{k+1,k}b_k. \end{aligned}$$

Betrachte die Matrix

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,k} \\ \vdots & \ddots & \vdots \\ \alpha_{k+1,1} & \cdots & \alpha_{k+1,k} \end{pmatrix} \in \mathbb{K}^{(k+1)\times k}.$$

Wegen Satz 23 in Verbindung mit Satz 24 gilt $\operatorname{Rang} A \leq k$. Damit gibt es $\beta_1, \ldots, \beta_{k+1} \in \mathbb{K}$, von denen nicht alle Null sind, so dass

$$\beta_1 \begin{pmatrix} \alpha_{1,1} \\ \vdots \\ \alpha_{1,k} \end{pmatrix} + \cdots + \beta_{k+1} \begin{pmatrix} \alpha_{k+1,1} \\ \vdots \\ \alpha_{k+1,k} \end{pmatrix} = 0,$$

d. h.

$$\begin{aligned} \beta_1\alpha_{1,1} + \cdots + \beta_{k+1}\alpha_{k+1,1} &= 0, \\ &\vdots \\ \beta_1\alpha_{1,k} + \cdots + \beta_{k+1}\alpha_{k+1,k} &= 0, \end{aligned}$$

d.h.

$$\begin{aligned}(\beta_1\alpha_{1,1} + \cdots + \beta_{k+1}\alpha_{k+1,1})b_1 &= 0,\\ &\vdots\\ (\beta_1\alpha_{1,k} + \cdots + \beta_{k+1}\alpha_{k+1,k})b_k &= 0.\end{aligned}$$

Addition dieser k Gleichungen liefert

$$0 = \beta_1 \underbrace{(\alpha_{1,1}b_1 + \cdots + \alpha_{1,k}b_k)}_{=v_1} + \cdots + \beta_{k+1}\underbrace{(\alpha_{k+1,1}b_1 + \cdots + \alpha_{k+1,k}b_k)}_{=v_{k+1}},$$

also ist $\{v_1, \ldots, v_{k+1}\}$ linear abhängig. ∎

Satz 41. Sei V ein $\mathbb{K}$-Vektorraum und B_1, B_2 seien Basen von V. Dann gilt: $|B_1| = |B_2|$.

Beweis. Da B_1 als Basis insbesondere ein Erzeugendensystem von V ist und B_2 als Basis insbesondere linear unabhängig ist, folgt aus dem vorherigen Satz $|B_1| \geq |B_2|$. Umgekehrt ist auch B_2 ein Erzeugendensystem und B_1 ist linear unabhängig, so dass auch $|B_2| \geq |B_1|$ gilt. ∎

Definition 32. Sei V ein $\mathbb{K}$-Vektorraum und B eine Basis von V. Dann heißt

$$\dim V := |B| \in \mathbb{N} \cup \{\infty\}$$

die *Dimension* von V.

Wegen des vorherigen Satzes hängt die Dimension nicht von der Wahl der Basis ab, sondern nur vom Vektorraum. Die Definition ist also in dieser Form zulässig.

Ist M irgendeine linear unabhängige Teilmenge von V, so gilt $|M| \leq \dim V$, und ist M irgendein Erzeugendensystem von V, so ist $|M| \geq \dim V$.

Wenn V sowohl als Vektorraum über $\mathbb{K}$ als auch als Vektorraum über einem anderen Körper $\mathbb{K}'$ aufgefasst werden kann, dann hängt die Dimension im allgemeinen davon ab, welchen Körper man zugrunde legt. Man schreibt deshalb auch $\dim_{\mathbb{K}} V$ statt $\dim V$, wenn der Körper nicht aus dem Zusammenhang klar ist.

Beispiel.

1. $\dim\{0\} = 0$
2. $\dim \mathbb{K} = 1$, wenn man $\mathbb{K}$ als Vektorraum über sich selbst auffasst.
3. $\dim \mathbb{K}^n = n$
4. $\dim \mathbb{K}^{n\times m} = nm$
5. Sei $v_1, \ldots, v_k \in \mathbb{K}^n$ und sei

   $$V = \{\, \alpha_1 v_1 + \cdots + \alpha_k v_k : \alpha_1, \ldots, \alpha_k \in \mathbb{K} \,\}$$

   die Menge aller Linearkombinationen von $v_1, \ldots, v_k$. Klarerweise ist $\{v_1, \ldots, v_k\}$ ein Erzeugendensystem von V, d. h. es gilt $V = \langle v_1, \ldots, v_k\rangle$.

Wegen Satz 38 gilt daher $\dim V \leq k$. Gleichheit gilt genau dann, wenn $\{v_1, \ldots, v_k\}$ linear unabhängig (und damit eine Basis) ist.

Um das zu überprüfen bzw. um eine Basis zu bestimmen, berechnet man eine Trepenform von $\begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix} \in \mathbb{K}^{k \times n}$. Die von 0 verschiedenen Zeilen bilden eine Basis von V. (Beweis: Übung.)

Insbesondere gilt: $\dim V = \text{Rang} \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix}$.

6. Eine Basis von $\mathbb{K}[X]$ ist $B = \{1, X, X^2, X^3, \ldots\}$. Es gilt also $\dim \mathbb{K}[X] = \infty$. Eine andere Basis von $\mathbb{K}[X]$ ist $\{1, X, X(X-1), X(X-1)(X-2), X(X-1)(X-2)(X-3), \ldots\}$. Es gilt sogar: Wenn $b\colon \mathbb{N} \to \mathbb{K}[X] \setminus \{0\}$ eine beliebige Folge von Polynomen ist mit der Eigenschaft $\deg b_n = n$ für alle $n \in \mathbb{N}$, dann ist $\{b_n : n \in \mathbb{N}\}$ eine Basis von $\mathbb{K}[X]$. Dabei bezeichnet $\deg b_n$ den Grad des Polynoms b_n.

7. Eine Basis von $\mathbb{K}[[X]]$ ist nicht bekannt. Aus dem nächsten Satz folgt aber, dass wegen $\mathbb{K}[X] \subseteq \mathbb{K}[[X]]$ und $\dim \mathbb{K}[X] = \infty$ der Vektorraum $\mathbb{K}[[X]]$ allenfalls eine unendliche Basis haben kann. Im darauffolgenden Satz 43 werden wir zeigen, dass jeder Vektorraum eine Basis hat. Es gilt also $\dim \mathbb{K}[[X]] = \infty$.

Satz 42. Sei V ein $\mathbb{K}$-Vektorraum und $U \subseteq V$ ein Unterraum von V. Dann gilt $\dim U \leq \dim V$.

Im Fall $\dim V < \infty$ gilt außerdem $\dim U = \dim V \iff U = V$.

Beweis. Sei B_U eine Basis von U und B_V eine Basis von V. Dann gilt $|B_U| = \dim U$ und $|B_V| = \dim V$ und wegen $U \subseteq V$ auch $B_U \subseteq V$. Als Basis von U ist B_U linear unabhängig, und als Basis von V ist B_V ein Erzeugendensystem von V. Aus Satz 40 folgt deshalb $|B_U| \leq |B_V|$. Damit ist die Ungleichung bewiesen.

Zur Gleichheit ist offensichtlich, dass $U = V \Rightarrow \dim U = \dim V$ gilt. Es bleibt also zu zeigen, dass $U \subseteq V \wedge \dim U = \dim V \Rightarrow U = V$ gilt. Angenommen nicht, d. h. angenommen es gilt $U \subsetneq V$. Dann gibt es also mindestens ein $v \in V$, das nicht in U liegt. Ist B eine Basis von U, so ist dann $B \cup \{v\}$ linear unabhängig, und damit $\dim V \geq \dim U + 1 > \dim U$. ∎

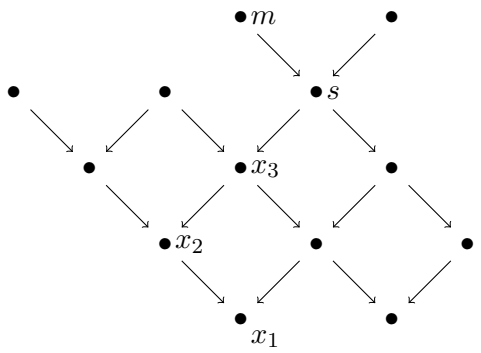
Als nächstes wollen wir beweisen, dass jeder Vektorraum eine Basis hat. Dazu brauchen wir zunächst ein weiteres Axiom aus der Mengenlehre.

Axiom. (Lemma von Zorn) Es sei M eine Menge, $\leq$ eine Halbordnung auf M (d. h. eine Relation, die reflexiv, transitiv und antisymmetrisch ist), und es gelte: Für jede Teilmenge $T \subseteq M$, so dass $\leq$ auf T eine Totalordnung ist (d. h. es gilt $\forall\, x, y \in T : x \leq y \vee y \leq x$) existiert ein $s \in M$, so dass für jedes $x \in T$ gilt $x \leq s$.

Dann gilt: $\exists\, m \in M\ \forall\, x \in M : m \leq x \Rightarrow m = x$.

Beispiel.

1.



2. $M = [0,1] \subseteq \mathbb{R}$. Es gilt: jede monoton steigende Folge $x_1 \leq x_2 \leq \cdots$ in $[0,1]$ hat eine obere Schranke in $[0,1]$, nämlich zum Beispiel $\sup\{x_1, x_2, \dots\}$. (Beachte: $[0,1]$ ist abgeschlossen.) Aus dem Axiom folgt die Existenz eines Maximums in $[0,1]$. Dieses Maximum ist natürlich das Element 1.

3. Sei A eine beliebige Menge. Betrachte $M = \mathcal{P}(A)$ und die Inklusion $\subseteq$ als Halbordnung. Es gilt: Für jede Teilmenge $T \subseteq M$ mit der Eigenschaft, dass $U \subseteq V$ oder $V \subseteq U$ für alle $U, V \in T$ gilt, es ein $S \in \mathcal{P}(A)$ mit $U \subseteq S$ für alle $U \in T$, nämlich zum Beispiel $S = \bigcup_{U \in T} U$.

   Aus dem Axiom folgt die Existenz einer Menge $U \in \mathcal{P}(A)$, die nicht in einer noch größeren Menge enthalten ist. (So eine Menge U ist zum Beispiel A selbst.)

Satz 43. (Basisergänzungssatz) Sei V ein $\mathbb{K}$-Vektorraum und $A \subseteq V$ linear unabhängig. Dann gibt es eine Basis B von V mit $A \subseteq B$.

Insbesondere gilt: Jeder Vektorraum hat eine Basis.

Beweis. Betrachte die Menge

$$M = \{\, U : A \subseteq U \subseteq V \text{ und } U \text{ ist linear unabhängig} \,\}$$

zusammen mit der Halbordnung $\subseteq$. Jede Teilmenge $T \subseteq M$ mit $\forall\, U, V \in T : U \subseteq V \vee V \subseteq U$ hat eine obere Schranke, nämlich zum Beispiel $S = \bigcup_{U \in T} U$. Klarerweise gilt $U \subseteq S$ für alle $U \in T$. Außerdem gilt auch $S \in M$ (d. h. S ist linear unabhängig), denn für jede Wahl $v_1, \dots, v_k \in S$ von endlich vielen Vektoren gibt es endlich viele $U_1, \dots, U_k \in T$, so dass $v_1 \in U_1, \dots, v_k \in U_k$. Da T total geordnet ist, enthält eines dieser U_i alle anderen, etwa $U_{i_1} \subseteq U_{i_2} \subseteq \cdots \subseteq U_{i_k}$ für gewisse $i_1, \dots, i_k$. Dann gilt also $v_1, \dots, v_k \in U_{i_k}$, und da U_{i_k} wie alle Elemente von T linear unabhängig ist, folgt aus $\alpha_1 v_1 + \cdots + \alpha_k v_k = 0$ dass $\alpha_1 = \cdots = \alpha_k = 0$. Da $v_1, \dots, v_k$ beliebige Elemente von S waren, ist S linear unabhängig.

Damit ist gezeigt, dass in M jede total geordnete Teilmenge eine obere Schranke hat. Aus dem Zornschen Lemma folgt deshalb die Existenz einer Menge $B \in M$, die maximal ist in dem Sinn, dass $B \cup \{v\}$ für jedes $v \in V \setminus B$ linear abhängig ist. Mit Satz 39 folgt, dass B eine Basis von V ist. ∎

Für endlich-dimensionale Vektorräume lässt sich der Satz durch ein weniger abstraktes Argument einsehen. Betrachte dazu folgenden „Algorithmus“:

$B := A$
solange B kein Erzeugendensystem von V ist:
  wähle ein $v \in V \setminus \langle B \rangle$
  setze $B := B \cup \{v\}$
return B.

Dabei bezeichnet $\langle B \rangle$ wie üblich den Unterraum von V, der von den Elementen von B erzeugt wird. Man überzeugt sich leicht durch Induktion, dass B während des gesamten Algorithmus linear unabhängig ist, also insbesondere auch am Ende. Ferner ist klar, dass B am Ende auch ein Erzeugendensystem ist, also der Algorithmus tatsächlich eine Basis von V produziert, die A enthält. Da sich in jedem Schleifendurchlauf die Dimension von $\langle B \rangle$ um eins erhöht, ist klar, dass der Algorithmus nach spätestens $\dim V$ $(< \infty)$ vielen Schritten terminiert. Im Fall $\dim V = \infty$ funktioniert dieses Argument nicht, weil dann der Algorithmus nicht terminiert.

Beispiel. $V = \mathbb{R}^4$,

$$A = \{ \underset{a_1}{\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}}, \underset{a_2}{\begin{pmatrix} 5 \\ 6 \\ 7 \\ 8 \end{pmatrix}} \}.$$

Um A zu einer Basis von V zu ergänzen, berechne eine Treppenform von $\binom{a_1}{a_2} \in \mathbb{R}^{2\times 4}$ und fülle diese mit Einheitsvektoren zu einer quadratischen Matrix von maximalem Rang auf. A bildet zusammen mit den hinzugefügten Einheitsvektoren eine Basis von V.

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix} \begin{matrix} \urcorner -5 \\ \hookleftarrow + \end{matrix} \leftrightarrow \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Daraus folgt, dass $A \cup \{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \}$ eine Basis von V ist.

Natürlich gibt es andere Möglichkeiten. Zum Beispiel ist auch $A \cup \{ \begin{pmatrix} 1 \\ 2 \\ 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \}$ eine Basis von V.

15 Konstruktionen

Wie man mit Vektoren rechnet, ist in Definition 29 festgelegt. Wir wollen jetzt mit ganzen Vektorräumen „rechnen“. Damit ist gemeint, dass wir aus gegebenen Vektorräumen neue Vektorräume

konstruieren wollen, und für diese Vektorräume wissen wollen, wie ihre Dimension mit der Dimension der ursprünglichen Vektorräume zusammenhängt, bzw. wie man aus bekannten Basen der ursprünglichen Räume eine Basis des neuen Raums ausrechnen kann.

Beispiel. Nach Satz 37 gilt: Sind U_1, U_2 Unterräume eines $\mathbb{K}$-Vektorraums V, so sind auch $U_1 \cap U_2$ und $U_1 + U_2$ Unterräume. Wie bekommt man eine Basis für diese Räume, wenn man Basen B_1, B_2 von U_1 und U_2 kennt? Nehmen wir zur Vereinfachung an $V = \mathbb{K}^n$.

1. $U_1 + U_2$: In diesem Fall ist $B_1 \cup B_2$ ein Erzeugendensystem, aber im allgemeinen keine Basis. Wie im Beispiel 5 nach Definition 32 gezeigt, kann man $B_1 \cup B_2$ zu einer Basis machen, indem man eine Treppenform der Matrix der Zeilenvektoren bestimmt und daraus die Nullzeilen streicht. Ist zum Beispiel $U_1 = \langle \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix} \rangle$, $U_2 = \langle \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\1\end{pmatrix} \rangle$, so gilt

$$\begin{aligned} U_1 + U_2 &= \langle \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\1\end{pmatrix} \rangle \\ &= \langle \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\1\end{pmatrix} \rangle. \end{aligned}$$

2. $U_1 \cap U_2$. Für $v \in V$ gilt $v \in U_1 \cap U_2$ genau dann, wenn sich v sowohl als Linearkombination von Elementen einer Basis B_1 von U_1 als auch als Linearkombination von Elementen einer Basis B_2 von U_2 schreiben lässt. Die Menge all dieser v lässt sich bestimmen, indem man ein lineares Gleichungssystem löst.

   Beispiel: $U_1 = \langle \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix} \rangle$, $U_2 = \langle \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\1\end{pmatrix} \rangle$.

$$\begin{aligned} & \alpha_1 \begin{pmatrix}1\\1\\0\\0\end{pmatrix} + \alpha_2 \begin{pmatrix}0\\1\\0\\0\end{pmatrix} = \beta_1 \begin{pmatrix}1\\0\\0\\0\end{pmatrix} + \beta_2 \begin{pmatrix}1\\0\\1\\1\end{pmatrix} \\ \Leftrightarrow \quad & \begin{pmatrix}1 & 0 & -1 & -1\\ 1 & 1 & 0 & 0\\ 0 & 0 & 0 & -1\\ 0 & 0 & 0 & -1\end{pmatrix} \begin{pmatrix}\alpha_1\\ \alpha_2\\ \beta_1\\ \beta_2\end{pmatrix} = 0. \\ \Leftrightarrow \quad & (\alpha_1, \alpha_2, \beta_1, \beta_2) \in \langle \begin{pmatrix}1\\-1\\1\\0\end{pmatrix} \rangle. \end{aligned}$$

Daraus folgt, dass $\{1 \cdot \begin{pmatrix}1\\1\\0\\0\end{pmatrix} + (-1) \cdot \begin{pmatrix}0\\1\\0\\0\end{pmatrix}\}$ ein Erzeugendensystem von $U_1 \cap U_2$ ist.

Satz 44. Seien U_1, U_2 endlich-dimensionale Unterräume eines $\mathbb{K}$-Vektorraums V. Dann gilt:

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).$$

Beweis. Sei $\{b_1, \ldots, b_k\}$ eine Basis von $U_1 \cap U_2$. Nach Satz 43 gibt es $b_{k+1}, \ldots, b_m \in U_1$ und $\tilde{b}_{k+1}, \ldots, \tilde{b}_\ell \in U_2$, so dass

$$\{b_1, \ldots, b_k, b_{k+1}, \ldots, b_m\}$$

eine Basis von U_1 und

$$\{b_1, \ldots, b_k, \tilde{b}_{k+1}, \ldots, \tilde{b}_\ell\}$$

eine Basis von U_2 ist. Betrachte

$$B := \{b_1, \ldots, b_k, b_{k+1}, \ldots, b_m, \tilde{b}_{k+1}, \ldots, \tilde{b}_\ell\}.$$

B ist ein Erzeugendensystem von $U_1 + U_2$, denn jedes $u \in U_1 + U_2$ lässt sich schreiben als $u = u_1 + u_2$ mit $u_1 \in U_1$ und $u_2 \in U_2$, und jedes $u_1 \in U_1$ ist eine Linearkombination von $\{b_1, \ldots, b_k, b_{k+1}, \ldots, b_m\} \subseteq B$ und jedes $u_2 \in U_2$ ist eine Linearkombination von $\{b_1, \ldots, b_k, \tilde{b}_{k+1}, \ldots, \tilde{b}_\ell\} \subseteq B$, und wenn also jedes $u_1 \in U_1$ und jedes $u_2 \in U_2$ eine Linearkombination von Elementen aus B ist, dann gilt dies auch für $u = u_1 + u_2$.
B ist auch linear unabhängig: Betrachte dazu $\alpha_1, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_m, \tilde{\alpha}_{k+1}, \ldots, \tilde{\alpha}_\ell \in \mathbb{K}$ mit

$$\alpha_1 b_1 + \cdots + \alpha_m b_m + \tilde{\alpha}_{k+1}\tilde{b}_{k+1} + \cdots + \tilde{\alpha}_\ell \tilde{b}_\ell = 0.$$

Zu zeigen: $\alpha_1 = \cdots = \alpha_m = \tilde{\alpha}_{k+1} = \cdots = \tilde{\alpha}_\ell = 0$. Aus der angenommenen Relation folgt

$$\underbrace{\alpha_1 b_1 + \cdots + \alpha_m b_m}_{\in U_1} = \underbrace{(-\tilde{\alpha}_{k+1})\tilde{b}_{k+1} + \cdots + (-\tilde{\alpha}_\ell)\tilde{b}_\ell}_{\in U_2}.$$

Beide Seiten liegen also in $U_1 \cap U_2$. Der Vektor auf den beiden Seiten der Gleichung hat deshalb auch eine Darstellung $\beta_1 b_1 + \cdots + \beta_k b_k$ für gewisse $\beta_1, \ldots, \beta_k \in \mathbb{K}$. Da $\{b_1, \ldots, b_k, \tilde{b}_{k+1}, \ldots, \tilde{b}_\ell\}$ eine Basis von U_2 und damit linear unabhängig ist, folgt aus

$$\beta_1 b_1 + \cdots + \beta_k b_k = (-\tilde{\alpha}_{k+1})\tilde{b}_{k+1} + \cdots + (-\tilde{\alpha}_\ell)\tilde{b}_\ell$$

zunächst, dass $\beta_1 = \cdots = \beta_k = \tilde{\alpha}_{k+1} = \cdots = \tilde{\alpha}_\ell = 0$ ist. Wir haben es also mit dem Nullvektor zu tun, und deshalb folgt als nächstes aus

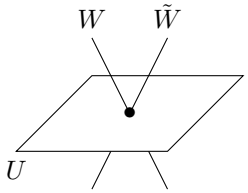
$$\alpha_1 b_1 + \cdots + \alpha_m b_m = 0$$

mit der linearen Unabhängigkeit der Basis $\{b_1, \ldots, b_m\}$ von U_1 auch $\alpha_1 = \cdots = \alpha_m = 0$.
Damit ist gezeigt, dass B eine Basis von $U_1 + U_2$ ist. Es folgt $\dim(U_1 + U_2) = |B| = m + \ell - k = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$. ∎

Gilt $U_1 \cap U_2 = \{0\}$, so schreibt man statt $U_1 + U_2$ auch $U_1 \oplus U_2$ und sagt, die Summe ist *direkt*. Im Fall einer direkten Summe gilt $\dim(U_1 \oplus U_2) = \dim U_1 + \dim U_2$, weil ja $\dim\{0\} = 0$ ist.

Wenn $U = U_1 + U_2$ ist, dann lässt sich jeder Vektor $u \in U$ schreiben als $u = u_1 + u_2$ für ein $u_1 \in U_1$ und ein $u_2 \in U_2$. Bei einer direkten Summe ist diese Darstellung eindeutig.

Aus dem Basisergänzungssatz folgt, dass es für jeden Unterraum U von V ein Unterraum W von V existiert mit $V = U \oplus W$. Einen solchen Raum W nennt man einen *Komplementärraum* von U. Der Komplementärraum von U ist im allgemeinen nicht eindeutig.



Satz 45. Seien U, W zwei $\mathbb{K}$-Vektorräume und B_U, B_W Basen von U bzw. W. Dann ist die Menge $V = U \times W$ zusammen mit

$$\underset{\text{in } V}{(u_1, w_1) + (u_2, w_2)} := \underset{\text{in } U}{(u_1 + u_2}, \underset{\text{in } W}{w_1 + w_2})$$

$$\underset{\text{in } V}{\alpha \cdot (u, w)} = (\underset{\text{in } U}{\alpha \cdot u}, \underset{\text{in } W}{\alpha \cdot w})$$

ein $\mathbb{K}$-Vektorraum, und $B = (B_U \times \{\underset{\in W}{0}\}) \cup (\{\underset{\in U}{0}\} \times B_W)$ ist eine Basis. Insbesondere gilt

$$\dim V = \dim U + \dim W.$$

Beweis. Dass V ein Vektorraum ist, zeigt man durch Nachrechnen der nötigen Gesetze. Die Dimensionsaussage folgt direkt aus der Aussage über die Basis. Wir zeigen: $B = (B_U \times \{0\}) \cup (\{0\} \times B_W)$ ist eine Basis von V.

B ist linear unabhängig: Seien $b_1, \ldots, b_k \in B$ und $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ mit $\alpha_1 b_1 + \cdots + \alpha_k b_k = 0$. Jedes b_i hat die Form $(b, 0)$ für ein $b \in B_U$ oder $(0, b)$ für ein $b \in B_W$. O.B.d.A. seien $b_1, \ldots, b_i$ von der ersten und $b_{i+1}, \ldots, b_k$ von der zweiten Form. Dann gilt

$$\alpha_1 b_1 + \cdots + \alpha_i b_i = 0 \quad \text{und} \quad \alpha_{i+1} b_{i+1} + \cdots + \alpha_k b_k = 0,$$

und da B_U und B_W linear unabhängig sind, folgt $\alpha_1 = \cdots = \alpha_i = 0$ und $\alpha_{i+1} = \cdots = \alpha_k = 0$.

B ist ein Erzeugendensystem: Sei $v \in V$. Dann ist $v = (u, w)$ für gewisse $u \in U$ und $w \in W$. Daraus folgt, dass es $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ und $u_1, \ldots, u_k \in B_U$ sowie $\beta_1, \ldots, \beta_\ell \in \mathbb{K}$ und $w_1, \ldots, w_\ell \in B_W$ gibt mit

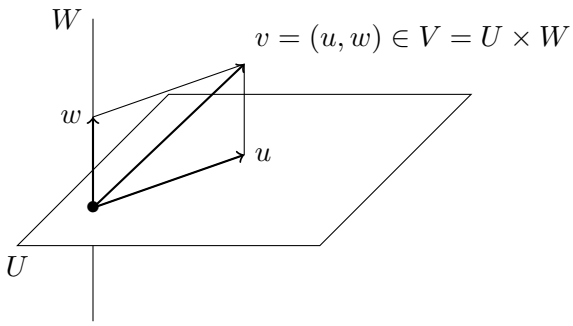
$$u = \alpha_1 u_1 + \cdots + \alpha_k u_k \quad \text{und} \quad w = \beta_1 w_1 + \cdots + \beta_\ell w_\ell.$$

Damit gilt

$$v = \begin{pmatrix} u \\ w \end{pmatrix} = \alpha_1 \begin{pmatrix} u_1 \\ 0 \end{pmatrix} + \cdots + \alpha_k \begin{pmatrix} u_k \\ 0 \end{pmatrix} + \beta_1 \begin{pmatrix} 0 \\ w_1 \end{pmatrix} + \cdots + \beta_\ell \begin{pmatrix} 0 \\ w_\ell \end{pmatrix}.$$

■

Beispiel. $U = \mathbb{R}^2$, $W = \mathbb{R}$



Wenn $\times$ eine Art „Multiplikation“ von Vektorräumen ist, wie müsste dann eine passende „Division“ aussehen? Wenn also ein Vektorraum V und ein Unterraum W von V gegeben ist, wie können wir dann sinnvoll erklären, was $U := V/W$ sein soll, damit diese Operation in gewisser Weise die Produktbildung $V = U \times W$ rückgängig macht?

Die Idee ist, dass man alle Vektoren $v \in V$, die den gleichen U-Anteil haben, als ein und denselben Vektor auffast, d. h. dass man die jeweiligen W-Anteile der Vektoren einfach ignoriert. Formal erreicht man das, indem man auf V die Äquivalenzrelation $\sim$ einführt mit

$$v_1 \sim v_2 \iff v_1 - v_2 \in W.$$

Dann gilt nämlich $v_1 \sim v_2$ genau dann, wenn v_1, v_2 den gleichen U-Anteil haben, denn dieser hebt sich dann bei der Bildung der Differenz heraus und es bleibt nur noch die Differenz der (möglicherweise unterschiedlichen) W-Anteile übrig. Es ist leicht nachzuprüfen, dass $\sim$ tatsächlich eine Äquivalenzrelation ist:

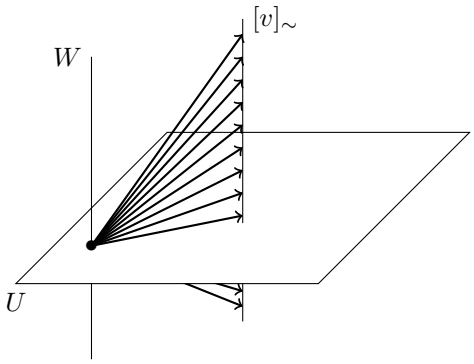
Reflexivität: Da W als Unterraum von V insbesondere ein Vektorraum ist, gilt $0 \in W$. Damit gilt für jedes $v \in V$, dass $v - v = 0 \in W$, also $v \sim v$.

Symmetrie: $v_1 \sim v_2 \Rightarrow v_1 - v_2 \in W \Rightarrow (-1) \cdot (v_1 - v_2) \in W \Rightarrow v_2 - v_1 \in W \Rightarrow v_2 \sim v_1$. Im zweiten Schritt wird wieder verwendet, dass W ein Vektorraum ist, und also abgeschlossen unter Skalarmultiplikation.

Transitivität:

$$\underbrace{\begin{array}{cc} v_1 \sim v_2 & v_2 \sim v_3 \\ \Downarrow & \Downarrow \\ v_1 - v_2 \in W & v_2 - v_3 \in W \end{array}}$$
$$\Downarrow$$
$$(v_1 - v_2) + (v_2 - v_3) = v_1 - v_3 \in W$$
$$\Downarrow$$
$$v_1 \sim v_3$$

Die Vektoren mit dem selben U-Anteil bilden genau die Äquivalenzklassen bezüglich $\sim$.



Wir zeigen als nächstes, dass sich die Menge der Äquivalenzklassen als Vektorraum auffassen lässt. Diesen Vektorraum nennt man dann den Quotientenraum V/W.

Satz 46. Sei V ein $\mathbb{K}$-Vektorraum, $W \subseteq V$ ein Unterraum von V. Für $v_1, v_2 \in V$ sei definiert $v_1 \sim v_2 \iff v_1 - v_2 \in W$. Die Menge $U := V/W := V/\sim$ bildet zusammen mit

$$\begin{aligned} +\colon U \times U \to U, &\quad [v_1]_\sim + [v_2]_\sim := [v_1 + v_2]_\sim \\ \cdot\colon \mathbb{K} \times U \to U, &\quad \alpha[v]_\sim := [\alpha v]_\sim \end{aligned}$$

einen $\mathbb{K}$-Vektorraum.

Beweis. Zu zeigen ist: (a) die Definitionen sind repräsentantenunabhängig, und (b) $(U, +, \cdot)$ ist ein Vektorraum.

(a) Addition:

$$\begin{array}{cc} v_1 \sim v_2 & \tilde{v}_1 \sim \tilde{v}_2 \\ \Downarrow & \Downarrow \\ v_1 - v_2 \in W & \tilde{v}_1 - \tilde{v}_2 \in W \end{array}$$
$$\Downarrow$$
$$(v_1 - v_2) + (\tilde{v}_1 - \tilde{v}_2) \in W$$
$$\Downarrow$$
$$v_1 + \tilde{v}_1 \sim v_2 + \tilde{v}_2$$

Skalarmultiplikation:

$$v_1 \sim v_2 \Rightarrow v_1 - v_2 \in W \Rightarrow \alpha(v_1 - v_2) = \alpha v_1 - \alpha v_2 \in W \Rightarrow \alpha v_1 \sim \alpha v_2.$$

(b) Die nötigen Gesetze sind erfüllt, weil sie nach Voraussetzung in V erfüllt sind und sich von dort übertragen. ∎

Definition 33. Der Vektorraum V/W aus obigem Satz heißt der *Quotientenraum* (engl. *quotient space*) oder *Faktorraum* von V nach W.

Satz 47. Sei V ein $\mathbb{K}$-Vektorraum, $W \subseteq V$ ein Unterraum, $U \subseteq V$ ein Komplementärraum von W (d. h. $V = U \oplus W$, d. h. $V = U + W$ und $U \cap W = \{0\}$). Weiter sei B eine Basis von U. Dann ist $\tilde{B} := \{\, [b]_\sim : b \in B \,\}$ eine Basis von V/W. Wenn U endlich-dimensional ist, gilt insbesondere $\dim V/W = \dim U$. (Wenn auch V und W endlich-dimensional sind, gilt weiters $\dim U = \dim V - \dim W$.)

Beweis. Die Dimensionsaussagen folgen unmittelbar aus der Basiseigenschaft. Daher ist nur zeigen: $\tilde{B}$ ist linear unabhängig und ein Erzeugendensystem.

(a) $\tilde{B}$ ist linear unabhängig: Seien $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ und $[b_1]_\sim, \dots, [b_k]_\sim \in \tilde{B}$ so, dass

$$\alpha_1[b_1]_\sim + \cdots + \alpha_k[b_k]_\sim = [0]_\sim.$$

Dann ist $[\alpha_1 b_1 + \cdots + \alpha_k b_k]_\sim = [0]_\sim$. Dann ist

$$\underbrace{\alpha_1 b_1 + \cdots + \alpha_k b_k}_{\in U} \in W.$$

Dann $\alpha_1 b_1 + \cdots + \alpha_k b_k = 0$, da $U \cap W = \{0\}$. Dann $\alpha_1 = \cdots = \alpha_k = 0$, da B linear unabhängig ist.

(b) $\tilde{B}$ ist ein Erzeugendensystem: Sei $[x]_\sim \in V/W$. Wegen $U + W = V$ lässt sich x schreiben als $x = u + w$ für gewisse $u \in U$ und $w \in W$. Dann lässt sich u schreiben als $u = \alpha_1 b_1 + \cdots + \alpha_k b_k$ für gewisse $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ und $b_1, \dots, b_k \in B$. Dann ist

$$[x]_\sim = [u]_\sim = [\alpha_1 b_1 + \cdots + \alpha_k b_k]_\sim = \alpha_1[b_1]_\sim + \cdots + \alpha_k[b_k]_\sim.$$

■

Beispiel. $V = \mathbb{R}^4$, $W = \langle \begin{pmatrix}1\\2\\3\\4\end{pmatrix}, \begin{pmatrix}5\\6\\7\\8\end{pmatrix} \rangle$. Nach dem Beispiel auf Seite 96 ist $U = \langle \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix} \rangle$ ein Komplementärraum von W, und also ist $\{[\begin{pmatrix}0\\0\\1\\0\end{pmatrix}]_\sim, [\begin{pmatrix}0\\0\\0\\1\end{pmatrix}]_\sim\}$ eine Basis von V/W.

Die nächste Konstruktion ist wieder eine Art „Multiplikation“ von Vektoräumen. Dabei nimmt man nicht das kartesische Produkt $U \times W$ der Räume selbst, sondern betrachtet den Vektorraum, der das kartesische Produkt einer Basis von U mit einer Basis von W als Basis hat.

Dazu beachte man zunächst, dass man jede beliebige Menge M als Basis eines $\mathbb{K}$-Vektorraums auffassen kann, nämlich des Raums $\mathbf{F}_\mathbb{K}(M)$ aller Funktionen $f\colon M \to \mathbb{K}$ mit $|\{x \in M : f(x) \neq 0\}| < \infty$ zusammen mit der naheliegenden Addition und Skalarmultiplikation. Man nennt $\mathbf{F}_\mathbb{K}(M)$ den *freien Vektorraum* über M.

Die Menge M ist insofern eine Basis von $\mathbf{F}_\mathbb{K}(M)$, als man jedes $m \in M$ identifizieren kann mit der Funktion $\tilde{m}\colon M \to \mathbb{K}$ mit $\tilde{m}(x) = 1$ falls $x = m$ und $\tilde{m}(x) = 0$ falls $x \neq m$. Für ein Element $v \in \mathbf{F}_\mathbb{K}(M)$ mit $v(m_1) = \alpha_1$, $v(m_2) = \alpha_2$ und $v(m) = 0$ für alle $m \in M \setminus \{m_1, m_2\}$ schreibt man $\alpha_1\, m_1 + \alpha_2\, m_2$, usw.

Beispiel.

1. Sei $\mathbb{K} = \mathbb{Q}$ und $M = \{a, b, c\}$. Dann besteht $\mathbf{F}_\mathbb{K}(M)$ aus allen „Linearkombinationen“ $\alpha a + \beta b + \gamma c$ mit $\alpha, \beta, \gamma \in \mathbb{K}$. Streng genommen hat diese Summe keine eigene Bedeutung sondern ist nur eine Kurzschreibweise für die Funktion $f\colon M \to \mathbb{K}$ mit $f(a) = \alpha$, $f(b) = \beta$, $f(c) = \gamma$. Allerdings sind die Addition und Skalarmultiplikation auf $\mathbf{F}_\mathbb{K}(M)$ genau so definiert, wie es die Notation suggeriert. Zum Beispiel gilt $2(5a - 3b + 8c) - (2a + 3b - c) = 8a - 3b + 17c$.

2. Für $n \in \mathbb{N}$ ist $\mathbb{K}^n = \mathbf{F}_{\mathbb{K}}(\{1, 2, \ldots, n\})$, wenn man die gewohnte Vektorschreibweise $a = (\alpha_1, \ldots, \alpha_n)$ für Elemente a von $\mathbb{K}^n$ interpretiert als eine Schreibweise für die Funktion $a\colon \{1, \ldots, n\} \to \mathbb{K}$ mit $a(i) = \alpha_i$ für $i = 1, \ldots, n$. Die Addition und Skalarmultiplikation entsprechen genau den gewohnten Vektor-Rechenregeln. In diesem Fall ist die Kurzschreibeweise $a = \alpha_1 1 + \alpha_2 2 + \cdots + \alpha_n n$ nicht zu empfehlen, weil die Symbole $1, \ldots, n$ sowohl Elemente in M als auch Elemente in $\mathbb{K}$ bezeichnen können. Wenn man aber statt $\{1, \ldots, n\}$ die Menge $\{e_1, \ldots, e_n\}$ nimmt, wobei $e_1, \ldots, e_n$ als neue Symbole verstanden werden, die nicht auch schon für gewisse Elemente von $\mathbb{K}$ stehen, dann ist die Schreibeweise $a = \alpha_1 e_1 + \cdots + \alpha_n e_n$ durchaus suggestiv. Sie passt insbesondere mit der gewohnten Schreibweise zusammen, wenn man die e_i nicht als formale Symbole interpretiert, sondern als Variablen, die für die Einheitsvektoren $(0, \ldots, 0, 1, 0, \ldots, 0)$ stehen.

3. Ähnlich wie im vorherigen Beispiel kann man sagen, dass $\mathbb{K}[X]$ nichts anderes ist als $\mathbf{F}_{\mathbb{K}}(\mathbb{N})$, wobei man für die Basiselemente $0, 1, 2, \ldots$ zur besseren Unterscheidbarkeit von Körperelementen $1, X, X^2, \ldots$ schreibt. Es sei noch einmal daran erinnert, dass auch bei unendlich großen Basen Linearkombinationen immer nur von endlich vielen Vektoren gebildet werden können. Deshalb ist z.B. $1 + X + X^2 + X^3 + \cdots$ kein Element von $\mathbb{K}[X]$.

Wenn nun U und W zwei Vektorräume sind und B_U ist eine Basis von U und B_W eine Basis von W, so kann man definieren $U \otimes W := \mathbf{F}_{\mathbb{K}}(B_U \times B_W)$. Man nennt diesen Vektorraum das *Tensorprodukt* (engl. *tensor product*) von U und W, seine Elemente heißen *Tensoren.*

Ein Tensor ist also eine Linearkombination von Paaren (b_U, b_W), wobei $b_U \in B_U$ und $b_V \in B_V$ ist. Sind $u \in U$, $w \in W$ beliebig, etwa $u = \alpha_1 b_U^{(1)} + \cdots + \alpha_k b_U^{(k)}$ für gewisse $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ und $b_U^{(1)}, \ldots, b_U^{(k)} \in B_U$, und $w = \beta_1 b_W^{(1)} + \cdots + \beta_m b_W^{(m)}$ für gewisse $\beta_1, \ldots, \beta_m \in \mathbb{K}$ und $b_W^{(1)}, \ldots, b_W^{(m)} \in B_W$, so definiert man

$$u \otimes w := \sum_{i=1}^{k} \sum_{j=1}^{m} \alpha_i \beta_j (b_U^{(i)}, b_V^{(j)}) \in U \otimes W.$$

Insbesondere gilt dann $b_U \otimes b_W = (b_U, b_W)$ für alle $b_U \in B_U$ und $b_W \in B_W$. Ferner gelten die Rechenregeln $\alpha(u \otimes w) = (\alpha u) \otimes w = u \otimes (\alpha w)$ und $(u_1 + u_2) \otimes w = (u_1 \otimes w) + (u_2 \otimes w)$ und $u \otimes (w_1 + w_2) = (u \otimes w_1) + (u \otimes w_2)$ und insbesondere $u \otimes 0 = 0 \otimes w = 0$ für alle $\alpha \in \mathbb{K}$, $u, u_1, u_2 \in U$ und $w, w_1, w_2 \in W$.

Beispiel. Im Fall $U = \mathbb{K}^n$ und $W = \mathbb{K}^m$ kann man sich $V = U \otimes W$ als den Raum $\mathbb{K}^{n \times m}$ der Matrizen vorstellen. Nimmt man für U und W jeweils die Standardbasis $\{e_1, \ldots, e_n\} \subseteq U$ bzw. $\{\tilde{e}_1, \ldots, \tilde{e}_m\} \subseteq W$, so entspricht $e_i \otimes \tilde{e}_j$ der Matrix, die eine 1 an der Stelle (i, j) hat, und überall sonst nur Nullen.

Allgemeiner: Sind $u = (u_1, \ldots, u_n) \in U$, $w = (w_1, \ldots, w_m) \in W$ zwei beliebige Vektoren, so entspricht der Tensor $u \otimes w \in U \otimes W$ der Matrix, die man erhält, wenn man u als Spaltenvektor mit w als Zeilenvektor multipliziert, also

$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} (w_1, w_2, \ldots, w_m) = \begin{pmatrix} u_1 w_1 & u_1 w_2 & \cdots & u_1 w_m \\ u_2 w_1 & u_2 w_2 & \cdots & u_2 w_m \\ \vdots & \vdots & \ddots & \vdots \\ u_n w_1 & u_2 w_n & \cdots & u_n w_m \end{pmatrix}.$$

Dass nicht jede Matrix diese Form hat, verdeutlicht, dass sich nicht jeder Tensor $v \in U \otimes W$

schreiben lässt als $v = u \otimes w$ für gewisse $u \in U, w \in W$. Es gibt zum Beispiel im Fall $n = m = 2$ keine $(u_1, u_2), (w_1, w_2) \in \mathbb{K}^2$, so dass $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} (w_1, w_2) = \begin{pmatrix} u_1 w_1 & u_1 w_2 \\ u_2 w_1 & u_2 w_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ist.

Die obige Definition von $U \otimes W$ ist etwas unbefriedigend, weil sie auf Basen von U und W zurückgreift. Wählt man statt B_U und B_W zwei andere Basen B'_U, B'_W, so ist $\mathbf{F}_{\mathbb{K}}(B_U \times B_W)$ streng genommen nicht dasselbe wie $\mathbf{F}_{\mathbb{K}}(B'_U \times B'_W)$. Es zeigt sich aber, dass diese Räume „im wesentlichen“ dieselben sind. Um das konkret zu machen, verwendet man den Begriff der Isomorphie von Vektorräumen, um den es im folgenden Abschnitt gehen wird.

16 Lineare Abbildungen und Isomorphie

Definition 34. Seien V, W zwei $\mathbb{K}$-Vektorräume.

1. $h\colon V \to W$ heißt *Homomorphismus* oder *lineare Abbildung,* falls gilt

$$\forall\, x, y \in V\ \forall\, \alpha, \beta \in \mathbb{K} : h(\alpha \cdot x \overset{\text{in } V}{+} \beta \cdot y) = \alpha \cdot h(x) \overset{\text{in } W}{+} \beta \cdot h(y).$$

   Ein Homomorphismus von V nach V heißt auch *Endomorphismus.*

2. Ein bijektiver Homomorphismus heißt *Isomorphismus.* Wenn ein solcher existiert, sagt man, V und W sind (zueinander) *isomorph.* Schreibweise: $V \cong W$.

   Ein Isomorphismus von V nach V heißt auch *Automorphismus.*

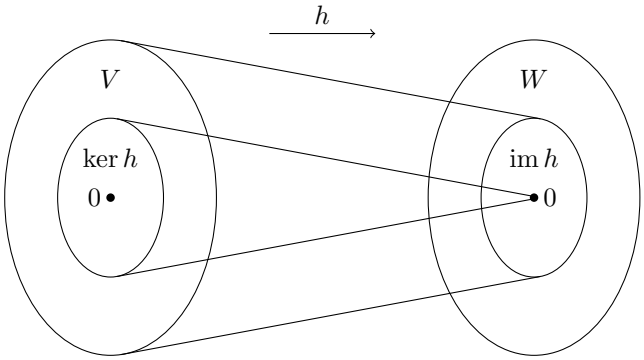
3. Ist $h\colon V \to W$ ein Homomorphismus, so heißt

$$\ker h := \{\, x \in V : h(x) = 0 \,\}$$

   der *Kern* (engl. *kernel*) und

$$\operatorname{im} h := h(V) = \{\, h(x) : x \in V \,\}$$

   das *Bild* (engl. *image*) von h.



Beispiel.

1. $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = 3x$ ist linear. Die Funktionen $g\colon \mathbb{R} \to \mathbb{R}$, $g(x) = 3x+5$ und $h\colon \mathbb{R} \to \mathbb{R}$, $h(x) = x^2$ sind nicht linear.

2. $V = \mathbb{K}^m$, $W = \mathbb{K}^n$. Für jede beliebige Matrix $A \in \mathbb{K}^{n\times m}$ ist

$$h\colon V \to W, \quad h(x) = Ax$$

eine lineare Abbildung. Es gilt ja

$$h(\alpha x + \beta y) = A(\alpha x + \beta y) = \alpha Ax + \beta Ay = \alpha h(x) + \beta h(y)$$

für alle $\alpha, \beta \in \mathbb{K}$ und alle $x, y \in V$.

3. $V = \mathbb{K}[X]$, $W = \mathbb{K}^{\mathbb{K}}$. Die Funktion

$$h\colon V \to W, \quad h(a_0 + a_1X + \cdots + a_nX^n) := (z \mapsto a_0 + a_1z + \cdots + a_nz^n),$$

die jedem Polynom die entsprechende Polynomfunktion zuordnet, ist linear.

4. Die Abbildung

$$h\colon \mathbb{K}[[X]] \to \mathbb{K}[X], \quad h(\sum_{k=0}^{\infty} a_nX^n) := \sum_{k=0}^{12} a_nX^n$$

ist linear.

5. Die Abbildung $h\colon \mathbb{K}^{n\times m} \to \mathbb{K}^n$, die dadurch definiert ist, dass $h(A)$ die erste Spalte von A ist, ist linear.

6. Sei $V = \mathbb{K}[[X]]$ und $N \in \mathbb{N}$ fix. Dann ist die Abbildung

$$[X^N]\colon V \to \mathbb{K}, \qquad [X^N]\sum_{n=0}^{\infty} a_nX^n := a_N,$$

die aus einer formalen Potenzreihe den N-ten Koeffizient extrahiert, linear.

7. Sei V der $\mathbb{R}$-Vektorraum aller konvergenten Folgen, $W = \mathbb{R}$. Dann ist

$$h\colon V \to W, \quad h(\,(a_n)_{n=0}^{\infty}\,) := \lim_{n\to\infty} a_n$$

eine lineare Abbildung. Ihr Kern ist der Raum aller Nullfolgen.

8. Die Funktion $h\colon \mathbb{Q} \to \mathbb{Q}(\sqrt{2})$ mit $h(x) = x + 0\sqrt{2}$ für alle $x \in \mathbb{Q}$ ist linear.

9. Die Funktionen $\mathrm{Re}, \mathrm{Im}\colon \mathbb{C} \to \mathbb{R}$, die jeder komplexen Zahl ihren Realteil bzw. ihren Imaginärteil zuordnen, sind linear, wenn man $\mathbb{C}$ als Vektorraum über $\mathbb{R}$ auffasst.

10. Sei $V = C^5([-1,1],\mathbb{R})$ die Menge aller fünf mal stetig differenzierbaren Funktionen $f\colon [-1,1] \to \mathbb{R}$. Die Funktion $h\colon V \to \mathbb{R}^6$ mit $h(f) = (f(0), f'(0), \ldots, f^{(5)}(0))$ ist linear.

11. Seien $v_2, \ldots, v_n \in \mathbb{K}^n$ fix und $h\colon \mathbb{K}^n \to \mathbb{K}$ definiert durch $h(x) := \det(x, v_2, \ldots, v_n)$. Dann ist h eine lineare Funktion. Dabei ist mit $(x, v_2, \ldots, v_n)$ die Matrix gemeint, deren Spalten $x, v_2, \ldots, v_n$ sind.

12. Sei V ein $\mathbb{K}$-Vektorraum, U ein Unterraum von V, $W = V/U$ und $h\colon V \to W$, $h(x) = [x]_\sim$. Dann ist h eine lineare Abbildung. Ihr Kern ist U.

13. Sind U, W zwei $\mathbb{K}$-Vektorräume und $V = U \times W$, so ist $\pi\colon V \to U$, $\pi(u, w) = u$ eine lineare Abbildung, die zwar surjektiv aber nicht injektiv ist. Ihr Kern ist $\ker \pi = \{0\} \times W$.

14. Sind U, W zwei $\mathbb{K}$-Vektorräume, so ist $h\colon U \times W \to U \otimes W$, $h(u, w) = u \otimes w$ eine lineare Abbildung, aber kein Isomorphismus, weil h zwar injektiv, aber nicht surjektiv ist.

    Für jedes fest gewählte $w \in W$ ist auch die Abbildung $h\colon U \to U \otimes W$, $h(u) = u \otimes w$ eine lineare Abbildung.

15. Wenn $V = U \oplus W$ gilt, so ist $U \cong V/W$.

16. $\mathbf{F}_\mathbb{K}(\{1, \dots, n\}) \cong \mathbb{K}^n$.

17. Ist V irgendein Vektorraum und B eine Basis von V, so gilt $\mathbf{F}_\mathbb{K}(B) \cong V$.

18. Für zwei $\mathbb{K}$-Vektorräume U, W gilt $(U \times W) \cong (W \times U)$ und $(U \otimes W) \cong (W \otimes U)$. Echte Gleichheit gilt in beiden Fällen nur, wenn $U = W$ ist.

19. Die Abbildung $\cdot^\top\colon \mathbb{K}^{n\times m} \to \mathbb{K}^{m\times n}$, die jeder Matrix $A \in \mathbb{K}^{n\times m}$ ihre Transponierte zuordnet, ist linear.

20. Es gilt $\mathbb{K}^n \otimes \mathbb{K}^m \cong \mathbb{K}^{n\times m}$.

21. Sei $V = C^\infty([0, 1], \mathbb{R})$. Dann ist die Abbildung $\frac{d}{dx}$, die jedem Element des Vektorraums dessen Ableitung zuordnet, linear. Ihr Kern ist die Menge der konstanten Funktionen.

    Für $V = \mathbb{K}[[X]]$ definiert man

    $$\frac{d}{dX}\colon V \to V, \quad \frac{d}{dX} \sum_{n=0}^{\infty} a_n X^n := \sum_{n=0}^{\infty} (n+1) a_{n+1} X^n.$$

    Auch diese (formale) „Ableitung“ ist eine lineare Abbildung.

22. Die Abbildung

    $$I\colon C([0, 1], \mathbb{R}) \to \mathbb{R}, \quad I(f) := \int_0^1 f(t)\, dt$$

    ist linear.

23. Sei $V = \mathbb{R}^{n+1}$ und $W = \mathbb{R}[X]$. Ferner seien $x_0, \dots, x_n \in \mathbb{R}$ fest gewählte paarweise verschiedene reelle Zahlen. Man kann zeigen, dass es dann für jede Wahl von $y_0, \dots, y_n \in \mathbb{R}$ genau ein Polynom $p \in \mathbb{R}[X]$ mit Grad höchstens n gibt, so dass $p(x_i) = y_i$ für alle i gilt. Mit $p(x_i)$ ist dabei die Auswertung der zu p gehörigen Polynomfunktion an der Stelle x_i gemeint. Man nennt p das *Interpolationspolynom* für $(x_0, y_0), \dots, (x_n, y_n)$.

    Die Abbildung, die jedem $(y_0, \dots, y_n) \in V$ dieses Polynom p zuordnet, ist linear.

Satz 48.

1. Die Verkettung linearer Abbildungen ist linear. Die Umkehrfunktion einer bijektiven linearen Abbildung ist linear.
2. Für je drei Vektorräume U, V, W gilt $U \cong U$, $U \cong V \Rightarrow V \cong U$, $U \cong V \wedge V \cong W \Rightarrow U \cong W$.
3. Sind V, W zwei $\mathbb{K}$-Vektorräume und ist $h\colon V \to W$ ein Homomorphismus, so ist $\ker h$ ein Unterraum von V und $\operatorname{im} h$ ein Unterraum von W.

Beweis.

1. Verkettung: Seien U, V, W drei $\mathbb{K}$-Vektorräume, $f\colon U \to V$ und $g\colon V \to W$ Homomorphismen und $h\colon U \to W$, $h(x) := g(f(x))$ deren Verkettung. Für $\alpha, \beta \in \mathbb{K}$ und $x, y \in V$ gilt dann:

   $$h(\alpha x+\beta y) = g(f(\alpha x+\beta y)) = g(\alpha f(x)+\beta f(y)) = \alpha g(f(x))+\beta g(f(y)) = \alpha h(x)+\beta h(y).$$

   Umkehrfunktion: Seien V, W zwei $\mathbb{K}$-Vektorräume, $f\colon V \to W$ ein Isomorphismus und $f^{-1}\colon W \to V$ seine Umkehrfunktion. Für $\alpha, \beta \in \mathbb{K}$ und $x, y \in W$ gilt dann

   $$f(\alpha f^{-1}(x) + \beta f^{-1}(y)) = \alpha f(f^{-1}(x)) + \beta f(f^{-1}(y)) = \alpha x + \beta y.$$

   Anwendung von f^{-1} auf beiden Seiten und Rückwärtslesen der Gleichung liefert

   $$f^{-1}(\alpha x + \beta y) = \alpha f^{-1}(x) + \beta f^{-1}(y).$$

2. Reflexivität: $U \cong U$ gilt, da die Identitätsfunktion id_U linear ist.

   Symmetrie: Folgt auch aus Teil 1 (Linearität der Umkehrfunktion).

   Transitivität: Folgt aus Teil 1 (Linearität der Verkettung).

3. Für beliebige $x, y \in \ker h$ und $\alpha, \beta \in \mathbb{K}$ gilt:

   $$\begin{aligned} h(x) = h(y) = 0 \quad &\Rightarrow \quad h(\alpha x + \beta y) = \alpha h(x) + \beta h(y) = \alpha 0 + \beta 0 = 0 \\ &\Rightarrow \quad \alpha x + \beta y \in \ker h. \end{aligned}$$

   Damit ist $\ker h$ ein Unterraum von V.

   Sind $x, y \in \operatorname{im} h$ beliebig, etwa $x = h(u)$, $y = h(v)$ für bestimmte $u, v \in V$, und sind $\alpha, \beta \in \mathbb{K}$ beliebig, so gilt

   $$\alpha x + \beta y = \alpha h(u) + \beta h(v) = h(\alpha u + \beta v),$$

   also $\alpha x + \beta y \in \operatorname{im} h$. Damit ist $\operatorname{im} h$ ein Unterraum von W. ∎

Beispiel. $V = \mathbb{K}^m$, $W = \mathbb{K}^n$, $h\colon V \to W$, $h(x) = Ax$ für eine bestimmte Matrix $A \in \mathbb{K}^{n\times m}$. Eine Basis für $\ker h = \ker A$ kann man berechnen wie in Abschnitt 9 erklärt.

Ein Erzeugendensystem für $\operatorname{im} h$ ist $\{\, h(b) : b \in B \,\}$, wenn B eine Basis von V ist. Wie man daraus eine Basis gewinnt, haben wir im Beispiel 5 nach Definition 32 gesehen.

Satz 49. Seien V, W zwei $\mathbb{K}$-Vektorräume, $h\colon V \to W$ ein Homomorphismus. Dann gilt: h ist genau dann injektiv, wenn $\ker h = \{0\}$ ist.

Beweis. „⇒“ Sei $x \in \ker h$. Dann gilt $h(x) = 0$. Es gilt aber auch $h(0) = 0$. Da h injektiv ist, folgt $x = 0$.
„⇐“ Seien $x, y \in V$ mit $h(x) = h(y)$. Dann gilt $h(x) - h(y) = 0$, also $h(x - y) = 0$, also $x - y \in \ker h$, also $x - y = 0$, also $x = y$. ∎

Satz 50. Seien V, W zwei $\mathbb{K}$-Vektorräume, B eine Basis von V, und $f\colon B \to W$ eine beliebige Funktion. Dann existiert genau eine lineare Abbildung $h\colon V \to W$ mit $h(b) = f(b)$ für alle $b \in B$.

Beweis. (a) Existenz: Jedes $x \in V$ lässt sich schreiben als $x = \alpha_1 b_1 + \cdots + \alpha_k b_k$ für gewisse $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ und $b_1, \ldots, b_k \in B$. Definiere $h(x) := \alpha_1 f(b_1) + \cdots + \alpha_k f(b_k)$. Da die Darstellung von x als Linearkombination von Basiselementen eindeutig ist, ist h wohldefiniert. Man überzeugt sich leicht, dass h linear ist und dass $h(b) = f(b)$ für alle $b \in B$ gilt.
(b) Eindeutigkeit: Für jedes $x \in V$ mit $x = \alpha_1 b_1 + \cdots + \alpha_k b_k$ muss gelten

$$\begin{aligned} h(x) &= h(\alpha_1 b_1 + \cdots + \alpha_k b_k) \\ &= \alpha_1 h(b_1) + \cdots + \alpha_k h(b_k) \\ &= \alpha_1 f(b_1) + \cdots + \alpha_k f(b_k). \end{aligned}$$

Eine andere Wahl von h ist also nicht möglich. ∎

Beispiel.

1. $V = \mathbb{Q}^4$, $B = \{e_1, \ldots, e_4\}$, $W = \mathbb{Q}^3$. Betrachte die Funktion $f\colon B \to W$ definiert durch

   $$\begin{aligned} e_1 &\mapsto (1, 2, 3), \quad & e_2 &\mapsto (4, 5, 6) \\ e_3 &\mapsto (7, 8, 9), \quad & e_4 &\mapsto (10, 11, 12). \end{aligned}$$

   Dann ist $h\colon V \to W$ die Abbildung, die $(x_1, x_2, x_3, x_4) \in V$ auf

   $$x_1 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + x_3 \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} + x_4 \begin{pmatrix} 10 \\ 11 \\ 12 \end{pmatrix}$$

   abbildet. Mit anderen Worten:

   $$h(x) = \begin{pmatrix} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{pmatrix} x.$$

2. Seien U, W zwei $\mathbb{K}$-Vektorräume, $B_U, \bar{B}_U$ und $B_W, \bar{B}_W$ je zwei Basen von U bzw. W.

   Die lineare Abbildung $h\colon \mathbf{F}_{\mathbb{K}}(B_U \times B_W) \to \mathbf{F}_{\mathbb{K}}(\bar{B}_U \times \bar{B}_W)$ sei dadurch definiert, dass jeder Basistensor $b_U \otimes b_W$ mit $b_U \in B_U, b_W \in B_W$ auf den entsprechenden Tensor $b_U \otimes b_W \in \mathbf{F}_{\mathbb{K}}(\bar{B}_U \times \bar{B}_W)$ abgebildet wird.

   Ferner sei die lineare Abbildung $\bar{h}\colon \mathbf{F}_{\mathbb{K}}(\bar{B}_U \times \bar{B}_W) \to \mathbf{F}_{\mathbb{K}}(B_U \times B_W)$ dadurch definiert, dass jeder Basistensor $\bar{b}_U \otimes \bar{b}_W$ mit $\bar{b}_U \in \bar{B}_U, \bar{b}_W \in \bar{B}_W$ auf den entsprechenden Tensor $\bar{b}_U \otimes \bar{b}_W \in \mathbf{F}_{\mathbb{K}}(B_U \times B_W)$ abgebildet wird.

Man kann nachrechnen, dass $h \circ \bar{h}$ die Identität ist, d.h. h und $\bar{h}$ sind bijektiv, d.h. $\mathbf{F}_{\mathbb{K}}(B_U \times B_W) \cong \mathbf{F}_{\mathbb{K}}(\bar{B}_U \times \bar{B}_W)$. Für alle $u \in U$ und $w \in W$ gilt $h(u \otimes w) = u \otimes w$. In diesem Sinn ist die Definition von $U \otimes W$ unabhängig von der Wahl der Basen von U und W.

Satz 51. Seien V, W zwei endlich-dimensionale $\mathbb{K}$-Vektorräume, $h\colon V \to W$ ein Homomorphismus. Dann gilt: $\dim V = \dim \ker h + \dim \operatorname{im} h$.

Beweis. Sei $\{b_1, \ldots, b_n\}$ eine Basis von $\ker h$. Nach Satz 43 gibt es $b_{n+1}, \ldots, b_m \in V$, so dass $\{b_1, \ldots, b_m\}$ eine Basis von V ist. Wir zeigen, dass $B = \{h(b_{n+1}), \ldots, h(b_m)\}$ eine Basis von $\operatorname{im} h$ ist. Daraus folgt die Behauptung.

(a) B ist linear unabhängig: Seien $\alpha_{n+1}, \ldots, \alpha_m \in \mathbb{K}$ so, dass $\alpha_{n+1}h(b_{n+1}) + \cdots + \alpha_m h(b_m) = 0$. Dann ist $h(\alpha_{n+1}b_{n+1} + \cdots + \alpha_m b_m) = 0$, also $\alpha_{n+1}b_{n+1} + \cdots + \alpha_m b_m \in \ker h$. Dann aber gibt es $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ so dass

$$\alpha_1 b_1 + \cdots + \alpha_n b_n = \alpha_{n+1}b_{n+1} + \cdots + \alpha_m b_m.$$

Da $\{b_1, \ldots, b_m\}$ als Basis von V linear unabhängig ist, folgt, dass alle α_i Null sind, insbesondere $\alpha_{n+1}, \ldots, \alpha_m$.

(b) B ist Erzeugendensystem: Sei $y \in \operatorname{im} h$. Dann gibt es ein $x \in V$ mit $y = h(x)$. Dann gibt es $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ mit $x = \alpha_1 b_1 + \cdots + \alpha_m b_m$. Und dann gilt

$$h(x) = \underbrace{\alpha_1 h(b_1) + \cdots + \alpha_n h(b_n)}_{= 0, \text{ da } \{b_1, \ldots, b_n\} \text{ Basis von } \ker h} + \alpha_{n+1}h(b_{n+1}) + \cdots + \alpha_m h(b_m).$$

∎

Mit Satz 51 und einigen früheren Resultaten können wir einiges über die vier Räume aussagen, die einer Matrix zugeordnet sind.

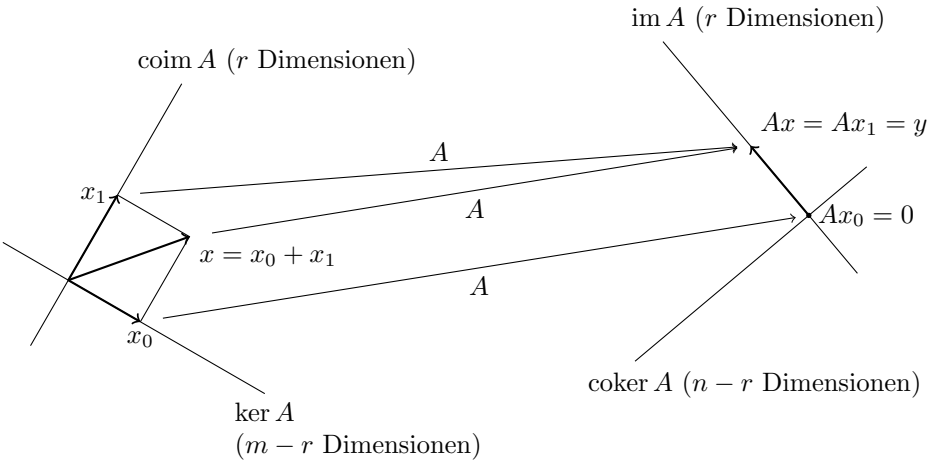
Sei $A \in \mathbb{K}^{n \times m}$ eine Matrix, $r = \operatorname{Rang} A$. Zunächst gilt für die Dimensionen:

- Zeilenraum: $\dim \operatorname{coim} A = r$ (wegen Satz 23)
- Spaltenraum: $\dim \operatorname{im} A = r$ (wegen $\dim \operatorname{coim} A = r$ und Satz 24)
- Kern: $\dim \ker A = m - r$ (wegen $\dim \operatorname{im} A = r$ und Satz 51)
- Ko-Kern: $\dim \operatorname{coker} A = n - r$ (wegen $\dim \operatorname{coim} A = r$ und den Sätzen 51 und 24).

Als nächstes gilt $\mathbb{K}^m = \ker A + \operatorname{coim} A$. Das sieht man ein, wenn man sich erinnert, wie man eine Basis von $\ker A$ aus der Treppennormalform bestimmt (vgl. Seite 56). Anhand der Lage der Treppenstufen erkennt man, dass die Zeilen von A die Basis des Lösungsraums zu einer Basis des gesamten Raums $\mathbb{K}^m$ ergänzen. Die Summe ist sogar direkt, wie mit Satz 44 leicht überprüft: $\ker A$ und $\operatorname{coim} A$ sind Unterräume von $\mathbb{K}^m$ und es gilt $m = \dim \mathbb{K}^m = \dim \ker A + \dim \operatorname{coim} A - \dim(\ker A \cap \operatorname{coim} A) = m - r + r - \dim(\ker A \cap \operatorname{coim} A)$, also $\dim(\ker A \cap \operatorname{coim} A) = 0$, also $\ker A \cap \operatorname{coim} A = \{0\}$.

Analog gilt $\mathbb{K}^n = \operatorname{im} A \oplus \operatorname{coker} A$, indem man dasselbe Argument auf $A^\top$ anwendet.

Die Zusammenhänge sind in der folgenden Zeichnung schematisch dargestellt. Man sieht links, dass jedes $x \in \mathbb{K}^m$ sich in einen $\operatorname{coim} A$-Anteil und einen $\ker A$-Anteil zerlegen lässt. Die Matrix A bildet x nach $\operatorname{im} A$ ab. Der $\ker A$-Anteil von x landet bei 0, und das Bild $y = Ax$ von x ist identisch mit dem Bild des $\operatorname{coim} A$-Anteils von x. Es gibt keine Vektoren, die auf $\operatorname{coker} A \setminus \{0\}$ abgebildet werden.



Als Anwendung können wir nun den noch ausstehenden Korrektheitsbeweis von Algorithmus 5 auf Seite 60 zu Ende bringen. Wir hatten dort eine Matrix $B \in \mathbb{K}^{k\times m}$ und eine Basis $\{a_1, \dots, a_n\} \subseteq \mathbb{K}^m$ von $\ker B$, d. h.

$$\{\, x \in \mathbb{K}^m : Bx = 0 \,\} = \{\, \alpha_1 a_1 + \cdots + \alpha_n a_n : \alpha_1, \dots, \alpha_n \in \mathbb{K} \,\}.$$

Die Behauptung war, dass dann auch gilt

$$\{\, x \in \mathbb{K}^m : Ax = 0 \,\} = \{\, \alpha_1 b_1 + \cdots + \alpha_k b_k : \alpha_1, \dots, \alpha_k \in \mathbb{K} \,\},$$

wobei $b_1, \dots, b_k$ die Zeilenvektoren von B sind und $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$. Die Richtung „$\supseteq$“ ist bereits gezeigt worden. Die rechte Seite ist also ein Unterraum des Vektorraums auf der linken Seite.

Die linke Seite ist $\ker A$, die rechte ist $\operatorname{coim} B$. Da $\{a_1, \dots, a_n\}$ eine Basis und damit linear unabhängig ist, gilt $\operatorname{Rang} A = n$ und daher $\dim \ker A = m - n$. Außerdem folgt $\dim \ker B = n$ und daher $\dim \operatorname{coim} B = m - n$. Damit sind $\ker A$ und $\operatorname{coim} B$ zwei Unterräume von $\mathbb{K}^m$ derselben Dimension. Wegen $\ker A \supseteq \operatorname{coim} B$ folgt aus Satz 42, dass $\ker A = \operatorname{coim} B$ gilt, was zu zeigen war.

Satz 52. Seien V, W zwei endlich-dimensionale $\mathbb{K}$-Vektorräume mit $\dim V = \dim W$, und sei $h\colon V \to W$ ein Homomorphismus. Dann gilt:

$$h \text{ ist injektiv} \iff h \text{ ist surjektiv} \iff h \text{ ist bijektiv.}$$

Beweis. Die zweite Äquivalenz folgt unmittelbar aus der ersten. Wir zeigen die erste.

„$\Rightarrow$“ h ist injektiv. Dann ist $\ker h = \{0\}$ nach Satz 49. Mit Satz 51 folgt dann $\dim V = 0 + \dim \operatorname{im} h$, und wegen $\dim V = \dim W$ also $\dim W = \dim \operatorname{im} h$. Nach Satz 42 folgt $\operatorname{im} h = W$, d.h. h ist surjektiv.

„$\Leftarrow$“ h ist surjektiv. Dann ist $\operatorname{im} h = W$, also $\dim \operatorname{im} h = \dim W$, also $\dim \ker h = \dim V - \dim W = 0$ nach Satz 51 und Voraussetzung $\dim V = \dim W$. Also gilt $\ker h = \{0\}$. Wegen Satz 49 folgt, dass h injektiv ist. ∎

Satz 53. Seien V, W zwei endlich-dimensionale $\mathbb{K}$-Vektorräume. Dann gilt:

1. $\dim V \leq \dim W \iff$ es gibt einen injektiven Homomorphismus $h\colon V \to W$.
2. $\dim V \geq \dim W \iff$ es gibt einen surjektiven Homomorphismus $h\colon V \to W$.
3. $\dim V = \dim W \iff$ es gibt einen bijektiven Homomorphismus $h\colon V \to W$.

Beweis. Wir zeigen den dritten Teil. Der Beweis für die ersten beiden Teile geht ähnlich.

„⇒“ Sei $A = \{a_1, \dots, a_n\}$ eine Basis von V und $B = \{b_1, \dots, b_n\}$ eine Basis von W. Nach Satz 50 existiert eine lineare Abbildung $h\colon V \to W$ mit $h(a_i) = b_i$ für $i = 1, \dots, n$. Diese lineare Abbildung ist injektiv, denn

$$\begin{aligned} 0 &= h(\alpha_1 a_1 + \cdots + \alpha_n a_n) \\ &= \alpha_1 h(a_1) + \cdots + \alpha_n h(a_n) \\ &= \alpha_1 b_1 + \cdots + \alpha_n b_n \\ \Rightarrow \qquad \alpha_1 &= \cdots = \alpha_n = 0, \end{aligned}$$

also $\ker h = \{0\}$.

h ist auch surjektiv, denn ist $y \in W$ beliebig, so ist $y = \beta_1 b_1 + \cdots + \beta_n b_n$ für gewisse $\beta_1, \dots, \beta_n \in W$, also

$$y = \beta_1 h(a_1) + \cdots + \beta_n h(a_n) = h(\beta_1 a_1 + \cdots + \beta_n a_n) \in \operatorname{im} h.$$

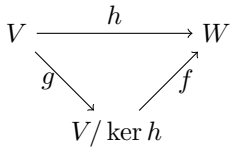
„⇐“ Ist $h\colon V \to W$ bijektiv und $A = \{a_1, \dots, a_n\}$ eine Basis von V, so ist

$$B = \{h(a_1), \dots, h(a_n)\}$$

eine Basis von W: die lineare Unabhängigkeit folgt aus der Injektivität von h, und dass B ein Erzeugendensystem ist, folgt aus der Surjektivität. ∎

Satz 54. (Homomorphiesatz für Vektorräume) Seien V, W zwei $\mathbb{K}$-Vektorräume, $h\colon V \to W$ ein Homomorphismus. Dann gibt es eine surjektive lineare Abbildung $g\colon V \to V/\ker h$ und eine injektive lineare Abbildung $f\colon V/\ker h \to W$ mit $h = f \circ g$.

Wenn h surjektiv ist, dann auch f. Insbesondere gilt $V/\ker h \cong \operatorname{im} h$.



Beweis. Da Vektorräume insbesondere auch abelsche Gruppen sind, und lineare Abbildungen insbesondere auch Gruppenhomomorphismen, haben wir die meisten Aussagen des Satzes bereits in Satz 10 bewiesen. Es bleibt hier nur noch zu zeigen, dass die dort angegebenen Funktionen

$$\begin{aligned} g\colon V \to V/\ker h, &\qquad g(x) = [x]_\sim \\ f\colon V/\ker h \to W, &\quad f([x]_\sim) = h(x) \end{aligned}$$

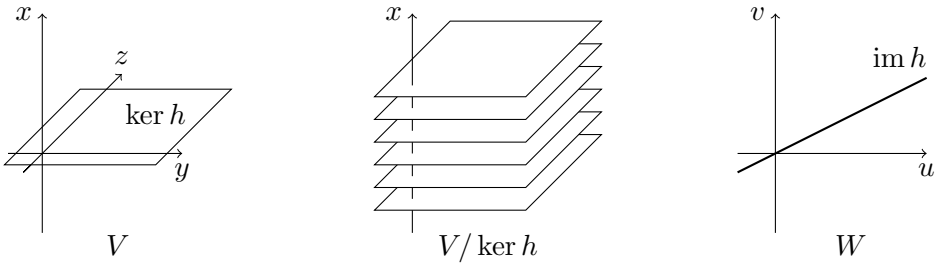
auch mit der Skalarmultiplikation verträglich sind. Und in der Tat gilt

$$\begin{aligned} g(\alpha x) &= [\alpha x]_\sim = \alpha [x]_\sim = \alpha\, g(x) \\ f(\alpha [x]_\sim) &= f([\alpha x]_\sim) = h(\alpha x) = \alpha\, h(x). \end{aligned}$$

∎

Beispiel. $V = \mathbb{R}^3$, $W = \mathbb{R}^2$, $h\colon V \to W$, $h(x,y,z) = (2x, x)$. Dann gilt $\ker h = \langle \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \rangle$ und $\operatorname{im} h = \langle \begin{pmatrix}2\\1\end{pmatrix} \rangle$. Geometrisch ist $\ker h$ die (y,z)-Ebene in $\mathbb{R}^3$ und $\operatorname{im} h$ eine Gerade in $\mathbb{R}^2$.

Die Elemente von $V/\ker h$ kann man sich vorstellen als die Ebenen, die parallel zu $\ker h$ liegen. Diese bilden einen Vektorraum, der genau wie $\operatorname{im} h$ eindimensional ist. Jede Ebene in $V/\ker h$ ist eindeutig charakterisiert durch ihren Schnittpunkt mit der x-Achse. Einen Isomorphismus zwischen $V/\ker h$ und $\operatorname{im} h$ erhält man, indem man die Ebene durch $(x,0,0)$ auf $(2x,x)$ abbildet. Beachte, dass zwei Punkte $(x,y,z) \in \mathbb{R}^3$ genau dann dasselbe Bild $h(x,y,z) \in \mathbb{R}^2$ haben, wenn sie zur selben Ebene gehören (vgl. die Beispiele zu Satz 6).

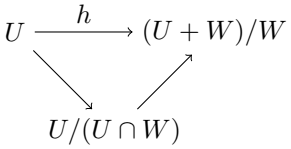


Satz 55. (Isomorphiesätze)

1. Seien U, W zwei Unterräume eines $\mathbb{K}$-Vektorraums V. Dann gilt
$$U/(U \cap W) \cong (U + W)/W.$$
2. Im Fall $U \subseteq W$ gilt außerdem $(V/U)/(W/U) \cong V/W$.

Beweis.

1. Nach Satz 43 gibt es einen Raum $\tilde{W} \subseteq U$ mit $(U \cap W) \oplus \tilde{W} = U$. Wähle so einen Raum $\tilde{W}$. Dann lässt sich jedes $u \in U$ schreiben als $u = u_1 + u_2$ für gewisse eindeutig bestimmte $u_1 \in U \cap W$ und $u_2 \in \tilde{W}$. Betrachte die Abbildung
$$h\colon U \to (U + W)/W, \quad h(u_1 + u_2) = [u_2]_\sim.$$
Diese Funktion ist linear und surjektiv und es gilt $\ker h = U \cap W$. Aus Satz 54 folgt deshalb die Behauptung.



2. Sei $\tilde{U} \subseteq W$ so, dass $U \oplus \tilde{U} = W$, und sei $\tilde{W} \subseteq V$ so, dass $W \oplus \tilde{W} = V$. Dann ist $V/U \cong \tilde{U} + \tilde{W}$ und $W/U \cong \tilde{U}$, und unter Verwendung von Teil 1 ist
$$(V/U)/(W/U) \cong (\tilde{U} + \tilde{W})/\tilde{U} \cong \tilde{W}/\underbrace{(\tilde{U} \cap \tilde{W})}_{=\{0\}} \cong \tilde{W} \cong V/W.$$

∎

17 Koordinatendarstellungen

Definition 35. Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum.

1. Ein Vektor $B = (b_1, \ldots, b_n) \in V^n$ heißt *geordnete Basis* von V, falls $\{b_1, \ldots, b_n\}$ eine Basis von V ist.

2. Sei $B = (b_1, \ldots, b_n) \in V^n$ eine geordnete Basis von V und $x \in V$. Sind $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ so, dass $x = \alpha_1 b_1 + \cdots + \alpha_n b_k$, dann heißt der Vektor $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ die *Koordinatendarstellung* von x bezüglich B. Für $i = 1, \ldots, n$ heißt dann α_i die i-te *Koordinate* von x bezüglich B.

Eine geordnete Basis ist nichts anderes als eine Basis, bei der eine bestimmte Reihenfolge für die Basiselemente festgelegt ist. Man beachte, dass $\{a, b, c\} = \{c, a, b\}$ aber $(a, b, c) \neq (c, a, b)$ ist.

Beim Teil 2 der Definition ist zu beachten, dass es zu jedem $x \in V$ genau eine passende Koordinatendarstellung $(\alpha_1, \ldots, \alpha_n)$ gibt, weil $\{b_1, \ldots, b_n\}$ nach Voraussetzung eine Basis ist.

Beispiel.

1. Vektoren $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ sind Koordinatendarstellungen von sich selbst bezüglich der Standardbasis $E = (e_1, \ldots, e_n)$.

2. Ist $B = (b_1, b_2, b_3, b_4)$ eine geordnete Basis eines Vektorraums V, so ist $\tilde{B} = (b_3, b_1, b_4, b_2)$ auch eine geordnete Basis von V, und zwar eine von B verschiedene. Ist (x_1, x_2, x_3, x_4) die Koordinatendarstellung eines bestimmten Vektors bezüglich B, so ist (x_3, x_1, x_4, x_2) die Koordinatendarstellung desselben Vektors bezüglich $\tilde{B}$.

3. $B = (1, X, X^2, X^3) \in K[X]^4$ ist eine geordnete Basis für den Unterraum von $K[X]$ bestehend aus allen Polynomen vom Grad höchstes drei. Eine andere geordnete Basis ist $\tilde{B} = (1, X, X(X-1), X(X-1)(X-2))$. Die Koordinatendarstellung von $2 + 3X - 7X^2 + 5X^3$ bezüglich B ist $(2, 3, -7, 5)$. Die Koordinatendarstellung desselben Polynoms bezüglich $\tilde{B}$ ist $(2, 1, 8, 5)$.

4. Im Fall $V = \mathbb{K}^n$ lässt sich eine geordnete Basis als Matrix auffassen. Per Konvention macht man die Basisvektoren zu den Spalten der Matrix. Dann ist zum Beispiel

$$B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$$

eine geordnete Basis von $\mathbb{Q}^3$. Ist $\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}$ die Koordinatendarstellung eines Vektors bezüglich dieser Basis B, so erhält man die Koordinatendarstellung desselben Vektors bezüglich der Standardbasis durch die Rechnung

$$3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + 4 \begin{pmatrix} 1 \\ 4 \\ 9 \end{pmatrix} = \begin{pmatrix} 9 \\ 23 \\ 45 \end{pmatrix}.$$

Ist umgekehrt $\begin{pmatrix}7\\4\\3\end{pmatrix}$ die Koordiantendarstellung eines Vektors bezüglich der Standardbasis, so bekommt man dessen Darstellung bezüglich der Basis B, indem man das lineare Gleichungssystem

$$\alpha_1 \begin{pmatrix}1\\1\\1\end{pmatrix} + \alpha_2 \begin{pmatrix}1\\2\\3\end{pmatrix} + \alpha_3 \begin{pmatrix}1\\4\\9\end{pmatrix} = \begin{pmatrix}7\\4\\3\end{pmatrix}$$

löst. Das Ergebnis lautet $\begin{pmatrix}12\\-6\\1\end{pmatrix}$.

Bei einem gegebenen Tupel von Körperelementen ist also nicht ohne weiteres klar, welcher Vektor damit beschrieben wird. Man muss immer auch die Basis wissen, bezüglich der die Koordinaten zu verstehen sind. Ist keine Basis angegeben, wird meistens die Standardbasis gemeint sein.

Satz 56. Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum, $B = (b_1, \dots, b_n) \in V^n$ eine geordnete Basis von V. Dann ist die Abbildung $h\colon V \to \mathbb{K}^n$, die jedem $x \in V$ dessen Koordinatendarstellung bezüglich B zuordnet, ein Isomorphismus.

Insbesondere sind alle endlich-dimensionalen $\mathbb{K}$-Vektorräume der gleichen Dimension zueinander isomorph.

Beweis. Übung. ∎

Im vorangegangenen Beispiel haben wir gesehen, wie man zwischen einer Koordinatendarstellung eines Vektors $v \in \mathbb{K}^n$ bezüglich der Standardbasis und der Koordinatendarstellung bezüglich einer anderen geordneten Basis hin und her wechseln kann: ist $v \in \mathbb{K}^n$ ein Vektor (und damit die Koordinatendarstellung von sich selbst bezüglich der Standardbasis), so ist Bv die Koordinatendarstellung desselben Vektors bezüglich B. Und ist $w \in \mathbb{K}^n$ die Koordinatendarstellung eines Vektors bezüglich B, so ist $B^{-1}w$ die Koordinatendarstellung desselben Vektors bezüglich der Standardbasis. (Matrizen, die geordnete Basen enthalten, sind immer invertierbar, weil ihre Spalten linear unabhängig sind.) Die folgende Definition verallgemeinert diesen Sachverhalt.

Definition 36. Seien V, W zwei endlich-dimensionale $\mathbb{K}$-Vektorräume, $A = (a_1, \dots, a_m) \in V^m$ eine geordnete Basis von V und $B = (b_1, \dots, b_n) \in W^n$ eine geordnete Basis von W. Weiter sei $h\colon V \to W$ eine lineare Abbildung. Die Matrix $M \in \mathbb{K}^{n\times m}$, deren (i,j)-ter Eintrag die i-te Koordinate von $h(a_j)$ bezüglich B ist, heißt die *Abbildungsmatrix* oder die *Koordinatendarstellung* von h bezüglich A und B.

Beispiel.

1. Sei $M \in \mathbb{K}^{n\times m}$, $V = \mathbb{K}^m$, $W = \mathbb{K}^n$ und sei $h\colon V \to W$, $h(x) = Mx$. Dann ist M die Abbildungsmatrix von h bezüglich der Standardbasis $\{e_1, \dots, e_m\}$ von V und der Standardbasis $\{e_1, \dots, e_n\}$ von W.

2. Sei $V = \mathbb{Q}[X]_{\le 3} = \{\, a_0 + a_1X + a_2X^2 + a_3X^3 : a_0, a_1, a_2, a_3 \in \mathbb{Q} \,\}$, und sei $W = \mathbb{Q}(\sqrt{2})$. Weiter sei

   $$h\colon V \to W, \quad h(a_0 + a_1X + a_2X^2 + a_3X^3) := a_0 + a_1(1+\sqrt{2}) + a_2(1+\sqrt{2})^2 + a_3(1+\sqrt{2})^3.$$

Dann ist h eine lineare Abbildung. $A = \{1, X, X^2, X^3\}$ ist eine Basis von V und $B = \{1, \sqrt{2}\}$ ist eine Basis von W. Die Abbildungsmatrix von h bezüglich A und B lautet

$$M = \begin{pmatrix} 1 & 1 & 3 & 7 \\ 0 & 1 & 2 & 5 \end{pmatrix}.$$

Zum Beispiel erklärt sich die letzte Spalte aus $(1+\sqrt{2})^3 = 7 + 5\sqrt{2}$.

3. Betrachte den Fall $V = W$, $h = \mathrm{id}$ mit zwei nicht notwendigerweise identische Basen $B_1, B_2 \in V^n$. Dann hat die Abbildungsmatrix M von h bezüglich B_1 und B_2 die Eigenschaft, dass sie Koordinatendarstellungen von $v \in V$ bezüglich B_1 in Koordinatendarstellungen bezüglich B_2 umwandelt. Man nennt M deshalb auch eine *Basiswechselmatrix* und schreibt $T_{B_1 \to B_2} := M$.

   Im Fall $V = \mathbb{K}^n$, wo sich B_1 und B_2 als invertierbare Matrizen auffassen lassen, und wo die Standardbasis $(e_1, \ldots, e_n)$ der Einheitsmatrix I_n entspricht, gilt $T_{I_n \to B_2} = B_2$, $T_{B_1 \to I_n} = B_1^{-1}$, $T_{B_1 \to B_2} = B_2 B_1^{-1}$, $T_{B_2 \to B_1} = B_1 B_2^{-1} = T_{B_1 \to B_2}^{-1}$. Außerdem gilt: ist B_3 eine dritte geordnete Basis von V, so ist $T_{B_1 \to B_3} = T_{B_2 \to B_3} T_{B_1 \to B_2}$.

Satz 57. Seien V, W zwei endlich-dimensionale $\mathbb{K}$-Vektorräume, $h\colon V \to W$ linear und $M \in \mathbb{K}^{n \times m}$ die Abbildungsmatrix von h bezüglich geordneter Basen A, B von V, W. Weiter seien $\pi_A\colon V \to \mathbb{K}^m$ und $\pi_B\colon W \to \mathbb{K}^n$ die Abbildungen, die jedem $x \in V$ bzw. jedem $y \in W$ die Koordinatendarstellungen dieser Vektoren bezüglich A bzw. B zuordnen. Dann gilt

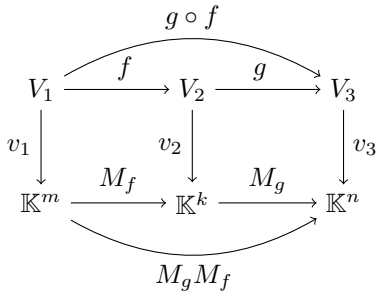
$$\pi_B(h(x)) = M \pi_A(x)$$

für alle $x \in V$.

$$\begin{array}{ccc} V & \xrightarrow{h} & W \\ {\scriptstyle \pi_A}\downarrow & & \downarrow{\scriptstyle \pi_B} \\ \mathbb{K}^m & \xrightarrow{M} & \mathbb{K}^n \end{array}$$

Beweis. Da alle beteiligten Abbildungen linear sind, genügt es, die Aussage für die Elemente der Basis von V zu zeigen. Ist a_j das j-te Element von A, so ist $\pi_A(a_j) = e_j$ und also $M\pi_A(a_j)$ die j-te Spalte von M. Nach Definition 36 ist das genau die Koordinatendarstellung von $h(a_j)$ bezüglich der gewählten Basis von W, also $\pi_B(h(a_j))$. ∎

Satz 58. Seien V_1, V_2, V_3 drei endlich-dimensionale $\mathbb{K}$-Vektorräume mit geordneten Basen B_1, B_2, B_3, seien $f\colon V_1 \to V_2$, $g\colon V_2 \to V_3$ lineare Abbildungen, und seien M_f, M_g die Abbildungsmatrix von f bezüglich B_1 und B_2 bzw. von g bezüglich B_2 und B_3. Dann ist $M_g M_f$ die Abbildungsmatrix von $g \circ f\colon V_1 \to V_3$ bezüglich B_1 und B_3.



Beweis. Nach Satz 57 gilt $v_2(f(x)) = M_f v_1(x)$ für alle $x \in V_1$, d. h. $f(x) = v_2^{-1}(M_f v_1(x))$. Außerdem gilt $v_3(g(y)) = M_g v_2(y)$ für alle $y \in V_2$, und damit auch

$$v_3(g(f(x))) = M_g v_2(f(x)) = M_g M_f v_1(x)$$

für alle $x \in V_1$. Da diese Gleichung dann insbesondere für die Basisvektoren von V_1 gilt, folgt die Behauptung. ∎

Aus dem Satz folgt, dass man eine Abbildungsmatrix M bezüglich zweier geordneter Basen B_1, B_2 in eine Abbildungsmatrix $\tilde{M}$ derselben Abbildung aber bezüglich zweier anderer geordneter Basen $\tilde{B}_1, \tilde{B}_2$ dadurch überführt, dass man sie mit den entsprechenden Basiswechselmatrizen multipliziert:

$$\tilde{M} = T_{B_2 \to \tilde{B}_2} M T_{\tilde{B}_1 \to B_1}.$$

Noch spezieller: Wenn $V_1 = V_2$, $B_1 = B_2$ und $\tilde{B}_1 = \tilde{B}_2$ ist, dann ist $T_{B_2 \to \tilde{B}_2} = T^{-1}_{\tilde{B}_1 \to B_1}$, d. h. die Basistransformation hat dann die einfache Form $\tilde{M} = T^{-1} M T$ für ein gewisses $T \in \mathrm{GL}(n, \mathbb{K})$. Es gilt dann

$$\begin{aligned} \det(\tilde{M}) &= \det(T^{-1} M T) \\ &= \det(T^{-1}) \det(M) \det(T) \\ &= \frac{1}{\det(T)} \det(M) \det(T) \\ &= \det(M). \end{aligned}$$

Das heißt, alle Abbildungsmatrizen einer linearen Abbildung $h \colon V \to V$ mit der gleichen Basis auf beiden Seiten haben dieselbe Determinante. Der Wert der Determinante hängt also nur von der linearen Abbildung h ab und nicht von der gewählten Basis für V. Ähnliches gilt für den Rang einer Matrix. Das gestattet die folgenden Definitionen:

Definition 37.

1. Sei $h \colon V \to V$ linear, B eine beliebige geordnete Basis von V und $M \in \mathbb{K}^{n \times n}$ die Abbildungsmatrix von h bezüglich B und B. Dann heißt $\det h := \det M$ die *Determinante* von h.

2. Sei $h \colon V \to W$ linear, A, B beliebige geordnete Basen von V, W, und sei $M \in \mathbb{K}^{n \times m}$ die Abbildungsmatrix von h bezüglich A und B. Dann heißt $\operatorname{Rang} h := \operatorname{Rang} M$ der *Rang* von h.

18 Der Dualraum

Definition 38.

1. Sind V und W zwei $\mathbb{K}$-Vektorräume, so wird die Menge aller linearen Abbildungen $h\colon V \to W$ mit $\operatorname{Hom}(V, W)$ bezeichnet.
2. Im Spezialfall $V = W$ schreibt man $\operatorname{End}(V) := \operatorname{Hom}(V, V)$. Für die Menge aller Automorphismen schreibt man $\operatorname{Aut}(V)$.
3. Im Spezialfall $W = \mathbb{K}$ schreibt man $V^* := \operatorname{Hom}(V, \mathbb{K})$ und nennt dies den *Dualraum* 04-18 von V. Die Elemente von V^* heißen *Funktionale.*

Beachte: $\operatorname{Hom}(V, W)$ bildet zusammen mit den Operationen

$$\underset{\text{in } \operatorname{Hom}(V,W)}{h_1} + \underset{\text{in } \operatorname{Hom}(V,W)}{h_2} := \big(\underset{\in V}{x} \mapsto \underset{\text{in } W}{h_1(x)} + \underset{\text{in } W}{h_2(x)} \big), \qquad \underset{\text{in } \operatorname{Hom}(V,W)}{\alpha} \cdot \underset{\text{in } \operatorname{Hom}(V,W)}{h} := \big(\underset{\in V}{x} \mapsto \underset{\text{in } W}{\alpha} \cdot \underset{\text{in } W}{h(x)} \big)$$

einen Vektorraum über $\mathbb{K}$.

Auch $\operatorname{End}(V)$ ist ein Vektorraum.

Die Menge $\operatorname{Aut}(V)$ ist **kein** Vektorraum, aber zusammen mit der Komposition eine Gruppe.

Beispiel.

1. $V = \mathbb{R}^3$, $h\colon \mathbb{R}^3 \to \mathbb{R}$, $h(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = x + y - z$. Dann ist $h \in V^*$.
2. $V = \mathbb{Q}[X]$, $e\colon \mathbb{Q}[X] \to \mathbb{Q}$ definiert durch

   $$e(a_0 + a_1 X + \cdots + a_n X^n) := a_0 + 3a_1 + 9a_2 + \cdots + 3^n a_n.$$

   Dann ist $e \in V^*$.
3. Einige weitere Beispiele für lineare Abbildungen $V \to \mathbb{K}$ befinden sich in der Liste nach Definition 34.

Satz 59. Sind V, W zwei endlich-dimensionale $\mathbb{K}$-Vektorräume, $\dim V = m$, $\dim W = n$, dann ist

$$\operatorname{Hom}(V, W) \cong \mathbb{K}^{n \times m}.$$

Insbesondere gilt $\dim \operatorname{Hom}(V, W) = nm$ und $V^* \cong \mathbb{K}^{1 \times m} \cong \mathbb{K}^m \cong V$ und $\dim V^* = \dim V$.

Beweis. Wir konstruieren einen Isomorphismus $h\colon \mathbb{K}^{n \times m} \to \operatorname{Hom}(V, W)$. Wähle dazu geordnete Basen $A = (a_1, \ldots, a_m)$ von V und $B = (b_1, \ldots, b_n)$ von W. Es sei $h_{i,j}\colon V \to W$ die nach Satz 50 eindeutig bestimmte lineare Abbildung mit $h_{i,j}(a_j) = b_i$ und $h_{i,j}(a_k) = 0$ für alle $k \neq j$.

Es bezeichne $e_{i,j} \in \mathbb{K}^{n \times m}$ die Matrix, die an Stelle (i, j) eine Eins und ansonsten nur Nullen enthält. Dann ist $\{ e_{i,j} : i = 1, \ldots, n, j = 1, \ldots, m \}$ eine Basis von $\mathbb{K}^{n \times m}$.

Es sei schließlich $h\colon \mathbb{K}^{n \times m} \to \operatorname{Hom}(V, W)$ die nach Satz 50 eindeutig bestimmte lineare Abbildung, die $e_{i,j}$ auf $h_{i,j}$ abbildet, für $i = 1, \ldots, m$ und $j = 1, \ldots, n$.

Diese Abbildung ist bijektiv: Die Umkehrabbildung $h^{-1}\colon \operatorname{Hom}(V,W) \to \mathbb{K}^{n\times m}$ ist die Funktion, die jedem $u \in \operatorname{Hom}(V,W)$ dessen Abbildungsmatrix bezüglich A und B zuordnet. ∎

Für $V = \mathbb{K}^m$ folgt aus dem Satz, dass jedes Element von V^* von der Form ist

$$x^*\colon V \to \mathbb{K}, \quad x^*\left(\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}\right) = \alpha_1 x_1 + \cdots + \alpha_m x_m = (\alpha_1, \ldots, \alpha_m) \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

für gewisse Konstanten $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$. Die („abstrakten“) Vektoren $x^* \in V^*$ lassen sich also auch durch („konkrete“) Vektoren $(\alpha_1, \ldots, \alpha_m) \in \mathbb{K}^m = V$ beschreiben. Die Abbildung

$$V \to V^*, \quad \underbrace{(\alpha_1, \ldots, \alpha_m)}_{\in V} \mapsto \underbrace{\big(\underset{\in V}{\underset{\uparrow}{x}} \mapsto (\alpha_1, \ldots, \alpha_m)x\big)}_{\in V^*}$$

ist ein Isomorphismus.

Im Fall $\dim V = \infty$ sind V und V^* nicht unbedingt isomorph. Zum Beispiel kann man zeigen, dass $\mathbb{K}[[X]]^* \cong \mathbb{K}[X] \ncong \mathbb{K}[[X]]$.

Satz 60. Ist $B = \{b_1, \ldots, b_m\}$ eine Basis von V, so bildet die Menge $B^* = \{b_1^*, \ldots, b_m^*\} \subseteq V^*$ mit

$$b_i^*(b_j) = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{sonst} \end{cases}$$

eine Basis von V^*. (Man nennt B^* die zu B *duale* Basis.)

Beweis. (a) B^* ist linear unabhängig: Seien $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ mit

$$\alpha_1 b_1^* + \cdots + \alpha_n b_n^* = 0.$$

Das bedeutet $\alpha_1 b_1^*(x) + \cdots + \alpha_n b_n^*(x) = 0$ für alle $x \in V$. Für beliebiges $i \in \{1, \ldots, n\}$ gilt daher insbesondere

$$0 = (\alpha_1 b_1^* + \cdots + \alpha_n b_n^*)(b_i) = \alpha_1 b_1^*(b_i) + \cdots + \alpha_n b_n^*(b_i) = \alpha_i.$$

(b) B^* ist ein Erzeugendensystem: Sei $x^* \in V^*$ beliebig. Sei $\alpha_i := x^*(b_i)$ für $i = 1, \ldots, n$. Für die lineare Abbildung $y^* := \alpha_1 b_1^* + \cdots + \alpha_n b_n^*$ gilt ebenfalls $\alpha_i = y^*(b_i)$ für $i = 1, \ldots, n$. Da nach Satz 50 eine lineare Abbildung eindeutig durch die Bilder auf den Basiselementen bestimmt ist, folgt $x^* = y^*$. Und da x^* beliebig war und y^* eine Linearkombination von $b_1^*, \ldots, b_n^*$ ist, folgt, dass B^* ein Erzeugendensystem von V^* ist. ∎

Definition 39. Seien V, W zwei $\mathbb{K}$-Vektorräume, $h \in \operatorname{Hom}(V,W)$. Dann wird die Abbildung $h^\top \in \operatorname{Hom}(W^*, V^*)$ definiert durch $h^\top(y^*) := (\ x \mapsto y^*(h(x))\ )$ für alle $y^* \in W^*$.

Beispiel. $V = \mathbb{R}^2$, $W = \mathbb{R}^3$, $h(x) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} x$. Dann ist $h^\top\colon W^* \to V^*$ die Abbildung, die eine Abbildung

$$y^*\colon W \to \mathbb{K}, \quad y^*\left(\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}\right) = (\beta_1, \beta_2, \beta_3) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

in die Abbildung

$$h^\top(y^*)\colon V \to \mathbb{K}, \quad h^\top(y^*)(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}) = (\beta_1, \beta_2, \beta_3) \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$
$$= (\beta_1 + 3\beta_2 + 5\beta_3)x_1 + (2\beta_1 + 4\beta_2 + 6\beta_3)x_2$$

überführt.
Einfacher lässt es sich ausdrücken, wenn man die Elemente der Dualräume durch die entsprechenden Vektoren in V bzw W ausdrückt. Dann ist nämlich einfach

$$h^\top(\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}) = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}.$$

Wie man sieht, ist die Abbildungsmatrix von $h^\top$ gerade die Transponierte der Abbildungsmatrix von h. Das ist natürlich kein Zufall:

Satz 61. Seien V, W zwei $\mathbb{K}$-Vektorräume.

1. Die Abbildung $\cdot^\top\colon \operatorname{Hom}(V,W) \to \operatorname{Hom}(W^*, V^*)$, $h \mapsto h^\top$ ist linear.
2. Seien V, W beide endlich-dimensional und seien A, B geordnete Basen von V bzw. W. Weiter seien A^* und B^* die dualen Basen von A und B. Dann gilt: Ist $M \in \mathbb{K}^{n\times m}$ die Abbildungsmatrix von $h \in \operatorname{Hom}(V,W)$ bezüglich A und B, so ist $M^\top \in \mathbb{K}^{m\times n}$ die Abbildungsmatrix von $h^\top$ bezüglich B^* und A^*.

Beweis.

1. Übung.
2. Sei $\tilde{M}$ die Abbildungsmatrix von $h^\top$ bezüglich B^* und A^*. Schreibe $A = \{a_1, \dots, a_m\}$, $B = \{b_1, \dots, b_n\}$, $A^* = \{a_1^*, \dots, a_m^*\}$, $B^* = \{b_1^*, \dots, b_n^*\}$.

   An Position (i,j) von $\tilde{M}$ steht die i-te Koordinate von $h^\top(b_j^*)$ bezüglich A^*. Zu zeigen: diese ist identisch mit der j-ten Koordinate von $h(a_i)$ bezüglich B.

   Aufgrund der Definiton der Dualbasis ist die i-te Koordinate eines Funktionals $x^* \in V^*$ gerade $x^*(a_i)$, und $b_j^*(y)$ ist die j-te Koordinate von $y \in W$. Daher:

   $$\begin{aligned} &i\text{-te Koordiante von } h^\top(b_j^*) \text{ bezüglich } A^* \\ &= h^\top(b_j^*)(a_i) \\ &= b_j^*(h(a_i)) \\ &= j\text{-te Koordinate von } h(a_i) \text{ bezüglich } B, \end{aligned}$$

   wie behauptet. ∎

Der Satz ist eine Verallgemeinerung der Matrixtransposition auf Homomorphismen zwischen beliebigen Vektorräumen. Es bietet sich deshalb an dieser Stelle an, auch die Begriffe Ko-Kern und Ko-Bild für beliebige Homomorphismen zu definieren. Für $h\colon V \to W$ war ja bereits definiert $\ker h = \{x \in V : h(x) = 0\}$ und $\operatorname{im} h = \{h(x) : x \in V\}$. Darüber hinaus definieren wir jetzt $\operatorname{coker} h = \{x \in W^* : h^\top(x) = 0\}$ und $\operatorname{coim} h = \{h^\top(x) : x \in W^*\}$. Man beachte, dass $\operatorname{coker} h$ ein Unterraum von W^* und $\operatorname{coim} h$ ein Unterraum von V^* ist.

Satz 62. Sei V ein $\mathbb{K}$-Vektorraum und $V^{**} := (V^*)^*$ der Dualraum des Dualraums von V (V^{**} ist der sogenannte *Bidualraum* von V). Die Abbildung

$$h\colon V \to V^{**}, \quad h(x) = (\ f \mapsto f(x)\ )$$

ist linear und injektiv, und im Fall $\dim V < \infty$ auch bijektiv.

Beweis. Linearität: Für alle $\alpha, \beta \in \mathbb{K}$ und alle $x, y \in V$ gilt:

$$\begin{aligned} h(\alpha x + \beta y) &= (\ f \mapsto f(\alpha x + \beta y)\ ) \\ &= (\ f \mapsto \alpha f(x) + \beta f(y)\ ) \\ &= \alpha(\ f \mapsto f(x)\ ) + \beta(\ f \mapsto f(y)\ ) \\ &= \alpha h(x) + \beta h(y). \end{aligned}$$

Injektivität: Für alle $x, y \in V$ gilt

$$\begin{aligned} h(x) = h(y) \quad &\Rightarrow \quad h(x-y) = 0 \\ &\Rightarrow \quad (\ f \mapsto f(x-y)\ ) = 0 \\ &\Rightarrow \quad \forall\ f \in V^* : f(x-y) = 0 \\ &\Rightarrow \quad x = y. \end{aligned}$$

Bijektivität im Fall $\dim V < \infty$ folgt aus Satz 52 und $\dim V^{**} = \dim V^* = \dim V$. ∎

Teil IV

Anwendungen

19 Affine und Projektive Geometrie

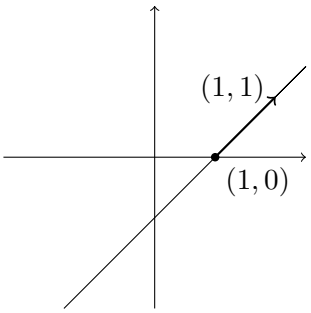
Definition 40. Sei $U \subseteq \mathbb{K}^n$ ein Untervektorraum und $x \in \mathbb{K}^n$. Dann heißt

$$x + U := \{\, x + u : u \in U \,\} \subseteq \mathbb{K}^n$$

ein *affiner Unterraum* (engl. *affine subspace*) von $\mathbb{K}^n$.
Im Fall $\dim U = 0$ / 1 / 2 / $n-1$ spricht man von einem *Punkt* / einer *Geraden* / einer *Ebene* / einer *Hyperebene*. (engl: *point*, *line*, *plane*, *hyper plane*)

Beispiel.

1. $G = \{ \begin{pmatrix}1\\0\end{pmatrix} + \lambda \begin{pmatrix}1\\1\end{pmatrix} : \lambda \in \mathbb{R} \} = \begin{pmatrix}1\\0\end{pmatrix} + \langle \begin{pmatrix}1\\1\end{pmatrix} \rangle \subseteq \mathbb{R}^2$ ist die Gerade durch den Punkt $\begin{pmatrix}1\\0\end{pmatrix}$ in Richtung $\begin{pmatrix}1\\1\end{pmatrix}$.



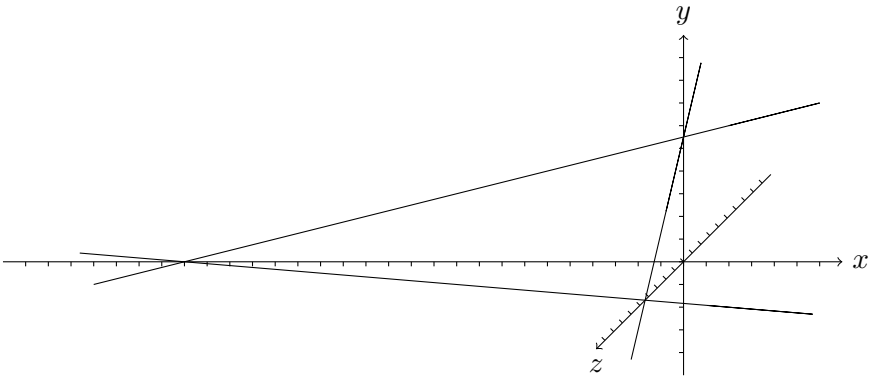
   $G' = \begin{pmatrix}0\\-1\end{pmatrix} + \langle \begin{pmatrix}-1\\-1\end{pmatrix} \rangle \subseteq \mathbb{R}^2$ ist eine andere Schreibeweise für dieselbe Gerade.

2. $E = \begin{pmatrix}1\\2\\3\end{pmatrix} + \langle \begin{pmatrix}1\\-1\\1\end{pmatrix}, \begin{pmatrix}1\\4\\-3\end{pmatrix} \rangle$ ist eine Ebene. Um sich die Lage der Ebene im dreidimensionalen Raum zu veranschaulichen, kann man die Schnittgeraden mit den drei Koordinatenebenen bestimmen:

$$E \cap \underbrace{\langle \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix} \rangle}_{(x,y)\text{-Ebene}} = \begin{pmatrix}2\\6\\0\end{pmatrix} + \langle \begin{pmatrix}4\\1\\0\end{pmatrix} \rangle$$

$$E \cap \underbrace{\langle \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \rangle}_{(x,z)\text{-Ebene}} = \begin{pmatrix}3\\0\\5\end{pmatrix} + \langle \begin{pmatrix}5\\0\\1\end{pmatrix} \rangle$$

$$E \cap \underbrace{\langle \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \rangle}_{(y,z)\text{-Ebene}} = \begin{pmatrix}0\\3\\2\end{pmatrix} + \langle \begin{pmatrix}0\\5\\-4\end{pmatrix} \rangle.$$



Den Schnitt zweier affiner Unterräume kann man berechnen, indem man ein inhomogenes lineares Gleichungssystem löst. Zum Beispiel bekommt man die erste Schnittgerade, indem man alle $\alpha, \beta, \lambda, \mu$ bestimmt, für die gilt:

$$\begin{pmatrix}1\\2\\3\end{pmatrix} + \alpha \begin{pmatrix}1\\-1\\1\end{pmatrix} + \beta \begin{pmatrix}1\\4\\-3\end{pmatrix} = \lambda \begin{pmatrix}1\\0\\0\end{pmatrix} + \mu \begin{pmatrix}0\\1\\0\end{pmatrix}.$$

Die Schnittgerade mit einer Koordinatenebene kann man auch einfacher ausrechnen. Dazu nutzt man aus, dass z. B. die (x, y)-Ebene gerade die Menge aller Punkte $(x, y, z) \in \mathbb{R}^3$ mit der Eigenschaft $z = 0$ ist. Der Schnitt dieser Ebene mit E ist also auch die Menge aller Punkte aus E, deren letzte Koordinate 0 ist. Um diese Punkte zu finden, braucht man nur das inhomogene Gleichungssystem

$$3 + \alpha \cdot 1 + \beta \cdot (-3) = 0$$

zu lösen.

Satz 63. Der Schnitt zweier affiner Unterräume von $\mathbb{K}^n$ ist entweder die leere Menge oder wieder ein affiner Unterraum von $\mathbb{K}^n$.

Beweis. Seien $x_1 + U_1$ und $x_2 + U_2$ zwei affine Unterräume von $\mathbb{K}^n$. Falls deren Schnitt leer ist, ist nichts zu zeigen. Anderenfalls gibt es ein $x \in \mathbb{K}^n$ mit $x \in (x_1 + U_1) \cap (x_2 + U_2)$, etwa $x = x_1 + u_1 = x_2 + u_2$ für gewisse $u_1 \in U_1$ und $u_2 \in U_2$. Wir zeigen, dass dann $(x_1 + U_1) \cap (x_2 + U_2) = x + (U_1 \cap U_2)$ gilt.

„$\subseteq$“ Sei $z \in (x_1 + U_1) \cap (x_2 + U_2)$. Zu zeigen: $z \in x + (U_1 \cap U_2)$, d. h. $x - z \in U_1 \cap U_2$. Nach Voraussetzung gilt $z \in x_1 + U_1$ und $z \in x_2 + U_2$, also $z = x_1 + \tilde{u}_1 = x_2 + \tilde{u}_2$ für gewisse $\tilde{u}_1 \in U_1$ und $\tilde{u}_2 \in U_2$. Also gilt

$$z - x = \begin{cases} u_1 - \tilde{u}_1 \in U_1 \\ u_2 - \tilde{u}_2 \in U_2 \end{cases}$$

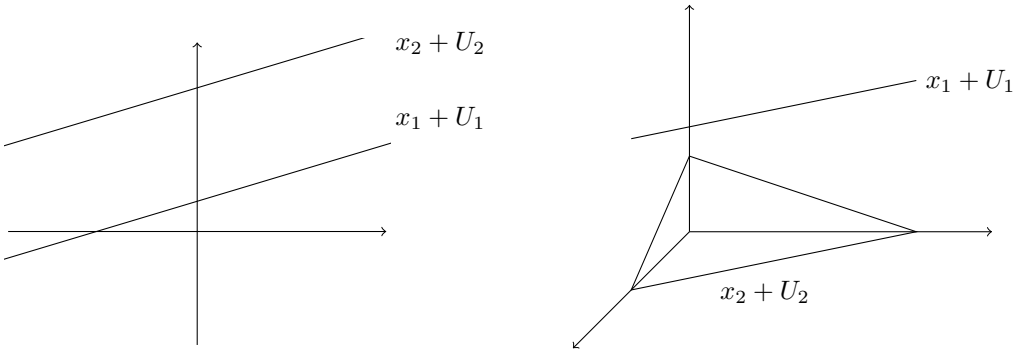
und damit $z - x \in U_1 \cap U_2$.

„$\supseteq$“ Sei $z \in x + (U_1 \cap U_2)$. Zu zeigen: $z \in x_1 + U_1$ und $z \in x_2 + U_2$. Aus Symmetriegründen genügt es, $z \in x_1 + U_1$ zu zeigen. Nach Voraussetzung gilt

$$z \in \underset{\substack{\uparrow \\ = x_1 + u_1}}{x} + \underbrace{(U_1 \cap U_2)}_{\subseteq U_1} \subseteq x_1 + \underbrace{u_1 + U_1}_{= U_1}.$$

■

Definition 41. Zwei affine Unterräume x_1+U_1, x_2+U_2 von $\mathbb{K}^n$ heißen (zueinander) *parallel,* falls $U_1 \subseteq U_2$ oder $U_2 \subseteq U_1$ ist.



Zwei Geraden im Raum $\mathbb{K}^2$, die sich nicht schneiden, sind stets zueinander parallel. (Beweis: Übung.) Zwei Geraden im Raum $\mathbb{K}^3$ können sich jedoch auch dann verfehlen, wenn sie nicht parallel sind.

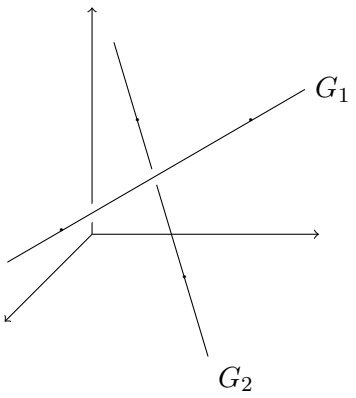
Beispiel. Betrachte die beiden Geraden

$$G_1 = \begin{pmatrix}1\\2\\3\end{pmatrix} + \langle \begin{pmatrix}2\\1\\-1\end{pmatrix} \rangle, \qquad G_2 = \begin{pmatrix}2\\5\\0\end{pmatrix} + \langle \begin{pmatrix}4\\-5\\5\end{pmatrix} \rangle.$$

Der Schnitt $G_1 \cap G_2$ ist leer, weil das inhomogene Gleichungssystem

$$\begin{pmatrix}1\\2\\3\end{pmatrix} + \alpha \begin{pmatrix}2\\1\\-1\end{pmatrix} = \begin{pmatrix}2\\5\\0\end{pmatrix} + \beta \begin{pmatrix}4\\-5\\5\end{pmatrix}$$

keine Lösung hat. Trotzdem sind diese Geraden nicht parallel, weil $\begin{pmatrix}2\\1\\-1\end{pmatrix}$ und $\begin{pmatrix}4\\-5\\5\end{pmatrix}$ linear unabhängig sind.



Zwei Ebenen im $\mathbb{K}^3$ sind entweder zueinander parallel oder sie schneiden sich in einer Gerade. Zwei Ebenen im $\mathbb{K}^4$ können als Schnittmenge eine Gerade, einen einzelnen Punkt, oder die leere Menge haben, und zwar auch dann, wenn sie nicht parallel sind.

Beispiel. Für die beiden Ebenen

$$E_1 = \begin{pmatrix}1\\2\\3\\4\end{pmatrix} + \langle \begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}, \begin{pmatrix}0\\1\\-1\\0\end{pmatrix} \rangle, \qquad E_2 = \begin{pmatrix}1\\2\\-2\\1\end{pmatrix} + \langle \begin{pmatrix}4\\3\\2\\1\end{pmatrix}, \begin{pmatrix}1\\2\\4\\8\end{pmatrix} \rangle$$

gilt

$$E_1 \cap E_2 = \{ \begin{pmatrix}5\\5\\0\\0\end{pmatrix} \}.$$

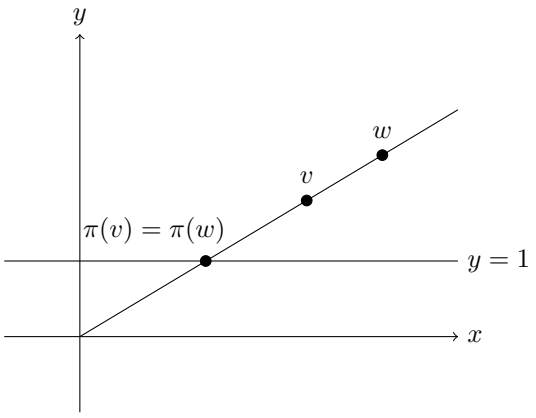
Definition 42. Für $v, w \in \mathbb{K}^{n+1} \setminus \{0\}$ sei definiert

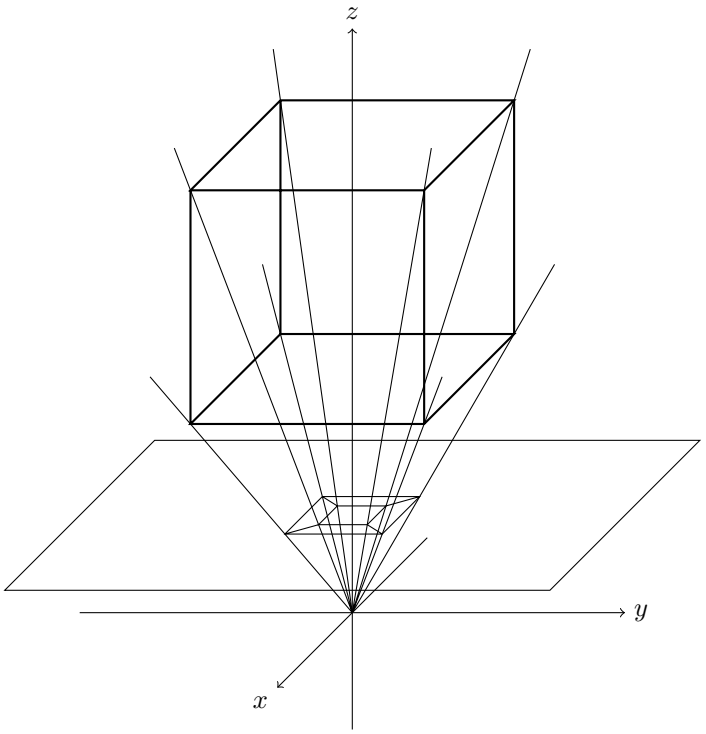
$$v \sim w \quad :\Longleftrightarrow \quad \exists\, \lambda \in \mathbb{K} \setminus \{0\} : v = \lambda w.$$

Dann heißt $\mathbb{P}^n := (\mathbb{K}^{n+1} \setminus \{0\})/\sim$ der *projektive Raum* der Dimension n über $\mathbb{K}$.

Statt $[(x_0, \ldots, x_n)]_\sim$ schreibt man $(x_0 : \cdots : x_n)$ und spricht von *projektiven Koordinaten.*

Anschauung: Das Auge eines Betrachters am Ursprung $(0, \ldots, 0)$ des Raums $\mathbb{K}^{n+1}$ kann zwei Punkte $v, w \in \mathbb{K}^{n+1}$ nicht voneinander unterscheiden, wenn $v \sim w$ gilt, d. h. wenn diese Punkte auf derselben Geraden durch den Ursprung liegen. Mit Ausnahme der Geraden auf der Grundfäche $x_{n+1} = 0$ schneiden alle diese Geraden den affinen Unterraum $\mathbb{K}^n \times \{1\} \subseteq \mathbb{K}^{n+1}$ in genau einem Punkt. Dieser Punkt bietet sich als kanonischer Repräsentant der jeweiligen Äquivalenzklassen an. Man kann sich den affinen Unterraum $\mathbb{K}^n \times \{1\}$ geometrisch als Projektionsfläche vorstellen. Betrachtet man ein geometrisches Objekt im $\mathbb{K}^{n+1}$ als ein Objekt in $\mathbb{P}^n$, so geht ein Teil der Information über Lage und Form des Objekts verloren. Was an Information übrig bleibt, entspricht genau dem Bild der Projektion auf der Projektionsfläche.





Die Projektion bildet den Punkt $(x_0 : \cdots : x_n) \in \mathbb{P}^n$ mit $x_n \neq 0$ auf den Punkt $(\frac{x_0}{x_n}, \ldots, \frac{x_{n-1}}{x_n}) \in \mathbb{K}^n$ ab. Die restlichen Punkte $(x_0 : \cdots : x_{n-1} : 0) \in \mathbb{P}^n$ bilden zusammengenommen eine Kopie von $\mathbb{P}^{n-1}$. Diese Punkte kann man sich anschaulich als Punkte vorstellen, die „unendlich weit“ vom Ursprung entfernt liegen. Im Fall $n = 1$ entsprechen die Punkte $(x_0 : x_1)$ mit $x_1 = 0$ genau den Punkten auf der x_0-Achse, die die zu ihr parallel verlaufende Projektionsgerade $x_0 = 1$ in einem gedachten unendlich fernen Punkt $(1 : 0) \in \mathbb{P}^1$ schneidet. Es gilt also „$\mathbb{P}^1 \cong \mathbb{K}^1 \cup \mathbb{P}^0$“. Im Fall $n = 2$ verlaufen die Grundebene $\{\, (x_0, x_1, x_2) : x_2 = 0 \,\}$ und die Projektionsebene $\{\, (x_0, x_1, x_2) : x_2 = 1 \,\}$ zueinander parallel. Sie schneiden sich in einer gedachten Gerade, die aus lauter unendlich fernen Punkten besteht, je einen für jede Richtung $(x_0 : x_1)$. Es gilt also „$\mathbb{P}^2 \cong \mathbb{K}^2 \cup \mathbb{P}^1$“.

Durch die Hinzunahme von unendlich fernen Punkte lassen sich die lästigen Fallunterscheidungen, die beim Rechnen mit affinen Räumen auftreten, vermeiden. Statt der inhomogenen Gleichungssysteme

$$x + \alpha_1 v_1 + \cdots + \alpha_m v_m = 0,$$

die da auftreten, hat man es im projektiven Raum nur mit homogenen Gleichungsystemen

$$\lambda x + \alpha_1 v_1 + \cdots + \alpha_m v_m = 0$$

zu tun.

In der Computergraphik wird grundsätzlich mit projektiven Koordinaten gearbeitet. Objekte in einem dreidimensionalen virtuellen Raum werden dargestellt durch Punkte im projektiven Raum $\mathbb{P}^3$. Positionsänderungen der Objekte lassen sich durch die Anwendung geeigneter 4×4-Matrizen ausdrücken. Für eine Matrix $A \in \mathbb{K}^{(n+1)\times(n+1)}$ und einen Punkt $[x]_\sim \in \mathbb{P}^n$ mit $x \in \mathbb{K}^{n+1} \setminus \{0\}$ definiert man $A[x]_\sim := [Ax]_\sim$. Eine solche Definition ist zulässig, weil das Ergebnis wegen $x' = Ax \iff \forall\, \lambda \neq 0 : \lambda x' = A(\lambda x)$ nicht von der Wahl des Repräsentanten abhängt.

So kann man zum Beispiel im projektiven Raum auch die Verschiebung eines Punktes $(x_0 : x_1 : x_2 :$

$x_3) \in \mathbb{P}^3$ um den Vektor $(y_0, y_1, y_2) \in \mathbb{K}^3$ als eine lineare Abbildung schreiben:

$$\begin{pmatrix} x'_0 \\ x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & y_1 \\ 0 & 1 & 0 & y_2 \\ 0 & 0 & 1 & y_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

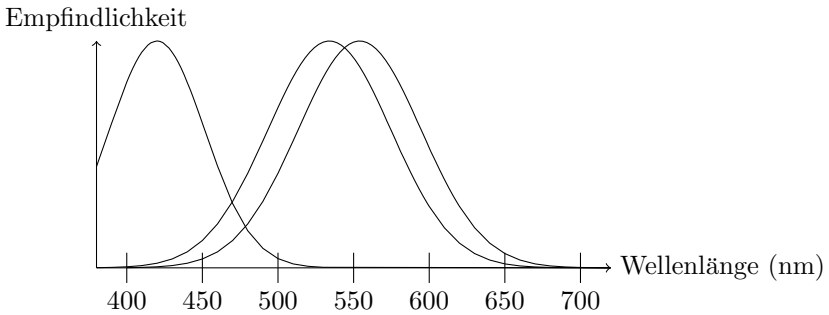
Im Raum $\mathbb{K}^3$ ist die Verschiebung um (y_1, y_2, y_3) dagegen keine lineare Abbildung.

Zur Darstellung der dreidimensionalen virtuellen Realiät auf einem zweidimensionalen Bildschirm betrachtet man die Bildschirmoberfläche als projektiven Raum $\mathbb{P}^2$ und verwendet eine Matrix $A \in \mathbb{K}^{3\times 4}$, um die Punkte in $\mathbb{P}^3$ auf Punkte im $\mathbb{P}^2$ abzubilden. Die Einträge von A codieren Lage, Blickrichtung, Drehwinkel und Brennweite (Zoom) der virtuellen Kamera. Wie das genau funktioniert, erfahren Sie in den einschlägigen Lehrveranstaltungen des Instituts für Angewandte Geometrie von Prof. Jüttler.

20 Farbräume

Das Spektrum des sichtbaren Lichts umfasst die elektromagnetische Strahlung mit Wellenlängen von ca. 400nm (violett) bis ca. 700nm (rot). Jede Wellenlänge entspricht einer bestimmten (Spektral-)Farbe, und jeder andere Farbeindruck einspricht einer bestimmten Überlagerung solcher Spektralfarben. Weisses Licht ergibt sich zum Beispiel aus der gleichmäßigen Mischung von Licht in allen Farben des Spektrums. In der Sprache der linearen Algebra kann man die Farben als einen Vektorraum auffassen, der von den Spektralfarben erzeugt wird. Jede Wellenlänge entspricht dann einem anderen Basisvektor, und jede Farbe ist eine Linearkombination, deren Koeffizienten angeben, wie stark der Lichtanteil der entsprechenden Spektralfarbe ist.

Wenn jede Wellenlänge einem Basisvektor entsprechen soll, dann ist der Farbraum offenbar unendlich dimensional. Das mag aus physikalischer Sicht auch eine adäquate Beschreibung der Natur des Lichts sein, aber für praktische Anwendungen sind so viele Dimensionen weder nützlich noch nötig. Das menschliche Auge kann so viele Farben gar nicht unterscheiden. Auf der Netzhaut des Auges gibt es vier verschiedene Arten von lichtempfindliche Zellen, von denen drei für das farbige Sehen zuständig sind. Jede dieser drei Zellarten reagiert auf Licht mit einer bestimmten Wellenlänge λ besonders stark, und auf Licht mit anderen Wellenlängen umso schwächer, je stärker die andere Wellenlänge von λ abweicht.



Der Farbeindruck, den wir subjektiv wahrnehmen, ergibt sich daraus, wie stark die drei verschiedenen Zelltypen vom eintreffenden Licht angeregt werden. Das Auge codiert also jede Farbe als einen Vektor (x, y, z), in dem die Koordinaten angeben, wie stark jeder der drei Zelltypen von der Farbe angeregt wird. Vereinfachend kann man sich vorstellen, dass die Koordinaten zwischen 0 (gar keine Anregung) und 1 (maximale Anregung) liegen.

Die Basis in diesem Modell bilden die drei „Grundfarben“ rot, grün und blau, für die jeweils einer der drei Zelltypen auf der Netzhaut maximal angeregt wird. Das Mischen von Farben entspricht der

Linearkombination von Vektoren. Wenn man dabei nichtnegative Koeffizienten wählt, die sich zu (höchstens) 1 aufaddieren, dann ist gewährleistet, dass der Ergebnisvektor wieder Koordinaten hat, die zwischen 0 und 1 liegen und damit einen gültigen Farbvektor darstellen. Solche Linearkombinationen $\alpha_1 a_1 + \cdots + \alpha_m a_m$ mit $\alpha_1, \ldots, \alpha_m \geq 0$ und $\alpha_1 + \cdots + \alpha_m = 1$ nennt man auch *Konvexkombinationen.*
Beispiele:

$(0,0,0) =$ ■, $(1,0,0) =$ ■, $(0,1,0) =$ ■, $(0,0,1) =$ ■,
$(0,1,1) =$ ■, $(1,0,1) =$ ■, $(1,1,0) =$ ■, $(1,1,1) =$ ■,
$(\frac{1}{3}, 1, \frac{2}{3}) =$ ■, $(\frac{1}{2}, 0, 0) =$ ■, $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}) =$ ■, $(0, \frac{1}{4}, 1) =$ ■.

Ausgabegeräte wie zum Beispiel Monitore erzeugen Farben, indem sie rotes, grünes und blaues Licht in genau dem Mischungsverhältnis ausstrahlen, das im Auge des Betrachters den gewünschten Farbeindruck hervorruft.

Ein Bild kann man als eine Funktion auffassen, die jedem Punkt der Ebene eine Farbe zuordnet. Farbige Bilder werden im Computer als Überlagerung dreier einfarbiger Bilder dargestellt, die den Anteil der drei Basisfarben am Gesamtbild angeben. Typischerweise verwendet man dabei rot, grün und blau (RGB) als Basisfarben:



Abhängig von der Anwendungssituation kann es von Vorteil sein, verschiedene Basen zu verwenden. Es kann zum Beispiel sein, dass es für die drei Basislichtquellen in einem Monitor technisch leichter ist, statt R, G und B drei andere Farben A, B und C zu verwenden. Die Farbinformation, die der Computer in der RGB-Basis an den Monitor leitet, muss dann vom Monitor in die ABC-Basis umgewandelt werden. Die Basistransformation geschieht natürlich durch Multiplikation des RGB-Vektors mit der 3×3-Matrix, deren Zeilen die Koordinatendarstellung der Farben in der RGB-Basis sind.

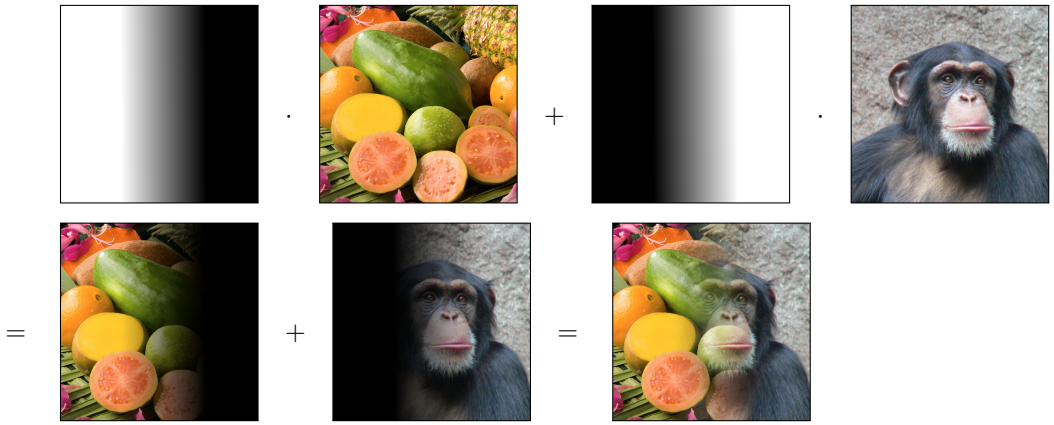
Personen, die an Rot-Grün-Blindheit leiden, haben auf ihrer Netzhaut nur zwei verschiedene Typen von farbempfindlichen Zellen. Für diese Personen ist der Farbraum deshalb nur zweidimensional, und es ergibt sich in etwa der folgende Eindruck:



Wenn Betroffene jetzt einwenden, dass für sie die Abbildung auf der linken Seite der Gleichung nicht identisch aussieht wie die Abbildung auf der linken Seite der dreidimensionalen Zerlegung, dann liegt

das daran, dass die Wellenlängen, bei denen die Farbrezeptoren im Auge maximal angeregt werden, von Mensch zu Mensch leicht verschieden sind. Den exakt gleichen Eindruck bekommt man nur dann, wenn man in der Rechnung ein rot und ein blau verwendet, die gemeisam den gleichen Untervektorraum aufspannen wie das rot und das blau, die die Zellen im Auge der betreffenden Person wahrnehmen. Darüber hinaus müsste sichergestellt sein, dass das verwendete Grafikprogram sowie der Drucker bzw. der Bildschirm exakt die gleichen Farbtöne als Basisfunktion verwenden, was technisch nur sehr schwer umzusetzen ist.

Die Skalierung eines Farbvektors ändert die Helligkeit: $1 \cdot v = v$ ist die Farbe v selbst, $0 \cdot v = 0$ ist schwarz, und z.B. $\frac{1}{2} \cdot v$ ist eine dunklere Variante der Farbe v. Die Zahen $\alpha \in [0, 1]$ kann man deshalb auch als Graustufen interpretieren (mit schwarz als 0 und weiß als 1). 03-03



Wenn Sie dieses Skriptum auf einem Schwarz-Weiss-Drucker ausgedruckt haben, dann hat Ihr Computer die Farben in den Abbildungen dieses Abschnitts in Grautöne umwandeln müssen. Ein Grauton ist eine reelle Zahl zwischen 0 (schwarz) und 1 (weiss). Um ein Farbbild in ein Graustufenbild umzuwandeln, wählt man Grauwerte $g_R, g_G, g_B \in [0, 1]$, auf die die Basisfarben R, G, B abgebildet werden sollen. Eine beliebige Farbe (c_R, c_G, c_B) des RGB-Raums wird dann abgebildet auf den Grauwert

$$(g_R, g_G, g_B) \begin{pmatrix} c_R \\ c_G \\ c_B \end{pmatrix}.$$

Die Abbildung, die jeder RGB-Farbe einen Grauwert zuordnet, ist also ein Funktional.

Umgekehrt: Vögel werden die Bilder, die auf unseren Monitoren angezeigt werden, vermutlich nicht besonders realistisch finden. Sie verfügen nämlich über eine Netzhaut mit vier verschiedenen Typen von Farbrezeptoren und genießen daher einen vierdimensionalen Farbraum. Sie können deshalb viele Farben voneinander unterscheiden, die für uns identisch sind.

21 Graphentheorie

Definition 43. Sei V eine Menge und E eine Relation auf V. Dann heißt das Paar $G = (V, E)$ ein *Graph*. Die Elemente von V heißen *Knoten* (engl. *vertex*) und die Elemente von E heißen *Kanten* (engl. *edge*) des Graphen. Eine Kante der Form (v, v) heißt *Schleife* (engl. *loop*).

Ein Graph in diesem Sinne hat nichts mit Funktionsgraphen zu tun. Es handelt sich viel mehr um eine Art Netzwerk, wie wir bereits im Abschnitt 2 als Beispiel für eine Relation gesehen haben. Graphen

kann man sich graphisch veranschaulichen, indem man für die Knoten irgendwelche Punkte in der Ebene wählt und je zwei Punkte $v_1, v_2 \in V$ genau dann durch einen Pfeil von v_1 nach v_2 verbindet, wenn $(v_1, v_2) \in E$ ist.

Um einen Graphen im Computer zu speichern, kann man einfach eine Liste mit den Elementen von V sowie eine Liste mit den Elementen von E speichern. Für manche Anwendungen ist es zweckmäßiger, die Kantenmenge E durch eine Matrix darzustellen, deren Einträge für je zwei Knoten v_1, v_2 angeben, ob es eine Kante von v_1 nach v_2 gibt oder nicht.

Definition 44. Es sei $G = (V, E)$ ein Graph mit $|V| = n < \infty$. Weiter sei $v\colon \{1, \dots, n\} \to V$ eine bijektive Funktion. Die *Adjazenzmatrix* von G (bezüglich v) ist definiert als die Matrix $A = ((a_{i,j}))_{i,j=1}^{n} \in \mathbb{Q}^{n \times n}$ mit

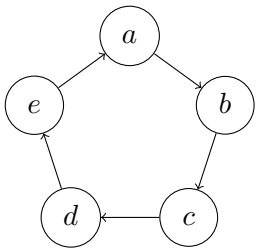
$$a_{i,j} = \begin{cases} 1 & \text{falls } (v(i), v(j)) \in E \\ 0 & \text{sonst} \end{cases}$$

für $i, j = 1, \dots, n$.

Die Funktion v in der obigen Definition bedeutet bloß, dass die Knoten in einer beliebigen aber bestimmten Weise angeordnet werden sollen. Welche Reihenfolge gewählt wird, ist egal, aber für zwei verschiedene Wahlen von v erhält man im allgemeinen verschiedene Adjazenzmatrizen.

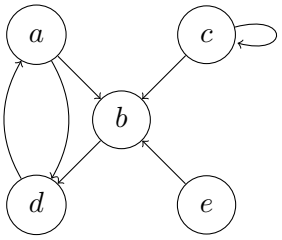
Beispiel. Im folgenden ist für die Adjazenzmatrix A immer angenommen, dass v die Knoten in der Reihenfolge durchnumeriert, in der sie notiert sind.

1. $G_1 = (V_1, E_1)$ mit $V_1 = \{a, b, c, d, e\}$, $E_1 = \{(a, b),\ (b, c),\ (c, d),\ (d, e),\ (e, a)\}$



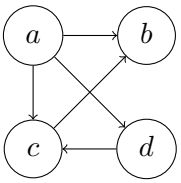
$$A_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

2. $G_2 = (V_2, E_2)$ mit $V_2 = \{a, b, c, d, e\}$,
   $E_2 = \{(a, b),\ (a, d),\ (c, c),\ (c, b),\ (e, b),\ (d, a),\ (b, d)\}$



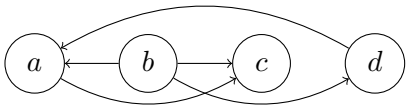
$$A_2 = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

3. $G_3 = (V_3, E_3)$ mit $V_3 = \{a, b, c, d\}$, $E_3 = \{(a, b),\ (a, c),\ (a, d),\ (c, b),\ (d, c)\}$



$$A_3 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

4. $G_4 = (V_4, E_4)$ mit $V_4 = \{a, b, c, d\}$, $E_4 = \{(a, c),\ (b, a),\ (b, c),\ (b, d),\ (d, a)\}$



$$A_4 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Definition 45. Seien $G = (V, E)$ und $G' = (V', E')$ zwei Graphen. Eine Funktion $h\colon V \to V'$ mit

$$\forall\ v_1, v_2 \in V : (v_1, v_2) \in E \iff (h(v_1), h(v_2)) \in E'$$

heißt *(Graphen-)Homomorphismus* von G nach G'. Ist h bijektiv, spricht man von einem *(Graphen-)Isomorphismus* und sagt, die Graphen G und G' seien (zueinander) *isomorph.* Schreibweise in diesem Fall: $G \cong G'$.

Beispiel. Von den vier Graphen im vorigen Beispiel sind nur der dritte und der vierte zueinander isomorph. Die Abbildung $h\colon V_3 \to V_4$, die durch folgende Wertetabelle definiert ist, ist ein Isomorphismus.

v	a	b	c	d
$h(v)$	b	c	a	d

Um das zu zeigen, rechnet man nach, dass die Mengen E_4 und

$$h(E_3) := \{\, (h(v), h(w)) : (v, w) \in E_3 \,\} \subseteq V_4^2$$

identisch sind. (Sie sind es.)

Dass G_1 und G_2 nicht isomorph sein können, sieht man zum Beispiel daran, dass der zweite Graph eine Schleife hat und der erste nicht. Ein Isomorphismus muss aber Schleifen (v, v) auf Schleifen $(h(v), h(v))$ abbilden.

Außerdem kann weder G_1 noch G_2 isomorph zu G_3 oder G_4 sein, weil ein Isomorphismus bijektiv sein muss und Graphen deshalb nur dann isomorph sein können, wenn sie die gleiche Anzahl von Knoten haben.

Anfänger glauben oft, dass zwei Graphen G_1 und G_2 isomorph sind, wenn für jedes $n \in \mathbb{N}$ die Zahl der Knoten in G_1, die mit genau n Knoten durch eine Kante verbunden sind, mit der Zahl der Knoten in G_2, die mit genau n durch eine Kante verbunden sind, übereinstimmt. Das ist aber falsch. Tatsächlich handelt es sich hierbei nur um eine notwendige Bedingung, aber nicht um eine hinreichende.

Satz 64. Seien $G = (V, E)$ und $G' = (V', E')$ zwei Graphen mit endlichen Kantenmengen V, V'. Sei A eine Adjazenzmatrix von G und A' eine Adjazenzmatrix von G'. Dann gilt: G und G' sind genau dann isomorph, wenn $|V| = |V'|$ gilt und es eine Permutationsmatrix $P \in \mathbb{Q}^{|V|\times|V|}$ gibt mit $A' = PAP^{-1}$.

Beweis. Seien $v\colon \{1, \dots, |V|\} \to V$ und $v'\colon \{1, \dots, |V'|\} \to V'$ die zu den gegebenen Adjazenzmatrizen A und A' gehörenden Bijektionen.

„$\Rightarrow$“ Sei $h\colon V \to V'$ ein Isomorphismus. Da h als Isomorphismus insbesondere bijektiv ist, ist $\pi := (v')^{-1} \circ h \circ v$ eine Bijektion von $\{1, \dots, |V|\}$ nach $\{1, \dots, |V'|\}$. Damit gilt $|V| = |V'|$ und somit ist π eine Permutation. Sei $P \in \mathbb{Q}^{|V|\times|V|}$ die zu π gehörende Permutationsmatrix. Wir zeigen $A' = PAP^{-1}$. Betrachte dazu eine beliebige Position $(i, j) \in \{1, \dots, |V|\}^2$. Bezeichnen wir mit $a_{i,j}$ und $a'_{i,j}$ den Eintrag von A bzw. A' an Position (i, j), so gilt

$$\begin{aligned} a_{i,j} = 1 &\iff (v(i), v(j)) \in E \\ &\iff (h(v(i)), h(v(j))) \in E' \\ &\iff a'_{(v')^{-1}(h(v(i))),(v')^{-1}(h(v(j)))} = 1 \\ &\iff a'_{\pi(i),\pi(j)} = 1. \end{aligned}$$

Damit folgt die Behauptung aus Teil 3 von Satz 19.

„$\Leftarrow$“ Sei $\pi\colon \{1, \dots, |V|\} \to \{1, \dots, |V|\}$ die zu P gehörige Permutation. Wir zeigen, dass $h\colon V \to V'$, $h = v' \circ \pi \circ v^{-1}$ ein Isomorphismus ist. Bezeichnen wir mit $a_{i,j}$ und $a'_{i,j}$ wieder den Eintrag von A bzw. A' an Position (i, j), so gilt nach Teil 3 von Satz 19 $a_{i,j} = 1 \iff a'_{\pi(i),\pi(j)} = 1$. Für zwei beliebige Knoten $u, w \in V$ gilt dann

$$\begin{aligned} (u, w) \in E &\iff a_{v^{-1}(u),v^{-1}(w)} = 1 \\ &\iff a'_{\pi(v^{-1}(u)),\pi(v^{-1}(w))} = 1 \\ &\iff a'_{(v')^{-1}(h(u)),(v')^{-1}(h(w))} = 1 \\ &\iff (h(u), h(w)) \in E', \end{aligned}$$

was zu zeigen war. ∎

Beispiel. Betrachten wir noch einmal die Graphen G_3 und G_4 aus dem Beispiel nach Definition 44. Die dort angegebenen Adjazenzmatrizen

$$A_3 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{und} \quad A_4 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

gelten beide bezüglich der Funktion $v\colon \{1, 2, 3, 4\} \to \{a, b, c, d\}$ mit $v(1) = a$, $v(2) = b$, $v(3) = c$, $v(4) = d$.

Ist $h\colon \{a, b, c, d\} \to \{a, b, c, d\}$ der Isomorphismus aus dem vorigen Beispiel, so ist die Permutation $\pi = v^{-1} \circ h \circ v$ gegeben durch

i	1	2	3	4
$\pi(i)$	2	3	1	4

.

Die zugehörige Permutationsmatrix lautet

$$P_\pi = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

In der Tat gilt

$$P_\pi^{-1} A_3 P_\pi = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = A_4.$$

Definition 46. Sei $G = (V, E)$ ein Graph. Ein Tupel $((v_0, v_1), (v_2, v_3), \ldots, (v_{m-1}, v_m)) \in E^m$ heißt *Pfad* (engl. *path*) von $v_1 \in V$ nach $v_m \in V$ von G der Länge m. Statt

$$((v_0, v_1), (v_1, v_2), \ldots, (v_{m-1}, v_m))$$

schreibt man einen Pfad auch in der Form $v_0 - v_1 - \cdots - v_m$.
Ein Pfad mit $v_0 = v_m$ heißt *geschlossener Pfad* oder *Zyklus* (engl. *cycle*) von G (der Länge m).

Beispiel.

1. Die Pfade der Länge 1 entsprechen genau den Kanten des Graphen, und die Zyklen der Länge 1 entsprechen genau seinen Schleifen.

2. Der Graph G_2 auf Seite 130 enthält den Pfad $c - c - b - d - a - b - d - a - d - a$. Dieser Pfad hat die Länge 9. Der Pfad ist kein Zyklus, aber z. B. die Pfade $a - b - d - a$ und $a - d - a$ sind Zyklen von G_2.

   Der Graph G_3 enthält dagegen keine Zyklen.

Satz 65. Sei $G = (V, E)$ ein Graph mit endlicher Knotenmenge, $v \colon \{1, \ldots, |V|\} \to V$ bijektiv und A die Adjazenzmatrix von G bezüglich v. Weiter seien $i, j \in \{1, \ldots, |V|\}$.

1. Der (i, j)-te Eintrag in der Matrix A^n ist genau die Anzahl der Pfade von $v(i)$ nach $v(j)$ der Länge n.

2. Der (i, j)-te Eintrag in der Matrix $A + A^2 + \cdots + A^n$ ist genau die Anzahl der Pfade von $v(i)$ nach $v(j)$ der Länge höchstes n.

3. Ein Pfad von $v(i)$ nach $v(j)$ existiert genau dann, wenn ein Pfad von $v(i)$ nach $v(j)$ der Länge höchstens $|V|$ existiert.

Beweis.

1. Induktion nach n. Für $n = 1$ folgt die Behauptung direkt aus Def. 44. Nehmen wir an, die Behauptung gilt für n. Sei $(i, j) \in \{1, \ldots, |V|\}^2$. Zu zeigen ist, dass der (i, j)-te Eintrag von A^{n+1} die Anzahl der Pfade von $v(i)$ nach $v(j)$ der Länge $n+1$ ist. Jeder solche Pfad lässt sich auffassen als ein Pfad von $v(i)$ nach $v(k)$ für ein $k \in \{1, \ldots, n\}$ gefolgt von der Kante $(v(k), v(j))$, falls diese existiert. Die Anzahl der Pfade von $v(i)$ nach $v(k)$ ist nach Voraussetzung der (i, k)-te Eintrag $a_{i,k}^{[n]}$ von A^n, und die Kante $(v(k), v(j))$ existiert nach Def. 44 genau dann, wenn $a_{k,j} = 1$ ist. Da $a_{k,j}$ im anderen Fall 0 ist, lässt sich die Anzahl der Pfade von $v(i)$ nach $v(j)$ der Länge $n + 1$ schreiben als

   $$a_{i,1}^{[n]} a_{1,j} + a_{i,2}^{[n]} a_{2,j} + \cdots + a_{i,|V|}^{[n]} a_{|V|,j}.$$

   Nach Definition der Matrizenmultiplikation ist dies genau der (i, j)-te Eintrag von $A^n A = A^{n+1}$.

2. Folgt direkt aus Teil 1.

3. Wir zeigen: aus jedem Pfad von u nach v der Länge $m > |V|$ lässt sich ein Pfad von u nach v mit einer Länge $< m$ konstruieren. Sei also P ein Pfad von u nach v der Länge $m > |V|$. Dann muss es in P mindestens einen Knoten w geben, der mehr als einmal besucht wird. Das heißt, wenn etwa $P = v_0 - v_1 - \cdots - v_m$ mit $v_0 = u$ und $v_m = v$ ist, dann muss es ein Paar (i, j) mit $i < j$ geben, so dass $v_i = v_j = w$ ist. In diesem Fall ist

   $$P' := v_0 - \cdots - v_{i-1} - v_i - v_{j+1} - v_{j+2} - \cdots - v_m$$

   ein Pfad von u nach v der Länge $m - (j - i) < m$. Beachte, dass $(v_i, v_{j+1}) = (v_j, v_{j+1})$ eine Kante von G ist, weil sie bereits im ursprünglichen Pfad P vorkommt. ∎

Beispiel.

1. Betrachten wir noch einmal den Graphen G_2 von Seite 130. Für $n = 7$ erhalten wir

   $$A_2^7 = \begin{pmatrix} 3 & 2 & 0 & 4 & 0 \\ 2 & 1 & 0 & 2 & 0 \\ 4 & 4 & 1 & 6 & 0 \\ 2 & 2 & 0 & 3 & 0 \\ 1 & 1 & 0 & 2 & 0 \end{pmatrix}.$$

   Der Eintrag 6 an Position $(3, 4)$ sagt aus, dass es genau 6 Möglichkeiten gibt, in genau sieben Schritten von c nach d zu gelangen. In der Tat sind das die folgenden Pfade:

   $$c{-}b{-}d{-}a{-}b{-}d{-}a{-}d,$$
   $$c{-}b{-}d{-}a{-}b{-}a{-}b{-}d,$$
   $$c{-}c{-}c{-}b{-}d{-}a{-}b{-}d,$$
   $$c{-}c{-}c{-}c{-}b{-}d{-}a{-}d,$$
   $$c{-}c{-}c{-}c{-}c{-}c{-}a{-}d.$$

2. Für den Graphen G_3 auf Seite 130 gilt $A_3^4 = 0$. Es gibt in diesem Graphen also keine Pfade der Länge 4. Längere Pfade kann es dann natürlich auch nicht geben. Die Gesamtzahl sämtlicher Pfade (beliebiger Länge) von einem Knoten zu einem anderen ist deshalb gegeben durch

$$A_3 + A_3^2 + A_3^3 = \begin{pmatrix} 0 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

Man kann sich überlegen, dass ein Graph $G = (V, E)$ mit Adjazenzmatrix A genau dann einen Zyklus enthält wenn $A^{|V|} \neq 0$ ist.

3. Statt der Anzahl der Pfade von u nach v kann man sich auch fragen, wie lang der kürzeste Pfad von u nach v ist. Man nennt die Länge eines kürzesten Pfades von u nach v die *Distanz* von u nach v, Schreibweise: $d(u, v)$. (Beachte: im allgemeinen gilt $d(u, v) \neq d(v, u)$.) Wenn es gar keinen Pfad von u nach v gibt, setzt man $d(u, v) = \infty$.

Wir wollen für alle Knotenpaare u, v die Distanz $d(u, v)$ bestimmen. Dazu ist es hilfreich, eine Hilfsgröße

$$d_m(u, v) = \begin{cases} d(u, v) & \text{falls } d(u, v) \leq m \\ \infty & \text{sonst} \end{cases}$$

einzuführen. Dann bedeutet $d_m(u, v) = \infty$, dass $d(u, v) > m$ ist, und $d_m(u, v) = k \in \mathbb{N}$, dass $d(u, v) = k$ ist. Nach Teil 3 von Satz 65 gilt für je zwei Knoten u, v eines Graphen $G = (V, E)$ stets $d(u, v) \leq |V|$ oder $d(u, v) = \infty$. Es genügt also, $d_{|V|}(u, v)$ zu berechnen.

Ist $V = \{v_1, \ldots, v_n\}$, so gilt offenbar

$$d_{m_1+m_2}(u, v) = \min\big(d_{m_1}(u, v_1) + d_{m_2}(v_1, v), \ldots,\ d_{m_1}(u, v_n) + d_{m_2}(v_n, v)\big),$$

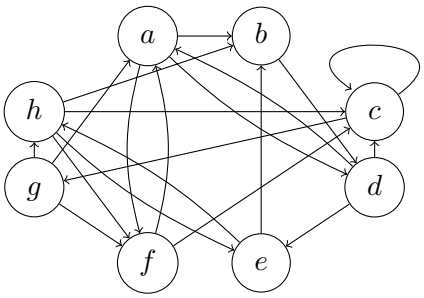
denn wenn man auf einem kürzesten Pfad P von u nach v nach m_1 Schritten an einem Knoten v_i anhält, dann kann es nicht sein, dass es von u nach v_i einen Pfad mit weniger als m_1 Schritten gibt, sonst ließe sich daraus ein Pfad von u nach v konstruieren, der noch kürzer ist als P. (Wir nehmen hier an, dass min und + auf $\mathbb{N} \cup \{\infty\}$ so definiert sind, dass $\min(\infty, \infty) = \infty$, $\min(k, \infty) = k$ und $k + \infty = \infty + k = \infty + \infty = \infty$ für alle $k \in \mathbb{N}$ gilt.)

Die obige Formel für $d_{m_1+m_2}(u, v)$ hat die gleiche Struktur wie die Formel für die Koeffizienten eines Matrixprodukts, wenn man + und · durch min und + ersetzt. Man kann also eine Entfernungstabelle für einen Graphen G wie folgt berechnen. Wähle $v\colon \{1, \ldots, |V|\} \to V$ und setze $D_1 = ((\, d_1(v(i), v(j))\,))_{i,j=0}^{n}$. Beachte dabei, dass

$$d_1(v(i), v(j)) = \begin{cases} 0 & \text{falls } i = j \\ 1 & \text{falls } i \neq j \text{ und } (v(i), v(j)) \in E \\ \infty & \text{sonst} \end{cases}$$

für alle i, j gilt. Berechne die Matrix-Potenz $D := D_1^{|V|+1}$, wobei min und + statt + und · verwendet werden. Der (i, j)-te Eintrag von D ist dann genau $d(v(i), v(j))$.

Beispiel: Für den Graphen



erhält man die folgende Entfernungstabelle:

	a	b	c	d	e	f	g	h
a	0	1	2	1	2	1	3	3
b	2	0	2	1	2	3	3	3
c	2	3	0	3	3	2	1	2
d	1	2	1	0	1	2	2	2
e	3	1	2	2	0	2	3	1
f	1	2	1	2	3	0	2	3
g	1	2	2	2	2	1	0	1
h	2	1	1	2	1	1	2	0

22 C-finite Folgen

Definition 47.

1. Eine Funktion $a\colon \mathbb{N} \to \mathbb{K}$ heißt *Folge* (engl. *sequence*) in $\mathbb{K}$. Statt a schreibt man auch $(a_n)_{n=0}^{\infty}$.

2. Eine Folge a heißt *C-finit,* falls es einen Vektor $(c_0, \ldots, c_r) \in \mathbb{K}^{r+1} \setminus \{0\}$ gibt, so dass für alle $n \in \mathbb{N}$ gilt

$$c_0 a(n) + c_1 a(n+1) + \cdots + c_r a(n+r) = 0.$$

Beispiel.

1. Die Folge $(F_n)_{n=0}^{\infty}$ der Fibonacci-Zahlen ist C-finit. Sie erfüllt die Rekurrenz

$$F_n + F_{n+1} - F_{n+2} = 0$$

   für alle $n \in \mathbb{N}$.

2. Die Folge $a\colon \mathbb{N} \to \mathbb{Q}$ mit $a(n) = n^2$ ist C-finit. Eine Rekurrenz lautet

$$a(n) - 3a(n+1) + 3a(n+2) - a(n+3) = 0.$$

3. Die Folge $a\colon \mathbb{N} \to \mathbb{Q}$ mit $a(n) = n!$ ist nicht C-finit. Um das zu zeigen, muss man zeigen, dass sie keine lineare Rekurrenz mit konstanten Koeffizienten erfüllt, egal wie groß die Ordnung r gewählt wird. Anders ausgedrückt ist zu zeigen, dass die Folgen $(n!)_{n=0}^{\infty}$, $((n+1)!)_{n=0}^{\infty}$, …, $((n+r)!)_{n=0}^{\infty}$ aufgefasst als Elemente des Vektorraums $\mathbb{Q}^{\mathbb{N}}$ linear unabhängig über $\mathbb{Q}$ sind. Dazu betrachte man Konstanten $c_0, \dots, c_r \in \mathbb{Q}$ mit

$$c_0 n! + c_1 (n+1)! + \cdots + c_r (n+r)! = 0$$

für alle $n \in \mathbb{N}$. Es ist zu zeigen, dass all diese Konstanten Null sein müssen. Wäre das nicht so, dann gäbe es zumindest ein i mit $c_i \neq 0$. O.B.d.A. können wir $i = r$ annehmen. Division durch $(n+r)!$ liefert die Gleichung

$$c_0 \frac{1}{(n+1)\cdots(n+r)} + c_1 \frac{1}{(n+2)\cdots(n+r)} + \cdots + c_{r-1} \frac{1}{n+r} + c_r = 0$$

für alle $n \in \mathbb{N}$. Auf der linken Seite steht nun eine Linearkombination von rationalen Ausdrücken. Aus der Analysis ist bekannt, dass $\lim_{n\to\infty} \frac{1}{p(n)} = 0$ ist für jedes Polynom p mit $\deg p > 0$. Der Grenzwert der linken Seite für $n \to \infty$ ist deshalb c_r. Der Grenzwert der rechten Seite ist aber 0. Also muss $c_r = 0$ sein, im Widerspruch zur Annahme.

Satz 66. Sei $(c_0, \dots, c_r) \in \mathbb{K}^{r+1}$ mit $c_r \neq 0$. Dann gilt:

1. $U := \{\, (a_n)_{n=0}^{\infty} \in \mathbb{K}^{\mathbb{N}} : \forall\ n \in \mathbb{N} : c_0 a_n + \cdots + c_r a_{n+r} = 0 \,\}$ ist ein Untervektorraum von $\mathbb{K}^{\mathbb{N}}$.

2. Die Abbildung
$$h\colon U \to \mathbb{K}^r, \quad h(a_0, a_1, \dots) := (a_0, \dots, a_{r-1})$$
ist ein Isomorphismus.

Beweis.

1. Zu zeigen: $0 \in U$ und für alle $u, v \in U$ und alle $\alpha, \beta \in \mathbb{K}$ gilt $\alpha u + \beta v \in U$. Dass 0 in U liegt, folgt aus
$$c_0 0 + c_1 0 + \cdots + c_r 0 = 0.$$
Sind $u, v \in U$, so gilt
$$\begin{aligned} c_0 u(n) + c_1 u(n+1) + \cdots + c_r u(n+r) &= 0, \\ c_0 v(n) + c_1 v(n+1) + \cdots + c_r v(n+r) &= 0 \end{aligned}$$
für alle $n \in \mathbb{N}$. Multiplikation der ersten Gleichung mit $\alpha \in \mathbb{K}$ und der zweiten mit $\beta \in \mathbb{K}$, und Addition der beiden resultierenden Gleichungen liefert
$$c_0\big(u(n) + v(n)\big) + c_1\big(u(n+1) + v(n+1)\big) + \cdots c_r\big(u(n+r) + v(n+r)\big) = 0$$
für alle $n \in \mathbb{N}$. Also ist auch $u + v \in U$.

2. Man überzeugt sich leicht, dass h ein Homomorphismus ist. Um zu zeigen, dass h bijektiv ist, genügt es zu zeigen, dass es für jeden Vektor $(a_0, \dots, a_{r-1}) \in \mathbb{K}^r$ genau eine Folge $a \in U$ mit $h(a) = (a_0, \dots, a_{r-1})$ gibt. Zunächst muss für jede solche Folge a offensichtlich zumindest $a(n) = a_n$ für $n = 0, \dots, r-1$ gelten. Ferner gilt für beliebiges $n \geq r$, dass es wegen der Rekurrenz für den Wert von $a(n)$ genau eine Möglichkeit gibt, nämlich

$$a(n) = -\frac{1}{c_r}\big(c_0 a(n-r) + c_1 a(n-r+1) + \cdots + c_{r-1} a(n-1)\big).$$

Durch Induktion nach n erhält man, dass alle Terme der Folge a eindeutig festgelegt sind. ∎

Beispiel. Die Lösungsmenge der Rekurrenz

$$f(n) + f(n+1) - f(n+2) = 0$$

ist ein zwei-dimensionaler Vektorraum. Jede Lösung ist eindeutig festgelegt durch ihre beiden Anfangswerte $f(0)$ und $f(1)$.

Insbesondere gibt es eine Lösung f_1 mit $f_1(0) = 0$ und $f_1(1) = 1$ und eine Lösung f_2 mit $f_2(0) = 1$ und $f_2(1) = 0$. Die Folge f_1 ist die Folge der Fibonacci-Zahlen. Diese beiden Lösungen sind linear unabhängig, weil die Vektoren $h(f_1) = (0, 1)$ und $h(f_2) = (1, 0)$ linear unabhängig sind. Es ist also $\{f_1, f_2\}$ eine Basis von U.

Alle anderen Lösungen f lassen sich damit als Linearkombination $\alpha f_1 + \beta f_2$ schreiben. Die ersten Terme von f ergeben sich zu

$$\beta,\ \alpha,\ \alpha+\beta,\ \alpha+2\beta,\ 2\alpha+3\beta,\ 3\alpha+5\beta,\ 5\alpha+8\beta, \dots$$

Satz 67. Es sei $a\colon \mathbb{N} \to \mathbb{K}$ eine C-finite Folge, die eine Rekurrenz der Ordnung r erfüllt. Dann gilt: Sind $c_0, \dots, c_r \in \mathbb{K}$ so, dass

$$c_0 a(n) + c_1 a(n+1) + \cdots + c_r a(n+r) = 0$$

für $n = 0, \dots, r-1$, so gilt

$$c_0 a(n) + c_1 a(n+1) + \cdots + c_r a(n+r) = 0$$

für alle $n \in \mathbb{N}$.

Beweis. Nach Annahme über a existieren gewisse Konstanten $\tilde{c}_0, \dots, \tilde{c}_r \in \mathbb{K}$, von denen wenigstens eine von Null verschieden ist, und für die gilt

$$\tilde{c}_0 a(n) + \tilde{c}_0 a(n+1) + \cdots + \tilde{c}_r a(n+r) = 0$$

für alle $n \in \mathbb{N}$. Betrachte den Vektorraum $U \subseteq \mathbb{K}^{\mathbb{N}}$ bestehend aus allen Folgen f mit

$$\tilde{c}_0 f(n) + \tilde{c}_0 f(n+1) + \cdots + \tilde{c}_r f(n+r) = 0$$

für alle $n \in \mathbb{N}$. Nach Satz 66 gilt $\dim U \leq r$. Ferner gilt neben $a \in U$ auch $(a(n+i))_{n\geq 0} \in U$ für jedes fest gewählte $i \in \mathbb{N}$, denn wenn

$$\tilde{c}_0 a(n) + \tilde{c}_0 a(n+1) + \cdots + \tilde{c}_r a(n+r) = 0$$

für alle $n \in \mathbb{N}$ gilt, dann gilt auch

$$\tilde{c}_0 a(n+i) + \tilde{c}_0 a(n+i+1) + \cdots + \tilde{c}_r a(n+i+r) = 0$$

für alle $n, i \in \mathbb{N}$. Da U ein Vektorraum ist, gehört auch die Linearkombination

$$b(n) := c_0 a(n) + c_1 a(n+1) + \cdots + c_r a(n+r)$$

zu U.

Nach Voraussetzung gilt nun $b(0) = b(1) = \cdots = b(r-1) = 0$. Aus Teil 2 von Satz 66 folgt, dass $b(n) = 0$ für alle $n \in \mathbb{N}$. ∎

Beispiel. Satz 67 besagt, dass man die Koeffizienten der Rekurrenz für eine C-finite Folge a rekonstruieren kann, wenn man die Ordnung r der Rekurrenz sowie die Terme $a(0), \ldots, a(2r-1)$ kennt. Man braucht dazu bloß ein lineares Gleichungssystem zu lösen.

Ist etwa bekannt, dass $a\colon \mathbb{N} \to \mathbb{Q}$ eine Rekurrenz der Ordnung zwei erfüllt, so sucht man Konstanten $c_0, c_1, c_2 \in \mathbb{Q}$ mit

$$c_0 a(n) + c_1 a(n+1) + c_2 a(n+2) = 0$$

für alle $n \in \mathbb{N}$. Nach dem Satz genügt schon, dass diese Bedingung für $n = 0, 1$ erfüllt ist. Wenn man außerdem weiß, dass a mit den Termen 0, 1, 5, 19 beginnt, dann konkretisiert sich die obige Gleichung für $n = 0, 1$ zu

$$\begin{pmatrix} 0 & 1 & 5 \\ 1 & 5 & 19 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = 0.$$

Der Lösungsraum dieses Gleichungssystems wird aufgespannt von $(6, -5, 1)$, und also erfüllt a nach Satz 67 die Rekurrenz

$$6a(n) - 5a(n+1) + a(n+2) = 0.$$

Wenn man die Ordnung r der Rekurrenz von a nicht kennt, dann kann man immer noch probieren, ob man für ein frei gewähltes r aus den bekannten Termen von a eine Rekurrenz rekonstruieren kann. Man macht dazu wie im vorherigen Beispiel einen Ansatz für die Koeffizienten und stellt ein lineares Gleichungssystem auf. Man sollte dabei r klein genug wählen, dass man aus den bekannten Termen von a ein überbestimmtes Gleichungssystem konstruieren kann (also eines mit mehr Gleichungen als Variablen). Solch ein System hat typischerweise keine Lösung außer 0. Wenn man trotzdem eine Lösung bekommt, dann ist das ein Indiz (aber natürlich noch kein Beweis), dass man die Koeffizienten der wahren Rekurrenz von a gefunden hat.

Beispiel. Betrachte die Folge $a\colon \mathbb{N} \to \mathbb{Q}$ mit $a(n) = n\,F_n$, wobei F_n die $n$te Fibonacci-Zahl ist. Die ersten 10 Terme von a lauten $0, 1, 2, 6, 12, 25, 48, 91, 168, 306, 550$. Ist a C-finit?

Versuchen wir es mit $r = 3$. Wenn

$$c_0 a(n) + c_1 a(n+1) + c_2 a(n+2) + c_3 a(n+3) = 0$$

für alle $n \in \mathbb{N}$ gelten soll, dann muss es auf jeden Fall auch für $n = 0, \ldots, 10 - 3$ gelten. Das führt auf das Gleichungssystem

$$\begin{pmatrix} 0 & 1 & 2 & 6 \\ 1 & 2 & 6 & 12 \\ 2 & 6 & 12 & 25 \\ 6 & 12 & 25 & 48 \\ 12 & 25 & 48 & 91 \\ 25 & 48 & 91 & 168 \\ 48 & 91 & 168 & 306 \\ 91 & 168 & 306 & 550 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = 0.$$

Diese Gleichungssystem hat nur die triviale Lösung $(0, 0, 0, 0)$. Daraus folgt, dass a ganz sicher keine Rekurrenz der Ordnung drei (oder kleiner) erfüllt.

Es folgt aber nicht, dass a nicht C-finit ist. Es kann immer noch eine Rekurrenz höherer Ordnung erfüllen. Zum Beispiel $r = 4$. In diesem Fall bekommt man mit den gleichen Daten das lineare Gleichungssystem

$$\begin{pmatrix} 0 & 1 & 2 & 6 & 12 \\ 1 & 2 & 6 & 12 & 25 \\ 2 & 6 & 12 & 25 & 48 \\ 6 & 12 & 25 & 48 & 91 \\ 12 & 25 & 48 & 91 & 168 \\ 25 & 48 & 91 & 168 & 306 \\ 48 & 91 & 168 & 306 & 550 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0.$$

Obwohl es sich wieder um ein überbestimmtes System handelt, gibt es eine nichttriviale Lösung, nämlich $(1, 2, -1, -2, 1)$. Das legt die Vermutung nah, dass a die Rekurrenz

$$a(n) + 2a(n+1) - a(n+2) - 2a(n+3) + a(n+4) = 0$$

erfüllt.

In der Tat ist nicht schwer zu zeigen, dass diese Vermutung korrekt ist, indem man $n\,F_n$ in die Rekurrenz einsetzt und den entstehenden Ausdruck mithilfe der bekannten Rekurrenz $F_{n+2} = F_{n+1} + F_n$ zu 0 vereinfacht.

Satz 68. Es seien $(a_n)_{n=0}^{\infty}$ und $(b_n)_{n=0}^{\infty}$ zwei C-finite Folgen. Dann gilt:

1. $(a_n + b_n)_{n=0}^{\infty}$ ist C-finit.
2. $(a_n b_n)_{n=0}^{\infty}$ ist C-finit.
3. für jedes fixe $u \in \mathbb{N}$ ist $(a_{un})_{n=0}^{\infty}$ C-finit.
4. $(\sum_{k=0}^{n} a_k)_{n=0}^{\infty}$ ist C-finit.

Beweis. Wir zeigen den zweiten Teil. Die Beweise für die anderen Teile gehen ähnlich.

Nach Voraussetzung sind $(a_n)_{n=0}^{\infty}$ und $(b_n)_{n=0}^{\infty}$ C-finit. Es gibt also $(\alpha_0, \ldots, \alpha_r) \in \mathbb{K}^{r+1} \setminus \{0\}$ und $(\beta_0, \ldots, \beta_s) \in \mathbb{K}^{s+1} \setminus \{0\}$ mit

$$\begin{aligned} \alpha_0 a_n + \alpha_1 a_{n+1} + \cdots + \alpha_r a_{n+r} &= 0 \\ \beta_0 b_n + \beta_1 b_{n+1} + \cdots + \beta_s b_{n+s} &= 0. \end{aligned}$$

Durch geeignete Wahl von r und s können wir o.B.d.A. annehmen, dass $\alpha_r \neq 0$ und $\beta_s \neq 0$ ist. Aus den Rekurrenzen folgt dann, dass sich die Folgen $(a_{n+r})_{n=0}^{\infty}$ und $(b_{n+s})_{n=0}^{\infty}$ als Linearkombinationen der Folgen $(a_n)_{n=0}^{\infty}, \ldots, (a_{n+r-1})_{n=0}^{\infty}$ bzw. $(b_n)_{n=0}^{\infty}, \ldots, (b_{n+s-1})_{n=0}^{\infty}$ darstellen lassen. Allgemeiner folgt durch Induktion, dass sich für jedes fixe $i \in \mathbb{N}$ sogar die Folgen $(a_{n+i})_{n=0}^{\infty}$ und $(b_{n+i})_{n=0}^{\infty}$ als Linearkombinationen dieser Folgen darstellen lassen.

Für den Vektorraum $V \subseteq \mathbb{K}^{\mathbb{N}}$, der von den Folgen $(a_{n+i}b_{n+j})_{n=0}^{\infty}$ mit $i, j \in \mathbb{N}$ erzeugt wird, gilt deshalb $\dim V \leq rs$; als Erzeugendensystem genügen nämlich die Folgen $(a_{n+i}b_{n+j})_{n=0}^{\infty}$ mit $0 \leq i < r$ und $0 \leq j < s$.

Wegen $\dim V \leq rs$ müssen je $rs+1$ viele Elemente von V linear abhängig sein. Insbesondere muss es eine lineare Abhängigkeit zwischen den Folgen $(a_n b_n)_{n=0}^{\infty}, \ldots, (a_{n+rs}b_{n+rs})_{n=0}^{\infty}$ geben. Es gibt also Konstanten $\gamma_0, \ldots, \gamma_{rs} \in \mathbb{K}$, von denen mindestens eine von Null verschieden ist, und für die gilt

$$\gamma_0 a_n b_n + \gamma_1 a_{n+1} b_{n+1} + \cdots + \gamma_s a_{n+rs} b_{n+rs} = 0$$

für alle $n \in \mathbb{N}$. ∎

Beispiel.

1. Sei $a = (F_n)_{n=0}^{\infty}$ die Folge der Fibonacci-Zahlen und $b = (n)_{n=0}^{\infty}$. Es sei bekannt, dass a die Rekurrenz $a(n) + a(n+1) - a(n+2) = 0$ und b die Rekurrenz $b(n) - 2b(n+1) + b(n+2) = 0$ erfüllt. Gesucht sei eine Rekurrenz für $c = (nF_n)_{n=0}^{\infty}$.

   Es gilt

   $$\begin{aligned} a(n+2) &= a(n) + a(n+1) & b(n+2) &= -b(n) + 2b(n+1) \\ a(n+3) &= a(n) + 2a(n+1) & b(n+3) &= -2b(n) + 3b(n+1) \\ a(n+4) &= 2a(n) + 3a(n+1) & b(n+4) &= -3b(n) + 4b(n+1) \end{aligned}$$

   für alle $n \in \mathbb{N}$. Für die Produkte ergibt sich daraus

   $$\begin{aligned} a(n)b(n) &= a(n)b(n) \\ a(n+1)b(n+1) &= a(n+1)b(n+1) \\ a(n+2)b(n+2) &= -a(n)b(n) - a(n+1)b(n) + 2a(n)b(n+1) + 2a(n+1)b(n+1) \\ a(n+3)b(n+3) &= -2a(n)b(n) - 4a(n+1)b(n) + 3a(n)b(n+1) + 6a(n+1)b(n+1) \\ a(n+4)b(n+4) &= -6a(n)b(n) - 9a(n+1)b(n) + 8a(n)b(n+1) + 12a(n+1)b(n+1). \end{aligned}$$

   Die Koeffizienten der Rekurrenz findet man durch Lösen des linearen Gleichungssystems

   $$\begin{pmatrix} 1 & 0 & -1 & -2 & -6 \\ 0 & 0 & -1 & -4 & -9 \\ 0 & 0 & 2 & 3 & 8 \\ 0 & 1 & 2 & 6 & 12 \end{pmatrix} \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix} = 0.$$

Dass das System eine Lösung haben muss, sieht man sofort daran, dass es mehr Variablen als Gleichungen hat. Und tatsächlich stellt sich heraus, dass der Lösungsraum eindimensional ist und vom Vektor $(1, 2, -1, -2, 1)$ aufgespannt wird. Daraus folgt, dass die Folge $c = (nF_n)_{n=0}^{\infty}$ die Rekurrenz

$$c(n) + 2c(n+1) - c(n+2) - 2c(n+3) + c(n+4) = 0$$

erfüllt.

2. Auf der Grundlage von Satz 68 kann man ein Computerprogramm schreiben, das Identitäten für C-finite Folgen automatisch beweist. Um zum Beispiel die Summationsformel

$$\sum_{k=0}^{n} F_k^2 = F_n F_{n+1}$$

zu beweisen, beginnt das Programm mit der Rekurrenz $F_n + F_{n+1} - F_{n+2} = 0$ für die Fibonacci-Zahlen, die entweder vom Benutzer eingegeben oder aus einer Datenbank gelesen wird. Dann berechnet es wie im vorherigen Beispiel je eine Rekurrenz für $(F_n^2)_{n=0}^{\infty}$ und für $(F_n F_{n+1})_{n=0}^{\infty}$. Aus der Rekurrenz für $(F_n^2)_{n=0}^{\infty}$ lässt sich als nächstes eine Rekurrenz für die Summe $(\sum_{k=0}^{n} F_k^2)_{n=0}^{\infty}$ berechnen, und daraus schließlich eine Rekurrenz für die Differenz $d(n) := (\sum_{k=0}^{n} F_k^2) - F_n F_{n+1}$ von linker und rechter Seite. Das Ergebnis dieser Rechnung lautet

$$d(n) - 2d(n+1) - 2d(n+2) + d(n+3) = 0$$

für alle $n \in \mathbb{N}$. Um zu zeigen, dass $d(n) = 0$ für alle $n \in \mathbb{N}$ gilt, braucht man nach Teil 2 von Satz 66 nur noch nachzurechnen, dass $d(n) = 0$ für $n = 0, 1, 2$ gilt.

23 Kodierungstheorie

Digitale Informationsverarbeitung beruht auf dem Prinzip, jede mögliche Information als eine Folge von Nullen und Einsen darzustellen. Diese Folgen sind immer endlich, aber möglicherweise sehr lang. Die Datei dieses Skriptums ist zum Beispiel mehrere Megabyte groß. Das bedeutet, sie wird im Computer dargestellt als eine Folge von mehreren Millionen Vektoren $b \in \{0, 1\}^8$. Einen solchen Vektor nennt man ein *Byte*. Die Koordinaten eines Bytes nennt man *Bit*.

Ein Bit kann je nach Kontext zum Beispiel Informationen wie an/aus, schwarz/weiss, 0/1, wahr/falsch, rechts/links, männlich/weiblich codieren. In einem Byte kann man z. B. die natürlichen Zahlen von 0 bis 255 codieren:

$$\begin{aligned}
(0,0,0,0,0,0,0,0) &= 0,\\
(0,0,0,0,0,0,0,1) &= 1,\\
(0,0,0,0,0,0,1,0) &= 2,\\
(0,0,0,0,0,0,1,1) &= 3,\\
(0,0,0,0,0,1,0,0) &= 4, \text{ etc.}
\end{aligned}$$

Auch die Zeichen in einem Text kann man mit Bytes codieren, indem man z. B. durch eine Tabelle festlegt, welches Byte welchem Buchstaben entsprechen soll. Die 256 verschiedenen Bytes sind mehr als ausreichend, um die 52 Groß- und Kleinbuchstaben, die 10 Ziffern, das Leerzeichen und einige Interpunktions- und Sonderzeichen unterzubringen.

Mit Folgen von Bytes läßt sich quasi alles codieren: Zahlen, Formeln, Texte, Computerprogramme, geometrische Objekte, Farben, Kundendaten, Fahrpläne, Musik, Fotos, Videos, Wetter- und Klimamodelle, usw. In der Kodierungstheorie beschäftigt man sich mit dem Design solcher Codierungen. Allgemeiner stellt man sich die Frage, wie man eine gegebene Codierung (d. h. eine gegebene Bitfolge) in eine andere umwandeln kann, die zum Übertragen oder Speichern der Information besser geeignet ist. Typische Fragestellungen sind in diesem Zusammenhang:

- *Datenkompression:* Wie kann ich eine gegebene Bitfolge in eine kürzere Bitfolge umwandeln, die die gleiche Information enthält? Oder: Kann ich aus einer gegebenen Bitfolge die „wesentliche" Information von der „unwesentlichen" trennen, so dass ich nur die wesentliche speichern oder übertragen muss?
- *Fehlererkennung:* Wie kann ich eine gegebene Bitfolge so abändern, dass sich nach der Übertragung feststellen lässt, ob ein Teil der Bitfolge fehlerhaft übertragen wurde? Lassen sich die Fehler nachträglich reparieren?
- *Kryptographie:* Wie kann ich eine gegebene Bitfolge verschlüsseln, d. h. so abändern, dass die darin codierte Information ohne geeignetes geheimes Zusatzwissen nicht zu rekonstruieren ist. Oder: wie kann ich eine gegebene Bitfolge signieren, d. h. so abändern, dass ein Empfänger verifizieren kann, dass die Nachricht wirklich von mir stammt?

Es gibt zu jedem dieser Punkte zahlreiche Techniken. Manche davon verwenden lineare Algebra. Als Beispiel erklären wir hier eine einfache Methode zur Fehlererkennung. Für Verallgemeinerungen und andere Methoden sei auf entsprechende Lehrveranstaltungen oder die Literatur verwiesen.

Definition 48.

1. Das *Gewicht* eines Vektors $a = (a_1, \ldots, a_n) \in \mathbb{Z}_2^n$ ist definiert als
$$w(a) := |\{\, i : a_i \neq 0 \,\}| \in \{0, \ldots, n\}.$$
2. Für $a = (a_1, \ldots, a_n), b = (b_1, \ldots, b_n) \in \mathbb{Z}_2^n$ heißt $d(a, b) := w(a - b)$ die *Hamming-Distanz* von a und b.
3. Ein Untervektorraum $U \subseteq \mathbb{Z}_2^n$ der Dimension k heißt *linearer Code* von Länge n und Rang k. Elemente von U heißen *Codewörter*. Die *Distanz* eines linearen Codes $U \subseteq \mathbb{Z}_2^n$ ist definiert als
$$d = \min\{\, w(u) : u \in U \setminus \{0\} \,\} \in \mathbb{N}.$$
4. Ist $U \subseteq \mathbb{Z}_2^n$ ein Code, dann heißt eine geordnete Basis $A \in \mathbb{Z}_2^{n \times k}$ von U eine *Erzeugermatrix* und eine Matrix $B \in \mathbb{Z}_2^{(n-k) \times n}$ mit $\ker B = U$ eine *Prüfmatrix* für U.

Beispiel.

1. $w(1, 0, 0, 1, 1, 0, 1, 0) = 4$, $w(0, 0, 1, 0, 1, 1, 1, 1) = 5$, $w(0, 1, 1, 0, 1, 0, 1, 1) = 5$, etc.

   $d(a, b)$ ist die Anzahl der Positionen, an denen sich a und b voneinander unterscheiden, zum Beispiel:
$$d\big((1, 1, 0, 1, 0, 1, 1),\ (1, 1, 0, 1, 0, 0, 1)\big) = 1$$
$$d\big((1, 1, 0, 1, 0, 1, 1),\ (1, 1, 0, 1, 1, 0, 1)\big) = 2$$
$$d\big((1, 1, 0, 1, 0, 1, 1),\ (0, 1, 0, 1, 1, 0, 1)\big) = 3$$

$$d\big((1,1,0,1,0,1,1),\ (0,1,1,1,1,0,1)\big) = 4$$
$$d\big((1,1,0,1,0,1,1),\ (0,1,1,0,1,0,1)\big) = 5$$

Es gilt $d(a,b) = 0$ genau dann, wenn $a = b$ ist.

2. Der Unterraum $U \subseteq \mathbb{Z}_2^4$, der von $u_1 = (1,0,1,1)$ und $u_2 = (0,1,1,0)$ erzeugt wird, besteht aus den folgenden Codewörtern:

$$0u_1 + 0u_2 = (0,0,0,0) \qquad 0u_1 + 1u_2 = (0,1,1,0)$$
$$1u_1 + 0u_2 = (1,0,1,1) \qquad 1u_1 + 1u_2 = (1,1,0,1).$$

Die Distanz dieses Codes ist 2.

Eine Erzeugermatrix für U ist $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$.

Eine Prüfmatrix für U ist $B = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$.

Angenommen, man will k Bits über einen Kanal übertragen. Rauschen auf dem Kanal kann jedes übertragene Bit mit einer bestimmten (kleinen) Wahrscheinlichkeit ändern. Wir wollen die k Bits so codieren, dass der Empfänger eine korrekte Übertragung von einer fehlerhaften unterscheiden kann, solange in der Übertragung weniger als d Fehler auftreten.

Man wählt dazu einen Code $U \in \mathbb{Z}_2^n$ vom Rang k und Distanz d. Statt des Codeworts $v \in \mathbb{Z}_2^k$ überträgt man $u := Av \in \mathbb{Z}_2^n$, wobei A eine Erzeugermatrix von U ist. Beim Empfänger kommt ein Vektor $\tilde{u} \in \mathbb{Z}_2^n$ an, den er sich vorstellt als eine Überlagerung $\tilde{u} = u + r$ des (ihm unbekannten) korrekten Codeworts $u \in \mathbb{Z}_2^n$ und einem (ihm ebenfalls unbekannten) Bitvektor $r \in \mathbb{Z}_2^n$, der dem Effekt des Rauschens entspricht.

Wenn $\tilde{u} \in U$ gilt, dann ist auch $r = \tilde{u} - u \in U$ und $w(r) \geq d$ oder $r = 0$, nach der Definiton der Distanz von U. Wenn der Empfänger also ein Codewort $\tilde{u}$ enthält, dann ist entweder die Übertragung fehlerfrei gewesen oder es ist zu mindestens d Fehlern gekommen. Ob $\tilde{u}$ ein Codewort ist, kann der Empfänger mit einer Prüfmatrix B für U überprüfen: dies ist genau dann der Fall, wenn $B\tilde{u} = 0$ gilt.

Beispiel. Betrachte den Code $H \in \mathbb{Z}_2^7$ mit

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}$$

als Erzeuger- bzw. Prüfmatrix. Dieser Code heißt *Hamming-Code.*

Die Länge ist $n = 7$, der Rang ist $k = 4$. Die Distanz ist nicht direkt offensichtlich, aber da es nur $2^k = 16$ Codewörter gibt, kann man leicht alle auflisten und sich vergewissern, dass $d = 3$ ist.

Bei der Übertragung zweier Codewörter u_1, u_2 durch einen verrauschten Kanal wurden die beiden Wörter $\tilde{u}_1 = (1,0,1,1,0,1,0)$ und $\tilde{u}_2 = (1,1,0,1,1,1,0)$ empfangen.

Wegen $B\tilde{u}_1 = (0,0,0)$ und $B\tilde{u}_2 = (1,1,1)$ wurde das zweite nicht fehlerfrei übertragen. Bei der Übertragung des ersten Worts gab es entweder keinen Fehler oder mehr als zwei.

Das zweite Wort enthält ein oder mehrere falsche Bits. Wenn wir annehmen, dass nur ein Bit verfälscht wurde, lässt sich das korrekte Codewort u_2 aus $\tilde{u}_2$ rekonstruieren. In diesem Fall ist $\tilde{u}_2 = u_2 + e_i$ für einen Einheitsvektor e_i. Wir haben $B\tilde{u}_2 = B(u_2 + e_i) = Bu_2 + Be_i = Be_i$. Da Be_i die i-te Spalte von B ist und $B\tilde{u}_2 = (1,1,1)$ gilt, muss $i = 4$ sein. Das gesendete Wort war also entweder $u_2 = (1,1,0,\mathbf{0},1,1,0)$, oder es gab bei der Übertragung mehr als einen Fehler.

Satz 69. Sei $U \subseteq \mathbb{Z}_2^n$ ein Code der Länge n mit Rang k und Distanz d. Dann gilt $k+d \leq n+1$.

Beweis. Wegen $|\mathbb{Z}_2| = 2$ und $\dim U = k$ gilt $|U| = 2^k$.

Nach Definition der Distanz gilt $w(u) \geq d$ für alle $u \in U \setminus \{0\}$. Für je zwei verschiedene Codewörter $u, v \in U$ gilt dann auch $d(u,v) = w(u-v) \geq d$. Wenn also $u = (u_1, \ldots, u_n)$ und $v = (v_1, \ldots, v_n)$ ist, dann muss auch $(u_d, \ldots, u_n) \neq (v_d, \ldots, v_n)$ gelten. Da es höchstens 2^{n-d+1} viele verschiedene Bitvektoren der Länge $n-d+1$ gibt, muss also $|U| \leq 2^{n-d+1}$ sein. Daraus folgt $2^k \leq 2^{n-d+1}$, und daraus $k \leq n-d+1$. ∎

Die Distanz d eines Codes sagt aus, wie viele Fehler auftreten dürfen, ohne dass dem Empfänger die Fehlerhaftigkeit des empfangenen Wortes entgehen kann. Die Differenz $n-k$ gibt an, wie viele zusätzliche Bits der Code einsetzt, um diese Fehlererkennung zu gewährleisten. In der Praxis möchte man d möglichst groß und n (bzw. $n-k$) möglichst klein haben. Der obige Satz setzt diesem Zielkonflikt eine Grenze.

Der Hamming-Code erreicht diese Grenze nicht: $k+d = 3+2 = 5 < 8 = 7+1 = n+1$. Tatsächlich gibt es überhaupt keine (brauchbaren) Codes, für die $k+d = n+1$ gilt, zumindest solange man über dem Körper $\mathbb{Z}_2$ arbeitet. Codes, die für Speichermedien oder in Mobilfunknetzen verwendet werden, verwenden daher Körper $\mathbb{K}$ mit z. B. $2^8 = 256$ Elementen. Für solche Körper kann man (brauchbare) lineare Codes konstruieren, für die $k+d = n+1$ gilt. Ein Beispiel sind die sogenannten Reed-Solomon-Codes, deren Erklärung an dieser Stelle allerdings zu weit führen würde.

24 Lineare Algebra in Maple, Mathematica und Sage

Mit Vektoren und Matrizen von Hand zu rechnen ist mühsam, zweitaufwendig, fehleranfällig – und zum Glück heute nicht mehr nötig. Es gibt eine ganze Reihe von Computerprogrammen, die diese Arbeit bequem, schnell, und zuverlässig übernehmen. Grob einteilen lassen sich diese Programme in *numerische* und *symbolische* Software. Numerische Software ist auf den Fall spezialisiert, dass der Grundkörper $\mathbb{R}$ oder $\mathbb{C}$ ist. In diesen Körpern kann man im allgemeinen nur approximativ rechnen. Das hat sowohl theoretische Gründe (eine reelle Zahl ist erst dann eindeutig festgelegt, wenn man all ihre potentiell unendlich vielen Nachkommastellen kennt – die passen aber nicht in den endlichen Speicher eines Computers) als auch praktische (wenn es sich um Messwerte einer physikalischen Anwendung handelt, wird man von einer gewissen Messungenauigkeit ausgehen müssen). Wenn man mit gerundeten Zahlen rechnet, stellt sich die Frage, wieviel Genauigkeit man durch die Rechnung verliert, d. h. wie viele Nachkommastellen des Ergebnisses korrekt sind. Man kann sich dann auch fragen, ob eine andere Rechnung, die mathematisch äquivalent ist, eventuell eine bessere Genauigkeit liefert. Damit wollen wir uns hier nicht beschäftigen. Antworten auf Fragen dieser Art und einen Überblick über die entsprechenden Softwaresysteme bekommen Sie in den Lehrveranstaltungen des Instituts für Numerik.

Symbolische Software ist auf die Fälle spezialisiert, wo man die Elemente des Grundkörpers exakt darstellen und ohne Genauigkeitsverlust mit ihnen rechnen kann. Das ist insbesondere der Fall für den

Körper der rationalen Zahlen und für endliche Körper, aber zum Beispiel auch für den Körper $\mathbb{Q}(X)$ der rationalen Funktionen über $\mathbb{Q}$. Wir geben hier für drei solche Systeme einige Beispiele, wie man einfache Rechnungen mit Vektoren und Matrizen in diesen Systemen durchführt. Man darf sich vorstellen, dass intern das gleiche passiert, was auch bei einer Handrechnung passieren würde, auch wenn das nicht unbedingt der Fall ist. Was diese Systeme wirklich tun, werden wir am Ende des zweiten Semesters besprechen. Mehr dazu erfahren Sie am Institut für Symbolisches Rechnen oder am Institut für Algebra.

Beispiel.

1. *Maple.* Die Funktionalität für das Rechnen mit Vektoren und Matrizen wird in Maple in einem Paket LinearAlgebra zusammengefasst, das man zunächst laden muss, wenn man Funktionen aus diesem Paket verwenden will. Danach hat man Funktionen zur Verfügung, mit denen man Vektoren und Matrizen erzeugen, auf deren Komponenten zugreifen, sie addieren und multiplizieren und Gleichungssysteme lösen kann. Es ist zu beachten, dass Maple zwischen Zeilen und Spaltenvektoren unterscheidet.

   with(LinearAlgebra) :
   $v :=$ Vector$([7, 4, -3])$;

   $$\begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

   $v[2]$

   $$4$$

   $5 \cdot v$

   $$\begin{bmatrix} 35 \\ 20 \\ -15 \end{bmatrix}$$

   $v +$ Vector$([1, 2, 3])$

   $$\begin{bmatrix} 8 \\ 6 \\ 0 \end{bmatrix}$$

   $A :=$ Matrix$([[1, 2, 3], [4, 5, 6], [7, 8, 9]])$

   $$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

   $A[2, 3]$

   $$6$$

   Multiply(A, v)

   $$\begin{bmatrix} 8 \\ 26 \\ 44 \end{bmatrix}$$

   Multiply(v, A)
   `Error, (in Multiply) cannot multiply a column Vector and a Matrix`

   $vt :=$ Transpose(v);

$$\begin{bmatrix} 7 & -4 & 3 \end{bmatrix}$$

Multiply(vt, A)

$$\begin{bmatrix} 12 & 18 & 24 \end{bmatrix}$$

$B :=$ Matrix$([[1,1,1],[1,2,3],[1,4,9]])$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix}$$

Multiply(A, B)

$$\begin{bmatrix} 6 & 17 & 34 \\ 15 & 38 & 73 \\ 24 & 59 & 112 \end{bmatrix}$$

$A\hat{}5$

$$\begin{bmatrix} 121824 & 149688 & 177552 \\ 275886 & 338985 & 402084 \\ 429948 & 528282 & 626616 \end{bmatrix}$$

$B\hat{}(-1)$

$$\begin{bmatrix} 3 & -\frac{5}{2} & \frac{1}{2} \\ -3 & 4 & -1 \\ 1 & -\frac{3}{2} & \frac{1}{2} \end{bmatrix}$$

NullSpace(A)

$$\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}$$

LinearSolve(B, v)

$$\begin{bmatrix} \frac{19}{2} \\ -2 \\ -\frac{1}{2} \end{bmatrix}$$

Standardmäßig werden Zahlen von Maple als rationale Zahlen interpretiert. Man kann aber auch mit Vektoren und Matrizen über einem endlichen Körper $\mathbb{Z}_p$ mit $p \in \mathbb{Z}$ prim rechnen. Funktionen dafür gibt es im Unterpaket LinearAlgebra/Modular.

Modular[Multiply]$(17, A, B)$

$$\begin{bmatrix} 6 & 0 & 0 \\ 15 & 4 & 5 \\ 7 & 8 & 10 \end{bmatrix}$$

Modular[Inverse]$(17, B)$

$$\begin{bmatrix} 3 & 6 & 9 \\ 14 & 4 & 16 \\ 1 & 7 & 9 \end{bmatrix}$$

2. *Mathematica.* Vektoren sind in Mathematica einfach die Listen ihrer Koordinaten. Eine Liste wird in Mathematica mit geschweiften Klammern geschrieben. Es handelt sich aber wirklich um Listen, nicht um Mengen, d. h. Reihenfolge und Vielfachheit von Elementen machen einen Unterschied. Vektoren kann man miteinander addieren oder mit einer Konstanten multiplizieren.

```
In[1]:= {1, 2, 3} == {1, 2, 3, 3}
Out[1]= False
In[2]:= {1, 2, 3} == {1, 3, 2}
Out[2]= False
In[3]:= v = {7, 4, -3}
Out[3]= {7, 4, -3}
In[4]:= Length[v]
Out[4]= 3
In[5]:= v[[2]]
Out[5]= 4
In[6]:= 5{1, 2, 3}
Out[6]= {5, 10, 15}
In[7]:= {1, 2, 3} + v
Out[7]= {8, 6, 0}
```

Matrizen werden in Mathematica als Listen von (Zeilen-)Vektoren dargestellt. Man kann Matrizen miteinander addieren und multiplizieren, oder mit Zahlen oder Vektoren multiplizieren. Außerdem gibt es Funktionen zum Transponieren, Invertieren, und zum Lösen von Gleichungssystemen.

```
In[8]:= A = {{1, 2, 3}, {4, 5, 6}, {7, 8, 9}}
Out[8]= {{1, 2, 3}, {4, 5, 6}, {7, 8, 9}}
In[9]:= A[[2, 3]]
Out[9]= 6
In[10]:= A.v
Out[10]= {6, 30, 54}
In[11]:= v.A
Out[11]= {2, 10, 18}
In[12]:= Transpose[A]
Out[12]= {{1, 4, 7}, {2, 5, 8}, {3, 6, 9}}
In[13]:= MatrixRank[A]
```

Out[13]= 2

In[14]:= **Nullspace[*A*]**

Out[14]= $\{\{1, -2, 1\}\}$

In[15]:= ***B* = {{1, 1, 1}, {1, 2, 3}, {1, 4, 9}}**

Out[15]= $\{\{1, 1, 1\}, \{1, 2, 3\}, \{1, 4, 9\}\}$

In[16]:= **2*A* − *B***

Out[16]= $\{\{1, 3, 5\}, \{7, 8, 9\}, \{13, 12, 9\}\}$

In[17]:= ***A*.*B***

Out[17]= $\{\{6, 17, 34\}, \{15, 38, 73\}, \{24, 59, 112\}\}$

In[18]:= **LinearSolve[*B*, *v*]**

Out[18]= $\{\frac{19}{2}, -2, -\frac{1}{2}\}$

In[19]:= ***B*.% − *v***

Out[19]= $\{0, 0, 0\}$

In[20]:= **Inverse[*B*]**

Out[20]= $\{\{3, -\frac{5}{2}, \frac{1}{2}\}, \{-3, 4, -1\}, \{1, -\frac{3}{2}, \frac{1}{2}\}\}$

Der Punkt beim Multiplizieren ist wesentlich: $A.B$ bedeutet Matrix-Matrix oder Matrix-Vektor-Multiplikation. Schreibt man $A * B$ oder AB, so werden die Einträge komponentenweise miteinander multipliziert – oder es gibt eine Fehlermeldung, falls das Format nicht passt. Vorsicht auch beim Potenzieren: Wenn man A^5 eingibt, nimmt Mathematica alle Einträge der Matrix einzeln hoch fünf, aber nicht die gesamte Matrix im Sinn der Matrixmultiplikation. Um das zu tun, verwendet man die Funktion **MatrixPower**

In[21]:= ***A* * *B* == *A*.*B***

Out[21]= False

In[22]:= ***A*^5**

Out[22]= $\{\{1, 32, 243\}, \{1024, 3125, 7776\}, \{16807, 32768, 59049\}\}$

In[23]:= **MatrixPower[*A*, 5]**

Out[23]= $\{\{121824, 149688, 177552\}, \{275886, 338985, 402084\}, \{429948, 528282, 626616\}\}$

Standardmäßig interpretiert Mathematica Zahlen als rationale Zahlen. Will man stattdessen in einem endlichen Körper $\mathbb{Z}_p$ mit $p \in \mathbb{Z}$ prim rechnen, muss man den Modulus p der jeweiligen Funktion als optionales Argument mitgeben.

In[24]:= **Inverse[*B*, Modulus → 17]**

Out[24]= $\{\{3, 6, 9\}, \{14, 4, 16\}, \{1, 7, 9\}\}$

In[25]:= **LinearSolve[*B*, *v*, Modulus → 17]**

Out[25]= $\{1, 15, 8\}$

In[26]:= $\boldsymbol{B.\% - v}$

Out[26]= $\{17, 51, 136\}$

In[27]:= $\mathbf{Mod[\%, 17]}$

Out[27]= $\{0, 0, 0\}$

3. *Sage.* Dieses System basiert auf der Programmiersprache Python. Korrekter Python-Code ist (fast) immer auch korrekter Sage-Code. Darüber hinaus hat Sage einige Sprachkonstrukte, mit denen man mathematische Objekte beschreiben kann. Jedes solche Sage-Objekt ist entweder ein „Parent“ – dabei handelt es sich zum Beispiel um Gruppen, Ringe, Körper, Vektorräume – oder ein „Element“ – das sind Objekte, die zu einem Parent gehören.

2.parent()

Integer Ring

(1/3).parent()

Rational Field

$v = \text{vector}(\text{QQ}, [7, 4, -3])$

v

$(7, 4, -3)$

v.parent()

Vector space of dimension 3 over Rational Field

$v[0]$

7

$A = \text{matrix}(\text{QQ}, [[1, 2, 3], [4, 5, 6], [7, 8, 9]])$

$B = \text{matrix}(\text{QQ}, [[1, 1, 1], [1, 2, 3], [1, 4, 9]])$

A.parent()

Full MatrixSpace of 3 by 3 dense matrices over Rational Field

$A[1, 2]$

6

$v * A$

$(2, 10, 18)$

$A * v$

$(6, 30, 54)$

$2 * A - B$

$$\begin{pmatrix} 1 & 3 & 5 \\ 7 & 8 & 9 \\ 13 & 12 & 9 \end{pmatrix}$$

```
A.transpose()
```

$$\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$$

```
A.right_kernel()
```

Vector space of degree 3 and dimension 1 over Rational Field
Basis matrix:
(1 −2 1)

```
A.rank()
```

2

```
A * B
```

$$\begin{pmatrix} 6 & 17 & 34 \\ 15 & 38 & 73 \\ 24 & 59 & 112 \end{pmatrix}$$

```
A^5
```

$$\begin{pmatrix} 121824 & 149688 & 177552 \\ 275886 & 338985 & 402084 \\ 429948 & 528282 & 626616 \end{pmatrix}$$

```
B^(-1)
```

$$\begin{pmatrix} 3 & -5/2 & 1/2 \\ -3 & 4 & -1 \\ 1 & -3/2 & 1/2 \end{pmatrix}$$

```
B.solve_right(v)
```

$(19/2, -2, -1/2)$

Um über einem endlichen Körper zu rechnen, muss man bei der Erzeugung der Vektoren und Matrizen GF(p) für $\mathbb{Z}_p$ ($p \in \mathbb{Z}$ prim) statt QQ für $\mathbb{Q}$ angeben. Die Kommandos für die Rechnungen selbst ändern sich nicht.

```
v = vector(GF(17), [7, 4, -3])
```

```
A = matrix(GF(17), [[1, 2, 3], [4, 5, 6], [7, 8, 9]])
```

```
A.v
```

$(6, 13, 3)$

```
A.right_kernel()
```

Vector space of degree 3 and dimension 1 over Finite Field of size 17
Basis matrix:
(1 15 1)

Systeme wie Maple, Mathematica und Sage stellen tausende mathematische Funktionen bereit. Allein die Funktionen, die etwas mit linearer Algebra zu tun haben, sind zu zahlreich, um sie alle hier zu erwähnen. Einen vollständigen Überblick finden Sie in den Dokumentationen dieser Systeme.

Teil V

Eigenwerte

25 Polynome

Zur Erinnerung: Ist $\mathbb{K}$ ein Körper, so bezeichnet man mit $\mathbb{K}[X]$ den Ring aller Polynome mit Koeffizienten in $\mathbb{K}$. Jedes Polynom hat die Form $p = p_0 + p_1 X + \cdots + p_d X^d$ für gewisse $p_0, \ldots, p_d \in \mathbb{K}$. Ist $p_d \neq 0$, so ist $\deg(p) := d$ der Grad des Polynoms. Für das Nullpolynom definiert man $\deg(0) = -\infty$. Man nennt $[x^i]p := p_i$ den i-ten Koeffizienten von p und insbesondere $\mathrm{lc}(p) := p_d$ den *Leitkoeffizienten* (engl. *leading coefficient*) von p.

Satz 70. Seien $a, b \in \mathbb{K}[X]$ mit $b \neq 0$. Dann gibt es genau ein Paar $(q, r) \in \mathbb{K}[X]^2$ mit $a = qb + r$ und $\deg(r) < \deg(b)$.

Beweis. Existenz: Zunächst ist klar, dass es Paare $(q, r) \in \mathbb{K}[X]^2$ mit $a = qb + r$ gibt, z.B. $(q, r) = (0, a)$. Betrachte von all diesen Paaren ein Paar (q, r), für das $\deg(r)$ minimal ist. Wäre $\deg(r) \geq \deg(b)$, so wäre (q', r') mit

$$r' = r - \frac{\mathrm{lc}(r)}{\mathrm{lc}(b)} x^{\deg r - \deg b} b, \quad q' = q + \frac{\mathrm{lc}(r)}{\mathrm{lc}(b)} x^{\deg r - \deg b}$$

ein Paar mit $q', r' \in \mathbb{K}[X]$ und $a = q'b + r'$ und $\deg(r') < \deg(r)$, im Widerspruch zur angenommenen Minimalität von $\deg(r)$.

Eindeutigkeit: Seien $(q, r), (q', r') \in \mathbb{K}[X]^2$ zwei Paare mit $a = qb + r = q'b + r'$ und $\deg(r), \deg(r') < \deg(b)$. Dann gilt $(q - q')b = r' - r$. Wäre $q - q' \neq 0$, so wäre $(q - q')b \neq 0$ und $\deg(q - q')b \geq \deg b$, während $\deg(r' - r) \leq \max\{\deg(r), \deg(r')\} < \deg(b)$. Da das nicht sein kann, muss $q - q' = 0$, d.h. $q = q'$ gelten. Dann aber folgt $0 = r' - r$ und also $r = r'$. ∎

Beispiel. Die Konstruktion aus dem Beweis des Satzes lässt sich als Rechenvorschrift zur Berechnung des Paars (q, r) für ein gegebenes Paar $(a, b) \in \mathbb{K}[X]^2$ interpretieren. Man bezeichnet diesen Prozess als *Polynomendivision*. Für $a = 3X^4 - 4X^2 + 8X - 1$ und $b = X^2 + X + 2$ erhält man zum Beispiel

$$\overbrace{3X^4 + 0X^3 - 4X^2 + 8X - 1}^{a} = \overbrace{(X^2 + X + 2)}^{b} \underbrace{(3X^2 - 3X - 7)}_{q} + \underbrace{(21X + 13)}_{r}$$

$$\begin{array}{r}
3X^4 + 3X^3 + 6X^2 \\
\hline
-3X^3 - 10X^2 + 8X - 1 \\
-3X^3 - 3X^2 - 6X \\
\hline
-7X^2 + 14X - 1 \\
-7X^2 - 7X - 14 \\
\hline
21X + 13
\end{array}$$

Bei der Berechnung arbeitet man sich in absteigender Reihenfolge durch die Koeffizienten von a. Zunächst teilt man den höchsten Term $3X^4$ von a durch den höchsten Term X^2 von b. Das Ergebnis $3X^4/X^2 = 3X^2$ ist der höchste Term von q. Nachdem man diesen Term notiert hat, schreibt man das Produkt dieses Terms mit b unter a; in diesem Fall $3X^4 + 3X^3 + 6X^2$. Dann zieht man dieses Polynom von a ab und erhält $-3X^3 - 10X^2 + 8X - 1$. Dessen höchsten Term $-3X^3$ teilt man durch den höchsten Term X^2 von b, um den nächsten Term von q zu erhalten, das ist in unserem Fall $-3X$. Das Verfahren lässt sich fortsetzen, bis eine Differenz entsteht, deren Grad kleiner ist als $\deg(b)$. Diese Differenz ist dann r.

Definition 49. Seien $a, b \in \mathbb{K}[X]$, $b \neq 0$.

1. Es sei $(q, r) \in \mathbb{K}[X]^2$ das Paar aus Satz 70. Dann heißen $\operatorname{quo}(a, b) := q$ und $\operatorname{rem}(a, b) := r$ 03-14
der *Quotient* und der *Rest* (engl. *remainder*) der Division von a durch b.

2. b heißt *Teiler* (engl. *divisor*) oder *Faktor* von a, falls $\operatorname{rem}(a, b) = 0$. In diesem Fall schreibt man $b \mid a$, anderenfalls $b \nmid a$. Ist $e \in \mathbb{N}$ so, dass $b^e \mid a$ und $b^{e+1} \nmid a$, so heißt e die *Vielfachheit* (engl. *multiplicity*) des Teilers b von a.

3. Ist $b = X - \alpha$ für ein $\alpha \in \mathbb{K}$ und gilt $b \mid a$, so heißt α eine *Nullstelle* (engl. *root*) von a. Die Vielfachheit der Nullstelle α ist die Vielfachheit des Faktors $X - \alpha$.

Beispiel.

1. Es gilt $\operatorname{quo}(3X^4 - 4X^2 + 8X - 1, X^2 + X + 2) = 3X^2 - 3X - 7$ und $\operatorname{rem}(3X^4 - 4X^2 + 8X - 1, X^2 + X + 2) = 21X + 13$. Da der Rest nicht Null ist, gilt also $X^2 + X + 2 \nmid 3X^4 - 4X^2 + 8X - 1$.

2. Für ein beliebiges Polynom $a = a_0 + a_1 X + \cdots + a_n X^n$ und $\alpha \in \mathbb{K}$ gilt

$$\operatorname{rem}(a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0, X - \alpha) = a_n \alpha^n + a_{n-1} \alpha^{n-1} + \cdots + a_0.$$

Dass $\alpha \in \mathbb{K}$ eine Nullstelle von a ist, heißt also, dass die zu a gehörende Polynomfunktion 03-03

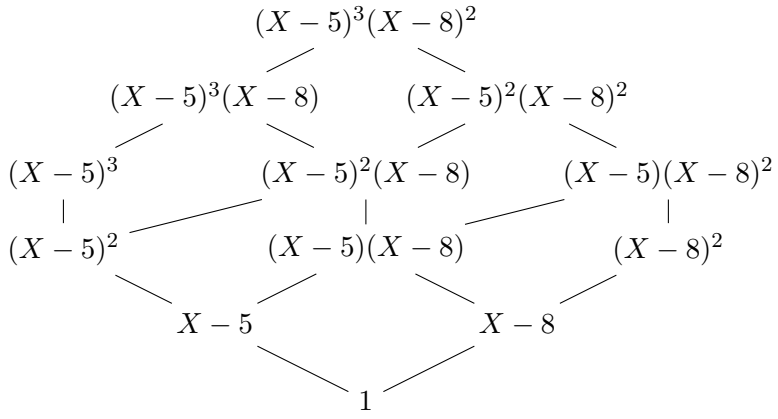
$$\mathbb{K} \to \mathbb{K}, \quad x \mapsto a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$$

an der Stelle α den Wert 0 hat.

3. $a = (X - 2)^4 (X - 3)^2 (X + 5)$ hat die drei Nullstellen $2, 3, -5$. Die Nullstelle 2 hat die Vielfachheit 4, die Nullstelle 3 hat die Vielfachheit 2 („3 ist eine doppelte Nullstelle“), und die Nullstelle -5 hat die Vielfachheit 1 („-5 ist eine einfache Nullstelle“).

4. Ob ein Polynom Nullstellen hat, hängt im allgemeinen davon ab, welchen Körper man zugrunde legt. Zum Beispiel hat $a = X^2 - 2$ als Element von $\mathbb{Q}[X]$ keine Nullstellen, aber als Element von $\mathbb{R}[X]$ schon (nämlich $\sqrt{2}$ und $-\sqrt{2}$).

5. Teilbarkeit ist eine transitive Relation auf $\mathbb{K}[X] \setminus \{0\}$, d.h. aus $u \mid v$ und $v \mid w$ folgt $u \mid w$. Für zwei Polynome $u, v \in \mathbb{K}[X] \setminus \{0\}$ gilt sowohl $u \mid v$ als auch $v \mid u$ genau dann, wenn es ein $\alpha \in \mathbb{K} \setminus \{0\}$ gibt mit $u = \alpha v$. (Beweis: Übung.) Man sagt dann, die Polynome u, v sind (zueinander) *assoziiert*. Jedes Polynom $a \in \mathbb{K}[X] \setminus \{0\}$ hat nur endlich viele nicht zueinander assoziierte Teiler. Die Teiler von $a = (X - 5)^3 (X - 8)^2$ sind im folgenden dargestellt:



Für fixes $p \in \mathbb{K}[X] \setminus \{0\}$ ist die Abbildung $R\colon \mathbb{K}[X] \to \mathbb{K}[X]$ mit $R(u) := \operatorname{rem}(u, p)$ ein Vektorraumhomomorphismus, d.h. es gilt $R(\alpha u + \beta v) = \alpha R(u) + \beta R(v)$ für alle $\alpha, \beta \in \mathbb{K}$ und $u, v \in \mathbb{K}[X]$. Es gilt

$$\ker R = p\mathbb{K}[X] = \{\, u \in \mathbb{K}[X] : p \mid u \,\}.$$

Eine Basis des Quotientenraums $\mathbb{K}[X]/p\mathbb{K}[X]$ ist $\{[1]_\sim, [X]_\sim, \ldots, [X^{\deg(p)-1}]_\sim\}$.

Der Vektorraum $\mathbb{K}[X]/p\mathbb{K}[X]$ wird mit den Verknüpfungen

$$[u]_\sim + [v]_\sim := [u+v]_\sim, \qquad [u]_\sim \cdot [v]_\sim := [uv]_\sim$$

zu einem Ring. Diese Definitionen sind repräsentantenunabhängig, denn wenn z.B. $u \sim u'$ und $v \sim v'$ gilt, so gilt $uv \sim uv + (u'-u)v = u'v \sim u'v + u'(v'-v) = u'v'$.

Man beachte die Ähnlichkeit zur Konstruktion endlicher Ringe $\mathbb{Z}_m$. In der Tat lässt sich vieles, was in diesem Abschnitt über den Ring $\mathbb{K}[X]$ gesagt wird, in ähnlicher Weise für den Ring $\mathbb{Z}$ formulieren. Dem Grad eines Polynoms entspricht in $\mathbb{Z}$ der Absolutbetrag. So gilt zum Beispiel, dass es zu jeder Wahl von $a, b \in \mathbb{Z}$ mit $b \neq 0$ genau ein Paar $(q, r) \in \mathbb{Z}^2$ gibt mit $a = bq + r$ und $|r| < |b|$. Der Begriff der Nullstelle eines Polynoms hat dagegen für ganze Zahlen keine Entsprechung.

Definition 50. Seien $p, q, g \in \mathbb{K}[X]$.

1. g heißt ein *gemeinsamer Teiler* (engl. *common divisor*) von p und q, falls es $u, v \in \mathbb{K}[X]$ gibt mit $gu = p$ und $gv = q$.
2. g heißt ein *größter gemeinsamer Teiler* (engl. *greatest common divisor*) von p und q, falls g ein gemeinsamer Teiler von p und q ist und für jeden anderen gemeinsamen Teiler $\tilde{g}$ von p und q gilt $\tilde{g} \mid g$.

Beispiel. $X-8$ ist ein gemeinsamer Teiler von $(X-5)^2(X-8)$ und $(X-5)(X-8)^2$. Es ist aber kein größter gemeinsamer Teiler. Ein größter gemeinsamer Teiler von $(X-5)^2(X-8)$ und $(X-5)(X-8)^2$ ist $(X-5)(X-8)$. Ein anderer größter gemeinsamer Teiler ist $23(X-5)(X-8)$.

Satz 71. Seien $p, q \in \mathbb{K}[X]$. Dann gilt:

1. Es gibt einen größten gemeinsamen Teiler g von p und q.
2. Zu je zwei größten gemeinsamen Teilern $g, \tilde{g}$ von p und q gibt es ein $\alpha \in \mathbb{K} \setminus \{0\}$ mit $g = \alpha\tilde{g}$.

Beweis.

1. Für Polynome $a, b \in \mathbb{K}[X]$ bezeichne $T(a,b)$ die Menge aller gemeinsamen Teiler von a und b, d.h. $T(a,b) = \{\, u \in \mathbb{K}[X] \setminus \{0\} : u \mid a, u \mid b \,\}$. Wir zeigen, dass $T(p,q)$ bezüglich der Relation $\mid$ ein maximales Element hat (und insbesondere nicht leer ist).

   Man überzeugt sich leicht, dass für je drei Polynome $a, b, u \in \mathbb{K}[X]$ gilt $T(a,b) = T(b,a)$ und $T(a,b) = T(a, b+ua)$. Außerdem ist leicht einzusehen, dass a ein maximales Element von $T(a,0)$ ist.

   Aus den ersten beiden Beobachtungen folgt, dass $T(p,q) = T(q, \operatorname{rem}(p,q))$ gilt. Definiert man $r = \operatorname{rem}(p,q)$, $r' = \operatorname{rem}(q,r)$, $r'' = \operatorname{rem}(r,r')$, $\ldots$, so gilt $\deg(r) > \deg(r') > \deg(r'') > \cdots$. Da die Grade eines Polynoms natürliche Zahlen sind, muss diese Folge irgendwann abbrechen. Das kann sie nur, indem nach endlich vielen Schritten das Nullpolynom auftaucht, denn für jedes andere Polynom ließe sich die Folge nach Satz 70 durch einen weiteren Divisionsschritt fortsetzen.

   Ist $r^{(k)}$ das letzte von Null verschiedene Polynom der Folge, so gilt also $T(p,q) = T(r^{(k)}, 0)$. Damit ist $r^{(k)}$ ein maximales Element von $T(p,q)$.

2. Sind $g, \tilde{g}$ zwei größte gemeinsame Teiler, so gilt $g \mid \tilde{g}$ und $\tilde{g} \mid g$. Also gibt es Polynome $u, \tilde{u} \in \mathbb{K}[X]$ mit $g = \tilde{g}\tilde{u}$ und $\tilde{g} = gu$. Ist $g = 0$, so ist auch $\tilde{g} = 0$ und wir können $\alpha = 1$ 03-03
   wählen. Wenn $g \neq 0$ ist, folgt $u \neq 0$ und $\tilde{u} \neq 0$, also $\deg(u), \deg(\tilde{u}) \geq 0$. Außerdem gilt dann $\deg(g) \geq 0$ und damit wegen $g = gu\tilde{u}$ auch $\deg(g) = \deg(g) + \deg(u) + \deg(\tilde{u})$, also $\deg(u) + \deg(\tilde{u}) = 0$. Diese Summe kann nur dann Null sein, wenn $\deg(u) = \deg(\tilde{u}) = 0$ 04-26
   ist, und dies ist gleichbedeutend mit $u, \tilde{u} \in \mathbb{K} \setminus \{0\}$. Wir können also $\alpha = \tilde{u}$ wählen.

Beispiel. Der Algorithmus im Beweis von Teil 1 heißt *euklidischer Algorithmus*. Für

$$\begin{aligned} p &= X^6 - 3X^5 + 3X^4 - 11X^3 + 14X^2 - 23X - 3 \\ q &= X^3 - X^2 - 3X - 9 \end{aligned}$$

erhält man 03-04

$$\begin{aligned} r &= \operatorname{rem}(p,q) = 4X^2 + X - 39 \\ r' &= \operatorname{rem}(q,r) = \frac{113}{16}X - \frac{339}{16} \\ r'' &= \operatorname{rem}(r,r') = 0. \end{aligned}$$

Also ist $r' = \frac{113}{16}X - \frac{339}{16}$ ein größter gemeisamer Teiler von p und q. Da der größte gemeinsame Teiler nur bis auf Multiplikation mit einem von Null verschiedenen Körperelement bestimmt ist, kann man mit $\frac{16}{113}$ multiplizieren und erhält, dass $X - 3$ ein größter gemeinsamer Teiler von p und q ist.

Man kann übrigens jedes der Polynome r, r', r'' vor dem Weiterrechnen mit einem Element von $\mathbb{K} \setminus \{0\}$ multiplizieren, ohne die Korrektheit des Algorithmus zu beeinträchtigen. Im Fall $\mathbb{K} = \mathbb{Q}$ ist das auch dringend empfohlen, weil sonst im Laufe der Rechnung sehr schnell sehr lange Brüche auftreten, die keine besondere Bedeutung haben und nur die Rechnung verlangsamen.

Definition 51.

1. Ein Polynom $p \in \mathbb{K}[X] \setminus \{0\}$ heißt *normiert* (engl. *monic*), falls $\mathrm{lc}(p) = 1$ gilt.
2. Für $p, q \in \mathbb{K}[X]$ mit $p \neq 0$ oder $q \neq 0$ sei $\gcd(p, q)$ der größte gemeisame Teiler g von p und q mit $\mathrm{lc}(g) = 1$. Im Fall $p = q = 0$ definiert man $\gcd(0, 0) = 0$.
3. Zwei Polynome $p, q \in \mathbb{K}[X]$ heißen *teilerfremd* (engl. *coprime*), falls $\gcd(p, q) = 1$ ist.

Der Begriff des größten gemeinsamen Teilers lässt sich analog für $\mathbb{Z}$ statt $\mathbb{K}[X]$ formulieren. Insbesondere übertragen sich Satz 71 und der euklidische Algorithmus. Es gibt aber zwischen dem gcd in $\mathbb{K}[X]$ und dem in $\mathbb{Z}$ keinen direkten Zusammenhang, sondern es handelt sich nur um eine analoge Begriffsbildung.

Insbesondere gilt im allgemeinen **nicht,** dass für Polynome $p, q \in \mathbb{Z}[X] \subseteq \mathbb{Q}[X]$ die Polynomfunktion von $\gcd(p, q)$ ausgewertet an $n = 1, 2, 3, \ldots$ gerade die größten gemeinsamen Teiler von $p(n)$ und $q(n)$ in $\mathbb{Z}$ liefert. Für $p = X + 3$ und $q = 2X - 1$ gilt zum Beispiel $\gcd(p, q) = 1$ in $\mathbb{Q}[X]$, aber $\gcd(p(4), q(4)) = \gcd(7, 7) = 7$ in $\mathbb{Z}$.

Satz 72. Für alle $p, q \in \mathbb{K}[X]$ existieren $u, v \in \mathbb{K}[X]$ mit $\gcd(p, q) = up + vq$.

Beweis. Im Fall $p = 0$ oder $q = 0$ ist die Aussage trivial. Betrachte also den Fall $p \neq 0 \neq q$. Nach dem euklidischen Algorithmus (im Beweis von Satz 71) taucht $\gcd(p, q)$ in der Folge
03-08 $r_0 = p$, $r_1 = q$, $r_k = \mathrm{rem}(r_{k-2}, r_{k-1})$ $(k \geq 2)$ auf. Es genügt also zu zeigen, dass für jedes k Polynome $u_k, v_k \in \mathbb{K}[X]$ mit $r_k = u_k p + v_k q$ existieren.

Induktion nach k.

Induktionsanfang: Für $k = 0$ und $k = 1$ können wir $u_0 = 1$, $v_0 = 0$, $u_1 = 0$, $v_1 = 1$ wählen.

Induktionsschritt $k - 2, k - 1 \to k$. Nach Annahme gibt es $u_{k-2}, v_{k-2}, u_{k-1}, v_{k-1} \in \mathbb{K}[X]$ mit $r_{k-2} = u_{k-2}p + v_{k-2}q$ und $r_{k-1} = u_{k-1}p + v_{k-1}q$. Nach Definition gilt

$$\begin{aligned} r_k &= \mathrm{rem}(r_{k-2}, r_{k-1}) \\ &= r_{k-2} - \mathrm{quo}(r_{k-2}, r_{k-1}) r_{k-1} \\ &= (u_{k-2}p + v_{k-2}q) - \mathrm{quo}(r_{k-2}, r_{k-1})(u_{k-1}p + v_{k-1}q) \\ &= \underbrace{\big(u_{k-2} - \mathrm{quo}(r_{k-2}, r_{k-1}) u_{k-1}\big)}_{\in \mathbb{K}[X]} p + \underbrace{\big(v_{k-2} - \mathrm{quo}(r_{k-2}, r_{k-1}) v_{k-1}\big)}_{\in \mathbb{K}[X]} q, \end{aligned}$$

wie gefordert. ∎

Beispiel. Betrachte $a = 3X^3 + 4X^2 - X + 2, b = X^2 + 2X + 3 \in \mathbb{Q}[X]$. Aus der Rechnung
03-03

k	$q_k = \mathrm{quo}(r_{k-2}, r_{k-1})$	$r_k = r_{k-2} - q_k r_{k-1}$	$u_k = u_{k-2} - q_k u_{k-1}$	$v_k = v_{k-2} - q_k v_{k-1}$
0	–	$3X^3 + 4X^2 - X + 2$	1	0
1	–	$X^2 + 2X + 3$	0	1
2	$3X - 2$	$-6X + 8$	1	$-3X + 2$
3	$-\frac{1}{6}X - \frac{5}{9}$	$\frac{67}{9}$	$\frac{1}{6}X + \frac{5}{9}$	$-\frac{1}{2}X^2 - \frac{4}{3}X + \frac{19}{9}$
4	$\frac{9}{67}(-6X + 8)$	0	–	–

folgt

$$\frac{67}{9} = \Big(\frac{1}{6}X + \frac{5}{9}\Big)a + \Big(-\frac{1}{2}X^2 - \frac{4}{3}X + \frac{19}{9}\Big)b,$$

bzw. nach Normierung

$$\gcd(a,b) = 1 = \underbrace{\frac{9}{67}\Big(\frac{1}{6}X + \frac{5}{9}\Big)}_{=u}\, a + \underbrace{\frac{9}{67}\Big(-\frac{1}{2}X^2 - \frac{4}{3}X + \frac{19}{9}\Big)}_{=v}\, b.$$

Definition 52. Ein Polynom $p \in \mathbb{K}[X] \setminus \{0\}$ heißt *irreduzibel,* falls gilt:

$$\forall\ a, b \in \mathbb{K}[X] : p = ab \Rightarrow a \in \mathbb{K} \vee b \in \mathbb{K}.$$

Beispiel.

1. Alle Polynome vom Grad 1 sind irreduzibel. Solche Polynome nennt man auch „linear“, obwohl die zugehörigen Polynomfunktionen im allgemeinen nicht linear im Sinn von Def. 34 sind.

2. $X^2 - 2 \in \mathbb{Q}[X]$ ist irreduzibel, aber $X^2 - 2 \in \mathbb{R}[X]$ ist nicht irreduzibel, da $X^2 - 2 = (X - \sqrt{2})(X + \sqrt{2})$ und $X - \sqrt{2}, X + \sqrt{2} \in \mathbb{R}[X]$.

3. $X^2 + 1$ ist irreduzibel als Polynom in $\mathbb{Q}[X]$ und $\mathbb{R}[X]$, aber nicht als Polynom in $\mathbb{C}[X]$, da $X^2 + 1 = (X + \mathrm{i})(X - \mathrm{i})$.

4. $X^2 - 2$ ist irreduzibel in $\mathbb{Z}_5[X]$, denn gäbe es $a, b \in \mathbb{Z}_5$ mit $X^2 - 2 = (X + a)(X + b) = X^2 + (a + b)X + ab$, so müsste $a + b = 0$ und $ab = -2$, also $a = -b$ und $b^2 = 2$ sein. Es gilt aber $0^2 = 0$, $1^2 = 1$, $2^2 = 4$, $3^2 = 4$, $4^2 = 1$ in $\mathbb{Z}_5$, und weitere Möglichkeiten für b gibt es nicht. 05-05

   Jedoch ist $X^2 - 2$ nicht irreduzibel als Element von $\mathbb{Z}_7[X]$, denn hier gilt $(X+3)(X+4) = X^2 + (3+4)X + 12 = X^2 + 0X + (-2) = X^2 - 2$. 05-05 05-05

Satz 73. Sei $p \in \mathbb{K}[X] \setminus \{0\}$. Dann sind folgende Aussagen äquivalent:

1. p ist irreduzibel.

2. Der Ring $\mathbb{K}[X]/p\mathbb{K}[X]$ ist ein Körper.

3. $\forall\ a, b \in \mathbb{K}[X] : p \mid ab \Rightarrow p \mid a \vee p \mid b$

Beweis.

(1)⇒(2). Zu zeigen: Für jedes $u \in \mathbb{K}[X]/p\mathbb{K}[X] \setminus \{0\}$ existiert ein $v \in \mathbb{K}[X]/p\mathbb{K}[X]$ mit $uv = 1$.

Sei $u \in \mathbb{K}[X]$ beliebig mit $p \nmid u$. Dann gilt zunächst $\gcd(u,p) = 1$, denn wäre $\gcd(u,p) = \frac{1}{\mathrm{lc}(p)}p$, so wäre $p \mid u$, was ausgeschlossen ist, und wäre $0 < \deg(\gcd(u,p)) < \deg(p)$, so wäre $\gcd(u,p)$ ein echter Teiler von p, im Widerspruch zur Irreduzibilität von p. Es bleibt also nur $\deg(\gcd(u,p)) = 0$ und also $\gcd(u,p) = 1$.

Nach Satz 72 gibt es nun $v, s \in \mathbb{K}[X]$ mit $1 = vu + sp$, d.h. $1 \sim_p uv$. 03-08

(2)⇒(3). Nach Satz 12 gilt in jedem Körper K die Formel $\forall\ a, b \in K : ab = 0 \Rightarrow a = 0 \vee b = 0$. Daraus folgt die Behauptung.

(3)⇒(1). Seien $u, v \in \mathbb{K}[X]$ mit $p = uv$. Zu zeigen: $u \in \mathbb{K}$ oder $v \in \mathbb{K}$.

Mit $p = uv$ gilt insbesondere $p \mid uv$. Nach Voraussetzung gilt $p \mid u$ oder $p \mid v$, damit $\deg(u) \geq \deg(p)$ oder $\deg(v) \geq \deg(p)$. Wegen $\deg(p) = \deg(u) + \deg(v)$ muss gelten $\deg(p) = \deg(u)$ (und dann $\deg(v) = 0$, also $v \in \mathbb{K}$) oder $\deg(p) = \deg(v)$ (und dann $\deg(u) = 0$, also $u \in \mathbb{K}$). ∎

03-04 **Beispiel.**

1. $p = X^2 - 2$ ist irreduzibel in $\mathbb{Q}[X]$. Betrachte die Körper $\mathbb{K}_1 = \mathbb{Q}[X]/p\mathbb{Q}[X]$ und $\mathbb{K}_2 = \mathbb{Q}(\sqrt{2}) = \{a+b\sqrt{2} : a, b \in \mathbb{Q}\} \subseteq \mathbb{R}$. Diese beiden Körper sind auch $\mathbb{Q}$-Vektorräume. Eine Basis von $\mathbb{K}_1$ ist $\{[1]_\sim, [X]_\sim\}$ und eine Basis von $\mathbb{K}_2$ ist $\{1, \sqrt{2}\}$. Die beiden Vektorräume sind isomorph, ein Isomorphismus ist $h\colon \mathbb{K}_1 \to \mathbb{K}_2$ mit $h([1]_\sim) = 1$ und $h([X]_\sim) = \sqrt{2}$. Dieser Vektorraum-Homomorphismus ist auch mit der Multiplikation verträglich, d.h. es gilt $h(ab) = h(a)h(b)$ für alle $a, b \in \mathbb{K}_1$. Es sind also $\mathbb{K}_1$ und $\mathbb{K}_2$ nicht nur als Vektorräume sondern auch als Körper isomorph, und im Körper $\mathbb{K}_1$ spielt das Element $[X]_\sim$ die Rolle, die $\sqrt{2}$ in $\mathbb{K}_2$ spielt. Insbesondere gilt $[X]_\sim^2 = [X^2]_\sim = [X^2 - (X^2 - 2)]_\sim = [2]_\sim$.

   Im gleichen Sinn sind $\mathbb{R}[X]/(X^2+1)\mathbb{R}[X]$ und $\mathbb{C}$ im wesentlichen die gleichen Körper.

2. Jedes lineare Polynom $aX - b \in \mathbb{K}[X]$ mit $a \neq 0$ ist irreduzibel, aber die zugehörigen Körper $\mathbb{K}[X]/(aX - b)\mathbb{K}[X]$ sind nicht besonders interessant. Sie sind isomorph zum ursprünglichen Körper $\mathbb{K}$.

3. Das Polynom $p = X^4 + 3X + 4$ ist in $\mathbb{Z}_5[X]$ irreduzibel. Der Körper $\mathbb{K} = \mathbb{Z}_5[X]/p\mathbb{Z}_5[X]$ hat als $\mathbb{Z}_5$-Vektorraum die Basis $\{[1]_\sim, [X]_\sim, [X^2]_\sim, [X^3]_\sim\}$, ist als solcher also isomorph zu $\mathbb{Z}_5^4$. Also ist $\mathbb{K}$ ein endlicher Körper mit $5^4 = 625$ Elementen. Beachten Sie, dass dieser Körper etwas anderes ist als $\mathbb{Z}_{625}$. Da 625 keine Primzahl ist, ist $\mathbb{Z}_{625}$ gar kein Körper.

Satz 74. Für jedes $p \in \mathbb{K}[X] \setminus \{0\}$ gibt es ein $c \in \mathbb{K} \setminus \{0\}$ und normierte, irreduzible,
03-08 paarweise verschiedene $p_1, \ldots, p_m \in \mathbb{K}[X]$ sowie $e_1, \ldots, e_m \in \mathbb{N} \setminus \{0\}$ mit

$$p = c\, p_1^{e_1} \cdots p_m^{e_m}.$$

05-05 Die Menge $\{(p_1, e_1), \ldots, (p_m, e_m)\}$ und das Element c sind eindeutig durch p bestimmt.

Beweis. Mit der Wahl von $c = \mathrm{lc}(p)$ kann man o.B.d.A. annehmen $\mathrm{lc}(p) = 1$.

Existenz: Induktion nach $\deg(p)$.

Induktionsanfang: Ist $\deg(p) = 1$, so ist $p = 1p^1$ bereits irreduzibel.

Induktionsvoraussetzung: Der Satz gilt für alle Polynome p mit $\deg(p) < n$.

Induktionsschluss: Sei $p \in \mathbb{K}[X]$ mit $\deg(p) = n$. Ist p irreduzibel, so ist $p = 1p^1$ eine Darstellung der gewünschten Form. Ist p nicht irreduzibel, so gibt es nach Def. 52 eine Zerlegung $p = uv$ mit $u, v \in \mathbb{K}[X] \setminus \mathbb{K}$, d.h. mit $\deg(u), \deg(v) > 0$, d.h. mit $\deg(u), \deg(v) < \deg(p) = n$. Nach Induktionsvorraussetzung haben u und v Zerlegungen der gewünschten Form. Deren Produkt ist eine Zerlegung von p.

03-08 Eindeutigkeit: Es seien $p = p_1^{e_1} \cdots p_m^{e_m} = \tilde{p}_1^{\tilde{e}_1} \cdots \tilde{p}_{\tilde{m}}^{\tilde{e}_{\tilde{m}}}$ zwei Zerlegungen von p. Da die p_i irreduzibel sind, muss es nach Satz 73 zu jedem $i \in \{1, \ldots, m\}$ ein $j \in \{1, \ldots, \tilde{m}\}$ geben mit $p_i \mid \tilde{p}_j$. Da aber auch die $\tilde{p}_j$ irreduzibel sind und sowohl die p_i als auch die $\tilde{p}_j$ normiert sind, muss

gelten $p_i = \tilde{p}_j$. Daraus folgt $\{p_1, \ldots, p_m\} \subseteq \{\tilde{p}_1, \ldots, \tilde{p}_{\tilde{m}}\}$, und aus Symmetriegründen sogar die Gleichheit dieser Mengen.

Da für jedes Polynom u offenbar genau ein $e \in \mathbb{N}$ existiert mit $u^e \mid p$ und $u^{e+1} \nmid p$, und da die p_i und die $\tilde{p}_j$ als paarweise verschieden angenommen sind, sind auch die Exponenten eindeutig bestimmt. ∎

Definition 53. Ein Körper $\mathbb{K}$ heißt *algebraisch abgeschlossen* (engl. *algebraically closed*), falls für jedes irreduzible Polynom $p \in \mathbb{K}[X]$ gilt $\deg(p) \leq 1$. 03-14

Beispiel.

1. $\mathbb{Q}$ und $\mathbb{R}$ und endliche Körper sind nicht algebraisch abgeschlossen.

2. Man kann zeigen, dass $\mathbb{C}$ algebraisch abgeschlossen ist. Das ist der *Fundamentalsatz der Algebra.* Es gilt also: jedes $p \in \mathbb{C}[X] \setminus \mathbb{C}$ lässt sich schreiben als

$$p = c\,(X - \xi_1)^{e_1} \cdots (X - \xi_m)^{e_m}$$

   für gewisse $e_1, \ldots, e_m \in \mathbb{N} \setminus \{0\}$ und $\xi_1, \ldots, \xi_m \in \mathbb{C}$.

   Anders gesagt: Jedes Polynom in $\mathbb{C}[X] \setminus \mathbb{C}$ hat mindestens eine Nullstelle in $\mathbb{C}$. Die entsprechende Aussage mit $\mathbb{R}$ anstelle von $\mathbb{C}$ ist falsch.

   Man verwendet auch die Sprechweise „jedes Polynom $p \in \mathbb{C}[x]$ zerfällt in Linearfaktoren“.

3. Man erhält als Folgerung aus dem Fundamentalsatz der Algebra auch eine Charakterisierung der irreduziblen Polynome in $\mathbb{R}[X]$. Es zeigt sich, dass dies genau die Polynome vom Grad 1 sowie die Polynome vom Grad 2 sind, die keine reelle Nullstelle haben. Letztere sind bekanntlich die Polynome $p = p_0 + p_1 X + p_2 X^2$ mit $p_2 \neq 0$ und $4p_0p_2 > p_1^2$.

   Um das zu sehen, verwendet man die *Konjugation* komplexer Zahlen: für $z = x + \mathrm{i}y \in \mathbb{C}$ mit $x, y \in \mathbb{R}$ definiert man $\bar{z} = x - \mathrm{i}y$. Man rechnet leicht nach, dass dann $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$ und $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$ für alle $z_1, z_2 \in \mathbb{C}$ gilt, sowie $z = \bar{z} \iff z \in \mathbb{R}$ für alle $z \in \mathbb{C}$. Für Polynome $p \in \mathbb{R}[X] \subseteq \mathbb{C}[X]$ folgt daraus $p(\bar{z}) = \overline{p(z)}$ für jedes $z \in \mathbb{C}$. Insbesondere gilt also: Wenn $z \in \mathbb{C}$ eine Nullstelle von p ist, dann auch $\bar{z}$.

   Ein Polynom $p \in \mathbb{R}[X]$ kann also reelle und nicht-reelle komplexe Nullstellen haben, aber die nicht-reellen Nullstellen treten immer als Paare zueinander konjugierter Zahlen $z, \bar{z}$ auf. Aufgefasst als Element von $\mathbb{C}[X]$ lässt sich p also faktorisieren in der Form

$$p = c(X - \xi_1)^{e_1} \cdots (X - \xi_m)^{e_m} (X - \zeta_1)^{\epsilon_1} (X - \bar{\zeta}_1)^{\epsilon_1} \cdots (X - \zeta_n)^{\epsilon_n} (X - \bar{\zeta}_n)^{\epsilon_n}$$

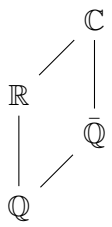
   für gewisse $\xi_1, \ldots, \xi_m \in \mathbb{R}$ und $\zeta_1, \ldots, \zeta_n \in \mathbb{C} \setminus \mathbb{R}$ und $e_1, \ldots, e_m, \epsilon_1, \ldots, \epsilon_n \in \mathbb{N} \setminus \{0\}$.

   Fasst man schließlich die jeweils zueinander gehörenden Faktoren $(X - \zeta)(X - \bar{\zeta})$ zusammen, so erhält man $X^2 - (\zeta + \bar{\zeta})X + \zeta\bar{\zeta}$. Das ist ein Polynom in $\mathbb{R}[X]$, denn ist $\zeta = u + \mathrm{i}v$ für reelle u, v, so sind $\zeta + \bar{\zeta} = (u + \mathrm{i}v) + (u - \mathrm{i}v) = 2u$ und $\zeta\bar{\zeta} = (u + \mathrm{i}v)(u - \mathrm{i}v) = u^2 - \mathrm{i}^2 v^2 = u^2 + v^2$ beide reell.

   Damit ist gezeigt, dass $\mathbb{R}$ zwar nicht algebraisch abgeschlossen ist, dass sich aber jedes $p \in \mathbb{R}[X]$ schreiben lässt als ein Produkt von Polynomen in $\mathbb{R}[X]$ vom Grad höchstens 2.

4. Die irreduziblen Polynome in $\mathbb{Q}[X]$ lassen sich nicht so leicht charakterisieren. Man kann aber zeigen, dass es für jedes $n \in \mathbb{N} \setminus \{0\}$ ein irreduzibles Polynom $p \in \mathbb{Q}[X]$ from Grad n gibt. Zum Beispiel ist $X^n - 5$ für jedes $n \in \mathbb{N} \setminus \{0\}$ irreduzibel als Element 03-08
von $\mathbb{Q}[X]$. Zur Berechnung der Zerlegung eines gegebenen Polynoms $p \in \mathbb{Q}[X]$ in irreduzible Faktoren gibt es Algorithmen. Wie diese funktionieren, wird in Vorlesungen über Computeralgebra erklärt.

   Eine Zahl $z \in \mathbb{C}$ heißt *algebraisch,* falls es ein Polynom $p \in \mathbb{Q}[X] \setminus \{0\}$ gibt, so dass z
03-08 eine Nullstelle von p ist (wobei p als Polynom in $\mathbb{C}[X]$ aufgefasst wird). Zahlen, die nicht algebraisch sind, nennt man *transzendent.* Zum Beispiel sind $\sqrt{2}$ und i (offensichtlich) algebraisch und π und e sind – das ist nicht offensichtlich – transzendent. Die Menge $\bar{\mathbb{Q}} \subseteq \mathbb{C}$ aller algebraischen Zahlen ist also eine echte Teilmenge von $\mathbb{C}$. Man kann zeigen, dass diese Teilmenge einen algebraisch abgeschlossenen Körper bildet.



   In einem bestimmten Sinn ist $\bar{\mathbb{Q}}$ der kleinste algebraisch abgeschlossene Körper, der $\mathbb{Q}$ enthält, und $\mathbb{C}$ ist der kleinste algebraisch abgeschlossene Körper, der $\mathbb{R}$ enthält. In weiterführenden Vorlesungen über Algebra werden Sie erfahren, dass es zu jedem Körper $\mathbb{K}$ einen solchen kleinstmöglichen algebraisch abgeschlossenen Körper $\bar{\mathbb{K}}$ gibt, der $\mathbb{K}$ enthält, und dass dieser im wesentlichen eindeutig durch $\mathbb{K}$ bestimmt ist. Man nennt $\bar{\mathbb{K}}$ den *algebraischen Abschluss* (engl. *algebraic closure*) von $\mathbb{K}$.

26 Diagonalisierbarkeit

Definition 54. Sei V ein $\mathbb{K}$-Vektorraum, $h \in \mathrm{End}(V)$ und $\lambda \in \mathbb{K}$.

1. $v \in V$ heißt *Eigenvektor* (engl. *eigenvector*) von h zu λ, falls $h(v) = \lambda v$ gilt.
2. λ heißt *Eigenwert* (engl. *eigenvalue*) von h, falls es einen von 0 verschiedenen Eigenvektor zu λ gibt.

Im Fall $V = \mathbb{K}^n$ und $h\colon V \to V$, $h(v) = Av$ für eine Matrix $A \in \mathbb{K}^{n\times n}$ spricht man auch von den Eigenvektoren bzw. Eigenwerten von A und meint damit die Eigenvektoren bzw. Eigenwerte von h.

Beispiel.

1. $v = 0$ ist Eigenvektor zu jedem $\lambda \in \mathbb{K}$, denn es gilt $h(0) = \lambda 0 = 0$. Ist $v \neq 0$, so kann v nicht Eigenvektor für zwei verschiedene $\lambda_1, \lambda_2 \in \mathbb{K}$ sein, denn aus $h(v) = \lambda_1 v$ und $h(v) = \lambda_2 v$ folgt $\lambda_1 v = \lambda_2 v$, also $(\lambda_1 - \lambda_2)v = 0$ und daraus wegen Teil 4 von Satz 35 $\lambda_1 = \lambda_2$.

2. $0 \in \mathbb{K}$ ist genau dann ein Eigenwert von h, wenn $\ker h \neq \{0\}$ ist. Die Vektoren in $\ker h$ sind genau die Eigenvektoren zu $\lambda = 0$, denn es gilt ja $v \in \ker h \iff h(v) = 0 = 0v$.

3. Ist $A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \in \mathbb{K}^{n\times n}$ eine Diagonalmatrix, so sind offenbar $\lambda_1, \ldots, \lambda_n$ Eigenwerte von A und die Einheitsvektoren e_i sind Eigenvektoren zu λ_i: Für jedes i gilt $Ae_i = \lambda_i e_i$.

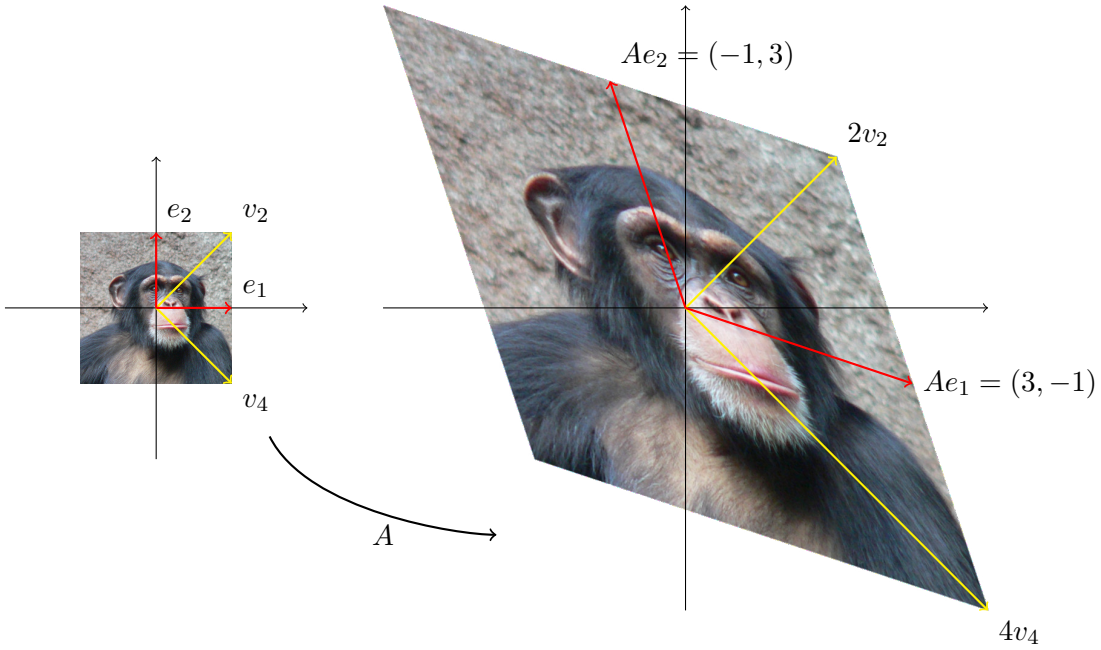
4. Betrachte $A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. Diese Matrix hat zwei Eigenwerte, nämlich 2 und 4. Ein Eigenvektor zum Eigenwert 2 ist $v_2 = \binom{1}{1}$:

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

Ein Eigenvektor zum Eigenwert 4 ist $v_4 = \binom{1}{-1}$:

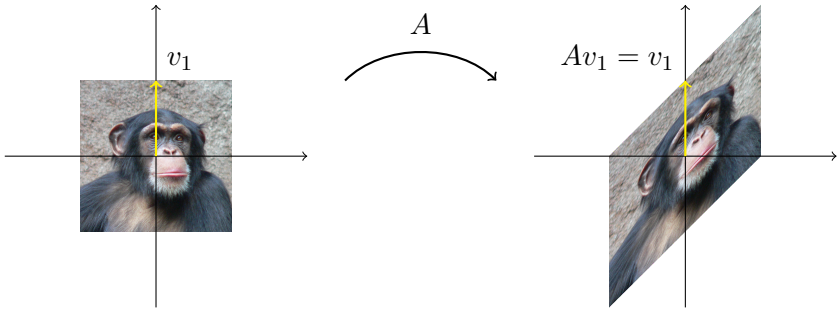
$$\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ -4 \end{pmatrix}.$$

Die Matrix A transformiert die Ebene $\mathbb{R}^2$, indem sie in Richtung $\binom{1}{1}$ ein Streckung um den Faktor 2 und in Richtung $\binom{1}{-1}$ eine Streckung um den Faktor 4 bewirkt:



Während die Standardkoordinatenrichtungen e_1 und e_2 von A auf andere Richtungen abgebildet werden, zeigen v_2 und v_4 nach Anwendung von A immer noch in die gleiche Richtung wie vorher.

5. Die Matrix $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ hat nur einen Eigenwert, nämlich $\lambda = 1$. Ein zugehöriger Eigenvektor ist $v_1 = \binom{0}{1}$.



In diesem Fall wird jeder Vektor $v \in \mathbb{R}^2 \setminus \langle v_1 \rangle$ von A auf einen Vektor abgebildet, der in eine andere Richtung als v zeigt.

05-05 6. Sei $V = C^\infty(\mathbb{R}, \mathbb{R})$ der $\mathbb{R}$-Vektorraum aller beliebig oft differenzierbaren Funktionen und $h\colon V \to V$, $h(f) = f'$ der Ableitungsoperator. Dann ist jedes $\lambda \in \mathbb{R}$ ein Eigenwert von h, und ein zu λ passender Eigenvektor ist die Funktion $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = \mathrm{e}^{\lambda x}$, denn für diese gilt ja $f'(x) = \lambda f(x)$.

Satz 75. Sei $A \in \mathbb{K}^{n \times n}$.

1. Sei $\lambda \in \mathbb{K}$ und sei $E_\lambda = \{\, v \in \mathbb{K}^n : Av = \lambda v \,\}$ die Menge aller Eigenvektoren von A zu λ. Dann gilt: $E_\lambda = \ker(A - \lambda I_n)$. Insbesondere bildet die Menge aller Eigenvektoren zu einem fixen $\lambda \in \mathbb{K}$ stets einen Untervektorraum von $\mathbb{K}^n$.

03-10 2. (Ergänzung zum Satz 26) A ist genau dann invertierbar, wenn 0 kein Eigenwert von A ist.

3. Ist A invertierbar, so ist $\lambda \in \mathbb{K}$ genau dann ein Eigenwert von A, wenn $1/\lambda$ ein Eigenwert von A^{-1} ist.

Beweis.

1. Für alle $v \in \mathbb{K}^n$ gilt

$$v \in E_\lambda \iff Av = \lambda v \iff Av - \lambda v = 0 \iff (A - \lambda I_n)v = 0 \iff v \in \ker(A - \lambda I_n).$$

2. Folgt aus Teil 1 und Satz 26.

3. Für alle $v \in \mathbb{K}^n$ gilt

$$Av = \lambda v \iff v = A^{-1}\lambda v \iff v = \lambda A^{-1} v \iff \lambda^{-1} v = A^{-1} v \iff A^{-1} v = \lambda^{-1} v. \qquad \blacksquare$$

Beispiel. Betrachte $A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \in \mathbb{Q}^{2\times 2}$.

1. Für $\lambda = 2$ haben wir

$$A - 2I_2 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{matrix} \\ \leftarrow\!\rfloor_{+} \end{matrix} \leftrightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$$

   und daher $E_2 = \langle \binom{-1}{-1} \rangle$.

2. Für $\lambda = 4$ haben wir

$$A - 4I_2 = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \begin{matrix} \mid \cdot(-1) \\ \leftarrow\!\rfloor_{+} \end{matrix} \leftrightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

   und daher $E_4 = \langle \binom{1}{-1} \rangle$.

3. $\lambda = 3$ ist kein Eigenwert von A, denn wegen

$$A - 3I_2 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{matrix} \leftarrow\rfloor \mid \cdot(-1) \\ \leftarrow\rfloor \mid \cdot(-1) \end{matrix} \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

   haben wir $E_3 = \{0\}$.

Die Eigenwerte von $A \in \mathbb{K}^{n\times n}$ sind nach Satz 75 genau jene $\lambda \in \mathbb{K}$, für die $\ker(A - \lambda I_n)$ nicht $\{0\}$ ist. Nach Satz 30 sind dies genau jene $\lambda \in \mathbb{K}$, für die $\det(A - \lambda I_n) = 0$ ist. Man kann deshalb die Eigenwerte einer gegebenen Matrix $A \in \mathbb{K}^{n\times n}$ dadurch bestimmen, dass man zunächst das Polynom $\det(A - X\, I_n) \in \mathbb{K}[X]$ ausrechnet und dann dessen Nullstellen bestimmt.

Beispiel.

1. $A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$. Wegen 03-08

$$\begin{aligned} \det(A - XI_2) &= \begin{vmatrix} 3-X & -1 \\ -1 & 3-X \end{vmatrix} \\ &= (3-X)^2 - (-1)^2 \\ &= X^2 - 6X + 9 - 1 = (X-2)(X-4) \end{aligned}$$

   hat A genau zwei Eigenwerte, nämlich 2 und 4.

2. $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Hier gilt

$$\det(A - XI_2) = \begin{vmatrix} 1-X & 1 \\ 0 & 1-X \end{vmatrix} = (1-X)^2.$$

   Man sagt in so einem Fall, dass $\lambda = 1$ ein doppelter Eigenwert von A ist.

3. $A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & -1 & 1 \end{pmatrix}$. Hier gilt

$$\begin{aligned} \det(A - XI_3) &= \begin{vmatrix} 2-X & 1 & 1 \\ 0 & 3-X & 1 \\ 0 & -1 & 1-X \end{vmatrix} \\ &= (2-X)\begin{vmatrix} 3-X & 1 \\ -1 & 1-X \end{vmatrix} \\ &= (2-X)\big((3-X)(1-X)+1\big) \\ &= (2-X)^3. \end{aligned}$$

In diesem Fall ist $\lambda = 2$ ein dreifacher Eigenwert von A. Es gilt $E_2 = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle$.

Definition 55. Sei $A \in \mathbb{K}^{n\times n}$.

1. $\chi := \det(A - XI_n) \in \mathbb{K}[X]$ heißt das *charakteristische Polynom* von A.
2. Ist χ das charakteristische Polynom von A und $\lambda \in \mathbb{K}$ eine Nullstelle von χ der Vielfachheit e, so heißt e die *Vielfachheit* des Eigenwerts λ von A.
3. $E_\lambda = \ker(A - \lambda I_n)$ heißt der *Eigenraum* (engl. *eigenspace*) von A zu λ.

Für ein vollständig faktorisiertes normiertes Polynom $\chi = (\lambda_1 - X)\cdots(\lambda_n - X)$ (wobei die $\lambda_1, \ldots, \lambda_n$ nicht notwendigerweise paarweise verschieden sein müssen) erhält man durch Ausmultiplizieren

$$\chi = (-X)^n + (\lambda_1 + \cdots + \lambda_n)(-X)^{n-1} + \cdots + (\lambda_1 \cdots \lambda_n).$$

Für das charakteristische Polynom χ einer Matrix $A = ((a_{i,j}))_{i,j=1}^n \in \mathbb{K}^{n\times n}$ gilt

$$\chi = (-X)^n + (a_{1,1} + \cdots + a_{n,n})(-X)^{n-1} + \cdots + \det(A).$$

Es gilt also: die Summe der Diagonalelemente von A ist genau die Summe der Eigenwerte von A (bei Berücksichtigung der Vielfachheiten), und die Determinante von A ist genau das Produkt der Eigenwerte von A (ebenfalls bei Berücksichtigung der Vielfachheiten). Die Summe der Diagonalelemente von A nennt man die *Spur* (engl. *trace*) von A.

Wenn das charakteristische Polynom χ einer Matrix $A \in \mathbb{K}^{n\times n}$ nicht in Linearfaktoren zerfällt, dann kann man sich $\mathbb{K}$ durch einen größeren algebraisch abgeschlossenen Körper $\bar{\mathbb{K}}$ ersetzt vorstellen. Dort zerfällt χ dann sicher in Linearfaktoren, und A hat in diesem Körper $\bar{\mathbb{K}}$ genau n (nicht notwendig paarweise verschiedene) Eigenwerte, deren Summe die Spur und deren Produkt die Determinante von A ist.

04-05 **Satz 76.** Seien $A \in \mathbb{K}^{n\times n}$ und $\chi \in \mathbb{K}[X]$ das charakteristische Polynom von A. Dann gilt:

1. Ein Element $\lambda \in \mathbb{K}$ ist genau dann ein Eigenwert von A, wenn λ eine Nullstelle von χ ist.
2. Ist $T \in \mathbb{K}^{n\times n}$ invertierbar, so ist χ auch das charakteristische Polynom von $T^{-1}AT$.

03-08 3. Die Dimension eines Eigenraums E_λ von A ist höchstens so groß wie die Vielfachheit des Eigenwerts λ.

Beweis.

1. Folgt direkt aus der vorangegangenen Diskussion.

2. Es gilt

$$\begin{aligned}\det(T^{-1}AT - XI_n) &= \det(T^{-1}AT - XT^{-1}T) = \det(T^{-1}(A - XI_n)T)\\ &= \det(T^{-1})\det(A - XI_n)\det(T)\\ &= \det(A - XI_n).\end{aligned}$$

3. Zu zeigen: gibt es d linear unabhängige Eigenvektoren von λ, so gilt $(X-\lambda)^d \mid \chi$.

   Seien also $v_1, \ldots, v_d \in \mathbb{K}^n$ linear unabhängig mit $Av_i = \lambda v_i$ für $i = 1, \ldots, d$.

   Wähle $w_{d+1}, \ldots, w_n \in \mathbb{K}^n$ so, dass $\{v_1, \ldots, v_d, w_{d+1}, \ldots, w_n\}$ eine Basis von $\mathbb{K}^n$ ist. Nach Satz 43 ist das möglich. Setze

$$T = (v_1, \ldots, v_d, w_{d+1}, \ldots, w_n) \in \mathbb{K}^{n\times n}$$

   und $B = T^{-1}AT$. Nach Teil 2 hat B das gleiche charakteristische Polynom wie A. Für $i = 1, \ldots, d$ gilt $Av_i = \lambda v_i$, also $T^{-1}Av_i = \lambda T^{-1}v_i$, also $T^{-1}ATT^{-1}v_i = \lambda T^{-1}v_i$, also $Bu_i = \lambda u_i$ für $u_i = T^{-1}v_i$.

   Nach Definition von T ist $u_i = e_i$ gerade der i-te Einheitsvektor. Deshalb hat B die Form

$$B = \begin{pmatrix} \lambda & & & * & \cdots & * \\ & \ddots & & \vdots & & \vdots \\ & & \lambda & \vdots & & \vdots \\ & & & \vdots & & \vdots \\ & & & * & \cdots & * \end{pmatrix}.$$

   Offensichtlich teilt $(X-\lambda)^d$ das charakteristische Polynom von B, und damit auch das von A. ∎

Definition 56.

1. Zwei Matrizen $A, B \in \mathbb{K}^{n\times n}$ heißen (zueinander) *ähnlich,* (engl. *similar*) falls es eine Matrix $T \in \mathrm{GL}(n, \mathbb{K})$ gibt mit $B = T^{-1}AT$. Notation: $A \sim B$.

2. Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *diagonalisierbar,* falls es eine Diagonalmatrix

$$D = \mathrm{diag}(d_1, \ldots, d_n) := \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}$$

   gibt mit $A \sim D$.

Satz 77. $\sim$ ist eine Äquivalenzrelation.

Beweis. Übung. ∎

Satz 78. Sei $A \in \mathbb{K}^{n\times n}$. Das charakteristische Polynom $\chi \in \mathbb{K}[X]$ von A habe eine Darstellung 03-10

$$\chi = (\lambda_1 - X)^{\mu_1} \cdots (\lambda_k - X)^{\mu_k}$$

für gewisse $\mu_1, \ldots, \mu_k \in \mathbb{N} \setminus \{0\}$ und paarweise verschiedene $\lambda_i \in \mathbb{K}$. Dann sind folgende Aussagen äquivalent:

1. A ist diagonalisierbar
2. $\dim E_{\lambda_i} = \mu_i$ für $i = 1, \ldots, k$
3. $\mathbb{K}^n = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$
4. $\mathbb{K}^n$ hat eine Basis, die nur Eigenvektoren von A enthält
5. $A \sim \begin{pmatrix} \boxed{\lambda_1 I_{\mu_1}} & & & \\ & \boxed{\lambda_2 I_{\mu_2}} & & \\ & & \ddots & \\ & & & \boxed{\lambda_k I_{\mu_k}} \end{pmatrix}$, wobei $\boxed{\lambda_i I_{\mu_i}}$ für einen Matrix-Block der Größe $\mu_i \times \mu_i$ steht, bei dem auf der Diagonale überall λ_i und sonst überall 0 steht.

Beweis. Wir zeigen (1)⇒(4)⇒(5)⇒(1) und (2)⇒(3)⇒(4)⇒(2).

(1)⇒(4). A ist diagonalisierbar, also existiert $T \in \mathrm{GL}(n, \mathbb{K})$ mit $T^{-1}AT = \mathrm{diag}(d_1, \ldots, d_n)$ für gewisse $d_1, \ldots, d_n \in \mathbb{K}$. Da T invertierbar ist, bilden die Spalten von T nach Satz 26 03-14 eine Basis von $\mathbb{K}^n$. Ist $v = Te_i$ die i-te Spalte von T, so ist $T^{-1}ATe_i = d_i e_i$, also $ATe_i = d_i Te_i$, also $Av = d_i v$. Da i beliebig war, ist jede Spalte von T ein Eigenvektor von A.

(4)⇒(5). Sei $\{v_1, \ldots, v_n\} \subseteq \mathbb{K}^n$ eine Basis bestehend aus Eigenvektoren von A. Dann ist $T = (v_1, \ldots, v_n) \in \mathbb{K}^{n\times n}$ invertierbar und für $B = T^{-1}AT$ und beliebiges $i \in \{1, \ldots, n\}$ gilt $Be_i = T^{-1}ATe_i = T^{-1}Av_i = T^{-1}\phi v_i = \phi T^{-1}v_i = \phi e_i$, wobei $\phi \in \{\lambda_1, \ldots, \lambda_k\}$ der Eigenwert zu v_i sei. Damit ist gezeigt, dass $T^{-1}AT$ eine Diagonalmatrix ist, wo auf der Diagonale nur Eigenwerte von A stehen. Dass dabei jeder Eigenwert mit der richtigen Vielfachheit auftreten muss, folgt aus Teil 2 von Satz 76. Die Form aus dem Satz er- 03-11 gibt sich schließlich, indem man T durch TP für eine geeignete Permutationsmatrix P ersetzt.

(5)⇒(1). Klar.

(2)⇒(3). Für $i \neq j$ gilt

$$E_{\lambda_i} \cap E_{\lambda_j} = \{\, x \in \mathbb{K}^n : Ax = \lambda_i x = \lambda_j x \,\} \subseteq \{\, x \in \mathbb{K}^n : \underbrace{(\lambda_i - \lambda_j)}_{\neq 0} x = 0 \,\} = \{0\}.$$

Damit ist die Summe direkt und wegen Satz 44 folgt $\dim(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}) = \dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}) = \mu_1 + \cdots + \mu_k = n$, weil $\deg(\chi) = n$. Wegen Satz 42 folgt die Behauptung.

(3)⇒(4). Klar.

(4)⇒(2). Nach Satz 76 gilt auf jeden Fall $\dim E_{\lambda_i} \leq \mu_i$ für alle i. Außerdem gilt natürlich immer $\dim E_{\lambda_i} \geq 0$. Wäre also $\dim E_{\lambda_i} \neq \mu_i$ für wenigstens ein i, so wäre $\dim E_{\lambda_i} < \mu_i$ und $\dim(E_{\lambda_1} + \cdots + E_{\lambda_k}) < n$. Nach Voraussetzung gibt es aber eine Basis aus lauter Eigenvektoren. Es müsste also einen Eigenvektor v geben, der nicht in $E_{\lambda_1} + \cdots + E_{\lambda_k}$ liegt. Der zugehörige Eigenwert kann dann keines der $\lambda_1, \ldots, \lambda_k$ sein, im Widerspruch zu Satz 76. ∎

Beispiel.

1. $A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$ hat die beiden Eigenwerte 2 und 4. Ein Eigenvektor zum Eigenwert 2 ist $\binom{1}{1}$ und ein Eigenvektor zu 4 ist $\binom{1}{-1}$. Für $T = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ gilt

$$T^{-1}AT = \begin{pmatrix} 2 & \\ & 4 \end{pmatrix}.$$

Besser merken kann man sich die Basiswechselgleichung vielleicht in der Form $AT = TD$, 05-06
wobei D für die Diagonalmatrix steht, denn hier wird links jede Spalte von T mit der Matrix A multipliziert, und rechts jede Spalte von T mit dem entsprechenden Eintrag der Diagonalmatrix. Diese Darstellung entspricht also der üblichen Gleichung $Av = \lambda v$ aus der Definition von Eigenwert und Eigenvektor.

2. $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ist nicht diagonalisierbar. Zwar ist $\lambda = 1$ ein doppelter Eigenwert, aber $\dim E_1 = 1 < 2$, und das reicht nicht.

3. Für $A = \begin{pmatrix} 0 & -2 & -2 \\ 0 & 2 & 0 \\ 1 & 1 & 3 \end{pmatrix}$ gilt $\det(A - XI_3) = (1-X)(2-X)^2$. Damit ist 1 ein einfacher und 2 ein doppelter Eigenwert von A. Die Eigenräume lauten

$$E_1 = \langle \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \rangle \quad \text{und} \quad E_2 = \langle \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle,$$

und wegen $\dim E_1 + \dim E_2 = 1 + 2 = 3$ ist A diagonalisierbar. In der Tat gilt

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

4. Wenn A diagonalisierbar ist, kann man leicht A^m für große $m \in \mathbb{N}$ ausrechnen, denn in diesem Fall gibt es ja eine invertierbare Matrix T und eine Diagonalmatrix D, so dass $A = T^{-1}DT$. Für beliebiges $m \in \mathbb{N}$ gilt dann 03-10

$$A^m = (T^{-1}DT)^m = T^{-1}DTT^{-1}DT \cdots T^{-1}DT = T^{-1}D^mT,$$

und das Potenzieren von Diagonalmatrizen ist einfach:

$$\operatorname{diag}(\lambda_1, \ldots, \lambda_n)^m = \operatorname{diag}(\lambda_1^m, \ldots, \lambda_n^m).$$

Betrachte als konkretes Beispiel die Matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. Für diese Matrix gilt

$$A \begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \begin{pmatrix} F_{n+1} \\ F_n + F_{n+1} \end{pmatrix} = \begin{pmatrix} F_{n+1} \\ F_{n+2} \end{pmatrix}$$

für alle $n \in \mathbb{N}$, wobei F_n für die $n$te Fibonacci-Zahl steht. Wegen $F_0 = 0$ und $F_1 = 1$ gilt also die Formel

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = A^n \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

für alle $n \in \mathbb{N}$. Die Matrix A ist diagonalisierbar, es gilt

$$A = \begin{pmatrix} \frac{\sqrt{5}-1}{2} & -\frac{1+\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \\ & \frac{1-\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{5}-1}{2} & \frac{1+\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}^{-1}.$$

Daraus folgt

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{5}-1}{2} & -\frac{1+\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & \\ & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix} \begin{pmatrix} \frac{\sqrt{5}-1}{2} & -\frac{1+\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

für alle $n \in \mathbb{N}$, und schließlich die sogenannte Formel von Binet:

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \right) \qquad (n \in \mathbb{N}).$$

27 Annihilierende Polynome

Definition 57. Sei $p = p_0 + p_1 X + \cdots + p_d X^d \in \mathbb{K}[X]$.

1. Für $\alpha \in \mathbb{K}$ sei $p(\alpha) := p_0 + p_1\alpha + \cdots + p_d\alpha^d \in \mathbb{K}$.
2. Für $A \in \mathbb{K}^{n\times n}$ sei $p(A) := p_0 I_n + p_1 A + \cdots + p_d A^d \in \mathbb{K}^{n\times n}$.
3. Ist V ein $\mathbb{K}$-Vektorraum und $h \in \operatorname{End}(V)$, so sei $p(h) := p_0 \operatorname{id}_V + p_1 h + \cdots + p_d h^d \in \operatorname{End} V$.

03-12

4. Für $q \in \mathbb{K}[X]$ sei $p(q) := p_0 + p_1 q + \cdots + p_d q^d \in \mathbb{K}[X]$.

03-12

Allgemeiner: Ist R zugleich ein Ring mit Eins und ein $\mathbb{K}$-Vektorraum, und zwar so, dass die Ring-Addition mit der Vektor-Addition übereinstimmt und die Skalarmultiplikation mit der Ring-Multiplikation durch die Regel

$$\forall\, \alpha \in \mathbb{K}\ \forall\, a, b \in R : (\alpha a)b = \alpha(ab) = a(\alpha b)$$

verbunden ist, so nennt man R eine $\mathbb{K}$-*Algebra.* Ist R eine solche $\mathbb{K}$-Algebra, so lässt sich jedes Polynom $p \in \mathbb{K}[X]$ in natürlicher Weise als eine Funktion $R \to R$ interpretieren. Wir wollen aber weiterhin die

Polynome selbst als algebraische Objekte auffassen, die von diesen Funktionen zu unterscheiden sind. Insbesondere ist für ein Polynom $p \in \mathbb{K}[X]$ a priori kein Definitionsbereich festgelegt.

Die saubere Unterscheidung ist angebracht, weil auf Polynome unter Umständen andere Aussagen zutreffen als auf die zugehörigen Polynomfunktionen. Betrachtet man zum Beispiel das Polynom $X^2 + X \in \mathbb{Z}_2[X]$, so ist dies offensichtlich verschieden vom Nullpolynom $0 \in \mathbb{Z}_2[X]$, während die zugehörige Polynomfunktion $\mathbb{Z}_2 \to \mathbb{Z}_2$, $x \mapsto x^2 + x$ identisch zur Nullfunktion ist (weil $1^2 + 1 = 1 + 1 = 0$ und $0^2 + 0 = 0$ in $\mathbb{Z}_2$ gilt).

Beispiel. Sei $p = 2 + 3X + X^2$ und $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Dann ist

$$A^2 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}$$

und

$$\begin{aligned} p(A) &= 2I_2 + 3A + A^2 \\ &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 3 & 6 \\ 9 & 12 \end{pmatrix} + \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} = \begin{pmatrix} 12 & 16 \\ 24 & 36 \end{pmatrix}. \end{aligned}$$

Satz 79. Sei V ein $\mathbb{K}$-Vektorraum und seien $p, q \in \mathbb{K}[X]$.

1. Für alle $h \in \operatorname{End}(V)$ gilt $(p+q)(h) = p(h) + q(h)$ und $(pq)(h) = p(h) \circ q(h)$.
2. Für alle $A \in \mathbb{K}^{n \times n}$ gilt $(p+q)(A) = p(A) + q(A)$ und $(pq)(A) = p(A)q(A)$.
3. Für alle $A \in \mathbb{K}^{n \times n}$ und alle $T \in \mathrm{GL}(n, \mathbb{K})$ gilt $p(TAT^{-1}) = Tp(A)T^{-1}$.

Beweis. Übung. ∎

Satz 80. Sei V ein $\mathbb{K}$-Vektorraum mit $\dim V = n < \infty$, und sei $h \in \operatorname{End}(V)$. Dann existiert ein $p \in \mathbb{K}[X] \setminus \{0\}$ mit $p(h) = 0$.

Beweis. Nach Satz 59 ist $\operatorname{End}(V)$ ein $\mathbb{K}$-Vektorraum der Dimension n^2. Darin müssen je $n^2 + 1$ viele Elemente linear abhängig sein (Satz 40). Insbesondere müssen $\mathrm{id}_V, h, h^2, \ldots, h^{n^2}$ linear abhängig sein, d.h. es gibt $p_0, \ldots, p_{n^2} \in \mathbb{K}$, von denen nicht alle Null sind, so dass $p_0 \,\mathrm{id}_V + \cdots + p_{n^2} h^{n^2} = 0$. Es ist also $p = p_0 + p_1 X + \cdots + p_{n^2} X^{n^2} \in \mathbb{K}[X] \setminus \{0\}$ ein Polynom mit der gewünschten Eigenschaft. ∎

Beispiel. Betrachte den Endomorphismus $h \colon \mathbb{R}^2 \to \mathbb{R}^2$, $h(x) = Ax$ mit $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Laut dem Beweis des vorherigen Satzes gibt es ein Polynom p vom Grad höchstens 4 mit $p(A) = 0$. 03-14
Tatsächlich gibt es sogar ein Polynom von Grad 2. Um dieses zu finden, macht man einen Ansatz mit unbestimmten Koeffizienten p_0, p_1, p_2. Die Forderung

$$\begin{aligned} &p_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + p_1 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + p_2 \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} \\ &= \begin{pmatrix} p_0 + p_1 + 7p_2 & 0p_0 + 2p_1 + 10p_2 \\ 0p_0 + 3p_1 + 15p_2 & p_0 + 4p_1 + 22p_2 \end{pmatrix} \overset{!}{=} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

führt durch Nullsetzen der vier Matrixeinträge auf das homogene lineare Gleichungssystem 03-11 04-11

$$\begin{pmatrix} 1 & 1 & 7 \\ 0 & 2 & 10 \\ 0 & 3 & 15 \\ 1 & 4 & 22 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Die Lösungsmenge dieses Gleichungssystems ergibt sich zu $\left\langle \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix} \right\rangle$, also ist $p = 2 + 5X - X^2$ ein Polynom mit der gewünschten Eigenschaft $p(h) = 0$.

Ein Polynom $p \in \mathbb{K}[X]$ mit $p(h) = 0$ bzw. $p(A) = 0$ heißt ein *annihilierendes Polynom* von h bzw. A. Mit solchen Polynomen kann man die Berechnung von $q(A)$ für Polynome q von hohem Grad erheblich vereinfachen. Schreibe dazu $q = \text{quo}(q,p)p + \text{rem}(q,p)$ und beachte, dass $\text{quo}(q,p)p$ ausgewertet bei A 04-11 Null ist, wenn p ein annihilierendes Polynom für A ist. Es gilt also $q(A) = \text{rem}(q,p)(A)$, und egal wie groß der Grad von q war, der Grad von $\text{rem}(q,p)$ ist immer kleiner als $\deg(p)$.

Für die spezielle Wahl $q = X^m$ ergibt sich eine weitere effiziente Methode zur Berechnung von Potenzen von Matrizen. Diese Methode funktioniert auch für Matrizen, die nicht diagonalisierbar sind. Sie liefert dabei allerdings keine allgemeinen Ausdrücke für die Einträge von A^m, wenn m eine Variable ist.

Wegen Teil 3 von Satz 79 gilt $p(h) = 0$ für ein $p \in \mathbb{K}[X]$ und ein $h \in \text{End}(V)$ genau dann wenn für jede Basis B von V gilt, dass $p(A) = 0$ für die Abbildungsmatrix von h bezüglich B und B. Es ist deshalb für die folgenden Überlegungen relativ egal, ob wir über Endomorphismen oder quadratische Matrizen sprechen.

Satz 81. Sei $A \in \mathbb{K}^{n\times n}$. Dann gilt:

1. Sind $p, q \in \mathbb{K}[X]$ zwei annihilierende Polynome von A, d.h. gilt $p(A) = q(A) = 0$, und ist $g = \gcd(p,q)$, so gilt auch $g(A) = 0$.
2. Es sei $M \subseteq \mathbb{K}[X]$ die Menge aller $p \in \mathbb{K}[X] \setminus \{0\}$ mit $p(A) = 0$ für die es kein $q \in \mathbb{K}[X] \setminus \{0\}$ mit $\deg(q) < \deg(p)$ und $q(A) = 0$ gibt. Dann gibt es in M genau ein normiertes Polynom.
3. Ist $m \in \mathbb{K}[X]$ das eindeutig bestimmte Polynom aus Teil 2, und ist $p \in \mathbb{K}[X]$ beliebig, so gilt $p(A) = 0 \iff m \mid p$.

Beweis.

1. Nach Satz 72 existieren Polynome $u, v \in \mathbb{K}[X]$ mit $g = up + vq$. Wegen Satz 79 folgt direkt $g(A) = (up + vq)(A) = u(A)p(A) + v(A)q(A) = 0$, weil nach Voraussetzung $p(A) = q(A) = 0$ ist.
2. Nach Satz 80 gibt es überhaupt Polynome $p \in \mathbb{K}[X] \setminus \{0\}$ mit $p(A) = 0$. Unter diesen muss es solche geben, für die der Grad minimal wird. Unter diesen wiederum muss dann wenigstens ein normiertes Polynom geben. Daraus folgt, dass M mindestens ein normiertes Polynom enthält.

   Es bleibt zu zeigen, dass M höchstes ein normiertes Polynom enthält. Gäbe es zwei verschiedene normierte Polynome $m_1, m_2 \in M$, so wäre $m_1 - m_2 \neq 0$ und wegen $\text{lc}(m_1) = \text{lc}(m_2) = 1$ wäre $\deg(m_1) = \deg(m_2) > \deg(m_1 - m_2)$. Mit $m_1(A) = m_2(A) = 0$ gilt aber auch $(m_1 - m_2)(A) = 0$, d. h. $m_1 - m_2$ wäre ein annihilierendes Polynom von A von kleinerem Grad als m_1, im Widerspruch zur Definition von M.

3. „⇒“ Sei $p \in \mathbb{K}[X]$ beliebig und $r = \operatorname{rem}(p, m)$. Dann gilt $r = p - qm$ für ein bestimmtes $q \in \mathbb{K}[X]$ und deshalb $p(A) = r(A)$ (weil ja $m(A) = 0$ ist).

   Ist $p(A) = 0$, so ist $r(A) = p(A) = 0$. Da $\deg(r) < \deg(m)$ und m von allen von Null verschiedenen annihilierenden Polynomen von A minimalen Grad hat, muss $r = 0$ gelten. Daraus folgt $m \mid p$.

   „⇐“ Aus $m \mid p$ folgt $p = qm$ für ein $q \in \mathbb{K}[X]$, und daraus $p(A) = q(A)m(A) = 0$. ∎

Beispiel. Für $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ gilt

$$p_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + p_1 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 0 \quad \iff \quad p_0 = p_1 = 0.$$

Davon kann man sich durch Lösen eines geeigneten linearen Gleichungssystems überzeugen. Das Polynom $m = X^2 - 5X + 2$ aus dem vorherigen Beispiel hat also minimalen Grad und es gilt

$$\{\, p \in \mathbb{K}[X] : p(A) = 0 \,\} = \{\, p \in \mathbb{K}[X] : m \mid p \,\}.$$

Definition 58. Seien $A \in \mathbb{K}^{n \times n}$ und $m \in \mathbb{K}[X] \setminus \{0\}$ wie im Teil 2 von Satz 81. Dann heißt m das *Minimalpolynom* (engl. *minimal polynomial*) von A.

Beispiel.

1. Das Minimalpolynom der Nullmatrix $0 \in \mathbb{K}^{n \times n}$ ist $1 \in \mathbb{K}[X]$.
2. Das Minimalpolynom der Identitätsmatrix $I_n \in \mathbb{K}^{n \times n}$ ist $X - 1 \in \mathbb{K}[X]$.
3. Das Minimalpolynom von $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathbb{Q}^{2 \times 2}$ ist $(X-1)^2$.

   Zum Beweis rechnet man zunächst nach, dass $m = (X-1)^2$ ein annihilierendes Polynom von A ist. In der Tat gilt $m(A) = A^2 - 2A + I_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

   Um zweitens zu zeigen, dass es kein annihilierendes Polynom von kleinerem Grad gibt, macht man einen Ansatz für ein solches Polynom und löst ein lineares Gleichungssystem für die Koeffizienten. Damit etwa $p = p_0 + p_1 X$ ein annihilierendes Polynom von A ist, muss gelten

$$p_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + p_1 \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} p_0 + p_1 & p_1 \\ 0 & p_0 + p_1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

   also $p_1 = p_0 + p_1 = 0$, und das ist offensichtlich nur dann der Fall, wenn $p_0 = p_1 = 0$ ist. Es gibt also kein annihilierendes Polynom von A vom Grad 0 oder 1.

   Damit ist gezeigt, dass $m = (X-1)^2$ das Minimalpolynom von A ist.

4. Ist $A \in \mathbb{K}^{n \times n}$ eine gegebene Matrix, so gibt es nach Satz 80 ein annihilierendes Polynom vom Grad n^2. Alle annihilierenden Polynome vom Grad $\leq n^2$ bilden einen $\mathbb{K}$-Vektorraum, für den man sich durch Ansatz und Koeffizientenvergleich und Lösen eines linearen Gleichungssystems eine Basis verschaffen kann. Schreibt man die Koeffizienten

der Basispolynome in die Zeilen einer Matrix und berechnet davon die Treppennormalform, so erscheinen die Koeffizienten des Minimalpolynoms als unterste Zeile in dieser Treppenform.

Alternativ kann man auch für $n = 1, 2, 3, \ldots$ nacheinander durch Lösen eines linearen Gleichungssystems überprüfen, ob es ein von Null verschiedenes annihilierendes Polynom vom Grad n gibt. Das erste, das man dabei findet, ist (ein $\mathbb{K}$-Vielfaches) des Minimalpolynoms.

Beispiel: Für die Matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ aus dem vorherigen Beispiel ist $\{X^2 - 5X - 2\}$ eine Basis des Vektorraum aller annihilierenden Polynome vom Grad ≤ 2. Insbesondere ist $X^2 - 5X - 2$ das Minimalpolynom von A.

Hätten wir stattdessen einen Ansatz für ein annihilierendes Polynom vom Grad 4 gemacht, so hätten wir möglicherweise die Basis $\{X^4 + X^2 - 150X - 56, X^4 + X^3 - 172X - 64, X^4 - 145X - 54\}$ gefunden. Daraus darf man nicht schließen, dass das Minimalpolynom den Grad vier hat. Stattdessen folgt aus

$$\begin{pmatrix} 1 & 0 & 1 & -150 & -56 \\ 1 & 1 & 0 & -172 & -64 \\ 1 & 0 & 0 & -145 & -54 \end{pmatrix} \leftrightarrow \begin{pmatrix} 1 & 0 & 0 & -154 & -54 \\ 0 & 1 & 0 & -27 & -10 \\ 0 & 0 & 1 & -5 & -2 \end{pmatrix},$$

dass das Minimalpolynom $X^2 - 5X - 2$ ist.

Satz 82. (Ergänzung zum Satz 26) Sei $A \in \mathbb{K}^{n \times n}$ und $m = m_0 + m_1 X + \cdots + m_d X^d \in \mathbb{K}[X]$ das Minimalpolynom von A. Dann gilt: A ist genau dann invertierbar wenn $m_0 \neq 0$, und in diesem Fall ist

$$A^{-1} = -\frac{1}{m_0}\big(m_1 + m_2 A + \cdots + m_d A^{d-1}\big).$$

03-13 **Beweis.** „⇒“ A ist invertierbar. Wäre $m_0 = 0$, so würde aus $m_1 A + \cdots + m_d A^d = 0$ durch Multiplikation beider Seiten mit A^{-1} (von rechts oder links) folgen $m_1 + \cdots + m_d A^{d-1} = 0$, im Widerspruch zur Minimalität von m.

„⇐“ $m_0 \neq 0$. Betrachte $B = -\frac{1}{m_0}(m_1 + m_2 A + \cdots + m_d A^{d-1})$. Dann gilt

$$\begin{aligned} AB = BA &= -\frac{1}{m_0}\big(m_1 A + m_2 A^2 + \cdots + m_d A^d\big) \\ &= -\frac{1}{m_0}\big(m_1 A + m_2 A^2 + \cdots + m_d A^d - m(A)\big) \\ &= -\frac{1}{m_0}\big(-m_0 I_n\big) = I_n. \end{aligned}$$

Damit ist A invertierbar und auch die behauptete Formel für die Inverse ist bewiesen. ∎

28 Der Satz von Cayley-Hamilton

Satz 83. Sei $A \in \mathbb{K}^{n \times n}$ und $m \in \mathbb{K}[X]$ das Minimalpolynom von A. Dann gilt: $\lambda \in \mathbb{K}$ ist genau dann ein Eigenwert von A, wenn $m(\lambda) = 0$ ist.

Beweis. Es sei $m = m_0 + m_1 X + \cdots + m_d X^d$ das Minimalpolynom von A und $p = m_0 + m_1(X+\lambda) + \cdots + m_d(X+\lambda)^d$. Dann ist p annihilierendes Polynom von $A - \lambda I_n$, denn $p(A - \lambda I_n) = m(A) = 0$. Es ist sogar das Minimalpolynom von $A - \lambda I_n$, denn gäbe es annihilierendes Polynom $q = q_0 + q_1 X + \cdots + q_{d-1} X^{d-1} \neq 0$ kleineren Grades von $A - \lambda I_n$, so wäre $q(X-\lambda) = q_0 + q_1(X-\lambda) + \cdots + q_{d-1}(X-\lambda)^{d-1}$ ein von Null verschiedenes annihilierendes Polynom von A mit einem kleineren Grad als m. Das steht im Widerspruch zur Minimalität von m.

Nun gilt:

$$\lambda \text{ ist Eigenwert von } A \overset{\text{Satz 75}}{\iff} \ker(A - \lambda I_n) \neq \{0\} \overset{\text{Satz 82}}{\iff} p(0) = 0 \iff m(\lambda) = 0. \qquad \blacksquare$$

Der Satz sagt, dass das Minimalpolynom dieselben Nullstellen hat wie das charakteristische Polynom aus Definition 55. Daraus folgt aber nicht, dass auch die Polynome identisch sind. Zum einen wird nichts ausgesagt über die irreduziblen Faktoren höheren Grades, und zum anderen können sich die Vielfachheiten der Nullstellen in den beiden Polynomen unterscheiden.

Der Satz von Cayley-Hamilton, um den es in diesem Abschnitt geht, besagt, dass das charakteristische Polynom einer Matrix A immer auch ein annihilierendes Polynom für A ist. Nach Satz 81 Teil 3 bedeutet das, dass das Minimalpolynom ein Teiler des charakteristischen Polynoms ist. Damit müssen alle irreduziblen Faktoren des Minimalpolynoms auch im charakteristischen Polynom vorkommen, und die Vielfachheit eines jeden irreduziblen Faktors des charakteristischen Polynoms muss mindestens so groß sein wie die Vielfachheit des entsprechenden Faktors des Minimalpolynoms. Außerdem folgt aus dem Satz, dass das Minimalpolynom einer Matrix $A \in \mathbb{K}^{n\times n}$ höchstens den Grad n haben kann, während Satz 80 nur die gröbere Gradschranke n^2 liefert.

Bevor wir zum Satz von Cayley-Hamilton kommen, beweisen wir zwei Hilfssätze, die wir in seinem Beweis verwenden werden.

Satz 84. Sei $a = a_0 + a_1 X + \cdots + a_{d-1} X^{d-1} + X^d \in \mathbb{K}[X]$ und

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{d-2} & -a_{d-1} \end{pmatrix} \in \mathbb{K}^{d\times d}.$$

Dann ist $(-1)^d a$ das charakteristische Polynom von A.

Beweis. Wir zeigen durch Induktion nach d die folgende etwas allgemeinere Aussage: für alle $a_0, \ldots, a_{d-1} \in \mathbb{K}(X)$ gilt

$$\begin{vmatrix} -X & 1 & & & \\ & -X & 1 & & \\ & & \ddots & \ddots & \\ & & & -X & 1 \\ -a_0 & -a_1 & \cdots & -a_{d-2} & -a_{d-1} - X \end{vmatrix} = (-1)^d (a_0 + a_1 X + \cdots + a_{d-1} X^{d-1} + X^d).$$

Für $d = 1$ ist das offensichtlich richtig.

Induktionsschluss $d-1 \to d$:

$$
\begin{vmatrix}
-X & 1 & & & \\
 & -X & 1 & & \\
 & & \ddots & \ddots & \\
 & & & -X & 1 \\
-a_0 & -a_1 & \cdots & -a_{d-2} & -a_{d-1}-X
\end{vmatrix}
\overset{X^{-1}\ +}{\underset{\text{Spaltenoperation}}{\curvearrowright}}
$$

$$
\begin{aligned}
&= \begin{vmatrix}
-X & 0 & & & \\
 & -X & 1 & & \\
 & & \ddots & \ddots & \\
 & & & -X & 1 \\
-a_0 & (-a_1 - a_0X^{-1}) & \cdots & -a_{d-2} & -a_{d-1}-X
\end{vmatrix} \\
&= (-X) \begin{vmatrix}
-X & 1 & & & \\
 & -X & 1 & & \\
 & & \ddots & \ddots & \\
 & & & -X & 1 \\
(-a_1 - a_0X^{-1}) & \cdots & -a_{d-2} & -a_{d-1}-X
\end{vmatrix} \\
&= (-X)(-1)^{d-1}\big((a_1 + a_0X^{-1}) + a_2X + \cdots + a_{d-1}X^{d-2} + X^{d-1}\big) \\
&= (-1)^d(a_0 + a_1X + \cdots + a_{d-1}X^{d-1} + X^d),
\end{aligned}
$$

wie gefordert. Dabei wurde im vorletzten Schritt die Induktionsannahme verwendet. ∎

Satz 85. Seien $A \in \mathbb{K}^{n\times n}$, $B \in \mathbb{K}^{m\times m}$, $C \in \mathbb{K}^{n\times m}$, und sei

$$M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \in \mathbb{K}^{(n+m)\times(n+m)}.$$

Dann gilt $\det(M) = \det(A)\det(B)$.

Insbesondere ist das charakteristische Polynom von M das Produkt der charakteristischen Polynome von A und B.

Beweis. Die Aussage über die charakteristischen Polynome folgt direkt aus der behaupteten Determinatenformel und der Definition des charakteristischen Polynoms.

Wir zeigen die Determinantenformel durch Induktion nach n. Schreibe $A = ((a_{i,j}))_{i,j=1}^n$.

Für $n = 0$ ist nichts zu zeigen. Zum Induktionsschluss $n-1 \to n$ entwickeln wir die Determinante nach der ersten Spalte (vgl. Satz 33). Es bezeichne $A^{(i)} \in \mathbb{K}^{(n-1)\times(n-1)}$ die Matrix, die aus A entsteht, wenn man die i-te Zeile und die erste Spalte löscht und $C^{(i)} \in \mathbb{K}^{(n-1)\times m}$ die
03-16 Matrix, die aus C entsteht, wenn man die i-te Zeile löscht. Dann gilt

$$
\begin{aligned}
\begin{vmatrix} A & C \\ 0 & B \end{vmatrix} &= a_{1,1}\begin{vmatrix} A^{(1)} & C^{(1)} \\ 0 & B \end{vmatrix} - a_{2,1}\begin{vmatrix} A^{(2)} & C^{(2)} \\ 0 & B \end{vmatrix} \pm \cdots + (-1)^n a_{n,1}\begin{vmatrix} A^{(n)} & C^{(n)} \\ 0 & B \end{vmatrix} + 0 + \cdots + 0 \\
&= a_{1,1}\det(A^{(1)})\det(B) - a_{2,1}\det(A^{(2)})\det(B) \pm \cdots + (-1)^n a_{n,1}\det(A^{(n)})\det(B)
\end{aligned}
$$

$$= \underbrace{\big(a_{1,1}\det(A^{(1)}) - a_{2,1}\det(A^{(2)}) \pm \cdots + (-1)^n a_{n,1}\det(A^{(n)})\big)}_{=\det(A)} \det(B).$$

Dabei wird im zweiten Schritt die Induktionshypothese verwendet. ∎

Satz 86. (Cayley-Hamilton) Sei $A \in \mathbb{K}^{n\times n}$ und sei $\chi \in \mathbb{K}[X]$ das charakteristische Polynom von A. Dann gilt $\chi(A) = 0$.

Beweis. Betrachte den Homomorphismus $h\colon \mathbb{K}^n \to \mathbb{K}^n$, $h(x) = Ax$. Zu zeigen: $\chi(h) = 0$, d.h. für alle $x \in \mathbb{K}^n$ gilt $\chi(h)(x) = 0$.

Sei $x \in \mathbb{K}^n$ beliebig. Dann gibt es ein maximales $d \in \mathbb{N}$, für das $x, h(x), \ldots, h^{d-1}(x)$ linear unabhängig sind. Dann sind $x, h(x), \ldots, h^d(x)$ linear abhängig und es gibt $u_0, \ldots, u_{d-1} \in \mathbb{K}$ mit

$$h^d(x) = u_0 x + u_1 h(x) + \cdots + u_{d-1}h^{d-1}(x).$$

Außerdem können wir $b_{d+1}, \ldots, b_n \in \mathbb{K}^n$ so wählen, dass

$$B = \{x, h(x), \ldots, h^{d-1}(x), b_{d+1}, \ldots, b_n\}$$

eine Basis von $\mathbb{K}^n$ ist.

Die Abbildungsmatrix M von h bezüglich B und B hat die Form

$$M = \begin{pmatrix} U & V \\ 0 & W \end{pmatrix}$$

mit

$$U = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ u_0 & \cdots & \cdots & \cdots & u_{d-1} \end{pmatrix} \in \mathbb{K}^{d\times d}, \quad V \in \mathbb{K}^{d\times(n-d)}, \quad W \in \mathbb{K}^{(n-d)\times(n-d)}.$$

Nach Teil 2 von Satz 76 haben M und A dasselbe charakteristische Polynom. Wegen Satz 84 ist

$$(-1)^d\big(X^d - u_{d-1}X^{d-1} - \cdots - u_1 X - u_0\big)$$

das charakteristische Polynom χ_U von U und damit wegen Satz 85 ein Teiler von χ.

Nach der Wahl von $u_0, \ldots, u_{d-1}$ gilt $\chi_U(h)(x) = 0$ und damit $\chi(h)(x) = 0$, was zu zeigen war. ∎

Beispiel.

1. Für $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ gilt

$$\chi = \det(A - XI_2) = \begin{vmatrix} 1-X & 2 \\ 3 & 4-X \end{vmatrix} = (1-X)(4-X) - 6 = X^2 - 5X - 2,$$

und in der Tat gilt

$$A^2 - 5A - 2I_2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} - \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

In diesem Fall ist das charakteristische Polynom zugleich auch das Minimalpolynom von A.

2. Für $A = I_n \in \mathbb{K}^{n\times n}$ gilt $\det(A - XI_n) = (1 - X)^n$. Das Minimalpolynom von A ist $X - 1$.

3. Sei $A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 20 \\ 21 & 22 & 23 & 24 & 25 \end{pmatrix}$.

   Das charakteristische Polynom von A lautet $X^3(X^2-65X-250)$. Das Minimalpolynom ist $X(X^2 - 65X - 250)$.

4. Die Vielfachheit von Faktoren im Minimalpolynom ist meistens eins. Aber nicht immer: Für ein beliebiges $\lambda \in \mathbb{K}$ betrachte die Matrix

$$A = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} \in \mathbb{K}^{n\times n}.$$

   Das charakteristische Polynom von A ist $(\lambda-X)^n$. Nach dem Satz von Cayley-Hamilton muss das Minimalpolynom die Form $(X - \lambda)^e$ für ein bestimmtes $e \in \{1, \ldots, n\}$ haben.

   Tatsächlich gilt $e = n$. Um das zu sehen, überlegt man sich, wie die Potenzen A^m von A aussehen. Induktiv zeigt man, dass all diese Potenzen die Form

$$A^m = \begin{pmatrix} a_{m,1}\lambda^m & a_{m,2}\lambda^{m-1} & \cdots & a_{m,n}\lambda^{m-n+1} \\ 0 & a_{m,1}\lambda^m & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{m,2}\lambda^{m-1} \\ 0 & \cdots & 0 & a_{m,1}\lambda^m \end{pmatrix}$$

   mit

$a_{m,1}$	$a_{m,2}$	$\cdots$	$a_{m,k}$	$\cdots$	$a_{m,m}$	$a_{m,m+1}$	$a_{m,m+2}$	$\cdots$	$a_{m,n}$
1	m	$\cdots$	$\binom{m}{k}$	$\cdots$	m	1	0	$\cdots$	0

   haben. Dabei ist $\binom{m}{k} = \frac{m(m-1)\cdots(m-k+1)}{k!}$ der Binomialkoeffizient. Die Einträge $a_{m,i}$ sind also genau jene Körperelemente, für die gilt

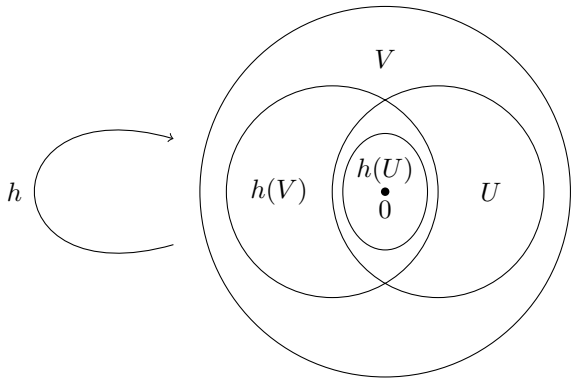
$$(X + 1)^m = a_{m,1}X^m + a_{m,2}X^{m-1} + \cdots + a_{m,m+1}.$$

Jedenfalls enthält für jedes $m < n$ die Matrix A^m an der Stelle $(1, m+1)$ eine Eins, während bei allen Matrizen $A^0, \dots, A^{m-1}$ an dieser Stelle eine Null steht. Daraus folgt, dass $A^0, A^1, \dots, A^m$ für jedes $m < n$ linear unabhängig sind. Daraus folgt schließlich, dass das Minimalpolynom (mindestens) den Grad n haben muss.

5. Für die Matrix $A = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & 0 & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix}$ ($\lambda \in \mathbb{K}$ wieder beliebig) lautet das charakteristische Polynom $(\lambda - X)^4$. Das Minimalpolynom ist in diesem Fall $(X - \lambda)^2$.

29 Invariante Unterräume

Definition 59. Sei V ein $\mathbb{K}$-Vektorraum und $h\colon V \to V$ ein Endomorphismus. Ein Unterraum $U \subseteq V$ heißt *h-invariant,* falls $h(U) \subseteq U$ ist.



Beispiel.

1. V und $\ker h$ und $\{0\}$ sind h-invariante Unterräume von V für jedes $h\colon V \to V$.
2. Ist λ ein Eigenwert von $h\colon V \to V$, dann ist E_λ ein h-invarianter Unterraum von V. Allgemeiner: Sind $\lambda_1, \dots, \lambda_k$ paarweise verschiedene Eigenwerte von h und $E_{\lambda_1}, \dots, E_{\lambda_k}$ die zugehörigen Eigenräume, so ist $E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k}$ ein h-invarianter Unterraum von V.
3. Sei $V = \mathbb{K}[[X]]$ und $h = \frac{d}{dX}\colon V \to V$ die (formale) Ableitung (vgl. S. 106). Dann ist $\mathbb{K}[X] \subseteq V$ ein h-invarianter Unterraum von V, weil die Ableitung jedes Polynoms wieder ein Polynom ist.

   Darüber hinaus ist für jedes fixe $d \in \mathbb{N}$ der Raum $\mathbb{K}[X]_{\leq d}$ aller Polynome vom Grad höchstens d ein h-invarianter Unterraum, weil die Ableitung den Grad eines Polynoms nicht erhöhen kann. Wir haben hier also eine Kette von ineinander enthaltenden h-invarianten Unterräumen

$$\{0\} \subseteq \mathbb{K}[X]_{\leq 1} \subseteq \mathbb{K}[X]_{\leq 2} \subseteq \dots \subseteq \mathbb{K}[X]_{\leq d} \subseteq \dots \subseteq \mathbb{K}[X] \subseteq \mathbb{K}[[X]].$$

   Natürlich gibt es auch Unterräume von $\mathbb{K}[[X]]$, die nicht h-invariant sind, zum Beispiel ist $U = \langle X^{17} + X \rangle$ nicht h-invariant, weil $h(X^{17} + X) = 17X^{16} + 1 \notin U$.

4. Sei $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ und $h\colon \mathbb{Q}^3 \to \mathbb{Q}^3$, $h(x) = Ax$. Da A nicht diagonalisierbar ist, lässt sich $\mathbb{Q}^3$ nicht als Summe von Eigenräumen schreiben. Die Eigenräume von A sind $E_1 = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \rangle$ und $E_2 = \langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \rangle$.

   Neben E_1 und E_2 ist auch $U = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rangle$ ein h-invarianter Unterraum von V, und wir haben die Zerlegung $\mathbb{Q}^3 = U \oplus E_2$. In diesem Sinn darf man sich unter dem Begriff des invarianten Unterraums eine Verallgemeinerung des Begriffs des Eigenraums vorstellen.

Wenn $h\colon V \to V$ ein Homomorphismus ist und U ein h-invarianter Unterraum von V, dann ist die Einschränkung $h|_U\colon U \to V$ ein Homomorphismus mit $\operatorname{im} h|_U \subseteq U$. Es ist deshalb zulässig, den Bildbereich von $h|_U$ auf U zu beschränken und $h|_U$ als eine Abbildung von U nach U aufzufassen. Damit wird $h|_U$ zu einem Endomorphismus.

Satz 87. Sei V ein $\mathbb{K}$-Vektorraum mit $\dim V = n < \infty$, und sei $h\colon V \to V$ ein Endomorphismus. Dann sind folgende Aussagen äquivalent:

1. Es gibt zwei h-invariante Unterräume U_1, U_2 von V mit $\dim U_1, \dim U_2 > 0$ und $V = U_1 \oplus U_2$
2. Es gibt eine geordnete Basis B von V bezüglich der die Abbildungsmatrix von h die Form

$$\begin{array}{l} \overbrace{}^{k} \\ \left.\begin{pmatrix} * & \cdots & * & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ * & \cdots & * & 0 & \cdots & 0 \\ 0 & \cdots & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \underbrace{* \quad \cdots \quad *}_{n-k} & & \end{pmatrix}\right\} n-k \end{array} \quad k$$

   für ein $k \in \{1, \ldots, n-1\}$ hat.

Beweis.

(1)⇒(2). Wähle eine geordnete Basis $B_1 = (b_1, \ldots, b_k)$ von U_1 und eine geordnete Basis $B_2 = (b_{k+1}, \ldots, b_n)$ von U_2 und setze $B = (b_1, \ldots, b_k, b_{k+1}, \ldots, b_n)$. Wegen $U_1 \oplus U_2 = V$ ist B eine Basis von V. Ist M_1 die Abbildungsmatrix von $h|_{U_1}$ bezüglich B_1 und M_2 die Abbildungsmatrix von $h|_{U_2}$ bezüglich B_2, dann ist

$$M = \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix}$$

die Abbildungsmatrix von h bezüglich B.

(2)⇒(1). Sei $B = (b_1, \dots, b_k, b_{k+1}, \dots, b_n)$ die geordnete Basis, bezüglich der die Abbildungsmatrix von h die angegebene Form hat. Betrachte $U_1 = \langle b_1, \dots, b_k \rangle$ und $U_2 = \langle b_{k+1}, \dots, b_n \rangle$. Dann ist $U_1 \oplus U_2 = V$, $\dim U_1 = k > 0$, $\dim U_2 = n - k > 0$ und wegen

$$\begin{pmatrix} * & \cdots & * & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ * & \cdots & * & 0 & \cdots & 0 \\ 0 & \cdots & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & * & \cdots & * \end{pmatrix} \begin{pmatrix} * \\ \vdots \\ * \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} * \\ \vdots \\ * \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} * & \cdots & * & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ * & \cdots & * & 0 & \cdots & 0 \\ 0 & \cdots & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & * & \cdots & * \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ * \\ \vdots \\ * \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ * \\ \vdots \\ * \end{pmatrix}$$

sind U_1 und U_2 beide h-invariant. ∎

Satz 88. Sei V ein $\mathbb{K}$-Vektorraum, $h\colon V \to V$ ein Endomorphismus, $p \in \mathbb{K}[X]$. Dann ist $\ker p(h) \subseteq V$ ein h-invarianter Unterraum.

Beweis. Zu zeigen: für alle $x \in \ker p(h)$ gilt $h(x) \in \ker p(h)$. Sei also $x \in \ker p(h)$. Dann gilt $p(h)(x) = 0$. Dann $(h \circ p(h))(x) = 0$. Dann $(Xp)(h)(x) = 0$. Dann $(pX)(h)(x) = 0$. Dann $(p(h) \circ h)(x) = 0$. Dann $p(h)(h(x)) = 0$. Dann $h(x) \in \ker p(h)$. ∎

Beispiel. Sei $V = C^\infty([0,1], \mathbb{R})$ der Vektorraum aller beliebig oft differenzierbaren Funktionen. Betrachte eine lineare Differentialgleichung mit konstanten Koeffizienten:

$$c_0 f(x) + c_1 f'(x) + \cdots + c_r f^{(r)}(x) = 0$$

mit $c_0, \dots, c_r \in \mathbb{R}$. Eine Funktion $f \in V$ ist offenbar genau dann eine Lösung dieser Gleichung, wenn $f \in \ker p(\frac{d}{dx})$ für $p = c_0 + c_1 X + \cdots + c_r X^r$ gilt.

Der Satz besagt also, dass die Lösungsmenge einer lineare Differentialgleichung mit konstanten Koeffizienten abgeschlossen unter Ableitung ist.

Satz 89. Seien V ein $\mathbb{K}$-Vektorraum, $h\colon V \to V$ ein Endomorphismus, $p, q \in \mathbb{K}[X]$ mit $\gcd(p, q) = 1$. Dann gilt $\ker p(h) \cap \ker q(h) = \{0\}$.

Beweis. Sei $x \in \ker p(h) \cap \ker q(h)$ beliebig. Zu zeigen: $x = 0$.

Aus $x \in \ker p(h) \cap \ker q(h)$ folgt $p(h)(x) = q(h)(x) = 0$. Wegen Satz 79 folgt daraus

$$(up + vq)(h)(x) = 0$$

für beliebige Polynome $u, v \in \mathbb{K}[X]$.

Aus Satz 72 folgt wegen $\gcd(p, q) = 1$, dass es Polynome $u, v \in \mathbb{K}[X]$ gibt mit $up + vq = 1$. Daher gilt $1(h)(x) = 0$, und wegen $1(h) = \mathrm{id}$ folgt $x = 0$. ∎

Beispiel. Wie im vorherigen Beispiel sei $V = C^\infty([0,1], \mathbb{R})$. Für zwei lineare Differentialgleichungen mit konstanten Koeffizienten,

$$p_0 f(x) + p_1 f'(x) + \cdots + p_r f^{(r)}(x) = 0$$
$$q_0 f(x) + q_1 f'(x) + \cdots + q_s f^{(s)}(x) = 0$$

mit $p_0, \dots, p_r, q_0, \dots, q_r \in \mathbb{R}$ sagt der Satz, dass es keine gemeinsamen Lösungen geben kann, wenn die Polynome $p_0 + p_1 X + \cdots + p_r X^r$ und $q_0 + q_1 X + \cdots + q_s X^s$ teilerfremd sind.

Im Fall $\dim V < \infty$ gibt es für jeden Endomorphismus $h\colon V \to V$ nach Satz 80 ein annihilierendes Polynom $a \in \mathbb{K}[X]$. Für dieses gilt $a(h) = 0$ und damit $\ker a(h) = V$. Für jedes zu a teilerfremde Polynom p muss deshalb gelten $\ker p(h) = \{0\}$. Allgemeiner gilt für jedes Polynom p und jedes annihilierende Polynom a von h die Beziehung $\ker p(h) \subseteq \ker \gcd(a,p)(h)$, denn für gewisse $u, v \in \mathbb{K}[X]$ gilt $\gcd(a,p) = ua + vp$, und daher

$$\gcd(a,p)(h) = (ua + vp)(h) = (vp)(h) = v(p(h)),$$

04-07 und daraus folgt $\gcd(a,p)(h)(x) = 0$ für alle $x \in V$ mit $p(h)(x) = 0$.

Am schärfsten wird diese Überlegung, wenn a das Minimalpolynom ist, weil das unter allen annihilierenden Polynomen von h das mit den wenigsten Teilern ist. Um nichttriviale h-invariante Unterräume von V zu finden, genügt es also, die Teiler des Minimalpolynoms (ggf. mit erhöhter Vielfachheit) zu betrachten.

Beispiel.

1. Betrachte $A = \begin{pmatrix} 4 & 1 & 0 & & \\ 0 & 4 & 1 & & \\ 0 & 0 & 4 & & \\ & & & 7 & 1 \\ & & & 0 & 7 \end{pmatrix}$ und $h\colon \mathbb{Q}^5 \to \mathbb{Q}^5$, $h(x) = Ax$.

   Das Minimalpolynom von h ist $m = (X-4)^3(X-7)^2$ und das charakteristische Polynom ist $\chi = (4-X)^3(7-X)^2$.

   Es gibt folgende h-lineare Unterräume:

$$\ker(h - 4\,\mathrm{id}) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\rangle, \qquad \ker(h - 7\,\mathrm{id}) = \left\langle \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle,$$

$$\ker(h - 4\,\mathrm{id})^2 = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\rangle, \qquad \ker(h - 7\,\mathrm{id})^2 = \left\langle \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle,$$

$$\ker(h - 4\,\mathrm{id})^3 = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle, \qquad \ker(h - 7\,\mathrm{id})^3 = \left\langle \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle,$$

$$\ker(h - 4\,\mathrm{id})^4 = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle, \qquad \ker(h - 7\,\mathrm{id})^4 = \left\langle \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle,$$

$\vdots$ $\vdots$

Für höhere Exponenten ändert sich nichts mehr, d.h. es gilt $\ker(h-4\,\mathrm{id})^m = \ker(h-4\,\mathrm{id})^3$ für alle $m \geq 3$ und $\ker(h-7\,\mathrm{id})^m = \ker(h-7\,\mathrm{id})^2$ für alle $m \geq 2$.

Alle h-invarianten Unterräume von V lassen sich schreiben als Summen $\ker(h-4\,\mathrm{id})^i \oplus \ker(h-7\,\mathrm{id})^j$ für gewisse $i, j \in \mathbb{N}$.

2. Betrachte $A = \begin{pmatrix} 4 & \mathbf{0} & 0 & & \\ 0 & 4 & 1 & & \\ 0 & 0 & 4 & & \\ & & & 7 & 1 \\ & & & 0 & 7 \end{pmatrix}$ und $h\colon \mathbb{Q}^5 \to \mathbb{Q}^5$, $h(x) = Ax$.

   Die Matrix unterscheidet sich von der Matrix im vorherigen Beispiel nur im Eintrag $(1,2)$. Für die invarianten Unterräume zum Eigenwert 4 ergibt sich nun:

$$\ker(h-4\,\mathrm{id}) = \langle \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix} \rangle,$$

$$\ker(h-4\,\mathrm{id})^2 = \langle \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} \rangle,$$

$$\ker(h-4\,\mathrm{id})^3 = \langle \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} \rangle,$$

$\vdots$

   Für den Eigenwert 7 ändert sich nichts.

Für jeden Teiler p des Minimalpolynoms gibt es eine Kette

$$\{0\} \subseteq \ker p(h) \subseteq \ker p(h)^2 \subseteq \ker p(h)^3 \subseteq \cdots \subseteq V$$

von h-invarianten Unterräumen. Wegen $\dim V < \infty$ muss es Exponenten $s \in \mathbb{N}$ geben mit $\ker p(h)^s = \ker p(h)^{s+1}$. Für ein solches s muss dann aber sogar schon gelten

$$\ker p(h)^{s+1} = \ker p(h)^{s+2} = \ker p(h)^{s+3} = \cdots .$$

(Beweis: Übung.) Die Kette stabilisiert sich also schon beim ersten Exponenten, bei dem nichts Neues dazukommt.

Um Basen für die Räume $\ker p(h)^s$ auszurechnen, ist es nicht nötig, zuerst die Matrix zu $p(h)^s$ auszurechnen und dann deren Kern zu bestimmen. Stattdessen kann man ausnutzen, dass $x \in \ker p(h)^s \iff$ 04-04

$p(h)(x) \in \ker p(h)^{s-1}$ gilt. Wenn man also schon eine Basis $\{b_1, \dots, b_k\}$ von $\ker p(h)^{s-1}$ kennt, dann kann man eine Basis für $\ker p(h)^s$ dadurch ausrechnen, dass man das lineare Gleichungssystem

$$p(h)x = \beta_1 b_1 + \cdots + \beta_k b_k$$

nach x und $\beta_1, \dots, \beta_k$ löst.

Beispiel. Im ersten Teil des vorherigen Beispiels war

$$\ker(h - 4\,\mathrm{id})^2 = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \rangle.$$

Wir wollen eine Basis von $\ker(h - 4\,\mathrm{id})^3$ bestimmen. Nach der obigen Bemerkung sind die Elemente $x \in \ker(h - 4\,\mathrm{id})^3$ genau jene Vektoren $x = (x_1, x_2, x_3, x_4, x_5)$, für die es $\beta_1, \beta_2 \in \mathbb{K}$ gibt mit

$$\begin{pmatrix} 0 & 1 & 0 & & \\ 0 & 0 & 1 & & \\ 0 & 0 & 0 & & \\ & & & 3 & 1 \\ & & & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \beta_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \beta_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Das lineare Gleichungssytem

$$\begin{pmatrix} 0 & 1 & 0 & & & -1 & 0 \\ 0 & 0 & 1 & & & 0 & -1 \\ 0 & 0 & 0 & & & 0 & 0 \\ & & & 3 & 1 & 0 & 0 \\ & & & 0 & 3 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ \beta_1 \\ \beta_2 \end{pmatrix} = 0$$

hat den Lösungsraum

$$\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \rangle.$$

Die gesuchte Basis von $\ker(h - 4\,\mathrm{id})^3$ erhält man, wenn man aus diesen Vektoren die letzten beiden Koordinaten (also die, die zu β_1, β_2 gehören) streicht.

30 Die Jordan-Normalform

Man kann die Überlegungen aus dem vorherigen Abschnitt noch weiter treiben. Betrachten wir einen endlich-dimensionalen $\mathbb{K}$-Vektorraum V und einen Endomorphismus $h\colon V \to V$, dessen charakteristisches Polynom χ nur Faktoren vom Grad 1 hat, d.h.

$$\chi = (\lambda_1 - X)^{e_1} \cdots (\lambda_k - X)^{e_k}$$

für gewisse paarweise verschiedene $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ und gewisse $e_1, \ldots, e_k \in \mathbb{N} \setminus \{0\}$.

Aus Satz 87 folgt, dass es eine Basis von V gibt, bezüglich der die Abbildungsmatrix von h die Form

$$M = \begin{pmatrix} \boxed{M_1} & & & \\ & \boxed{M_2} & & \\ & & \ddots & \\ & & & \boxed{M_k} \end{pmatrix}$$

hat, wobei $M_i \in \mathbb{K}^{e_i \times e_i}$ eine Matrix mit charakteristischem Polynom $(\lambda_i - X)^{e_i}$ ist $(i = 1, \ldots, k)$.

Den Blöcken $M_1, \ldots, M_k$ von M entsprechen h-invariante Unterräume von $\mathbb{K}^n$. Man kann sich jetzt fragen, ob sich diese Räume noch weiter zerlegen lassen, und wenn ja, wie sich diese Zerlegbarkeit auf die möglichen Darstellungen der Blöcke M_i auswirkt.

Die Einzelheiten der weiteren Herleitung sind relativ technisch und ansonsten nicht mehr allzu interessant. Wir geben deshalb ohne Beweis direkt das Endresultat an.

Satz 90. (Jordan-Normalform) Sei V ein $\mathbb{K}$-Vektorraum, $\dim V = n < \infty$, $h\colon V \to V$ ein Endomorphismus, dessen charakteristisches Polynom χ sich schreiben lässt als

$$\chi = (\lambda_1 - X)^{e_1} \cdots (\lambda_k - X)^{e_k}$$

für gewisse paarweise verschiedene $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ und gewisse $e_1, \ldots, e_k \in \mathbb{N} \setminus \{0\}$.

Dann gibt es eine geordnete Basis B von V, so dass die Abbildungsmatrix M von h bezüglich B die Form 03-11

$$J = \begin{pmatrix} \boxed{J_1} & & & \\ & \boxed{J_2} & & \\ & & \ddots & \\ & & & \boxed{J_k} \end{pmatrix}$$

hat, wobei jedes $J_i \in \mathbb{K}^{e_i \times e_i}$ die Form

$$J_i = \begin{pmatrix} \boxed{J_{i,1}} & & & \\ & \boxed{J_{i,2}} & & \\ & & \ddots & \\ & & & \boxed{J_{i,\ell_i}} \end{pmatrix}$$

hat, wobei jedes $J_{i,j}$ die Form

$$J_{i,j} = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}$$

hat. (Man nennt $J_{i,j}$ ein *Jordan-Kästchen* zum Eigenwert λ_i.)

Dabei gilt:

1. Die Anzahl der Jordan-Kästchen der Größe $m \times m$ im Block J_i ist genau 04-07

$$-\dim\ker(h-\lambda_i\,\mathrm{id})^{m+1}+2\dim\ker(h-\lambda_i\,\mathrm{id})^{m}-\dim\ker(h-\lambda_i\,\mathrm{id})^{m-1}$$

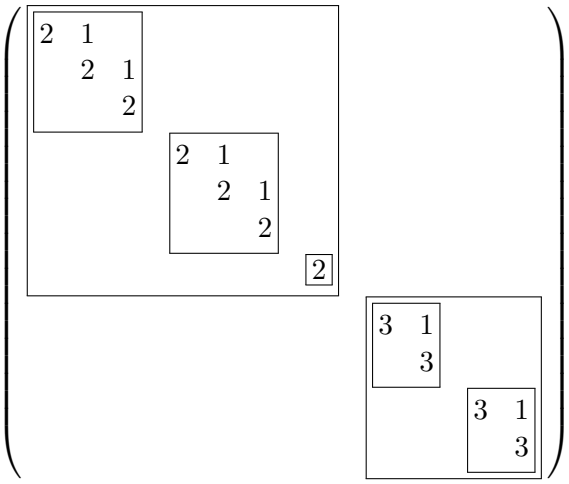
   $(m \in \{1,\dots,e_i\})$.

2. Die Gesamtzahl ℓ_i der Jordan-Kästchen im Block J_i ist $\dim E_{\lambda_i}$.

3. Ist s_i die Vielfachheit von $X-\lambda_i$ im Minimalpolynom von h, so gibt es in M_i mindestens
03-11 ein Jordan-Kästchen der Größe $s_i \times s_i$ und es gibt kein größeres.

4. Sind B, B' zwei Basen, für die die Abbildungsmatrizen J, J' diese Form haben, dann unterscheiden sich J, J' höchstens in der Reihenfolge der Jordan-Kästchen.

Beispiel. Die Matrix



hat die in Satz 90 beschriebene Form. Der Block zum Eigenwert 2 hat drei Jordan-Kästchen: zwei der Größe 3×3 und eins der Größe 1×1. Der Block zum Eigenwert 3 hat zwei Jordan-Kästchen, beide der Größe 2×2.

Das Minimalpolynom lässt sich direkt ablesen zu $(X-2)^3(X-3)^2$. Das charakteristische Polynom lautet $(2-X)^7(3-X)^4$.

Für die Dimensionen der Eigenäume gilt $\dim E_2 = 3$ und $\dim E_3 = 2$.

Die Form aus Satz 90 heißt *Jordan-Normalform* von h. Mit der Jordan-Normalform einer Matrix $A \in \mathbb{K}^{n\times n}$ ist eine Jordan-Normalform des Endomorphismus $h\colon \mathbb{K}^n \to \mathbb{K}^n$, $h(x) = Ax$ gemeint.

Nach dem Satz existiert eine Jordan-Normalform für jede Matrix, deren charakteristisches Polynom vollständig faktorisiert. Wenn $\mathbb{K}$ algebraisch abgeschlossen ist, gilt das für jede Matrix.

Um zu einer gegebenen Matrix $A \in \mathbb{K}^{n\times n}$ eine geordnete Basis $B \in \mathbb{K}^{n\times n}$ zu bestimmen, für die
04-05 $B^{-1}AB$ eine Jordan-Normalform ist, kann man folgenden Algorithmus verwenden.

Algorithmus 6. Eingabe: $A = ((a_{i,j}))_{i,j=1}^{n} \in \mathbb{K}^{n\times n}$, so dass das charakteristische Polynom von A die Form $\chi = (\lambda_1 - X)^{e_1}\cdots(\lambda_k - X)^{e_k}$ hat für $e_1,\dots,e_k \in \mathbb{N}\setminus\{0\}$ und paarweise verschiedene $\lambda_1,\dots,\lambda_k \in \mathbb{K}$.

Ausgabe: Eine Jordan-Normalform $J \in \mathbb{K}^{n\times n}$ sowie eine invertierbare Matrix $B \in \mathbb{K}^{n\times n}$ mit
04-05 $B^{-1}AB = J$

1 setze J auf die leere Matrix und B auf die leere Liste.

2 für $i = 1, \ldots, k$:

3 bestimme $E_{\lambda_i}^{(m)} := \ker(A - \lambda_i I_n)^m$ für $m = 0, \ldots, e_i + 1$. 04-05

4 sei m maximal mit $\dim E_{\lambda_i}^{(m)} > \dim E_{\lambda_i}^{(m-1)}$

5 solange $m > 0$, wiederhole:

6 bestimme eine Basis $\{b_1, \ldots, b_d\}$ eines Raums $U \subseteq \mathbb{K}^n$ mit der Eigenschaft

$$E_{\lambda_i}^{(m)} = U \oplus \big((A - \lambda_i I_n)E_{\lambda_i}^{(m+1)} + E_{\lambda_i}^{(m-1)}\big).$$

7 für $j = 1, \ldots, d$:

8 ergänze J um ein Jordan-Kästchen für λ_i der Größe $m \times m$

9 ergänze B um $(A - \lambda_i I_n)^{m-1} b_j,\ (A - \lambda_i I_n)^{m-2} b_j,\ \ldots,\ (A - \lambda_i I_n) b_j,\ b_j$

10 $m = m - 1$

11 gib J und B als Ergebnis aus.

Beispiel. Wir wenden den Algorithmus auf die Matrix

$$A = \begin{pmatrix} 3 & 1 & -3 & -5 & -9 & -4 \\ 1 & 4 & 5 & 2 & 7 & 3 \\ -3 & -1 & 10 & 12 & 23 & 10 \\ -3 & 2 & 15 & 18 & 35 & 15 \\ 0 & -1 & -8 & -7 & -14 & -7 \\ 6 & 1 & -7 & -14 & -26 & -9 \end{pmatrix} \in \mathbb{Q}^{6\times 6}$$

an. Das charakteristische Polynom von A ist $(2 - X)^6$, also brauchen wir die in Zeile 2 beginnende Schleife nur einmal zu durchlaufen.

Als invariante Teilräume erhalten wir

$$E_2^{(0)} = \{0\}$$

$$E_2^{(1)} = \langle \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 29 \\ 16 \\ -36 \\ 0 \\ 17 \\ 0 \end{pmatrix}, \begin{pmatrix} 11 \\ 9 \\ -16 \\ 0 \\ 0 \\ 17 \end{pmatrix} \rangle$$

$$E_2^{(2)} = \langle \begin{pmatrix} -1 \\ 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 2 \end{pmatrix} \rangle$$

$$E_2^{(3)} = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \rangle = \mathbb{Q}^6.$$

Ohne weitere Rechnung ergibt sich an dieser Stelle $E_2^{(3)} = \cdots = E_2^{(6)} = E_2^{(7)} = \mathbb{Q}^6$.

Nach Satz 90 muss es $\dim E_2 = \dim E_2^{(1)} = 3$ Jordan-Kästchen geben, und mindestens eines davon muss die Länge $m = 3$ haben. Da sich die Kästchenlängen insgesamt zu 6 aufaddieren müssen, ist an dieser Stelle bereits klar, wie die Jordan-Normalform aussehen wird: Sie hat je ein Kästchen der Längen 3, 2 und 1. Wenn wir nicht auch an der Basis B interessiert wären, könnten wir also schon aufhören.

Um die Basis B zu berechnen, führen wir die Schleife aus, die in Zeile 5 beginnt. Es ist

$$\begin{aligned}(A-2I_6)E_2^{(4)} + E_2^{(2)} &= \langle \underbrace{\begin{pmatrix}1\\1\\-3\\-3\\0\\6\end{pmatrix}, \ldots, \begin{pmatrix}-4\\3\\10\\15\\-7\\-11\end{pmatrix}}_{\text{von } (A-2I_6)E_2^{(4)}}, \underbrace{\begin{pmatrix}-1\\2\\0\\0\\0\\0\end{pmatrix}, \ldots, \begin{pmatrix}-1\\0\\0\\0\\0\\2\end{pmatrix}}_{\text{von } E_2^{(2)}} \rangle \\ &= \langle \begin{pmatrix}-1\\2\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}-3\\0\\2\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix} \begin{pmatrix}-1\\0\\0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\0\\0\\0\\2\end{pmatrix} \rangle\end{aligned}$$

und wir brauchen einen Komplementärraum dieses Raums in $E_2^{(3)}$. Im vorliegenden Fall ist das einfach, weil $E_2^{(3)} = \mathbb{Q}^6$ schon der ganze Raum ist. Eine mögliche Wahl ist

$$E_2^{(3)} = \langle \underbrace{\begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix}}_{=b_1} \rangle \oplus \big((A-2I_6)E_2^{(4)} + E_2^{(2)}\big).$$

Daraus ergeben sich die ersten drei Basisvektoren wie folgt:

$$(A-2I_6)^2\, b_1 = \begin{pmatrix}2\\0\\-4\\-4\\2\\4\end{pmatrix}, \quad (A-2I_6)^1\, b_1 = \begin{pmatrix}1\\1\\-3\\-3\\0\\6\end{pmatrix}, \quad (A-2I_6)^0\, b_1 = \begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix}.$$

Nächster Schleifendurchlauf: $m = 2$. Es ist

$$(A - 2I_6)E_2^{(3)} + E_2^{(1)} = \langle \underbrace{\begin{pmatrix} 1 \\ 1 \\ -3 \\ -3 \\ 0 \\ 6 \end{pmatrix}, \ldots, \begin{pmatrix} -4 \\ 3 \\ 10 \\ 15 \\ -7 \\ -11 \end{pmatrix}}_{\text{von } (A-2I_6)E_2^{(3)}}, \underbrace{\begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 29 \\ 16 \\ -36 \\ 0 \\ 17 \\ 0 \end{pmatrix}, \begin{pmatrix} 11 \\ 9 \\ -16 \\ 0 \\ 0 \\ 17 \end{pmatrix}}_{\text{von } E_2^{(1)}} \rangle$$

$$= \langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 5 \\ -12 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ -3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 4 \\ -11 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -2 \\ 4 \end{pmatrix} \rangle$$

und wir brauchen einen Komplementärraum dieses Raums in $E_2^{(2)}$. Der Raum $E_2^{(2)}$ ist echt kleiner als $\mathbb{Q}^6$, aber wir können einen geeigneten Raum U finden, indem wir zunächst wie üblich einen Komplementärraum durch Ergänzung zu einer Basis von $\mathbb{Q}^6$ bestimmen, und dann den Schnitt dieses Komplementärraums mit $E_2^{(2)}$ nehmen.

Aus der oben angegebenen Basis lässt sich ablesen, dass $\{e_5, e_6\}$ eine Basis für einen Komplementärraum in $\mathbb{Q}^6$ ist. Durch den Schnitt dieses Raums mit $E_2^{(2)}$ erhalten wir

$$E_2^{(2)} = \langle \underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ -2 \end{pmatrix}}_{=b_1} \rangle \oplus \big((A - 2I_6)E_2^{(3)} + E_2^{(1)}\big).$$

Daraus ergeben sich die nächsten beiden Basiselemente wie folgt:

$$(A - 2I_6)^1 b_1 = \begin{pmatrix} -1 \\ 1 \\ 3 \\ 5 \\ -2 \\ -4 \end{pmatrix}, \quad (A - 2I_6)^0 b_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ -2 \end{pmatrix}.$$

Letzter Schleifendurchlauf: $m = 1$. Es ist

$$(A - 2I_6)E_2^{(2)} + E_2^{(0)} = \langle \begin{pmatrix} -9 \\ 5 \\ 23 \\ 33 \\ -14 \\ -28 \end{pmatrix}, \begin{pmatrix} -10 \\ 6 \\ 26 \\ 38 \\ -16 \\ -32 \end{pmatrix}, \begin{pmatrix} -5 \\ 2 \\ 12 \\ 16 \\ -7 \\ -14 \end{pmatrix}, \begin{pmatrix} -9 \\ 7 \\ 25 \\ 39 \\ -16 \\ -32 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 1 \\ 7 \\ -2 \\ -4 \end{pmatrix} \rangle = \langle \begin{pmatrix} 1 \\ 0 \\ -2 \\ -2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 3 \\ -1 \\ -2 \end{pmatrix} \rangle.$$

Wegen

$$E_2^{(1)} = \langle \begin{pmatrix} 0 \\ 0 \\ 3 \\ 5 \\ 2 \\ -13 \end{pmatrix} \rangle \oplus \left((A - 2I_6)E_2^{(2)} + E_2^{(0)}\right)$$

ist $(0, 0, 3, 5, 2, -13)$ eine mögliche Wahl für das letzte Basiselement.
Als Endergebnis erhalten wir

$$B = \begin{pmatrix} 2 & 1 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ -4 & -3 & 0 & 3 & 0 & 3 \\ -4 & -3 & 0 & 5 & 0 & 5 \\ 2 & 0 & 0 & -2 & 1 & 2 \\ 4 & 6 & 0 & -4 & -2 & -13 \end{pmatrix}$$

04-05 und als Jordan-Normalform ergibt sich, wie erwartet,

$$J = B^{-1}AB = \begin{pmatrix} \boxed{\begin{matrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{matrix}} & & \\ & \boxed{\begin{matrix} 2 & 1 \\ 0 & 2 \end{matrix}} & \\ & & \boxed{2} \end{pmatrix}.$$

Für diagonalisierbare Matrizen A hatten wir gesehen, dass sie sich leicht potenzieren lassen. Mit Hilfe der Jordan-Normalform lässt sich das verallgemeinern. Ist nämlich J die Jordan-Normalform der Matrix A und ist B so, dass $A = B^{-1}JB$, dann gilt $A^n = B^{-1}J^nB$. Zur Potenzierung von J überlegt man sich zunächst, dass für beliebige Blockdiagonalmatrizen und für jedes $n \in \mathbb{N}$ gilt

$$\begin{pmatrix} \boxed{M_1} & & & \\ & \boxed{M_2} & & \\ & & \ddots & \\ & & & \boxed{M_k} \end{pmatrix}^n = \begin{pmatrix} \boxed{M_1^n} & & & \\ & \boxed{M_2^n} & & \\ & & \ddots & \\ & & & \boxed{M_k^n} \end{pmatrix}.$$

Es genügt deshalb, sich zu überlegen, wie die n-te Potenz eines Jordan-Kästchens aussieht. Das haben wir bereits im Beispiel nach Satz 86 getan.

Satz 91. Sei $A \in \mathbb{K}^{n\times n}$ so, dass das charakteristische Polynom keine irreduziblen Faktoren vom Grad ≥ 2 enthält. Dann gilt: A ist genau dann diagonalisierbar, wenn das Minimalpolynom von A keine mehrfachen Nullstellen hat.

Beweis. Folgt direkt aus Satz 90. ∎

Teil VI

Skalarprodukte

31 Metrische, normierte und euklidische Räume

In diesem und in den folgenden Abschnitten betrachten wir nur Vektorräume über dem Körper $\mathbb{K} = \mathbb{R}$. Die meisten Aussagen lassen sich (evtl. mit kleinen Anpassungen) auf Vektorräume über $\mathbb{K} = \mathbb{C}$ übertragen. Wir werden von der Betragsfunktion $|\cdot|$ auf $\mathbb{R}$ (bzw. $\mathbb{C}$) Gebrauch machen und von der Eigenschaft, dass $\mathbb{R}$ ein angeordneter Körper ist (d.h. es ist auf $\mathbb{R}$ eine Totalordnung $\leq$ erklärt, die mit den Körperoperationen in gewohnter Weise verträglich ist). Für $z = x + \mathrm{i}y$ mit $x, y \in \mathbb{R}$ ist $|z| := \sqrt{x^2 + y^2}$ definiert. Man beachte, dass auch für komplexe Zahlen $z \in \mathbb{C}$ der Betrag $|z|$ stets reell und nicht-negativ ist.

Eine nützliche Eigenschaft von $\mathbb{R}$ ist, dass für jedes $x \in \mathbb{R}$ mit $x \geq 0$ und jedes $p \in \mathbb{N} \setminus \{0\}$ genau ein $y \in \mathbb{R}$ mit $y \geq 0$ und $y^p = x$ existiert. Man bezeichnet dieses $y := \sqrt[p]{x}$ bekanntlich als die p-te *Wurzel* (engl. *root*) von x.

Definition 60.

04-07 1. Sei M eine Menge und $d\colon M \times M \to \mathbb{R}$ eine Funktion, so dass für alle $x, y, z \in M$ gilt:

$$\begin{aligned} &d(x,y) \geq 0 \\ &d(x,y) = 0 \iff x = y \\ &d(x,y) = d(y,x) \\ &d(x,y) + d(y,z) \geq d(x,z). \end{aligned}$$

Dann heißt d eine *Metrik* auf M und das Paar (M, d) heißt ein *metrischer Raum.*

2. Ein metrischer Raum (M, d), bei dem M ein Vektorraum über $\mathbb{R}$ ist, heißt *metrischer Vektorraum.*

3. Sei V ein $\mathbb{R}$-Vektorraum und $\|\cdot\|\colon V \to \mathbb{R}$ eine Funktion, so dass für alle $x, y \in V$ und
04-07 alle $\alpha \in \mathbb{R}$ gilt:

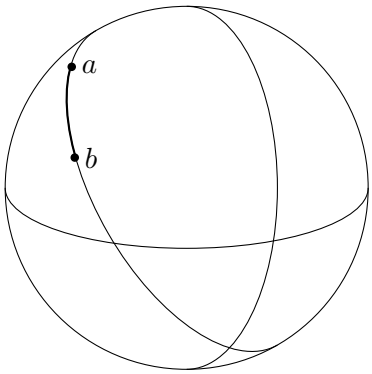
$$\begin{aligned} &\|x\| = 0 \iff x = 0 \\ &\|\alpha x\| = |\alpha| \, \|x\| \\ &\|x + y\| \leq \|x\| + \|y\|. \end{aligned}$$

Dann heißt $\|\cdot\|$ eine *Norm* auf V und das Paar $(V, \|\cdot\|)$ heißt ein *normierter Raum.*

Beispiel.

1. Die Hamming-Distanz aus Abschnitt 23 ist eine Metrik auf $M = \mathbb{Z}_2^n$. Sie macht $(\mathbb{Z}_2^n, d)$ zu einem metrischen Raum, aber nicht zu einem metrischen Vektorraum, weil $\mathbb{Z}_2^n$ zwar ein Vektorraum, aber kein $\mathbb{R}$-Vektorraum ist.

2. Sei $K = \{\, (x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1 \,\}$ die Oberfläche der Einheitskugel. Für je zwei
04-07 verschiedene Punkte $a, b \in K$ gibt es genau einen Kreis mit Mittelpunkt $(0, 0, 0)$ durch a und b. Die Punkte a und b teilen diese Kreislinie in zwei Kreisbogenstücke. Definiert man $d(a, b)$ als die Länge des kürzeren Bogenstücks, so ist $d\colon K \times K \to \mathbb{R}$ eine Metrik.



3. Ist V ein $\mathbb{R}$-Vektorraum und d eine Metrik auf V, so kann man $d(x, y)$ als den Abstand zwischen x und y interpretieren. Für $V = \mathbb{R}^2$ ist zum Beispiel durch

$$d(\ (x_1, x_2), (y_1, y_2)\ ) := \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

eine Metrik erklärt.

4. Ist V ein $\mathbb{R}$-Vektorraum und $\|\cdot\|$ eine Norm auf V, so kann man $\|x\|$ als die Länge des Vektors $x \in V$ interpretieren. Für $V = \mathbb{R}^2$ ist zum Beispiel durch 04-05

$$\|(x_1, x_2)\| := \sqrt{x_1^2 + x_2^2}$$

eine Norm definiert.

Für jede Norm $\|\cdot\|$ auf V ist $d\colon V \times V \to \mathbb{R}$, $d(x, y) := \|x - y\|$ eine Metrik. Allerdings kommt nicht jede Metrik in dieser Weise von einer Norm.

5. Sei $(V, \|\cdot\|)$ ein normierter Raum und

$$d\colon V \times V \to \mathbb{R}, \quad d(x, y) := \begin{cases} 0 & \text{falls } x = y \\ \|x\| + \|y\| & \text{sonst} \end{cases}$$

Dann ist d eine Metrik auf V, aber es gibt keine Norm $\|\cdot\|'$ auf V, so dass $d(x, y) = \|x - y\|'$ für alle $x, y \in V$ gelten würde, denn für eine solche Norm müsste gelten

$$\tfrac{3}{2}\|x\| = \|x\| + \tfrac{1}{2}\|x\| = d(x, \tfrac{1}{2}x) = \|x - \tfrac{1}{2}x\|' = d(x - \tfrac{1}{2}x, 0) = \tfrac{1}{2}\|x\|,$$

was für $x \neq 0$ nicht stimmen kann.

6. Sei $V \subseteq \mathbb{R}^{\mathbb{N}}$ der Vektorraum aller Folgen $(a_n)_{n=0}^{\infty}$, für die die Reihe $\sum_{n=0}^{\infty} a_n^2$ konvergiert. Dann wird durch

$$\|(a_n)_{n=0}^{\infty}\| := \sum_{n=0}^{\infty} a_n^2$$

eine Norm auf V erklärt.

7. Für jedes $p \in \mathbb{N} \setminus \{0\}$ wird durch

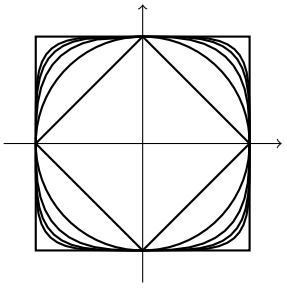
$$\|(x_1, \ldots, x_n)\|_p := \sqrt[p]{|x_1|^p + \cdots + |x_n|^p}$$

eine Norm auf $\mathbb{R}^n$ erklärt. Auch

$$\|(x_1, \ldots, x_n)\|_\infty := \max(|x_1|, \ldots, |x_n|)$$

definiert eine Norm.

Für eine Norm $\|\cdot\|$ auf einem $\mathbb{R}$-Vektorraum V nennt man die Menge $K := \{x \in V : \|x\| = 1\}$ die *Einheitskugel(schale)* (engl. *unit sphere*) in V bezüglich der Norm $\|\cdot\|$. Für $V = \mathbb{R}^2$ und die Normen $\|\cdot\|_1, \ldots, \|\cdot\|_5$ und $\|\cdot\|_\infty$ haben diese Mengen die folgende Gestalt:



8. Für jedes $p \in \mathbb{N} \setminus \{0\}$ wird auf dem Raum $V = C([0,1], \mathbb{R})$ durch

$$\|f\|_p := \sqrt[p]{\int_0^1 |f(x)|^p dx}$$

eine Norm definiert.

9. Sei $V = \mathbb{R}^{n\times m}$. Sei $\|\cdot\|_n$ eine Norm auf $\mathbb{R}^n$ und $\|\cdot\|_m$ eine Norm auf $\mathbb{R}^m$. Für $A \in V$ sei definiert

$$\|A\| = \sup\{\, \|Ax\|_n : x \in K \,\},$$

wobei K die Einheitskugelschale in $\mathbb{R}^m$ bezüglich der Norm $\|\cdot\|_m$ ist. $\|A\|$ ist also die Länge des längsten Vektors im Zielraum $\mathbb{R}^n$, der das Bild eines Vektors auf der Einheitskugel in $\mathbb{R}^m$ ist.

Im Sinn der Analysis ist die Menge K kompakt und die Abbildung $x \mapsto \|Ax\|_m$ stetig. Deshalb ist das Supremum sogar ein Maximum und es gibt einen Vektor $v \in \mathbb{R}^m$ mit $\|v\|_m = 1$ und $\|Av\|_n = \|A\|$.

Die Abbildung $\|\cdot\|$ ist eine Norm auf V.

Definition 61. Sei V ein $\mathbb{R}$-Vektorraum. Eine Funktion $\langle\cdot|\cdot\rangle\colon V \times V \to \mathbb{R}$ heißt *Skalarprodukt* (engl. *scalar product*) auf V, falls für alle $x, y, x_1, x_2 \in V$ und alle $\alpha_1, \alpha_2 \in \mathbb{R}$ gilt:

1. $\langle x|x\rangle \geq 0$ und $\langle x|x\rangle = 0 \iff x = 0$
2. $\langle x|y\rangle = \langle y|x\rangle$
3. $\langle \alpha_1 x_1 + \alpha_2 x_2|y\rangle = \alpha_1\langle x_1|y\rangle + \alpha_2\langle x_2|y\rangle$

Ist $\langle\cdot|\cdot\rangle$ ein Skalarprodukt auf V, so heißt $(V, \langle\cdot|\cdot\rangle)$ ein *Skalarproduktraum* oder ein *euklidischer (Vektor-)Raum.*

Statt $(V, \langle\cdot|\cdot\rangle)$ schreibt man auch einfach V.

Beispiel.

1. Im Raum $\mathbb{R}^n$ wird durch

$$\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} | \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rangle := x_1 y_1 + \cdots + x_n y_n$$

ein Skalarprodukt erklärt. Man nennt es das *Standardskalarprodukt.* Wenn in konkreten Beispielen in $\mathbb{R}^n$ das Skalarprodukt nicht explizit angegeben ist, wird meist dieses gemeint sein.

Für $x, y \in \mathbb{R}^n$ ist das Standardskalarprodukt gerade die Matrixmultiplikation xy, wobei x als Zeilen- und y als Spaltenvektor aufgefasst wird. Wir werden für dieses Skalarprodukt deshalb auch die Notationen xy und $x \cdot y$ verwenden.

2. Auf $\ell^2 := \{ (a_n)_{n=0}^\infty \in \mathbb{R}^{\mathbb{N}} : \sum_{n=0}^\infty a_n^2 < \infty \}$ wird durch

$$\langle (a_n)_{n=0}^\infty | (b_n)_{n=0}^\infty \rangle := \sum_{n=0}^\infty a_n b_n$$

ein Skalarprodukt definiert. (Beachte, dass auf der rechten Seite stets eine konvergente Reihe steht.)

3. Sei $V = \mathbb{R}^3$ und $A = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Durch

$$\langle x | y \rangle := xAy$$

wird auf V ein Skalarprodukt erklärt, das vom Standardskalarprodukt verschieden ist. Im Ausdruck xAy ist dabei x als Zeilenvektor und y als Spaltenvektor zu verstehen, so dass $xAy = x \cdot A \cdot y$ im Sinn der Matrixmultiplikation eine Zahl ergibt.

Dass es sich bei $\langle \cdot | \cdot \rangle$ um ein Skalarprodukt handelt, erkennt man durch Nachrechnen der nötigen Gesetze: die Linearität ist offensichtlich, und die Symmetrie folgt direkt daraus, dass A eine symmetrische Matrix ist (d.h. $A = A^\top$). Es ist auch klar, dass $\langle 0|0 \rangle = 0$. Weniger offensichtlich ist die Eigenschaften $\langle x|x \rangle \geq 0$ und $\langle x|x \rangle = 0 \Rightarrow x = 0$. Dass diese auch erfüllt sind, folgt aus der Beobachtung

$$\begin{aligned} &(x_1, x_2, x_3)\ A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= 4x_1^2 + 4x_1x_2 + 2x_2^2 + 2x_1x_3 + 2x_2x_3 + x_3^2 \\ &= (x_1 + x_2 + x_3)^2 + (x_1 + x_2)^2 + 2x_1^2, \end{aligned}$$

denn eine Summe von Quadraten ist immer nicht-negativ und wird nur dann Null, wenn alle Summanden Null sind. Und offenbar gilt $x_1 = 0 \wedge x_1 + x_2 = 0 \wedge x_1 + x_2 + x_3 = 0$ nur für $x_1 = x_2 = x_3 = 0$.

4. Sei $w \in C([-1,1],\mathbb{R})$ eine beliebige Funktion mit $w(x) > 0$ für alle $x \in [-1,1]$. Dann wird durch
$$\langle f|g\rangle := \int_{-1}^{1} w(x)f(x)g(x)dx$$
ein Skalarprodukt auf $C([-1,1],\mathbb{R})$ erklärt.

04-11 Bei Vektorräumen über $\mathbb{C}$ statt über $\mathbb{R}$ können Skalarprodukte im allgemeinen komplexe Werte annehmen. Man verlangt aber, dass $\langle x|x\rangle$ für alle x reell ist. Nur dann kann man die erste Bedingung überhaupt formulieren, weil auf $\mathbb{C}$ ja gar keine Ordnungsrelation erklärt ist.

Statt der Symmetrie $\langle x|y\rangle = \langle y|x\rangle$ fordert man die Eigenschaft $\langle x|y\rangle = \overline{\langle y|x\rangle}$, wobei $\overline{u+\mathrm{i}v} := u - \mathrm{i}v$ die komplex-konjugierte Zahl zu $u+\mathrm{i}v \in \mathbb{C}$ ist. Man sagt, die Abbildung $\langle\cdot|\cdot\rangle$ ist *hermitesch.*

Wegen $(u+\mathrm{i}v)(\overline{u+\mathrm{i}v}) = u^2+v^2$ ist für jede komplexe Zahl $z \in \mathbb{C}$ das Produkt $z\bar{z}$ eine nichtnegative reelle Zahl. Für $z \in \mathbb{R} \subseteq \mathbb{C}$ gilt $z = \bar{z}$, so dass sich Hermiteschheit als Verallgemeinerung der üblichen Symmetrie auffassen lässt.

04-07 Das Standardskalarprodukt in $\mathbb{C}^n$ ist definiert durch
$$\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} | \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rangle = x_1\bar{y}_1 + \cdots + x_n\bar{y}_n.$$

Für Skalarprodukte über reellen Räumen folgt aus der Linearität im ersten Argument und der Symmetrie, dass sie auch linear im zweiten Argument sind: $\langle x|\beta_1 y_1 + \beta_2 y_2\rangle = \beta_1\langle x|y_1\rangle + \beta_2\langle x|y_2\rangle$. Für komplexe Skalarproduktträume gilt das nicht! Stattdessen gilt hier nur $\langle x|y_1+y_2\rangle = \langle x|y_1\rangle + \langle x|y_2\rangle$ und $\langle x|\beta y\rangle = \bar{\beta}\langle x|y\rangle$. Das erste und das zweite Argument verhalten sich also leicht unterschiedlich. Dabei ist es eine Konventionsfrage, welches der beiden Argumente das erste und welches das zweite Argument ist. Statt $\langle \alpha x|y\rangle = \alpha\langle x|y\rangle$, $\langle x|\beta y\rangle = \bar{\beta}\langle x|y\rangle$ findet man manchmal auch die Variante $\langle \alpha x|y\rangle = \bar{\alpha}\langle x|y\rangle$, $\langle x|\beta y\rangle = \beta\langle x|y\rangle$.

Wir werden hier weder die eine noch die andere Version propagieren sondern uns auf den Fall reeller Skalarproduktträume beschränken. Im folgenden ist mit einem Skalarproduktraum immer ein reeller Skalarproduktraum im Sinn von Definition 61 gemeint.

Satz 92. Sei V ein Skalarproduktraum und $\|\cdot\|: V \times V \to \mathbb{R}$ definiert durch $\|x\| := \sqrt{\langle x|x\rangle}$. Dann ist $\|\cdot\|$ eine Norm und für alle $x, y \in V$ gilt:

1. $\langle x|y\rangle^2 \leq \langle x|x\rangle\,\langle y|y\rangle$ (Cauchy-Schwarz-Ungleichung)
2. $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$ (Parallelogrammgleichung)
3. $4\langle x|y\rangle = \|x+y\|^2 - \|x-y\|^2$
4. $\langle x|y\rangle = 0 \iff \|x\|^2 + \|y\|^2 = \|x+y\|^2$ (Pythagoras)

Beweis. Dass $\|\cdot\|$ eine Norm ist, kann man selbst überprüfen. Seien $x, y \in V$.

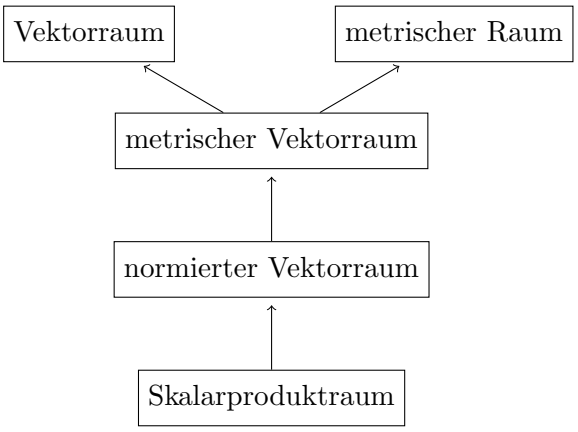
1. Falls $y = 0$ ist, ist die Aussage offensichtlich wahr. Wir betrachten den Fall $y \neq 0$. Dann gilt $\langle y|y\rangle \neq 0$, und für alle $\alpha \in \mathbb{R}$ gilt
$$0 \leq \langle x+\alpha y|x+\alpha y\rangle = \langle x|x\rangle + 2\alpha\langle x|y\rangle + \alpha^2\langle y|y\rangle.$$

Für die spezielle Wahl $\alpha = -\langle x|y\rangle/\langle y|y\rangle$ folgt daraus

$$\langle x|x\rangle - \frac{\langle x|y\rangle^2}{\langle y|y\rangle} \geq 0,$$

und daraus die Behauptung.

2. $\|x+y\|^2 + \|x-y\|^2 = \langle x+y|x+y\rangle + \langle x-y|x-y\rangle = \langle x|x\rangle + 2\langle x|y\rangle + \langle y|y\rangle + \langle x|x\rangle - 2\langle x|y\rangle + \langle y|y\rangle = 2\langle x|x\rangle + 2\langle y|y\rangle = 2\|x\|^2 + 2\|y\|^2$.
3. $\|x+y\|^2 - \|x-y\|^2 = \langle x+y|x+y\rangle - \langle x-y|x-y\rangle = \langle x|x\rangle + 2\langle x|y\rangle + \langle y|y\rangle - \langle x|x\rangle + 2\langle x|y\rangle - \langle y|y\rangle = 4\langle x|y\rangle$.
4. folgt aus $\|x+y\|^2 - \|x\|^2 - \|y\|^2 = \langle x+y|x+y\rangle - \langle x|x\rangle - \langle y|y\rangle = 2\langle x|y\rangle$. ∎



Konvention: Sofern nichts anderes vereinbart ist, werden wir bei Skalarproduktträumen die Notation $\langle\cdot|\cdot\rangle$ für das Skalarprodukt und die Notation $\|\cdot\|$ für die zugehörige Norm verwenden (also $\|x\| = \sqrt{\langle x|x\rangle}$). 04-15

Satz 93. Ist V ein $\mathbb{R}$-Vektorraum und $\langle\cdot|\cdot\rangle\colon V \times V \to \mathbb{R}$ ein Skalarprodukt auf V, dann gibt es ein $v^* \in (V \otimes V)^*$, so dass für alle $x, y \in V$ gilt $\langle x|y\rangle = v^*(x \otimes y)$.

Beweis. Sei B eine Basis von V. Dann ist $B \times B$ eine Basis von $V \otimes V$. Jedes Funktional $v^* \in (V \otimes V)^*$ ist eindeutig festgelegt durch die Werte $v^*(b_1 \otimes b_2)$ für alle $b_1, b_2 \in B$, und für jede Wahl von Werten gibt es genau ein passendes Funktional (Satz 50). Definiere also

$$v^*\colon V \otimes V \to \mathbb{R}$$

durch $v^*(b_1 \otimes b_2) := \langle b_1|b_2\rangle$ für alle $b_1, b_2 \in B$.
Sind dann $x, y \in V$ beliebig, etwa $x = \sum_{i=1}^n \alpha_i b_i$ und $y = \sum_{i=1}^n \beta_i b_i$ für gewisse $b_1, \ldots, b_n \in B$ und $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \mathbb{R}$, so gilt 04-07

$$\begin{aligned} v^*(x \otimes y) &= v^*\Big(\sum_{i=1}^n \alpha_i b_i \ \otimes \ \sum_{i=1}^n \beta_i b_i\Big) = v^*\Big(\sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j (b_i \otimes b_j)\Big) \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j v^*(b_i \otimes b_j) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j \langle b_i|b_j\rangle \end{aligned}$$

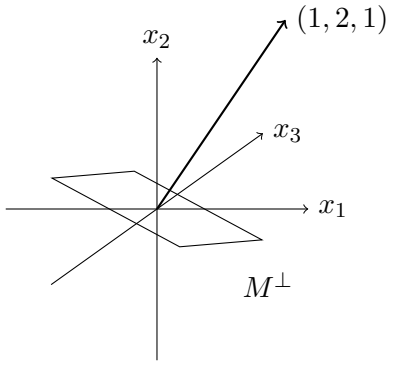
$$= \sum_{i=1}^{n} \alpha_i \langle b_i | \sum_{j=1}^{n} \beta_j b_j \rangle = \langle \sum_{i=1}^{n} \alpha_i b_i | \sum_{j=1}^{n} \beta_j b_j \rangle$$
$$= \langle x|y \rangle. \qquad \blacksquare$$

Definition 62. Sei V ein Skalarproduktraum.

1. Zwei Elemente $x, y \in V$ heißen (zueinander) *orthogonal* oder (aufeinander) *senkrecht* (engl. *perpendicular*), falls $\langle x|y \rangle = 0$ gilt. In diesem Fall schreibt man auch $x \perp y$.
2. Zwei Teilmengen $X, Y \subseteq V$ heißen (zueinander) *orthogonal* oder (aufeinander) *senkrecht,* falls für alle $x \in X$ und alle $y \in Y$ gilt $x \perp y$. In diesem Fall schreibt man $X \perp Y$.
3. Sei M eine Teilmenge von V. Dann heißt
$$M^\perp := \{\, x \in V \mid \forall\, y \in M : x \perp y \,\}$$
das *orthogonale Komplement* von M.
Für $x \in V$ schreibt man statt $\{x\}^\perp$ auch einfach $x^\perp$.

Beispiel.

1. Für alle $x \in V$ gilt $\langle x|0 \rangle = 0$, d.h. der Nullvektor steht senkrecht auf jeden Vektor. Für $x \in V \setminus \{0\}$ gilt $\langle x|x \rangle \neq 0$, d.h. außer dem Nullvektor steht kein Vektor senkrecht auf sich selbst.
Vorsicht: Zwei Ebenen im $\mathbb{R}^3$, die sich „rechtwinklig“ schneiden, wie z.B. die (x, y)-Ebene und die (x, z)-Ebene, stehen **nicht** im Sinn von Teil 2 der Definition senkrecht aufeinander, weil dazu alle Vektoren in der Schnittgeraden (im Bsp. die Punkte auf der x-Achse) senkrecht auf sich selbst stehen müssten. Der einzige Vektor, der auf sich selbst senkrecht steht, ist aber der Nullvektor. Es gilt also $X \perp Y \Rightarrow X \cap Y \subseteq \{0\}$ für alle $X, Y \subseteq V$.
2. Sei $V = \mathbb{R}^3$ und sei $\langle \cdot | \cdot \rangle$ das Standardskalarprodukt. Betrachte $M = \{(1, 2, 1)\}$. Wegen
$$\langle \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \rangle = 0 \iff x_1 + 2x_2 + x_3 = 0 \iff \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \langle \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \rangle$$
ist $M^\perp$ die Ebene, die von $(1, 0, -1)$ und $(-2, 1, 0)$ aufgespannt wird.



3. Umgekehrt: Betrachte die Ebene

$$E = \langle \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \rangle \subseteq \mathbb{R}^3.$$

Ein Vektor $(x_1, x_2, x_3) \in \mathbb{R}$ gehört genau dann zum orthogonalen Komplement $E^\perp$, wenn für alle $\alpha_1, \alpha_2 \in \mathbb{R}$ gilt

$$\langle \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid \alpha_1 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \rangle = 0.$$

Wegen der Linearität des Skalarprodukts sind diese Vektoren (x_1, x_2, x_3) genau jene, für die gilt

$$\begin{aligned} & \langle \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \rangle = \langle \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \rangle = 0 \\ \iff & \begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \iff & \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \langle \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \rangle. \end{aligned}$$

Daraus folgt, dass $E^\perp$ die Gerade durch $(0, 0, 0)$ und $(-1, 1, 2)$ ist.

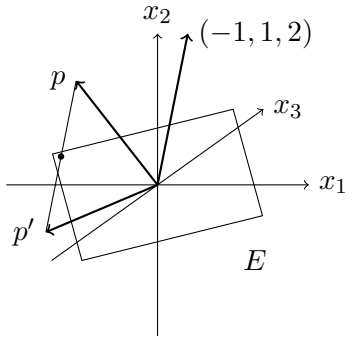
Diese Information kann man zum Beispiel dazu verwenden, einen gegebenen Punkt $p \in \mathbb{R}^3$ an der Ebene E zu spiegeln. Man nutzt aus, dass die Vektoren $(1, -1, 1)$ und $(2, 0, 1)$, die E aufspannen, zusammen mit dem Vektor $(-1, 1, 2)$, der $E^\perp$ aufspannt, eine Basis von $\mathbb{R}^3$ bilden. Aus einer Koordinatendarstellung des Punkts p bezüglich dieser Basis lässt sich leicht die Koordinatendarstellung des an E gespiegelten Punktes berechnen: Ist

$$p = \alpha_1 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix},$$

so ist

$$p' = \alpha_1 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} - \alpha_3 \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$$

der Bildpunkt.



4. Der Raum $\mathbb{R}^4$ sei versehen mit dem Skalarprodukt

$$\langle x|y\rangle = x \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 3 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix} y.$$

Dann gilt

$$\{\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix}\} \quad \perp \quad \{\begin{pmatrix} 6 \\ -2 \\ -4 \\ 5 \end{pmatrix}, \begin{pmatrix} 6 \\ 2 \\ -6 \\ 1 \end{pmatrix}\}.$$

Beachte, dass diese Mengen nicht bezüglich des Standardskalarprodukts senkrecht auf-
04-11 einander stehen. Zum Beispiel gilt

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

5. Im Raum ℓ^2 stehen die Folgen $(a_n)_{n=0}^{\infty}$ und $(b_n)_{n=0}^{\infty}$ mit

$$a_n = 3(1/2)^n - 4(1/3)^n, \qquad b_n = 9(1/2)^n - 8(1/3)^n$$

aufeinander senkrecht, denn es gilt

$$\begin{aligned} \langle (a_n)_{n=0}^{\infty} | (b_n)_{n=0}^{\infty} \rangle &= \sum_{n=0}^{\infty} a_n b_n \\ &= \lim_{n\to\infty} \sum_{k=0}^{n} \big(3(1/2)^k - 4(1/3)^k\big)\big(9(1/2)^n - 8(1/3)^n\big) \\ &= \lim_{n\to\infty} -\frac{(2^{n+1} - 3^{n+1})^2}{6^{2n}} = 0. \end{aligned}$$

6. Im Raum $C([0, 2\pi], \mathbb{R})$ mit dem Skalarprodukt

$$\langle f|g\rangle := \int_0^{2\pi} f(x)g(x)dx$$

stehen die Funktionen sin und cos senkrecht aufeinander.

7. Betrachte die Räume $\mathbb{R}^n$ und $\mathbb{R}^m$ zusammen mit dem Standardskalarprodukt. Sei $A \in \mathbb{R}^{n\times m}$. Dann gilt $\ker A \perp \operatorname{coim} A$ und $\operatorname{coker} A \perp \operatorname{im} A$. Dieser Sachverhalt wurde in der schematischen Zeichnung auf Seite 110 schon angedeutet.

   Um das einzusehen, betrachtet man am besten ein Beispiel. Ist zum Beispiel

$$T = \begin{pmatrix} 1 & 0 & 5 & 0 & 9 \\ 0 & 1 & 7 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

   die Treppennormalform von $A \in \mathbb{R}^{4\times 5}$, so bilden die von Null verschiedenen Zeilen von T eine Basis von $\operatorname{coim} A$. Das sind in diesem Fall

$$\begin{pmatrix} 1 \\ 0 \\ 5 \\ 0 \\ 9 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 7 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 3 \end{pmatrix}.$$

   Basisvektoren von $\ker A$ erhält man durch Auffüllen von T mit negativen Einheitsvektoren als zusätzlichen Zeilen und Extraktion der Spalten, in denen die neuen -1-Einträge zu liegen kommen. Im vorliegenden Fall wären das

$$\begin{pmatrix} 5 \\ 7 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 9 \\ 2 \\ 0 \\ 3 \\ -1 \end{pmatrix}.$$

   Wie man sieht, steht jeder Vektor aus der Basis von $\operatorname{coim} A$ senkrecht auf jedem Vektor in der Basis von $\ker A$. Wegen der Linearität des Skalarprodukts muss dann auch jeder Vektor von $\operatorname{coim} A$ senkrecht zu jedem Vektor in $\ker A$ sein.

8. Betrachte den Raum $\mathbb{R}^n$ zusammen mit dem Standardskalarprodukt. Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch, seien λ_1, λ_2 zwei verschiedene Eigenwerte von A. Dann gilt $E_{\lambda_1} \perp E_{\lambda_2}$.

   Zum Beweis betrachte man beliebige $v_1 \in E_{\lambda_1}$ und $v_2 \in E_{\lambda_2}$. Aus $Av_1 = \lambda_1 v_1$ und $Av_2 = \lambda_2 v_2$ folgt $v_2 A v_1 = \lambda_1 v_2 v_1$ und $v_1 A v_2 = \lambda_2 v_1 v_2$. Weil A symmetrisch ist, gilt $v_2 A v_1 = v_1 A v_2$, und weil außerdem $v_2 v_1 = v_1 v_2$ gilt, folgt $\lambda_1 v_1 v_2 = \lambda_2 v_1 v_2$, also $(\lambda_1 - \lambda_2) v_1 v_2 = 0$. Wegen $\lambda_1 \neq \lambda_2$ folgt $v_1 v_2 = 0$, wie behauptet.

Satz 94. Sei V ein Skalarproduktraum und $M \subseteq V$. Dann gilt:

1. $M^\perp$ ist ein Untervektorraum von V.
2. $\langle M \rangle \subseteq (M^\perp)^\perp$.
3. $M^\perp = \bigcap_{m\in M} \{m\}^\perp$.

Beweis.

1. Zu zeigen: $M^\perp \neq \emptyset$ und für alle $x, y \in M^\perp$ und $\alpha, \beta \in \mathbb{R}$ gilt $\alpha x + \beta y \in M^\perp$. 04-13

   Für alle $m \in M$ gilt $\langle m|0\rangle = 0$, also ist auf jeden Fall $0 \in M^\perp$ und damit $M \neq \emptyset$.

   Seien nun $x, y \in M^\perp$ und $\alpha, \beta \in \mathbb{R}$ beliebig.

   Nach Definition von $M^\perp$ gilt dann $\langle x|m\rangle = \langle y|m\rangle = 0$ für alle $m \in M$. Wegen der Linearität des Skalarprodukts folgt daraus

   $$\langle \alpha x + \beta y|m\rangle = \alpha\langle x|m\rangle + \beta\langle y|m\rangle = 0$$

   für alle $m \in M$, und also $\alpha x + \beta y \in M$.

2. Sei $x = \alpha_1 m_1 + \cdots + \alpha_k m_k$ für gewisse $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ and $m_1, \ldots, m_k \in M$.

   04-11 Zu zeigen: $x \in (M^\perp)^\perp$, d.h. für alle $v \in M^\perp$ gilt $\langle x|v\rangle = 0$. Sei $v \in M^\perp$. Nach Definition von $M^\perp$ gilt $\langle v|m\rangle = 0$ für alle $m \in M$, insbesondere für $m_1, \ldots, m_k$. Deshalb gilt

   $$\begin{aligned}\langle v|x\rangle = \langle x|v\rangle &= \langle \alpha_1 m_1 + \cdots + \alpha_k m_k|v\rangle \\ &= \alpha_1\langle m_1|v\rangle + \cdots + \alpha_k\langle m_k|v\rangle \\ &= \alpha_1\langle v|m_1\rangle + \cdots + \alpha_k\langle v|m_k\rangle \\ &= 0.\end{aligned}$$

   Daraus folgt die Behauptung.

3. Sei $x \in V$ beliebig. Dann gilt

   $$x \in M^\perp \iff \forall\, m \in M : x \perp m \iff \forall\, m \in M : x \in \{m\}^\perp \iff x \in \bigcap_{m \in M} \{m\}^\perp. \quad \blacksquare$$

32 Positivdefinitheit

Im Beispiel zu Definition 61 haben wir gesehen, dass es Matrizen $A \in \mathbb{R}^{n\times n}$ gibt, für die durch $\langle x|y\rangle := xAy$ ein Skalarprodukt erklärt wird. Aber offenbar kommen für eine solche Definition nicht alle Matrizen $A \in \mathbb{R}^{n\times n}$ infrage. Zwar gilt die Linearität für jede Wahl von A, aber die Bedingung $\langle x|y\rangle = \langle y|x\rangle$ in Definition 61 erzwingt, dass $A = A^\top$ gelten muss, denn aus der Bedingung $xAy = yAx$ folgt für die spezielle Wahl $x = e_i$, $y = e_j$, dass der (i, j)-Eintrag von A identisch mit dem (j, i)-Eintrag von A sein muss.

Als drittes soll nach Definition 61 gelten, dass für alle $x \in \mathbb{R}^n \setminus \{0\}$ gilt $xAx > 0$. Auch das ist nicht für jede Matrix $A \in \mathbb{R}^{n\times n}$ der Fall, zum Beispiel gilt

$$(0, 1)\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix} = -1 < 0.$$

Definition 63.

1. $A \in \mathbb{R}^{n\times n}$ heißt *symmetrisch,* falls $A = A^\top$ gilt.
2. Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ heißt
   - *positiv definit,* falls $\forall\, x \in \mathbb{R}^n \setminus \{0\} : xAx > 0$.
   - *negativ definit,* falls $\forall\, x \in \mathbb{R}^n \setminus \{0\} : xAx < 0$.
   - *positiv semidefinit,* falls $\forall\, x \in \mathbb{R}^n : xAx \geq 0$.
   - *negativ semidefinit,* falls $\forall\, x \in \mathbb{R}^n : xAx \leq 0$.
   - *indefinit* in allen anderen Fällen.

Beispiel.

1. $A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ ist symmetrisch. Für jeden Vektor $x = (x_1, x_2) \in \mathbb{R}^2$ gilt

$$\begin{aligned}(x_1, x_2)\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= (x_1, x_2)\begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + 5x_2 \end{pmatrix} \\ &= x_1(x_1 + 2x_2) + x_2(2x_1 + 5x_2) \\ &= x_1^2 + 4x_1x_2 + 5x_2^2 \\ &= (x_1 + 2x_2)^2 + x_2^2 \geq 0.\end{aligned}$$

   Damit ist A zumindest positiv semidefinit. Wegen

$$\begin{aligned} & (x_1 + 2x_2)^2 + x_2^2 = 0 \\ \iff\quad & x_1 + 2x_2 = 0 \wedge x_2 = 0 \\ \iff\quad & x_1 = x_2 = 0\end{aligned}$$

   ist A sogar positiv definit.

2. Für jede symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ gilt: A ist positiv (semi-)definit $\iff$ $-A$ ist negativ (semi-)definit.

   Außerdem gilt: Jede positiv definite Matrix ist positiv semidefinit, aber nicht jede positiv semidefinite Matrix ist positiv definit.

3. $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ ist indefinit, denn

$$\begin{aligned}(1, 0)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} &= 1 > 0 \\ \text{und}\quad (0, 1)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} &= -1 < 0.\end{aligned}$$

Satz 95. Jede symmetrische Matrix in $\mathbb{R}^{n\times n}$ ist diagonalisierbar.

Beweis. Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Zunächst ist zu zeigen, dass das charakteristische Polynom χ von A in $\mathbb{R}[X]$ vollständig faktorisiert. Sicher tut es das, wenn man es als Polynom 04-07
in $\mathbb{C}[X]$ auffasst, weil $\mathbb{C}$ algebraisch abgeschlossen ist. Ist aber $\lambda = a + \mathrm{i}b \in \mathbb{C}$ ein Eigenwert 04-11
von A und ist $v \in \mathbb{C}^n \setminus \{0\}$ ein Eigenvektor zu λ, so lässt sich v schreiben als $v = u + \mathrm{i}w$ für
gewisse $u, w \in \mathbb{R}^n$. Dann gilt 04-11

$$\begin{aligned} Av &= \lambda v \\ A(u+\mathrm{i}w) &= (a+\mathrm{i}b)(u+\mathrm{i}w) \\ Au + \mathrm{i}Aw &= (au - bw) + \mathrm{i}(bu + aw), \end{aligned}$$

04-11 also $Au = au - bw$ und $Aw = bu + aw$, also

$$u(Aw) = b\,uu + a\,uw \quad \text{und} \quad w(Au) = a\,uw - b\,ww.$$

Wegen der Symmetrie von A gilt $uAw = wAu$, und deshalb folgt aus den obigen beiden
04-15 Gleichungen $b\,uu = -b\,ww$, also $b(uu + ww) = 0$. Wegen $v \neq 0$ ist $u \neq 0$ oder $v \neq 0$, daher $uu \neq 0$ oder $ww \neq 0$, jedenfalls aber $uu \geq 0$ und $ww \geq 0$, also muss $b = 0$ sein.

Damit ist gezeigt, dass für jeden komplexen Eigenwert $z = a + \mathrm{i}b$ der Imaginärteil b Null ist, d.h. alle Eigenwerte sind reell, d.h. das charakteristische Polynom χ von A in $\mathbb{R}[X]$ faktorisiert vollständig.

Nach Satz 91 bleibt zu zeigen, dass das Minimalpolynom keine mehrfachen Nullstellen hat. Dazu genügt es zu zeigen, dass für jeden Eigenwert λ gilt $\ker(A - \lambda I_n)^2 \subseteq \ker(A - \lambda I_n)$.

Sei also $x \in \ker(A - \lambda I_n)^2$. Dann gilt

$$\begin{aligned} & (A-\lambda I_n)^2 x = 0 \\ \Rightarrow \quad & x(A-\lambda I_n)(A-\lambda I_n)x = 0 \\ \Rightarrow \quad & \big(x(A-\lambda I_n)^\top\big)\big((A-\lambda I_n)x\big) = 0 \\ \Rightarrow \quad & \big((A-\lambda I_n)x\big)\big((A-\lambda I_n)x\big) = 0 \\ \Rightarrow \quad & (A-\lambda I_n)x = 0 \\ \Rightarrow \quad & x \in \ker(A-\lambda I_n). \end{aligned}$$

∎

Satz 96. Sei $A = ((a_{i,j}))_{i,j=1}^n \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Dann sind folgende Aussagen äquivalent:

1. Alle Eigenwerte von A sind positiv.
2. Durch $\langle x|y\rangle := xAy$ wird auf $\mathbb{R}^n$ ein Skalarprodukt erklärt.
3. A ist positiv definit.
4. Für alle $k = 1, \ldots, n$ ist $((a_{i,j}))_{i,j=1}^k \in \mathbb{R}^{k\times k}$ positiv definit.
5. Für alle $k = 1, \ldots, n$ gilt $\det((a_{i,j}))_{i,j=1}^k > 0$.
6. Es gibt eine Matrix $B \in \mathrm{GL}(n, \mathbb{R})$ mit $A = B^\top B$ (Cholesky-Zerlegung).

Beweis.

(1)⇒(2). Zu zeigen: $\forall\, x \in \mathbb{R}^n \setminus \{0\} : xAx > 0$.

Da A symmetrisch ist, ist A nach Satz 95 diagonalisierbar. Daraus folgt, dass $\mathbb{R}^n = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$, wobei $\lambda_1, \ldots, \lambda_k$ die Eigenwerte von A sind. Diese sind nach Annahme positiv.

Sei $x \in \mathbb{R}^n \setminus \{0\}$ beliebig. Dann gibt es $x_1 \in E_{\lambda_1}, \ldots, x_k \in E_{\lambda_k}$ mit $x = x_1 + \cdots + x_k$. Daher

$$xAx = (x_1 + \cdots + x_k)A(x_1 + \cdots + x_k) = \sum_{i=1}^{k} \lambda_i x_i x_i + \sum_{i \neq j} \lambda_i x_i x_j.$$

Der erste Term ist sicher positiv, weil alle λ_i positiv, alle $x_i x_i$ nicht-negativ und wenigstens ein x_i nicht Null ist. Der zweite Term ist Null, weil Eigenräume bezüglich des Standardskalarprodukts senkrecht aufeinander stehen (vgl. das letzte Beispiel nach Definition 62).

(2)⇒(3). klar.

(3)⇒(4). Zu $k \in \{1, \ldots, n\}$ sei $A_k := ((a_{i,j}))_{i,j=1}^{k} \in \mathbb{R}^{k \times k}$.

Zu zeigen: $\forall\, x \in \mathbb{R}^k \setminus \{0\} : xA_kx > 0$.

Sei $x = (x_1, \ldots, x_k) \in \mathbb{R}^k \setminus \{0\}$ beliebig. Für den Vektor $\tilde{x} = (x_1, \ldots, x_k, 0, \ldots, 0) \in \mathbb{R}^n \setminus \{0\}$ gilt dann $xA_kx = \tilde{x}A\tilde{x}$. Da A nach Annahme positiv definit ist, folgt $xA_kx > 0$.

(4)⇒(5). Wir schreiben wieder $A_k = ((a_{i,j}))_{i,j=1}^{k}$ für $k = 1, \ldots, n$.

Die Determinante $\det A_k$ ist das Produkt der Eigenwerte von A_k. Ist das nicht positiv, so ist zumindest ein Eigenwert nicht positiv, etwa der Eigenwert $\lambda \leq 0$. Ist dann $x \in \mathbb{R}^k \setminus \{0\}$ ein Eigenvektor zu λ, so wäre $xA_kx = x\lambda x = \lambda xx \leq 0$, im Widerspruch zur Positiv-Definitheit von A_k. 04-14

(5)⇒(6). Wir schreiben wieder $A_k = ((a_{i,j}))_{i,j=1}^{k}$ für $k = 1, \ldots, n$ und beweisen die Behauptung durch Induktion nach k.

Für $k = 1$ ist $A_1 = a_{1,1} = \det(A_1) > 0$, d.h. wir können $B = \sqrt{a_{1,1}}$ nehmen.

Induktionsschluss $k - 1 \to k$: Sei $a_k \in \mathbb{R}^{k-1}$ so, dass $A_k = \begin{pmatrix} A_{k-1} & a_k \\ a_k & a_{k,k} \end{pmatrix}$ ist. Nach Induktionsvoraussetzung gibt es eine Matrix $B_{k-1} \in \mathrm{GL}(k-1, \mathbb{R})$ mit $A_{k-1} = B_{k-1}^{\top} B_{k-1}$. Insbesondere ist A_{k-1} invertierbar. Betrachte die Matrix

$$T = \begin{pmatrix} I_{k-1} & -A_{k-1}^{-1} a_k \\ 0 & 1 \end{pmatrix}.$$

Dann gilt $\det(T) = 1$ und

$$T^{\top} A_k T = \begin{pmatrix} A_{k-1} & 0 \\ 0 & a_{k,k} - a_k A_{k-1}^{-1} a_k \end{pmatrix}.$$

Für $\alpha = a_{k,k} - a_k A_{k-1}^{-1} a_k$ gilt

$$\alpha = \frac{\det(T^{\top} A_k T)}{\det(A_{k-1})} = \frac{\det(A_k)}{\det(A_{k-1})} > 0.$$

Wir können deshalb 04-15

$$B_k = \begin{pmatrix} B_{k-1} & 0 \\ 0 & \sqrt{\alpha} \end{pmatrix} T^{-1}$$

04-15 setzen und erhalten wie gefordert

$$B_k^\top B_k = (T^{-1})^\top \begin{pmatrix} B_{k-1}^\top & \\ & \sqrt{\alpha} \end{pmatrix} \begin{pmatrix} B_{k-1} & \\ & \sqrt{\alpha} \end{pmatrix} T^{-1} = (T^\top)^{-1} \begin{pmatrix} A_{k-1} & 0 \\ 0 & \alpha \end{pmatrix} T^{-1} = A_k$$

04-15 (6)⇒(1). Es sei $\|\cdot\|$ die zum Standardskalarprodukt gehörige Norm. Sei $\lambda \in \mathbb{R}$ ein Eigenwert von A und $x \neq 0$ ein Eigenvektor zu λ. Nach Voraussetzung gilt $A = B^\top B$ für eine invertierbare Matrix B. Aus $Ax = \lambda x$ folgt $xAx = \lambda \|x\|^2$. Andererseits gilt

$$xAx = xB^\top Bx = (Bx)(Bx) = \|Bx\|^2,$$

also $\|Bx\|^2 = \lambda \|x\|^2$. Wegen $x \neq 0$ gilt $\|x\|^2 > 0$ und da B invertierbar ist, gilt auch $\|Bx\|^2 > 0$. Daraus folgt $\lambda > 0$. ∎

Beispiel. Die Matrix $A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ ist positiv definit, weil

$$\det(3) = 3 > 0, \qquad \det \begin{pmatrix} 3 & 2 \\ 2 & 2 \end{pmatrix} = 6 - 4 = 2 > 0, \qquad \det \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} = 1 > 0.$$

Die Eigenwerte von A sind die Nullstellen von $-X^3 + 6X^2 - 5X + 1$. Diesem Polynom lässt sich nicht ohne weiteres ansehen, dass all seine Nullstellen (reell und) positiv sind.

33 Orthogonalsysteme

Definition 64. Sei V ein Skalarproduktraum und $B \subseteq V \setminus \{0\}$.

1. B heißt *Orthogonalsystem* (OS), falls $\forall\, b, b' \in B : b \neq b' \Rightarrow \langle b|b' \rangle = 0$
2. B heißt *Orthonormalsystem* (ONS), falls B ein Orthogonalsystem ist und zusätzlich gilt $\forall\, b \in B : \langle b|b \rangle = 1$.
3. B heißt *Orthogonalbasis* (OB) von V, falls B zugleich Orthogonalsystem und Basis von V ist.
4. B heißt *Orthonormalbasis* (ONB) von V, falls B zugleich Orthonormalsystem und Basis von V ist.

Beispiel.

1. Für $V = \mathbb{R}^n$ und das Standardskalarprodukt ist die Standardbasis $\{e_1, \ldots, e_n\}$ eine ONB.

2. Für $V = \mathbb{R}^2$ und das Standardskalarprodukt ist auch

$$B = \{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}\}$$

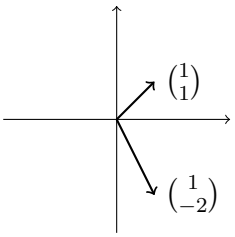
   eine ONB.

3. Ist B ein OS eines Vektorraums V, so ist $\{\frac{1}{\sqrt{\langle b|b\rangle}}b : b \in B\}$ ein ONS.

4. Sei $V = \mathbb{R}^2$. Bezüglich des Skalarprodukts

$$\langle x|y\rangle := x\begin{pmatrix}2 & 0\\0 & 1\end{pmatrix}y$$

   ist $\{\binom{1}{1}, \binom{1}{-2}\}$ eine Orthogonalbasis.



   Es sei hier noch einmal darauf hingewiesen, dass Orthogonalität von der Wahl des Skalarprodukts abhängt. Die übliche Orthogonalität bezieht sich aufs Standardskalarprodukt.

5. Im Raum $V = \ell^2$ bildet die Menge aller Folgen

$$e_i \colon \mathbb{N} \to \mathbb{R}, \quad e_i(n) = \begin{cases} 1 & \text{falls } i = n \\ 0 & \text{sonst} \end{cases}$$

   ein Orthonormalsystem (aber keine Basis!).

6. Im Raum $V = C([-1, 1], \mathbb{R})$ mit dem Skalarprodukt

$$\langle f|g\rangle = \int_{-1}^{1} f(x)g(x)\,dx$$

   ist z.B. $\{x \mapsto 1, x \mapsto x, x \mapsto \frac{1}{2}(3x^2 - 1), x \mapsto \frac{1}{2}(5x^3 - 3x)\}$ ein Orthogonalsystem. Davon überzeugt man sich leicht durch Nachrechnen.

   Ein anderes Orthogonalsystem ist

$$\{x \mapsto \sin(n\pi x) : n \in \mathbb{N} \setminus \{0\}\} \cup \{x \mapsto \cos(n\pi x) : n \in \mathbb{N}\}.$$

7. Die Folge $(T_n)_{n=0}^{\infty}$ in $\mathbb{R}[X]$ ist rekursiv definiert durch

$$T_0 = 1, \quad T_1 = X, \quad T_{n+2} = 2XT_{n+1} - T_n \quad (n \in \mathbb{N}).$$

Die Polynome T_n heißen *Chebyshev-Polynome* (erster Art). Man zeigt leicht durch Induktion, dass $\deg T_n = n$ für alle $n \in \mathbb{N}$ gilt. Daraus folgt, dass $\{T_n : n \in \mathbb{N}\}$ eine Basis von $\mathbb{R}[X]$ ist.

Auf $\mathbb{R}[X]$ wird durch

$$\langle p|q\rangle := \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} p(x)q(x)dx$$

ein Skalarprodukt erklärt.

Man kann dann durch Induktion nachrechnen, dass für alle $n, m \in \mathbb{N}$ gilt

$$\langle T_n|T_m\rangle = \begin{cases} 0 & \text{falls } n \neq m \\ \pi & \text{falls } n = m = 0 \\ \pi/2 & \text{falls } n = m \neq 0 \end{cases}$$

Also bildet $\{T_n : n \in \mathbb{N}\}$ eine Orthogonalbasis von $\mathbb{R}[X]$ bezüglich des Skalarprodukts $\langle\cdot|\cdot\rangle$.

Satz 97. Jedes Orthogonalsystem ist linear unabhängig.

04-15 **Beweis.** Sei M ein OS und seien $x_1, \ldots, x_n \in M$ paarweise verschieden, $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ so, dass $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$. Zu zeigen: $\alpha_1 = \cdots = \alpha_n = 0$.
Sei $i \in \{1, \ldots, n\}$ beliebig. Dann gilt

$$\begin{aligned} 0 &= \langle 0|x_i\rangle \\ &= \langle \alpha_1 x_1 + \cdots + \alpha_n x_n|x_i\rangle \\ &= \alpha_1\langle x_1|x_i\rangle + \cdots + \alpha_n\langle x_n|x_i\rangle \\ &= \alpha_i\langle x_i|x_i\rangle. \end{aligned}$$

Da x_i als Element eines Orthogonalsystems nicht Null ist, ist $\langle x_i|x_i\rangle \neq 0$. Darum muss $\alpha_i = 0$ sein. ∎

Satz 98. Sei V ein Skalarproduktraum und B eine ONB von V. Dann gilt:

1. Für jedes $x \in V$ gibt es höchstens endlich viele $b \in B$ mit $\langle x|b\rangle \neq 0$.
2. Für alle $x \in V$ gilt $x = \sum_{b\in B}\langle x|b\rangle b$.
3. Ist $n = \dim V < \infty$, sind $x, y \in V$ und sind $\bar{x}, \bar{y}$ die Koordinatendarstellungen von x, y bezüglich B, so gilt $\langle x|y\rangle = \bar{x}\bar{y}$, d.h. das Skalarprodukt auf V entspricht dem Standard-Skalarprodukt in $\mathbb{R}^n$.

Beweis.

1. Sei $x \in V$ beliebig. Dann existieren $b_1, \ldots, b_n \in B$ und $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ mit $x = \alpha_1 b_1 + \cdots + \alpha_n b_n$, weil B eine Basis von V ist. Für alle $b \in B \setminus \{b_1, \ldots, b_n\}$ gilt dann
$$\langle x|b\rangle = \alpha_1\langle b_1|b\rangle + \cdots + \alpha_n\langle b_n|b\rangle = 0,$$
weil B ein OS ist.

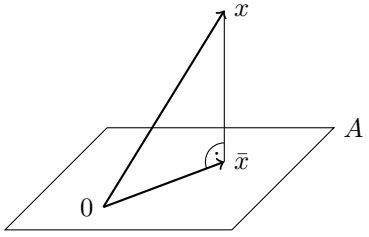
2. Sind $x \in V$ und $b_1, \ldots, b_n \in B$, $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ wie oben, so gilt
$$\langle x|b_i\rangle = \alpha_1\langle b_1|b_i\rangle + \cdots + \alpha_n\langle b_n|b_1\rangle = \alpha_i\langle b_i|b_i\rangle = \alpha_i,$$
weil B ein ONS ist. Zusammen mit $\langle x|b\rangle = 0$ für $b \in B \setminus \{b_1, \ldots, b_n\}$ folgt
$$x = \sum_{i=1}^{n} \alpha_i b_i = \sum_{i=1}^{n} \langle x|b_i\rangle b_i = \sum_{b\in B} \langle x|b\rangle b.$$

3. Sei $B = \{b_1, \ldots, b_n\}$ die ONB. Seien $x, y \in V$ beliebig, etwa $x = \alpha_1 b_1 + \cdots + \alpha_n b_n$ und $y = \beta_1 b_1 + \cdots + \beta_n b_n$ für gewisse $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \mathbb{R}$. Dann gilt: 04-15
$$\langle x|y\rangle = \langle \sum_{i=1}^{n} \alpha_i b_i | \sum_{j=1}^{n} \beta_j b_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\beta_j\langle b_i|b_j\rangle = \sum_{k=1}^{n} \alpha_k\beta_k = \begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}\begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix}. \qquad \blacksquare$$

Eine Folgerung von Teil 2 ist die sogenannte *Parsevalsche Gleichung*
$$\|x\|^2 = \sum_{b\in B} \langle x|b\rangle^2,$$
die immer gilt, wenn B eine ONB ist.
Ist B nur ein ONS und ist B abzählbar, so gilt immer noch die sogenannte *Besselsche Ungleichung*
$$\|x\|^2 \geq \sum_{b\in B} \langle x|b\rangle^2$$
für alle $x \in V$. Dabei kann die Summe auf der rechten Seite unter Umständen unendlich viele Terme haben, aber sie ist dann immer konvergent.

Satz 99. Sei V ein Skalarproduktraum, A ein endlich-dimensionaler Unterraum von V, $\{a_1, \ldots, a_n\}$ eine ONB von A, $x \in V$ beliebig und $\bar{x} = \sum_{i=1}^{n} \langle x|a_i\rangle a_i$. Dann gilt:

1. $\forall\ a \in A : x - \bar{x} \perp a$.
2. $\forall\ a \in A : \|x - \bar{x}\| \leq \|x - a\|$.



Beweis.

1. Sei $a = \alpha_1 a_1 + \cdots + \alpha_n a_n \in A$ beliebig.

   Nach Definition von $\bar{x}$ gilt $\langle \bar{x} | a_i \rangle = \langle x | a_i \rangle$ für alle i. Daraus folgt

   $$\langle \bar{x} | a \rangle = \sum_{i=1}^{n} \alpha_i \langle \bar{x} | a_i \rangle = \sum_{i=1}^{n} \alpha_i \langle x | a_i \rangle = \langle x | a \rangle,$$

   und daraus $\langle x - \bar{x} | a \rangle = 0$, wie behauptet.

2. Sei $a \in A$ beliebig. Dann gilt

   $$\begin{aligned} \langle x - a | x - a \rangle &= \langle x - \bar{x} + \bar{x} - a | x - \bar{x} + \bar{x} - a \rangle \\ &= \langle x - \bar{x} | x - \bar{x} \rangle + \underbrace{\langle \underbrace{\bar{x} - a}_{\in A} | x - \bar{x} \rangle + \langle x - \bar{x} | \underbrace{\bar{x} - a}_{\in A} \rangle}_{=0} + \underbrace{\langle \bar{x} - a | \bar{x} - a \rangle}_{\geq 0} \\ &\geq \langle x - \bar{x} | x - \bar{x} \rangle. \end{aligned}$$

   Daraus folgt die Behauptung. ∎

Beispiel.

1. Sei $V = \mathbb{R}^3$ ausgestattet mit dem Standardskalarprodukt, $A = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \rangle \subseteq V$,

05-02 $x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Die angegebene Basis von A ist orthogonal aber nicht orthonormal. Eine ONB ist z.B. $\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \}$. Damit bekommen wir

   $$\begin{aligned} \bar{x} &= \langle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} | \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \rangle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \langle \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} | \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \rangle \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix} \end{aligned}$$

   Gemäß Teil 1 ist $\bar{x}$ der Schnittpunkt der Ebene A mit der Geraden durch x, die senkrecht auf dieser Ebene steht. Gemäß Teil 2 gibt es in der Ebene A keinen Punkt, der näher an x liegt als $\bar{x}$.

2. Sei $V = C([-1, 1], \mathbb{R})$ ausgestattet mit dem Skalarprodukt

   $$\langle f | g \rangle = \int_{-1}^{1} f(x) g(x) dx.$$

Wir haben im Beispiel nach Definition 64 gesehen, dass $\{x \mapsto 1,\ x \mapsto x,\ x \mapsto \frac{1}{2}(3x^2 - 1),\ x \mapsto \frac{1}{2}(5x^3 - 3x)\}$ ein Orthogonalsystem bezüglich $\langle\cdot|\cdot\rangle$ ist. Durch Skalierung dieser Funktionen kommt man zum Orthonormalsystem

$$\{x \mapsto \frac{1}{\sqrt{2}}, x \mapsto \sqrt{\frac{3}{2}}x, x \mapsto \sqrt{\frac{5}{8}}(3x^2 - 1), x \mapsto \sqrt{\frac{7}{8}}(5x^3 - 3x)\}.$$

Es sei $A \subseteq V$ der Unterraum, der von diesen Funktionen erzeugt wird. Die Funktion $f = (x \mapsto \mathrm{e}^x) \in V$ liegt nicht in diesem Unterraum. Die Orthogonalprojektion $\bar{f}$ von f auf A lautet 04-16

$$\begin{aligned}
\bar{f}(x) = &\int_{-1}^{1} \frac{1}{\sqrt{2}}\ \mathrm{e}^t dt\ \frac{1}{\sqrt{2}} \\
&+ \int_{-1}^{1} \sqrt{\frac{3}{2}}t\ \mathrm{e}^t dt\ \sqrt{\frac{3}{2}}x \\
&+ \int_{-1}^{1} \sqrt{\frac{5}{8}}(3t^2 - 1)\ \mathrm{e}^t dt\ \sqrt{\frac{5}{8}}(3x^2 - 1) \\
&+ \int_{-1}^{1} \sqrt{\frac{7}{8}}(5t^3 - 3t)\ \mathrm{e}^t dt\ \sqrt{\frac{7}{8}}(5x^3 - 3x) \\
= &\frac{1}{4\mathrm{e}}\big(-3(\mathrm{e}^2 - 11) + 15(7\mathrm{e}^2 - 51)x + 15(\mathrm{e}^2 - 7)x^2 + 35(-5\mathrm{e}^2 + 37)x^3\big).
\end{aligned}$$

Bezüglich der von $\langle\cdot|\cdot\rangle$ induzierten Norm gibt es kein Polynom vom Grad ≤ 3, das näher an f liegt als $\bar{f}$. Der Abstand ist

$$\|f - \bar{f}\| = \sqrt{\int_{-1}^{1} \mathrm{e}^x \bar{f}(x) dx} \approx 0.0047$$

3. Seien V und $\langle\cdot|\cdot\rangle$ wie zuvor und betrachte noch einmal $f = (x \mapsto \mathrm{e}^x) \in V$. Jetzt sei 04-16

$$A = \langle (x \mapsto \sin(n\pi x)) : n \in \mathbb{N} \setminus \{0\} \rangle.$$

Für beliebiges $n \in \mathbb{N} \setminus \{0\}$ gilt

$$\begin{aligned}
\langle f | (x \mapsto \sin(n\pi x)) \rangle &= \int_{-1}^{1} \sin(n\pi x)\mathrm{e}^x dx \\
&= \frac{(1 - \mathrm{e}^2)n\cos(n) + (1 + \mathrm{e}^2)\sin(n)}{(n^2 + 1)\mathrm{e}} =: a_n.
\end{aligned}$$

Damit erhalten wir eine Folge von Approximationen

$$f_n \colon [-1, 1] \to \mathbb{R}, \qquad f_n(x) = \sum_{k=1}^{n} a_n \sin(n\pi x),$$

die sich immer näher an f annähern. Es ist nämlich f_n die Projektion von f auf den Raum

$$A_n = \langle (x \mapsto \sin(k\pi x)) : k = 1, \ldots, n \rangle,$$

und wegen $A_1 \subseteq A_2 \subseteq \cdots$ und Teil 2 des Satzes kann die Qualität der Approximation bei steigendem n nicht schlechter werden.

Damit ist aber noch nicht gesagt, dass die Funktionenfolge $(f_n)_{n=0}^{\infty}$ gegen f konvergiert. Tatsächlich kann sie das nicht, denn für alle $n \in \mathbb{N} \setminus \{0\}$ gilt auf jeden Fall $f_n(0) = 0 \neq 1 = f(0)$. Die Grenzfunktion $\lim_{n\to\infty} f_n$ (falls sie denn existiert) ist nur die beste Näherung an f, die mit den Funktionen in A möglich ist.

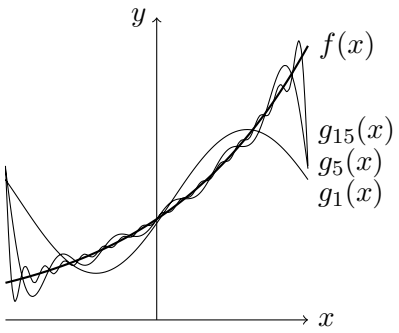
Um eine wirklich gute Approximation zu bekommen, muss man A so wählen, dass sich jedes Element von V erreichen lässt als Grenzfunktion einer Funktionenfolge in A. Eine bessere Wahl ist zum Beispiel der Raum A, der von

$$\{(x \mapsto \frac{1}{\sqrt{2}})\} \cup \{(x \mapsto \sin(\pi n x)) : n \in \mathbb{N} \setminus \{0\}\} \cup \{(x \mapsto \cos(\pi n x)) : n \in \mathbb{N} \setminus \{0\}\}$$

erzeugt wird. Für die Folge $(g_n)_{n=0}^{\infty}$ definiert durch

$$\begin{aligned} g_n \colon [-1,1] \to \mathbb{R}, \quad g_n(x) = {} & \frac{\mathrm{e} - \mathrm{e}^{-1}}{2} \\ & + \sum_{k=1}^{n} \langle f | (x \mapsto \sin(\pi n x)) \rangle \sin(\pi n x) \\ & + \sum_{k=1}^{n} \langle f | (x \mapsto \cos(\pi n x)) \rangle \cos(\pi n x) \end{aligned}$$

bekommt man deutlich bessere Approximationen, zumindest im Inneren des Definitionsbereichs.



Satz 100. (Gram-Schmidt) Sei V ein Skalarproduktraum und $\{b_1, b_2, \dots\}$ linear unabhängig. Setze

$$\begin{aligned} u_1 &= b_1, & v_1 &= \frac{1}{\|u_1\|} u_1, \\ u_2 &= b_2 - \langle b_2 | v_1 \rangle v_1, & v_2 &= \frac{1}{\|u_2\|} u_2, \\ &\vdots & &\vdots \\ u_n &= b_n - \sum_{k=1}^{n-1} \langle b_n | v_k \rangle v_k, & v_n &= \frac{1}{\|u_n\|} u_n, \end{aligned}$$

$\vdots$ $\vdots$

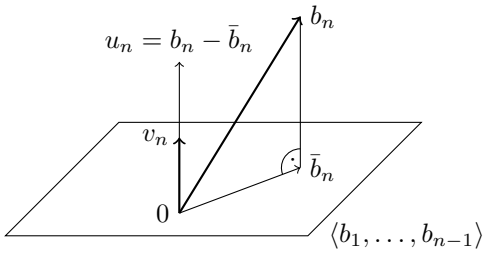
Dann ist $\{v_1, v_2, \dots\}$ eine ONB des Unterraums $\langle b_1, b_2, \dots \rangle$ von V.
Insbesondere gilt: Ist $\dim V < \infty$, so hat V eine ONB.

Beweis. Wir zeigen induktiv, dass $\langle b_1, \dots, b_n \rangle = \langle v_1, \dots, v_n \rangle$ für alle $n \in \mathbb{N}$, und dass $\{v_1, \dots, v_n\}$ ein ONS ist.
Für $n = 1$ ist das klar. Induktionsschritt $n - 1 \to n$:

1. Aus der Definition von u_n und v_n folgt $b_n \in \langle u_n \rangle + \langle v_1, \dots, v_{n-1} \rangle = \langle v_1, \dots, v_n \rangle$ und wegen der Induktionsvoraussetzung auch $u_n \in \langle b_n \rangle + \langle v_1, \dots, v_{n-1} \rangle = \langle b_1, \dots, b_n \rangle$. Damit gilt $\langle v_1, \dots, v_n \rangle = \langle b_1, \dots, b_n \rangle$.

2. Aus der Definition von u_n und Satz 99 Teil 1 folgt $u_n \perp v_i$ für alle $i = 1, \dots, n-1$. Aus der Definition von v_n folgt dann $v_n \perp v_i$ und $\|v_n\| = 1$ für alle i. ∎

Das Verfahren aus Satz 100 hat eine anschauliche geometrische Interpretation. Aus dem n-ten Basisvektor b_n wird der n-te Vektor v_n der ONB in zwei Schritten gewonnen. Im ersten Schritt projiziert man b_n orthogonal auf den Unterraum $\langle b_1, \dots, b_{n-1} \rangle$, der von den schon behandelten Vektoren aufgespannt wird. Satz 99 sagt, wie das geht. Ist $\bar{b}_n$ das Resultat dieser Projektion, so muss $u_n := b_n - \bar{b}_n$ im orthogonalen Komplement von $\langle b_1, \dots, b_{n-1} \rangle$ liegen. Im zweiten Schritt skaliert man u_n auf einen Vektor, der in die gleiche Richtung wie u_n zeigt und die Länge 1 hat.



Beispiel.

1. Betrachte $V = \mathbb{R}^3$ mit dem Standardskalarprodukt und

$$B = \{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \}.$$

Wir berechnen eine ONB von V mit dem Verfahren aus Satz 100:

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

$$u_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} | \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \rangle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \qquad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

$$u_3 = \begin{pmatrix}1\\2\\3\end{pmatrix} - \langle \begin{pmatrix}1\\2\\3\end{pmatrix} | \begin{pmatrix}1\\0\\0\end{pmatrix}\rangle \begin{pmatrix}1\\0\\0\end{pmatrix} - \langle \begin{pmatrix}1\\2\\3\end{pmatrix} | \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\end{pmatrix}\rangle \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\end{pmatrix} = \begin{pmatrix}0\\-1/2\\1/2\end{pmatrix},$$

$$v_3 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\-1\\1\end{pmatrix}.$$

Man erhält also die ONB

$$\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}0\\-1\\1\end{pmatrix}\}.$$

2. Bei der Berechnung von v_n aus u_n ist zu beachten, dass die Norm verwendet wird, die zum Skalarprodukt gehört, also $\|x\| = \sqrt{\langle x|x\rangle}$ $(x \in V)$. Ist zum Beispiel $V = \mathbb{R}^3$ mit dem Skalarprodukt

$$\langle x|y\rangle = x \begin{pmatrix}2&1&0\\1&2&0\\0&0&3\end{pmatrix} y$$

ausgestattet, und ist B wie oben, so liefert das Verfahren

$$u_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \qquad v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\end{pmatrix},$$

$$u_2 = \begin{pmatrix}1\\1\\1\end{pmatrix} - \underbrace{\langle \begin{pmatrix}1\\1\\1\end{pmatrix} | \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\end{pmatrix}\rangle \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\end{pmatrix}}_{=\frac{3}{2}\begin{pmatrix}1\\0\\0\end{pmatrix}} = \begin{pmatrix}-1/2\\1\\1\end{pmatrix}, \qquad v_2 = \frac{1}{3}\sqrt{2}\begin{pmatrix}-1/2\\1\\1\end{pmatrix},$$

$$u_3 = \begin{pmatrix}1\\2\\3\end{pmatrix} - \underbrace{\langle \begin{pmatrix}1\\2\\3\end{pmatrix} | \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\end{pmatrix}\rangle \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\end{pmatrix}}_{=\begin{pmatrix}2\\0\\0\end{pmatrix}} - \underbrace{\langle \begin{pmatrix}1\\2\\3\end{pmatrix} | \frac{1}{3}\sqrt{2}\begin{pmatrix}-1/2\\1\\1\end{pmatrix}\rangle \frac{1}{3}\sqrt{2}\begin{pmatrix}-1/2\\1\\1\end{pmatrix}}_{=\frac{1}{3}\begin{pmatrix}1\\-2\\1\end{pmatrix}}$$

$$= \frac{1}{3}\begin{pmatrix}1\\-2\\1\end{pmatrix}, \qquad v_3 = \frac{1}{3}\begin{pmatrix}1\\-2\\1\end{pmatrix}.$$

3. Sei $V = \mathbb{R}[X]$ mit dem Skalarprodukt

$$\langle p|q\rangle := \int_{-1}^{1} p(x)q(x)dx.$$

Ausgehend von $B = \{1, X, X^2, \dots\}$ berechnen wir eine ONB von V bezüglich dieses Skalarprodukts mit dem Verfahren des Satzes. Dabei ergibt sich:

$$\begin{aligned}
v_1 &= \frac{1}{\sqrt{2}},\\
v_2 &= \sqrt{\frac{3}{2}}\, X,\\
v_3 &= \sqrt{\frac{5}{8}}\, (3X^2 - 1),\\
v_4 &= \sqrt{\frac{7}{8}}\, (5X^3 - 3X),\\
v_5 &= \sqrt{\frac{9}{128}}\, (35X^4 - 20X^2 + 3),\\
v_6 &= \sqrt{\frac{11}{128}}\, (63X^5 - 70X^3 + 15X),\\
&\vdots
\end{aligned}$$

Die Polynome $1, X, \frac{1}{2}(3X^2 - 1), \frac{1}{2}(5X^2 - 3X), \dots$, von denen $v_1, v_2, \dots$ skalare Vielfache erscheinen, heißen *Legendre-Polynome*. Das n-te Legendre-Polynom wird mit $P_n(x)$ bezeichnet. Es gilt die Rekurrenz 05-02

$$(n+2)P_{n+2}(x) - (2n+3)xP_{n+1}(x) + (n+1)P_n(x) = 0, \quad P_0(x) = 1, P_1(x) = x.$$

4. Wenn $V = \mathbb{R}[X]$ mit dem Skalarprodukt

$$\langle p|q\rangle := \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} p(x)q(x)dx$$

ausgestattet ist und das Verfahren des Satzes auf die Basis $B = \{1, X, X^2, \dots\}$ angewendet wird, findet man

$$\begin{aligned}
v_1 &= \frac{1}{\sqrt{\pi}},\\
v_2 &= \sqrt{\frac{2}{\pi}} X,\\
v_3 &= \sqrt{\frac{2}{\pi}} (2X^2 - 1),\\
v_4 &= \sqrt{\frac{2}{\pi}} (4X^3 - 3X),\\
v_5 &= \sqrt{\frac{2}{\pi}} (8X^4 - 8X^2 + 1),\\
v_6 &= \sqrt{\frac{2}{\pi}} (16X^5 - 20X^3 + 5X),\\
&\vdots
\end{aligned}$$

Diese Polynome sind skalare Vielfache der Chebyshev-Polynome T_n.

34 Adjungierte Abbildungen

Definition 65. Seien V, W zwei Skalarproduktträume. Sei $\langle\cdot|\cdot\rangle_V$ das Skalarprodukt auf V und $\langle\cdot|\cdot\rangle_W$ das Skalarprodukt auf W. Zwei lineare Abbildungen $f \in \mathrm{Hom}(V, W)$ und $g \in \mathrm{Hom}(W, V)$ heißen (zueinander) *adjungiert,* falls gilt

$$\forall\, v \in V\ \forall\, w \in W : \langle f(v)|w\rangle_W = \langle v|g(w)\rangle_V.$$

Beispiel.

1. Seien $V = \mathbb{R}^3, W = \mathbb{R}^2$, beide ausgestattet mit dem jeweiligen Standardskalarprodukt.
04-21 Die beiden Abbildungen

$$f\colon \mathbb{R}^3 \to \mathbb{R}^2 \qquad f(x) = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} x$$

$$g\colon \mathbb{R}^2 \to \mathbb{R}^3 \qquad g(x) = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} x$$

sind zueinander adjungiert, denn für alle $v \in V$ und $w \in W$ gilt

$$\begin{aligned} \langle f(v)|w\rangle &= \langle \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} v|w\rangle \\ &= \left(v \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}\right) w \\ &= v \left(\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} w \right) \\ &= \langle v|g(w)\rangle. \end{aligned}$$

2. Sei $W = \ell^2$ und

$$V = \{\, (a_n)_{n=0}^{\infty} \in \ell^2 : \exists\, N \in \mathbb{N}\ \forall\, n \geq N : a_n = 0 \,\}.$$

Dann ist V ein echter Teilraum von W. Beide Räume seien mit dem Skalarprodukt von ℓ^2 versehen.

Betrachte $f\colon V \to W$, $f(v) = v$. Wir wollen zeigen, dass es zu f keine adjungierte
04-15 Abbildung gibt. Angenommen, $g\colon W \to V$ wäre eine adjungierte Abbildung zu f. Dann gilt

$$\forall\, v \in V\ \forall\, w \in W : \underbrace{\langle f(v)|w\rangle}_{=\langle v|w\rangle} = \langle v|g(w)\rangle,$$

also

$$\forall\, v \in V\ \forall\, w \in W : \langle v|w - g(w)\rangle = 0,$$

04-20 also

$$\forall\, w \in W : w - g(w) \in V^\top.$$

Jedoch ist $V^\top = \{0\}$, denn ist $(a_n)_{n=0}^\infty \in W \setminus \{0\}$, so gibt es mindestens ein $m \in \mathbb{N}$ mit $a_m \neq 0$, und für die Folge $(b_n)_{n=0}^\infty$ mit $b_m = 1$ und $b_n = 0$ für alle $n \neq m$ gilt $(b_n)_{n=0}^\infty \in V$ und $\langle (a_n)_{n=0}^\infty | (b_n)_{n=0}^\infty \rangle = a_m \neq 0$, also $(a_n)_{n=0}^\infty \notin V^\top$.

Aus $\forall\, w \in W : w - g(w) \in \{0\}$ folgt, dass g die Identitätsabbildung ist. Das ist aber wegen $g\colon W \to V \subsetneq W$ nicht möglich.

Satz 101. Seien U, V, W Skalarprodukträume. Dann gilt:

1. Zu $f \in \mathrm{Hom}(V, W)$ gibt es höchstens eine adjungierte Abbildung $f^* \in \mathrm{Hom}(W, V)$.

2. Falls zu $f_1, f_2 \in \mathrm{Hom}(V, W)$ die adjungierten Abbildungen $f_1^*, f_2^* \in \mathrm{Hom}(W, V)$ existieren, so existiert für alle $\alpha_1, \alpha_2 \in \mathbb{R}$ auch die adjungierte Abbildung $(\alpha_1 f_1 + \alpha_2 f_2)^*$ zu $\alpha_1 f_1 + \alpha_2 f_2$, und es gilt $(\alpha_1 f_1 + \alpha_2 f_2)^* = \alpha_1 f_1^* + \alpha_2 f_2^*$.

3. Falls zu $f \in \mathrm{Hom}(V, W)$ die adjungierte Abbildung $f^* \in \mathrm{Hom}(W, V)$ existiert, so existiert auch die adjungierte Abbildung $(f^*)^* \in \mathrm{Hom}(V, W)$ von f^*, und es gilt $(f^*)^* = f$.

4. Existiert die adjungierte Abbildung $f_1^* \in \mathrm{Hom}(V, U)$ zu $f_1 \in \mathrm{Hom}(U, V)$ sowie die adjungierte Abbildung $f_2^* \in \mathrm{Hom}(W, V)$ zu $f_2 \in \mathrm{Hom}(V, W)$, so existiert auch die adjungierte Abbildung $(f_2 \circ f_1)^* \in \mathrm{Hom}(W, U)$ von $f_2 \circ f_1 \in \mathrm{Hom}(U, W)$ und es gilt $(f_2 \circ f_1)^* = f_1^* \circ f_2^*$.

Beweis.

1. Sei $\langle \cdot | \cdot \rangle_V$ das Skalarprodukt auf V und $\langle \cdot | \cdot \rangle_W$ das Skalarprodukt auf W. Sei $f \in \mathrm{Hom}(V, W)$ und seien $f^*, f^\circ \in \mathrm{Hom}(W, V)$ zwei adjungierte Abbildungen zu f. Dann gilt 04-21, 04-28

$$\begin{aligned}
& \forall\, v \in V\ \forall\, w \in W : \langle f(v) | w \rangle_W = \langle v | f^*(w) \rangle_V = \langle v | f^\circ(w) \rangle_V, \\
\Rightarrow\ & \forall\, v \in V\ \forall\, w \in W : \langle v | (f^* - f^\circ)(w) \rangle_V = 0, \\
\Rightarrow\ & \forall\, w \in W : \langle (f^* - f^\circ)(w) | (f^* - f^\circ)(w) \rangle_V = 0, \\
\Rightarrow\ & \forall\, w \in W : (f^* - f^\circ)(w) = 0, \\
\Rightarrow\ & f^* = f^\circ.
\end{aligned}$$

2., 3., 4. Übung. ∎

Satz 102. (Riesz) Sei V ein Skalarproduktraum, $\dim V < \infty$. Dann existiert für jedes $v^* \in V^*$ genau ein $v \in V$ so dass

$$\forall\, x \in V : v^*(x) = \langle x | v \rangle.$$

Beweis. Existenz: Wähle eine ONB $\{b_1, \ldots, b_n\}$ von V und betrachte $v = v^*(b_1)b_1 + \cdots + v^*(b_n)b_n$. Für jedes $x = \alpha_1 b_1 + \cdots + \alpha_n b_n$ gilt dann

$$\begin{aligned}\langle x|v\rangle &= \langle \sum_{i=1}^{n} \alpha_i b_i | \sum_{j=1}^{n} v^*(b_j) b_j \rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i v^*(b_j) \langle b_i | b_j \rangle \\ &= \alpha_1 v^*(b_1) + \cdots + \alpha_n v^*(b_n) \\ &= v^*(\alpha_1 b_1 + \cdots + \alpha_n b_n) = v^*(x).\end{aligned}$$

Eindeutigkeit: Sind $v, v' \in V$ so, dass für alle $x \in V$ gilt $\langle x|v\rangle = \langle x|v'\rangle$, so gilt für alle $x \in V$ auch $\langle x|v - v'\rangle = 0$, insbesondere gilt dann $\langle v - v'|v - v'\rangle = 0$. Daraus folgt $v - v' = 0$, d.h. $v = v'$. ∎

Beispiel. Sei $V = \mathbb{R}^3$ ausgestattet mit dem Standardskalarprodukt. Das Funktional $v^* \in V^*$ sei definiert durch

$$v^*\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = 3x - 2y + 8z.$$

Der eindeutig bestimmte Vektor v aus dem Satz ist dann $v = (3, -2, 8)$.

Ist V dagegen ausgestattet mit dem Skalarprodukt

$$\langle x|y\rangle = x \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} y,$$

so lautet der eindeutig bestimmte Vektor aus dem Satz $v = (\frac{8}{3}, -\frac{7}{3}, \frac{8}{3})$.

Satz 103. Seien V, W zwei endlichdimensionale Skalarprodukträume. Dann existiert zu jedem $f \in \mathrm{Hom}(V, W)$ eine adjungierte Abbildung $f^* \in \mathrm{Hom}(W, V)$.

Beweis. Sei $\langle\cdot|\cdot\rangle_V$ das Skalarprodukt auf V und $\langle\cdot|\cdot\rangle_W$ das Skalarprodukt auf W. Sei $f \in \mathrm{Hom}(V, W)$.

Für jedes feste $w \in W$ ist die Abbildung $v^*\colon V \to \mathbb{R}$, $x \mapsto \langle f(x)|w\rangle_W$ ein Element von V^*. Nach Satz 102 gibt es genau ein $v \in V$, so dass für alle $x \in V$ gilt $v^*(x) = \langle x|v\rangle$.

Es sei $f^*\colon W \to V$ die Funktion, die jedem $w \in W$ das in dieser Weise eindeutig bestimmte v zuordnet. Dann gilt

$$\forall\, x \in V\ \forall\, w \in W : \langle f(x)|w\rangle_W = \langle x|f^*(w)\rangle_V.$$

Es bleibt zu zeigen, dass f^* linear ist.

Dazu betrachte $x \in V$, $\alpha_1, \alpha_2 \in \mathbb{R}$, $w_1, w_2 \in W$. Es gilt

$$\begin{aligned}\langle x|f^*(\alpha_1 w_1 + \alpha_2 w_2)\rangle &= \langle f(x)|\alpha_1 w_1 + \alpha_2 w_2\rangle \\ &= \alpha_1\langle f(x)|w_1\rangle + \alpha_2\langle f(x)|w_2\rangle \\ &= \alpha_1\langle x|f^*(w_1)\rangle + \alpha_2\langle x|f^*(w_2)\rangle \\ &= \langle x|\alpha_1 f^*(w_1) + \alpha_2 f^*(w_2)\rangle.\end{aligned}$$

Da x beliebig war, folgt daraus $f^*(\alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 f^*(w_1) + \alpha_2 f^*(w_2)$, wie gewünscht. ∎

Satz 104. Seien V, W zwei endlichdimensionale Skalarproduktträume. Sei A eine geordnete ONB von V und B eine geordnete ONB von W. Dann gilt: Ist M die Abbildungsmatrix von $f \in \mathrm{Hom}(V,W)$ bezüglich A und B, so ist $M^\top$ die Abbildungsmatrix der adjungierten Abbildung $f^* \in \mathrm{Hom}(W,V)$ bezüglich B und A.

Beweis. Sei $\langle\cdot|\cdot\rangle_V$ das Skalarprodukt auf V und $\langle\cdot|\cdot\rangle_W$ das Skalarprodukt auf W. Wir schreiben $A = (a_1, \ldots, a_n)$ und $B = (b_1, \ldots, b_m)$.
Sei $f \in \mathrm{Hom}(V,W)$ und $M = ((m_{i,j}))_{i=1,j=1}^{m,n}$ die Abbildungsmatrix von f bezüglich A und B, und sei $N = ((n_{i,j}))_{i=1,j=1}^{n,m}$ die Abbildungsmatrix von $f^* \in \mathrm{Hom}(W,V)$ bezüglich B und A. Für alle i, j gilt dann 04-21

$$\begin{aligned}\langle f(a_j)|b_i\rangle &= \langle\sum_{k=1}^{m} m_{k,j} b_k | b_i\rangle = \sum_{k=1}^{m} m_{k,j}\langle b_k|b_i\rangle = m_{i,j}\\ &\quad\|\\ \langle a_j|f^*(b_i)\rangle &= \langle a_j|\sum_{k=1}^{n} n_{k,i} a_k\rangle = \sum_{k=1}^{n} n_{k,i}\langle a_j|a_k\rangle = n_{j,i}.\end{aligned}$$ ∎

Definition 66. Sei V ein Skalarproduktraum. Ein Endomorphismus $f \in \mathrm{End}(V)$ heißt *selbstadjungiert* oder *symmetrisch,* falls gilt

$$\forall\, x, y \in V : \langle f(x)|y\rangle = \langle x|f(y)\rangle,$$

d.h. wenn f eine adjungierte Abbildung f^* besitzt und $f = f^*$ gilt.

Beispiel.

1. Betrachte $V = \mathbb{R}^2$ mit dem Standardskalarprodukt. Die Abbildung

   $$f\colon V \to V, \quad f(x) = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} x$$

   ist selbstadjungiert, da für alle $x, y \in V$ gilt:

   $$\langle f(x)|y\rangle = (\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} x)y = (x \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}^\top)y = x(\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}^\top y) = \langle x|f(y)\rangle.$$

2. Betrachte den Vektorraum $V = \{\, f \in C^\infty([0,1],\mathbb{R}) : f(0) = f'(0) = \cdots = 0,\ f(1) = f'(1) = \cdots = 0 \,\}$ mit dem Skalarprodukt 04-22

   $$\langle f|g\rangle = \int_0^1 f(t)g(t)dt.$$

   Dann ist $F\colon V \to V$, $F(f) = -f''$ selbstadjungiert, denn für alle $f, g \in V$ gilt

   $$\begin{aligned}\langle F(f)|g\rangle &= \int_0^1 (-f''(t))g(t)dt = -\big[f'(t)g(t)\big]_0^1 - \int_0^1 (-f'(t))g'(t)dt\\ &= 0 + \int_0^1 f'(t)g'(t)dt = \big[f(t)g'(t)\big]_0^1 - \int_0^1 f(t)g''(t)dt = \langle f|F(g)\rangle.\end{aligned}$$

Satz 105. (Spektralsatz) Sei V ein endlichdimensionaler Skalarproduktraum und $f\colon V \to V$ ein Endomorphismus. Dann gilt: f ist genau dann symmetrisch wenn V eine ONB aus Eigenvektoren von f hat.

Beweis. „⇒“ Sei B eine Orthonormalbasis von V. Dann entspricht das Skalarprodukt auf V dem Standardskalarprodukt für die Koordinatenvektoren bezüglich B. Aus $\langle f(x)|y\rangle = \langle x|f(y)\rangle$ für alle $x, y \in V$ folgt $(Ax)y = x(Ay)$ für alle $x, y \in \mathbb{R}^n$ ($n = \dim V$). Daraus folgt, dass A symmetrisch ist. Nach Satz 95 ist jede symmetrische Matrix diagonalisierbar. Nach Satz 96 ist $\mathbb{R}^n$ die direkte Summe der Eigenräume von A. Da ein Vektor $x \in V$ genau dann ein Eigenvektor von f ist, wenn seine Koordinatendarstellung ein Eigenvektor von A ist, folgt, dass V die direkte Summe der Eigenräume f ist. Nach Satz 100 hat jeder Eigenraum eine ONB. Wie im Beispiel nach Def. 62 zeigt man, dass Eigenräume symmetrischer Abbildungen senkrecht aufeinander stehen. Damit ist die Vereinigung der ONBs der Eigenräume von f eine ONB von V. 04-23

„⇐“ Sei $\{b_1, \ldots, b_n\}$ eine ONB von V bestehend aus Eigenvektoren von f. Seien $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ die zugehörigen Eigenwerte. (Es wird nicht angenommen, dass diese paarweise verschieden sind.) Zu zeigen:

$$\forall\, x, y \in V : \langle f(x)|y\rangle = \langle x|f(y)\rangle.$$

Wegen der Linearität des Skalarprodukts genügt es, die Gleichung für Basiselemente zu überprüfen. In der Tat gilt

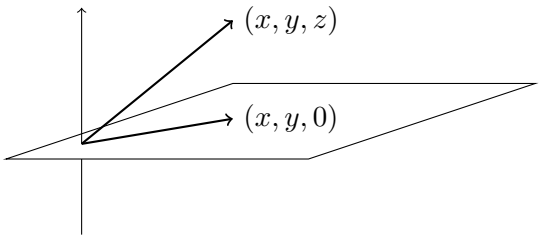
$$\langle f(b_i)|b_j\rangle = \langle \lambda_i b_i|b_j\rangle = \left\{ \begin{array}{ll} \lambda_i & \text{falls } i = j \\ 0 & \text{sonst} \end{array} \right\} = \langle b_i|\lambda_j b_j\rangle = \langle b_i|f(b_j)\rangle. \qquad \blacksquare$$

35 Projektionen und Isometrien

Definition 67. Sei V ein Vektorraum. Eine lineare Abbildung $h\colon V \to V$ heißt *Projektion,* falls gilt $h^2 = h$.

Beispiel.

1. $h\colon \mathbb{R}^3 \to \mathbb{R}^3$ mit $h(x, y, z) = (x, y, 0)$ ist eine Projektion. Ihr Bild ist die (x, y)-Ebene.



2. Sei $A = \begin{pmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{pmatrix}$. Dann ist $h\colon \mathbb{R}^3 \to \mathbb{R}^3$ mit $h(x) = Ax$ eine Projektion, denn es gilt $A^2 = A$.

3. $h\colon \mathbb{K}[[X]] \to \mathbb{K}[[X]]$ mit $h(\sum_{n=0}^{\infty} a_n X^n) = \sum_{n=0}^{20} a_n X^n$ ist eine Projektion.

4. Sei V ein $\mathbb{K}$-Vektorraum, U ein Unterraum von V und W ein Komplementärraum von U. Dann hat jedes $v \in V$ eine eindeutige Darstellung $v = u + w$ mit $u \in U$ und $w \in W$. Die lineare Abbildung $h\colon V \to V$, die jedem v den Vektor $u \in U$ aus dieser Darstellung zuordnet, ist eine Projektion.

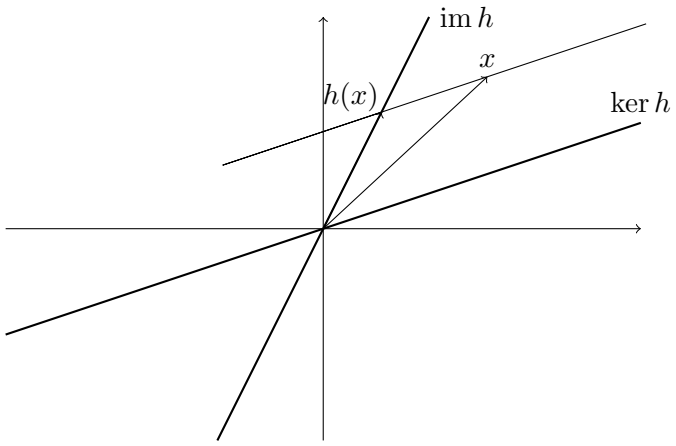
Satz 106. Sei V ein Vektorraum und $h\colon V \to V$ eine Projektion. Dann gilt $V = \ker h \oplus \operatorname{im} h$.

Beweis. Zunächst gilt $\ker h \cap \operatorname{im} h = \{0\}$, denn für ein beliebiges $v \in \ker h \cap \operatorname{im} h$ gilt $h(v) = 0$ und $v = h(w)$ für ein $w \in V$. Dann $0 = h(v) = h(h(w)) = h(w)$, da h eine Projektion ist. Dann ist $w \in \ker h$, und dann ist $v = h(w) = 0$. Damit ist die Summe direkt.

Es bleibt zu zeigen, dass zu jedem $v \in V$ ein $u \in \ker h$ und ein $w \in \operatorname{im} h$ existiert, so dass $v = u + w$. Sei $v \in V$ beliebig und wähle $u = v - h(v)$ und $w = h(v)$. Dann gilt $v = u + w$ und $w \in \operatorname{im} h$ und wegen $h(u) = h(v - h(v)) = h(v) - h(h(v)) = h(v) - h(v) = 0$ auch $u \in \ker h$. ∎ 04-22

Beispiel.

1. Betrachte die Projektion $h\colon \mathbb{R}^2 \to \mathbb{R}^2$, $h(x) = Ax$ mit $A = \frac{1}{5}\begin{pmatrix} -1 & 3 \\ -2 & 6 \end{pmatrix}$. Es gilt $\ker A = \langle \begin{pmatrix} 3 \\ 1 \end{pmatrix} \rangle$ und $\operatorname{im} A = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix} \rangle$. Geometrisch handelt es sich bei einer Projektion h immer um eine Parallelprojektion in Richtung von $\ker h$ auf den Unterraum $\operatorname{im} h$. 05-06



2. Dass aus $V = \ker h \oplus \operatorname{im} h$ im allgemeinen **nicht** folgt, dass h eine Projektion ist, sieht man an dem einfachen Beispiel $V = \mathbb{R}$, $h\colon V \to V$, $h(x) = 2x$. Hier gilt $\ker h = \{0\}$ und $\operatorname{im} h = \mathbb{R}$, also $V = \ker h \oplus \operatorname{im} h$. Aber h ist keine Projektion, denn $h(h(x)) = 4x \neq 2x = h(x)$ für $x \neq 0$.

Definition 68. Sei V ein Skalarproduktraum. Eine Projektion $h\colon V \to V$ heißt *Orthogonalprojektion,* falls für alle $x, y \in V$ gilt $x - h(x) \perp h(y)$. 04-21

Beispiel. 04-21

1. Wenn $\mathbb{R}^3$ mit dem Standardskalarprodukt versehen ist, ist die Abbildung $h\colon \mathbb{R}^3 \to \mathbb{R}^3$ mit $h(x,y,z) = (x,y,0)$ eine Orthogonalprojektion, denn für alle $x = (x_1,x_2,x_3) \in \mathbb{R}^3$ und alle $y = (y_1,y_2,y_3) \in \mathbb{R}^3$ gilt $x - h(x) = (0,0,x_3)$ und $h(y) = (y_1,y_2,0)$, also $\langle x - h(x) | h(y)\rangle = 0$, wie gefordert.

05-06 2. Die lineare Abbildung $h\colon \mathbb{R}^2 \to \mathbb{R}^2$ mit $h(x) = \frac{1}{5}\begin{pmatrix} -1 & 3 \\ -2 & 6 \end{pmatrix} x$ aus dem vorherigen Beispiel ist zwar eine Projektion, aber keine Orthogonalprojektion (zumindest nicht bezüglich des Standardskalarprodukts). Zum Beispiel für $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ gilt $\langle h(x) - x | y\rangle = \langle -\frac{1}{5}\binom{6}{2} | \frac{1}{5}\binom{3}{6}\rangle = -\frac{6}{5} \neq 0$.

3. Sei V ein Skalarproduktraum und $A = \{a_1,\dots,a_n\} \subseteq V$ ein ONS. Dann folgt aus Satz 99, dass die lineare Abbildung

$$h\colon V \to V, \quad h(x) = \sum_{i=1}^{n} \langle x | a_i\rangle a_i$$

eine Orthogonalprojektion ist.

Satz 107. Sei V ein Skalarproduktraum und $h\colon V \to V$ eine Projektion. Dann gilt: h ist genau dann eine Orthogonalprojektion, wenn $\ker h \perp \operatorname{im} h$.

Beweis. „⇒“ h ist eine Orthogonalprojektion. Seien $v \in \ker h$ und $w \in \operatorname{im} h$. Dann gilt
04-21 $h(v) = 0$ und $h(w) = w$. Daraus folgt: $\langle v | w\rangle = \langle v - h(v) | w\rangle = \langle v - h(v) | h(w)\rangle = 0$.

„⇐“ Sei $x \in V$. Dann gilt $h(x) \in \operatorname{im} h$, und weil h eine Projektion ist, gilt $h(x - h(x)) = h(x) - h(h(x)) = h(x) - h(x) = 0$, also $x - h(x) \in \ker h$. Wegen $\ker h \perp \operatorname{im} h$ folgt $x - h(x) \perp h(y)$
04-21 für alle $y \in V$. ∎

04-27 **Satz 108.** Sei V ein Skalarproduktraum und $U \subseteq V$ ein Unterraum mit $\dim U < \infty$. Dann existiert genau eine Orthogonalprojektion $h\colon V \to V$ mit $\operatorname{im} h = U$.

Beweis. Existenz: Nach Satz 100 hat U eine ONB, etwa $B = \{b_1,\dots,b_n\}$. Setzen wir $h(x) = \sum_{i=1}^{n} \langle x | b_i\rangle b_i$, so ist h nach Satz 99 eine Orthogonalprojektion, und für diese gilt $h(b_i) = b_i$
04-27 $(i = 1,\dots,n)$ und $h(x) \in U$ für alle $x \in V$. Darum ist $\operatorname{im} h = U$.

Eindeutigkeit: Seien $h, \tilde{h}\colon V \to V$ zwei Orthogonalprojektionen mit $\operatorname{im} h = \operatorname{im} \tilde{h} = U$. Zu
04-27 zeigen: für alle $v \in V$ gilt $h(v) = \tilde{h}(v)$. Sei $v \in V$ beliebig. Dann gilt $v - h(v) \in \ker h$ und
04-21 $v - \tilde{h}(v) \in \ker \tilde{h}$, weil h und $\tilde{h}$ Projektionen sind. Nach Satz 107 gilt $v - h(v) \perp \operatorname{im} h = U$ und $v - \tilde{h}(v) \perp \operatorname{im} \tilde{h} = U$. Wegen $h(v) - \tilde{h}(v) \in U$ folgt

$$\langle v - h(v) | h(v) - \tilde{h}(v)\rangle = 0,$$
$$\langle v - \tilde{h}(v) | h(v) - \tilde{h}(v)\rangle = 0.$$

Subtraktion dieser beiden Gleichungen liefert $\langle h(v) - \tilde{h}(v) | h(v) - \tilde{h}(v)\rangle = 0$, also $h(v) = \tilde{h}(v)$.
∎

Definition 69. Sei V ein Skalarproduktraum. Eine lineare Abbildung $h\colon V \to V$ heißt *Isometrie,* falls für alle $x \in V$ gilt $\|x\| = \|h(x)\|$.

Beispiel. Sei $V = \mathbb{R}^2$ versehen mit dem Standardskalarprodukt, $\phi \in \mathbb{R}$,

$$A = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix}$$

und $h\colon V \to V$, $h(x) = Ax$. Für jedes $x = (x_1, x_2) \in \mathbb{R}^2$ gilt dann

$$\begin{aligned} \|h(x)\| &= \| \begin{pmatrix} x_1\cos(\phi) + x_2\sin(\phi) \\ -x_1\sin(\phi) + x_2\cos(\phi) \end{pmatrix} \| \\ &= \sqrt{x_1^2\cos(\phi)^2 + x_2^2\sin(\phi)^2 + x_1^2\sin(\phi)^2 + x_2^2\cos(\phi)^2} \\ &= \sqrt{x_1^2 + x_2^2} = \|x\|. \end{aligned}$$

Also ist h eine Isometrie.

Es handelt sich um die Abbildung, die die Ebene im Uhrzeigersinn um den Winklel ϕ um den Ursprung dreht. (Für $\phi = \pi/2$ gilt zum Beispiel $h(1,0) = (0,-1)$.)

Satz 109. Sei V ein Skalarproduktraum und $h\colon V \to V$ eine lineare Abbildung. Dann gilt: h ist genau dann eine Isometrie, wenn für alle $v, w \in V$ gilt $\langle v|w\rangle = \langle h(v)|h(w)\rangle$.

Beweis. „⇒“ h ist eine Isometrie. Seien $v, w \in V$. Dann gilt $\|v - w\| = \|h(v - w)\|$, also $\langle v-w|v-w\rangle = \langle h(v-w)|h(v-w)\rangle$, also $\langle v|v\rangle - 2\langle v|w\rangle + \langle w|w\rangle = \langle h(v)|h(v)\rangle - 2\langle h(v)|h(w)\rangle + \langle h(w)|h(w)\rangle$. Wegen $\langle v|v\rangle = \|v\|^2 = \|h(v)\|^2 = \langle h(v)|h(v)\rangle$ und $\langle w|w\rangle = \langle h(w)|h(w)\rangle$ folgt 04-21
daraus $\langle v|w\rangle = \langle h(v)|h(w)\rangle$, wie gefordert.

„⇐“ Gilt $\langle v|w\rangle = \langle h(v)|h(w)\rangle$ für alle $v, w \in V$, so gilt insbesondere $\langle v|v\rangle = \langle h(v)|h(v)\rangle$ für alle $v \in V$. Daraus folgt $\|v\| = \|h(v)\|$ für alle $v \in V$. ∎

Beispiel. Sei $V = \mathbb{R}^n$ mit dem Standardskalarprodukt. Sei $B = (b_1, \ldots, b_n) \in \mathbb{R}^{n\times n}$ eine geordnete ONB von V und $h\colon V \to V$, $h(x) = Bx$. Dann ist h eine Isometrie, denn für je zwei Vektoren $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ und $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ gilt

$$\langle h(x)|h(y)\rangle = \langle \sum_{i=1}^{n} x_i b_i | \sum_{j=1}^{n} y_j b_j \rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j \langle b_i|b_j\rangle = x_1y_1 + \cdots + x_ny_n = \langle x|y\rangle.$$

Definition 70. Eine Matrix in $\mathbb{R}^{n\times n}$, deren Spalten bezüglich des Standardskalarprodukts eine ONB von $\mathbb{R}^n$ bilden, heißt *orthogonal.*

Satz 110. Sei $A \in \mathbb{R}^{n\times n}$.

1. A ist genau dann orthogonal, wenn A invertierbar ist und $A^{-1} = A^\top$ gilt.
2. Ist A orthogonal, so gilt $\det A \in \{-1, 1\}$ und für jeden Eigenwert $\lambda \in \mathbb{R}$ von A gilt $\lambda \in \{-1, 1\}$.

Beweis.

1. „⇐“ Aus $A^{-1} = A^\top$ folgt $AA^\top = I_n$. Wenn also $a_1, \dots, a_n \in \mathbb{R}^n$ die Spalten von A sind, dann gilt 04-21

$$\langle a_i | a_j \rangle = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{sonst} \end{cases}$$

für alle i, j. Damit ist $\{a_1, \dots, a_n\}$ eine ONB bezüglich des Standardskalarprodukts.

04-21 „⇒“ Ist $A = (a_1, \dots, a_n) \in \mathbb{R}^{n \times n}$ orthogonal, so gilt $AA^\top = I_n$ und $A^\top A = I_n$, und daraus folgt sofort die Behauptung.

2. Für die Determinante gilt

$$\det(A) \underset{\text{Satz 28}}{\uparrow} \det(A^\top) \underset{\text{Teil 1}}{\uparrow} \det(A^{-1}) \underset{\text{Satz 31}}{\uparrow} \frac{1}{\det(A)},$$

also $\det(A)^2 = 1$, also $\det(A) \in \{-1, 1\}$.

Sei $\lambda \in \mathbb{R}$ ein Eigenwert von A und $v \in \mathbb{R}^n \setminus \{0\}$ ein Eigenvektor von A zum Eigenwert λ. Da A orthogonal ist, ist $h\colon \mathbb{R}^n \to \mathbb{R}^n$, $h(x) = Ax$ eine Isometrie. Damit gilt $\|v\| = \|Av\| = \|\lambda v\| = |\lambda| \, \|v\|$, und wegen $\|v\| \neq 0$ folgt daraus $|\lambda| = 1$, also $\lambda \in \{-1, 1\}$. ∎

Satz 111. Sei V ein Skalarproduktraum, $\dim V = n < \infty$ und $h\colon V \to V$ eine lineare Abbildung. Dann sind folgende Aussagen äquivalent:

1. h ist eine Isometrie.
2. Für jede ONB $\{b_1, \dots, b_n\}$ von V ist $\{h(b_1), \dots, h(b_n)\}$ eine ONB von V.
3. Es gibt eine ONB $\{b_1, \dots, b_n\}$ von V, so dass $\{h(b_1), \dots, h(b_n)\}$ eine ONB von V ist.
4. Für jede ONB B von V ist die Abbildungsmatrix von h bezüglich B und B orthogonal.
5. Es gibt eine ONB B von V, so dass die Abbildungsmatrix von h bezüglich B und B orthogonal ist.

Beweis. Wir verwenden die Notation

$$\delta_{i,j} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{sonst} \end{cases}$$

für $i, j \in \mathbb{N}$.

(1)⇒(2). Für alle $i, j = 1, \dots, n$ gilt $\langle h(b_i) | h(b_j) \rangle = \langle b_i | b_j \rangle = \delta_{i,j}$. Damit ist $\{h(b_1), \dots, h(b_n)\}$ ein ONS, und aus Dimensionsgründen auch eine ONB.

(2)⇒(4). Sei $B = (b_1, \dots, b_n) \in \mathbb{R}^{n \times n}$ eine geordnete ONB und $M = ((m_{i,j}))_{i,j=1}^n$ die Abbildungsmatrix von h bezüglich B und B. Die Spalten von M sind dann die Koordinatenvektoren von $h(b_i)$ bezüglich B, d.h. für alle $j = 1, \dots, n$ gilt $h(b_j) = m_{1,j} b_1 + \dots + m_{n,j} b_n$. Nach Annahme gilt $\langle h(b_i) | h(b_j) \rangle = \delta_{i,j}$ für alle i, j, also

$$\langle m_{1,i} b_1 + \dots + m_{n,i} b_n | m_{1,j} b_1 + \dots + m_{n,j} b_n \rangle = \delta_{i,j},$$

also

$$m_{1,i}m_{1,j} + \cdots + m_{n,i}m_{n,j} = \delta_{i,j}$$

für alle i, j. Das bedeutet, dass M orthogonal ist.

(4)⇒(5). Klar, weil V nach Satz 100 mindestens eine ONB hat.

(5)⇒(3). Sei $B = (b_1, \ldots, b_n)$ die ONB und $M = ((m_{i,j}))_{i,j=1}^n$ die Abbildungsmatrix von M. Wir zeigen $\langle h(b_i)|h(b_j)\rangle = \delta_{i,j}$ für alle i, j. Seien also i, j beliebig. Dann gilt

$$\begin{aligned}\langle h(b_i)|h(b_j)\rangle &= \langle m_{1,i}b_1 + \cdots + m_{n,i}b_n | m_{1,j}b_1 + \cdots + m_{n,j}b_n\rangle \\ &= m_{1,i}m_{1,j} + \cdots + m_{n,i}m_{n,j} = \delta_{i,j},\end{aligned}$$

weil M nach Annahme orthogonal ist.

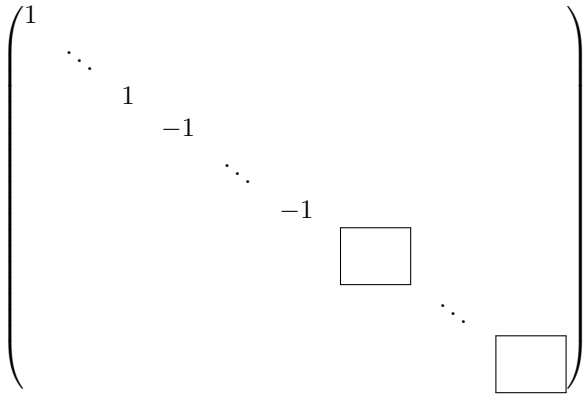
(3)⇒(1). Sei $\{b_1, \ldots, b_n\}$ eine ONB, so dass auch $\{h(b_1), \ldots, h(b_n)\}$ eine ONB ist.

Sei $x \in V$ beliebig, etwa $x = \alpha_1 b_1 + \cdots + \alpha_n b_n$ für gewisse $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$. Dann gilt

$$\begin{aligned}\|h(x)\|^2 &= \|\alpha_1 h(b_1) + \cdots + \alpha_n h(b_n)\|^2 \\ &= \alpha_1^2 + \cdots + \alpha_n^2 \\ &= \|\alpha_1 b_1 + \cdots + \alpha_n b_n\|^2 = \|x\|^2.\end{aligned}$$

Daraus folgt die Behauptung. ∎

Durch Satz 111 werden Isometrien recht genau charakterisiert. Man kann sich etwas mehr Mühe machen und orthogonale Matrizen genauer charakterisieren. Es lässt sich nämlich zeigen, dass eine 04-27
Matrix $A \in \mathbb{R}^{n\times n}$ genau dann orthogonal ist, wenn es eine orthogonale Matrix $B \in \mathbb{R}^{n\times n}$ gibt, so dass $B^\top AB$ die folgende Form



hat, wobei jedes $\square$ für einen 2×2-Block der Form $\begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix}$ für ein $\phi \in \mathbb{R}$ steht. Wie im Beispiel nach Defintion 69 gezeigt, entspricht ein solcher 2×2-Block einer Drehung in einer Ebene um den Winkel ϕ.

Eine Isometrie h mit $\det h = 1$ nennt man deshalb auch *Drehung* (engl. *rotation*). Isometrien h mit $\det h = -1$ heißen *Drehspiegelungen.*

Im Fall $n = 3$ hat eine Isometrie bezüglich einer geeigneten ONB (b_1, b_2, b_3) typischerweise eine Darstellungsmatrix der Form

$$\begin{pmatrix} \pm 1 & & \\ & \cos(\phi) & \sin(\phi) \\ & -\sin(\phi) & \cos(\phi) \end{pmatrix}.$$

Dabei findet die Drehung in der Ebene $\langle b_2, b_3\rangle$ statt, d.h. um die darauf senkrecht stehende Gerade $\langle b_1\rangle$ als Drehachse. Wenn der Eintrag oben links eine -1 ist, dann bewirkt die lineare Abbildung zusätzlich zur Drehung eine Spiegelung an der Ebene $\langle b_2, b_3\rangle$.

Bemerkenswert ist auch, dass die Verkettung zweier Isometrien $h_1, h_2 \colon V \to V$ wieder eine Isometrie ist, weil natürlich $\|x\| = \|h_1(x)\| = \|h_2(h_1(x))\|$ für alle $x \in V$ gilt. Da orthogonale Matrizen nach Satz 110 auch invertierbar sind, folgt daraus für $V = \mathbb{R}^n$, dass die Menge

$$\mathrm{O}(n) := \{\, A \in \mathbb{R}^{n\times n} : A \text{ ist orthogonal} \,\}$$

eine Untergruppe von $\mathrm{GL}(n, \mathbb{R})$ ist. Man nennt sie die *orthogonale Gruppe* der Größe n. Die Untergruppe

$$\mathrm{SO}(n) := \mathrm{O}(n) \cap \mathrm{SL}(n, \mathbb{R}) = \{\, A \in \mathbb{R}^{n\times n} : A \text{ ist orthogonal und } \det A = 1 \,\}$$

heißt die *spezielle orthogonale Gruppe* der Größe n.

36 Die Singulärwertzerlegung

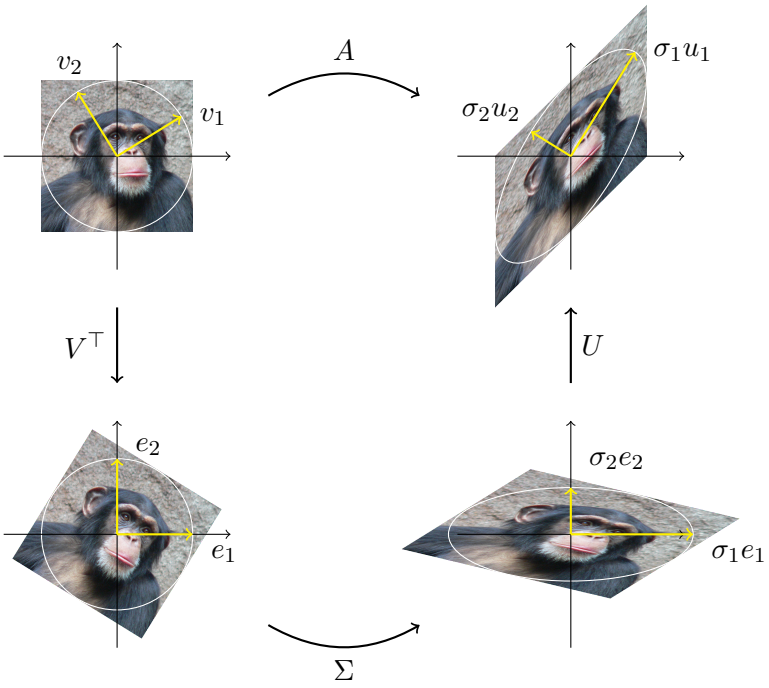
Diagonalisierbare Matrizen entsprechen Transformationen des Raumes, die sich als Streckungen in bestimmte Richtungen (nämlich die der Eigenvektoren) charakterisieren lassen. Dass nicht alle Matrizen diagonalisierbar sind, bedeutet also geometrisch, dass es lineare Abbildungen gibt, die nicht einfach einer Kombination von Streckungen in bestimmte Richtungen entsprechen. Das Resultat dieses Abschnitts lässt sich so zusammenfassen, dass sich jede lineare Abbildung $h \colon \mathbb{R}^n \to \mathbb{R}^n$ darstellen lässt als eine Kombination von Drehungen und Streckungen. Die Singulärwerte sind dabei die Streckungsfaktoren. Sie sind im allgemeinen verschieden von den Eigenwerten.

Beispiel. Die Matrix $A = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ ist nicht diagonalisierbar. Ihr einziger Eigenwert ist 1 und der dazugehörige Eigenraum ist $E_1 = \langle \binom{0}{1} \rangle \subsetneq \mathbb{R}^2$.

Die Matrix A lässt sich zerlegen als $A = U\Sigma V^\top$ mit

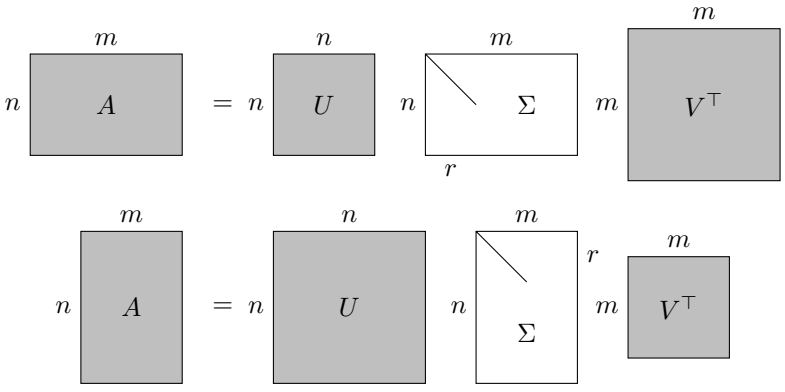
$$\begin{aligned}
U &= \frac{1}{\sqrt{10}} \begin{pmatrix} \sqrt{5-\sqrt{5}} & -\sqrt{5+\sqrt{5}} \\ \sqrt{5+\sqrt{5}} & \sqrt{5-\sqrt{5}} \end{pmatrix} = (u_1, u_2), \\
\Sigma &= \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{3+\sqrt{5}} & 0 \\ 0 & \sqrt{3-\sqrt{5}} \end{pmatrix} = \mathrm{diag}(\sigma_1, \sigma_2), \\
V &= \frac{1}{\sqrt{10}} \begin{pmatrix} \sqrt{5+\sqrt{5}} & -\sqrt{5-\sqrt{5}} \\ \sqrt{5-\sqrt{5}} & \sqrt{5+\sqrt{5}} \end{pmatrix} = (v_1, v_2).
\end{aligned}$$

Dabei sind U und V orthogonale Matrizen, bewirken also Drehungen der Ebene um den Ursprung. Die Matrix Σ ist eine Diagonalmatrix, bewirkt also Streckungen der Ebene in Richtung der Koordinatenachsen.



Zur Vereinfachung der Notation wollen wir in diesem Abschnitt annehmen, dass alle Räume $\mathbb{R}^n$ mit dem Standardskalarprodukt ausgestattet sind.

Definition 71. Sei $A \in \mathbb{R}^{n\times m}$. $U \in \mathbb{R}^{n\times n}$ und $V \in \mathbb{R}^{m\times m}$ seien orthogonale Matrizen und $\Sigma = ((\sigma_{i,j}))_{i,j=1}^{n,m} \in \mathbb{R}^{n\times m}$ sei so, dass $\sigma_{i,j} = 0$ für $i \neq j$ und $\sigma_{1,1} \geq \sigma_{2,2} \geq \cdots \geq 0$ gilt. Wenn $A = U\Sigma V^\top$ gilt, heißt (U, Σ, V) eine *Singulärwertzerlegung* (engl. *singular value decomposition,* SVD) von A.



Satz 112. Sei $A \in \mathbb{R}^{n\times m}$ und (U, Σ, V) eine Singulärwertzerlegung von A,

$$\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{pmatrix} \in \mathbb{R}^{n\times m}$$

Dann gilt $\|A\| = \sigma_1$.

Zur Erinnerung: Die Norm $\|A\|$ einer Matrix $A \in \mathbb{R}^{n\times m}$ ist definiert als der maximale Wert, der von $\|Ax\|$ erreicht wird, wenn x durch die Einheitskugelschale von $\mathbb{R}^m$ läuft (vgl. das Beispiel zu Def. 60). Aus analytischen Gründen gibt es für jede Matrix $A \in \mathbb{R}^{n\times m}$ einen Vektor $x \in \mathbb{R}^m$ mit $\|x\| = 1$ und $\|A\| = \|Ax\|$. Das werden wir im folgenden Beweis verwenden.

Beweis. Wegen $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ gilt $\|\Sigma\| = \sigma_1$. Da U, V orthogonale Matrizen sind, sind die zugehörigen linearen Abbildungen Isometrien, d.h. für jedes $x \in \mathbb{R}^m$ mit $\|x\| = 1$ gilt $\|V^\top x\| = 1$, damit $\|\Sigma V^\top x\| \leq \sigma_1$, damit $\|Ax\| = \|U\Sigma V^\top x\| \leq \sigma_1$. Daraus folgt $\|A\| \leq \sigma_1$.
Umgekehrt: Ist v die erste Spalte von V, so gilt $\|v\| = 1$ und $V^\top v = e_1$, weil V orthogonal ist. Dann ist $\Sigma V^\top v = \sigma_1 e_1$ und $\|Av\| = \|U\Sigma V^\top v\| = \sigma_1$, weil auch U orthogonal ist. Daraus folgt $\|A\| \geq \sigma_1$. ∎

Satz 113. Sei $A \in \mathbb{R}^{n\times m}$. Dann gilt: A hat eine Singulärwertzerlegung, und für je zwei Singulärwertzerlegungen (U, Σ, V), (U', Σ', V') von A gilt $\Sigma = \Sigma'$.

Beweis. Es ist leicht einzusehen, dass (U, Σ, V) genau dann eine Singulärwertzerlegung von A
04-29 ist, wenn $(V^\top, \Sigma^\top, U^\top)$ eine Singulärwertzerlegung von $A^\top$ ist. Wir können deshalb o.B.d.A. annehmen, dass $n \leq m$ gilt.
Für die Nullmatrix $A = 0$ ist $(I_n, 0, I_m)$ eine Singulärwertzerlegung, und eine andere Wahl für Σ ist dann nicht möglich, weil U, V als orthogonale Matrizen insbesondere invertierbar sind.
Für von Null verschiedene Matrizen zeigen wir die Behauptung durch Induktion nach n.
$n = 1$. In diesem Fall ist $A \in \mathbb{R}^{1\times m}$ ein Zeilenvektor und wir können $U = I_1$, $\Sigma =$
04-28 $(\|A\|, 0, \ldots, 0)$ und für $V \in \mathbb{R}^{m\times m}$ eine beliebige orthogonale Matrix wählen, deren erste Spalte $\frac{1}{\|A\|}A$ ist. Da $\|U\| = \|V\| = 1$ gelten muss, bleibt für Σ auch keine andere Möglichkeit.
$n-1 \to n$. Nach Satz 112 ist $\|A\|$ die einzige mögliche Wahl für den Eintrag σ_1 von Σ. Wähle einen Vektor $v_1 \in \mathbb{R}^m$ mit $\|A\| = \|Av_1\|$ und setze $u_1 = \frac{1}{\|A\|}Av_1$. (Wegen $A \neq 0$ ist $\|A\| \neq 0$.)
Ergänze $\{v_1\}$ zu einer ONB $\{v_1, \ldots, v_m\}$ von $\mathbb{R}^m$ und $\{u_1\}$ zu einer ONB $\{u_1, \ldots, u_n\}$ von $\mathbb{R}^n$. Setze $V_1 = (v_1, \ldots, v_m)$ und $U_1 = (u_1, \ldots, u_n)$.
Dann gilt $AV_1 = (\|A\|u_1, *, \ldots, *)$, also

$$U_1^\top A V_1 = \begin{pmatrix} \|A\| & a \\ 0 & A_2 \end{pmatrix} =: B$$

für gewisse $a \in \mathbb{R}^{m-1}$ und $A_2 \in \mathbb{R}^{(n-1)\times(m-1)}$. Aus

$$\| \begin{pmatrix} \|A\| & a \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} \|A\| \\ a \end{pmatrix} \| \geq \|A\|^2 + \langle a|a\rangle = \sqrt{\|A\|^2 + \langle a|a\rangle}\ \| \begin{pmatrix} \|A\| \\ a \end{pmatrix} \|$$

05-05 folgt $\|B\| \geq \sqrt{\|A\|^2 + \langle a|a\rangle}$. Andererseits gilt $\|B\| \leq \|A\|$, weil U_1 und V_1 orthogonal sind. Daraus folgt $\langle a|a\rangle = 0$, und daraus $a = 0$ und $\|A\| = \|B\|$. Aus $\|B\| = \|A\|$ folgt außerdem $\|A\| \geq \|A_2\|$.
Nach Induktionsvoraussetzung gibt es $U_2 \in \mathbb{R}^{(n-1)\times(n-1)}$, $V_2 \in \mathbb{R}^{(m-1)\times(m-1)}$ und eindeutig bestimmte $\sigma_2 \geq \sigma_3 \geq \cdots \geq 0$ mit $\|A\| \geq \sigma_2$ und

$$A_2 = U_2 \begin{pmatrix} \sigma_2 & & \\ & \sigma_3 & \\ & & \ddots \end{pmatrix} V_2^\top.$$

Wir erhalten mit

$$A = \underbrace{U_1 \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}}_{=:U} \underbrace{\begin{pmatrix} \|A\| & & \\ & \sigma_2 & \\ & & \ddots \end{pmatrix}}_{=:\Sigma} \underbrace{\left(V_1 \begin{pmatrix} 1 & 0 \\ 0 & V_2 \end{pmatrix} \right)^\top}_{=:V}$$

eine Darstellung der gewünschten Form. ∎

Beispiel. Sei $A = \begin{pmatrix} 4 & 0 \\ 3 & 5 \end{pmatrix}$. Um eine Singulärwertzerlegung von A zu berechnen, geht man vor wie im Beweis des Satzes. Als erstes braucht man einen Vektor $v_1 \in \mathbb{R}^2$ mit $\|v_1\| = 1$, für den $\|Av_1\|$ maximal wird. Wenn t durch die reellen Zahlen läuft, durchläuft $v_1 = (\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2})$ die Punkte auf dem Einheitskreis außer $(-1, 0)$. Für diese Punkte ist 04-28

$$\|Av_1\| = \sqrt{\left(4\frac{1-t^2}{1+t^2}\right)^2 + \left(3\frac{1-t^2}{1+t^2} + 5\frac{2t}{1+t^2}\right)^2} = \frac{\sqrt{5(5t^4 - 12t^3 + 10t^2 + 12t + 5)}}{1+t^2}.$$

Dieser Ausdruck wird maximal für $t = -1 \pm \sqrt{2}$, und für den isolierten Punkt $v = (-1, 0)$ ist $\|Av\|$ nicht größer. Daraus ergibt sich $\sigma_1 = \|A\| = 2\sqrt{10}$ und $v_1 = \pm\frac{1}{\sqrt{2}}(1, 1)$. Für $v_1 = \frac{1}{\sqrt{2}}(1, 1)$ folgt $u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{5}}(1, 2)$.
Als Ergänzung zu Orthonormalbasen wählen wir $v_2 = \frac{1}{\sqrt{2}}(1, -1)$ und $u_2 = \frac{1}{\sqrt{5}}(2, -1)$. Setze $V_1 = (v_1, v_2)$, $U_1 = (u_1, u_2)$. Wir haben dann

$$U_1^\top A V_1 = \begin{pmatrix} 2\sqrt{10} & \\ & \sqrt{10} \end{pmatrix}.$$

Damit ist (U, Σ, V) mit

$$U = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}, \qquad \Sigma = \sqrt{10} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \qquad V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

eine Singulärwertzerlegung von A.

Definition 72. Sei $A \in \mathbb{R}^{n \times m}$ und (U, Σ, V) eine Singulärwertzerlegung von A. Dann heißen die Einträge $\sigma_1, \ldots, \sigma_{\min(n,m)}$ auf der Diagonalen von Σ die *Singulärwerte* von A.

Satz 114. Sei $A \in \mathbb{R}^{n \times m}$ und sei (U, Σ, V) eine Singulärwertzerlegung von A. Es seien $\sigma_1, \ldots, \sigma_{\min(n,m)}$ die Singulärwerte von A, und es sei $r \in \{1, \ldots, \min(n, m)\}$ maximal mit $\sigma_r \neq 0$. Schreibe $U = (u_1, \ldots, u_n)$, $V = (v_1, \ldots, v_m)$. Dann gilt:

1. $\operatorname{Rang}(A) = r$.
2. $\{u_1, \ldots, u_r\}$ ist eine ONB von $\operatorname{im} A$.
3. $\{u_{r+1}, \ldots, u_n\}$ ist eine ONB von $\operatorname{coker} A$.
4. $\{v_1, \ldots, v_r\}$ ist eine ONB von $\operatorname{coim} A$. 04-29
5. $\{v_{r+1}, \ldots, v_m\}$ ist eine ONB von $\ker A$.

Insbesondere gilt $\ker A \perp \operatorname{coim} A$ und $\operatorname{coker} A \perp \operatorname{im} A$.

Beweis. Übung. ∎

Satz 115. Sei $A \in \mathbb{R}^{n\times m}$. Dann gilt: $\sigma \in \mathbb{R} \setminus \{0\}$ ist genau dann ein Singulärwert von A wenn σ^2 ein Eigenwert von $AA^\top \in \mathbb{R}^{n\times n}$ ist.

Beweis. Ist (U, Σ, V) eine Singulärwertzerlegung von A, so gilt $A = U\Sigma V^\top$ und $AA^\top = U\Sigma V^\top V\Sigma^\top U^\top = U\Sigma\Sigma^\top U^{-1}$. Die Matrix $\Sigma\Sigma^\top \in \mathbb{R}^{n\times n}$ ist eine Diagonalmatrix, auf deren Diagonale die Quadrate der Diagonaleinträge von Σ stehen. Im Fall $n > m$ ist die Diagonale von $\Sigma\Sigma^\top$ länger als die von Σ; die zusätzlichen Diagonaleinträge sind Null. ∎

05-06 **Beispiel.** Für $A = \begin{pmatrix} 4 & 0 \\ 3 & 5 \end{pmatrix}$ gilt $AA^\top = \begin{pmatrix} 16 & 12 \\ 12 & 34 \end{pmatrix}$. Die Eigenwerte von $AA^\top$ sind 40 und 10, die Singulärwerte von A sind daher $2\sqrt{10}$ und $\sqrt{10}$.
Wir können darüber hinaus auch U und V berechnen. Da $AA^\top$ symmetrisch ist, ist die Matrix nach Satz 105 diagonalisierbar und es gibt eine ONB aus Eigenvektoren. Tatsächlich gilt

$$AA^\top = U \begin{pmatrix} 40 & 0 \\ 0 & 10 \end{pmatrix} U^\top$$

für $U = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$.
Gesucht ist jetzt eine orthogonale Matrix V mit $A = U\Sigma V^\top$, wobei $\Sigma = \operatorname{diag}(2\sqrt{10}, \sqrt{10})$ ist. Im vorliegenden Beispiel sind A und U und Σ invertierbar, so dass die einzige mögliche Wahl für $V^\top$ die Matrix $\Sigma^{-1}U^{-1}A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. Also ist $V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$.

Satz 116. Sei $A \in \mathbb{R}^{n\times m} \setminus \{0\}$, $r = \operatorname{Rang} A$. Dann gibt es Vektoren $v_1, \dots, v_r \in \mathbb{R}^m$ und $u_1, \dots, u_r \in \mathbb{R}^n$, so dass

$$A = u_1 v_1 + \cdots + u_r v_r,$$

wobei mit $u_i v_i$ das Matrixprodukt des Spaltenvektors u_i (als $(n \times 1)$-Matrix) mit dem Zeilenvektor v_i (als $(1 \times m)$-Matrix) gemeint ist.

Beweis. Sei (U, Σ, V) eine Singulärwertzerlegung von A. Schreibe $U = (u_1, \dots, u_n)$, $V = (v_1, \dots, v_m)$, und $\sigma_1 \geq \cdots \geq \sigma_r > 0$ für die Singulärwerte von A. Dann gilt $A = U\Sigma V^\top$, also $A = u_1\sigma_1 v_1 + \cdots + u_r\sigma_r v_r$. ∎

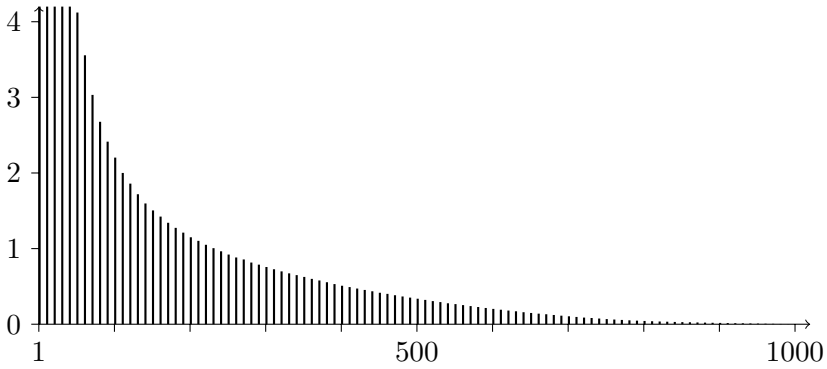
Wenn r viel kleiner ist als $\min(n, m)$, dann braucht die Darstellung von A aus Satz 116 wesentlich weniger Speicher als die gewöhnliche Darstellung, nämlich nur $r(n + m)$ Zahlen statt nm Zahlen. In praktischen Anwendungen gilt zwar meistens $\operatorname{Rang}(A) = \min(n, m)$, aber man kann A zu einer Matrix mit kleinerem Rang „runden“, indem man kleine Singulärwerte durch Null ersetzt. Je näher die gestrichenen Singulärwerte bei 0 liegen, desto weniger Information über A geht bei der Streichung verloren.

Beispiel. Ein Schwarz-Weiß-Bild lässt sich durch eine Matrix codieren, deren (i, j)-Eintrag den Grauwert des Bildes am Bildpunkt (i, j) beinhaltet. Betrachte das Bild

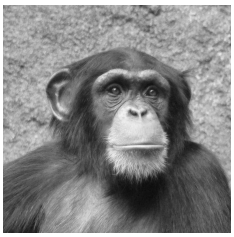
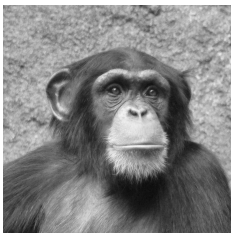
$$A = \left(\right. \quad \left. \right) \in \mathbb{R}^{1000\times 1000},$$

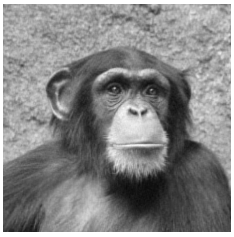
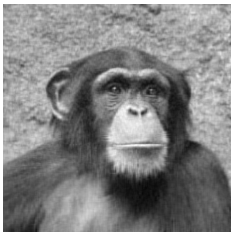
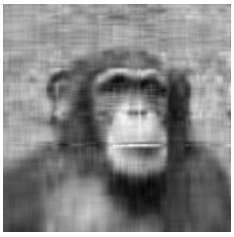


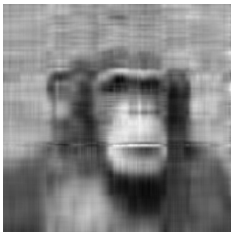
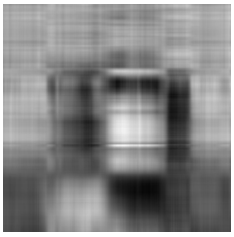
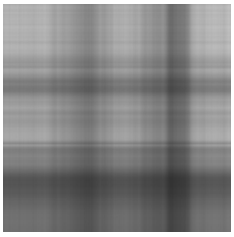
wobei schwarz durch 0 und weiß durch 1 codiert wird. Im folgenden Diagramm ist jeder zehnte Singulärwert von A dargestellt. Die größten Singulärwerte lauten 513.293, 106.563, 61.0108, 52.4002, 34.7057, …; die zugehörigen Balken sind im Diagramm abgeschnitten. Die kleisten sind 0.000253564, 0.000220873, 0.0000831866; auch sie sind kaum zu sehen.



Sei $A = U\Sigma V^\top$ eine Singulärwertzerlegung von A. Für $k \in \{1, \dots, 1000\}$ sei Σ_k die Diagonalmatrix, die man aus $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \dots, \sigma_{1000})$ erhält, indem man die Einträge $\sigma_{k+1}, \dots, \sigma_{1000}$ auf Null setzt. Dann lässt sich die Matrix $A_k := U\Sigma_k V^\top \in \mathbb{R}^{1000\times 1000}$ gemäß Satz 116 durch $2000k$ Zahlen codieren. Je kleiner k gewählt wird, desto größer wird die Einsparung gegenüber der Abspeicherung aller $1000 \times 1000 = 10^6$ Einträge von A. Allerdings sollte man k auch nicht zu klein wählen, weil sonst die Bildqualität sichtbar beeinträchtigt wird.

$A_{700} =$ 
$A_{400} =$ 
$A_{100} =$ 

$A_{70} =$ 
$A_{40} =$ 
$A_{10} =$ 

$A_7 =$ 
$A_4 =$ 
$A_1 =$ 

Teil VII

Modultheorie

37 Grundbegriffe

Zur Erinnerung: Ein Ring ist eine algebraische Struktur $R = (R, +, \cdot)$ mit zwei Verknüpfungen $+, \cdot \colon R \times R \to R$, die die gleichen Gesetze erfüllen wie bei Körpern, außer dass $\cdot$ nicht kommutativ sein muss und dass es nicht zu jedem von Null verschiedenen Element des Rings ein multiplikatives Inverses geben muss. Typische Beispiele für Ringe sind $\mathbb{Z}$, $\mathbb{K}[X]$ und $\mathbb{K}^{n\times n}$.

In Def. 15 haben wir noch Ringe mit und ohne Einselement unterschieden. Ab jetzt betrachten wir nur noch Ringe mit Einselement, ohne immer explizit „mit Eins“ dazuzusagen.

Definition 73. Sei R ein Ring, $(M, +)$ eine abelsche Gruppe und

$$\cdot \colon R \times M \to M$$

eine Abbildung, so dass für alle $r, r_1, r_2 \in R$ und alle $m, m_1, m_2 \in M$ gilt

1. $(r_1 + r_2) \cdot m = r_1 \cdot m + r_2 \cdot m$ 05-05
2. $(r_1 r_2) \cdot m = r_1 \cdot (r_2 \cdot m)$
3. $r \cdot (m_1 + m_2) = r \cdot m_1 + r \cdot m_2$
4. $1 \cdot m = m$

Dann heißt $(M, +, \cdot)$ ein *Modul* (engl. *module*) über R, oder kurz: ein *R-Modul.*

Wenn M zugleich ein R-Modul und auch eine Teilmenge von R ist und die auf M erklärte Addition mit der Addition von R übereinstimmt, dann heißt M ein *Ideal* (engl. *ideal*) von R. Mit anderen Worten: Ein Ideal von R ist eine Teilmenge $I \subseteq R$, so dass $a + b \in I$ für alle $a, b \in I$ und $ra \in I$ für alle $a \in I$ und alle $r \in R$ gilt. 05-05 05-05

Die Definition unterscheidet sich von der Definition eines Vektorraums (Def. 29) nur dadurch, dass statt eines Körpers ein Ring zugrunde gelegt wird. Der Begriff des Moduls ist also eine Verallgemeinerung des Begriffs des Vektorraums.

Zur Sprechweise: Im Sinn von Def. 73 spricht man von „dem Modul“ (männlich, Betonung auf der ersten Silbe, Plural: die Moduln). Es handelt sich dabei lexikalisch um ein anderes Wort als „das Modul“ (sächlich, Betonung auf der zweiten Silbe, Plural: die Module).

Beispiel.

1. Da jeder Körper insbesondere ein Ring ist, ist jeder Vektorraum insbesondere ein Modul.
2. Ist R ein beliebiger Ring, so ist R ein Modul über sich selbst.

   Allgemeiner: die Menge $R^n := \{(r_1, \ldots, r_n) : r_1, \ldots, r_n \in R\}$ wird mit den Verknüpfungen

   $$\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} + \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} := \begin{pmatrix} r_1 + s_1 \\ \vdots \\ r_n + s_n \end{pmatrix}, \qquad a \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} := \begin{pmatrix} ar_1 \\ \vdots \\ ar_n \end{pmatrix}$$

   zu einem R-Modul. Ebenso die Menge $R^{n\times m} = (R^n)^m$ aller $n \times m$-Matrizen mit Einträgen in R und die Menge $R[X]$ aller Polynome mit Koeffizienten in R.

Noch allgemeiner: Ist S irgendeine Menge, so wird die Menge R^S aller Funktionen $f\colon S \to R$ zu einem Modul, wenn man definiert

$$\begin{aligned} f+g\colon S &\to R, \qquad (f+g)(x) := f(x)+g(x) \\ af\colon S &\to R, \qquad (af)(x) := af(x). \end{aligned}$$

3. Jede abelsche Gruppe $(G,+)$ lässt sich auffassen als $\mathbb{Z}$-Modul, wenn man definiert

$$m \cdot g := \begin{cases} \overbrace{g+g+\cdots+g}^{m \text{ Terme}} & \text{falls } m > 0 \\ 0 & \text{falls } m = 0 \\ \underbrace{(-g)+\cdots+(-g)}_{|m| \text{ Terme}} & \text{falls } m < 0 \end{cases}$$

   Zum Beispiel ist $(\mathbb{Z}_6,+)$ ein $\mathbb{Z}$-Modul.

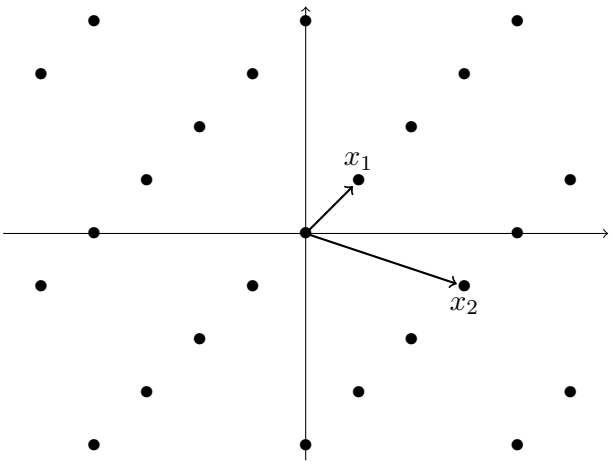
   Auch $(\mathbb{Z}_7 \setminus \{[0]_{\equiv_7}\},\cdot)$ ist eine abelsche Gruppe und damit ein $\mathbb{Z}$-Modul. In diesem Fall muss man aufpassen, dass man sich nicht von der Notation verwirren lässt, weil die Gruppenoperation nicht mit $+$ sondern mit $\cdot$ geschrieben wird. Zum Beispiel gilt hier $3 \cdot [2]_{\equiv_7} = [2]^3_{\equiv_7} = [2^3]_{\equiv_7} = [1]_{\equiv_7}$, und nicht etwa $3 \cdot [2]_{\equiv_7} = [6]_{\equiv_7}$.

   $(\mathbb{R},+)$ und $(\mathbb{R}\setminus\{0\},\cdot)$ sind auch abelsche Gruppen und damit $\mathbb{Z}$-Moduln.

4. Sind $x_1, x_2 \in \mathbb{R}^2$ beliebige Vektoren, so ist

$$\mathbb{Z}x_1 + \mathbb{Z}x_2 = \{\, a_1x_1 + a_2x_2 : a_1, a_2 \in \mathbb{Z} \,\} \subseteq \mathbb{R}^2$$

   zusammen mit der Addition und der Skalarmultiplikation von $\mathbb{R}^2$ ein $\mathbb{Z}$-Modul. Einen solchen Modul nennt man ein *Gitter* (engl. *lattice*).



5. Sind $x_1, x_2, x_3 \in \mathbb{R}$ beliebige Zahlen, so ist

$$\{\, (e_1,e_2,e_3) \in \mathbb{Z}^3 : e_1x_1 + e_2x_2 + e_3x_3 = 0 \,\} \subseteq \mathbb{Z}^3$$

   ein $\mathbb{Z}$-Modul.

Allgemeiner: Ist R ein Ring, M ein R-Modul, und sind $x_1, \dots, x_n \in M$, so ist

$$\operatorname{Syz}(x_1, \dots, x_n) := \{\, (e_1, \dots, e_n) \in R^n : e_1 x_1 + \cdots + e_n x_n = 0 \,\} \subseteq R^n$$

ein R-Modul, der sogenannte *Syzygienmodul* von $x_1, \dots, x_n$.

Variante: Sind $x_1, x_2, x_3 \in \mathbb{R} \setminus \{0\}$ beliebige Zahlen, so ist auch

$$\{\, (e_1, e_2, e_3) \in \mathbb{Z}^3 : x_1^{e_1} x_2^{e_2} x_3^{e_3} = 1 \,\} \subseteq \mathbb{Z}^3$$

ein $\mathbb{Z}$-Modul.

6. Ist $U \subseteq \mathbb{Q}^n$ ein Untervektorraum von $\mathbb{Q}^n$, so ist $U \cap \mathbb{Z}^n$ ein $\mathbb{Z}$-Modul.

7. Ideale:

   - Jeder Ring R ist ein Ideal in sich selbst. 05-05
   - $6\mathbb{Z} = \{\dots, -6, 0, 6, 12, \dots\} \subseteq \mathbb{Z}$ ist ein $\mathbb{Z}$-Modul und zugleich eine Untergruppe von $(\mathbb{Z}, +)$. Es handelt sich also um ein Ideal.
   - $\{\, p \in \mathbb{Q}[X] : X^2 + 3 \mid p \,\} \subseteq \mathbb{Q}[X]$ ist ein Ideal von $\mathbb{Q}[X]$.
   - Nach Satz 80 gilt: Für jede Matrix $A \in \mathbb{K}^{n \times n}$ ist die Menge aller annihilierenden Polynome von A ein Ideal des Rings $\mathbb{K}[X]$.
   - $\{\, 2a + Xb : a, b \in \mathbb{Z}[X] \,\}$ ist ein Ideal von $\mathbb{Z}[X]$.

8. Sei $p \in \mathbb{Z}$ eine Primzahl und $\mathbb{Z}_{(p)} := \{\frac{u}{v} \in \mathbb{Q} : \gcd(v, p) = 1\}$. Dann ist $\mathbb{Z}_{(p)}$ ein Ring (der sogenannte Ring der p-adischen ganzen Zahlen). Die additiven Gruppen von $\mathbb{Q}$, $\mathbb{Q}(\sqrt{2})$, $\mathbb{R}$ und $\mathbb{C}$ lassen sich als $\mathbb{Z}_{(p)}$-Moduln auffassen.

   Allgemeiner: Sind R_1, R_2 zwei Ringe mit $R_1 \subseteq R_2$, wobei die Addition und Multipliation von R_1 mit der von R_2 übereinstimmt, und ist M ein R_2-Modul, dann ist M auch ein R_1-Modul.

   Noch allgemeiner: Sind R_1, R_2 zwei Ringe, M ein R_2-Modul und $h\colon R_1 \to R_2$ ein Ring-Homomorphismus (d.h. eine Abbildung mit $h(a+b) = h(a)+h(b)$ und $h(ab) = h(a)h(b)$ für alle $a, b \in R_1$), so wird M durch die Operation

$$\begin{array}{ccccc} & R_1\text{-Moduloperation} & & R_2\text{-Moduloperation} & \\ & \downarrow & & \downarrow & \\ r & \cdot & m & := & \underbrace{h(r)}_{\in R_2} \cdot m \\ \uparrow & & \uparrow & & \\ \in R_1 & & \in M & & \end{array}$$

   zu einem R_1-Modul. Die Eigenschaften dieses Moduls hängen ganz wesentlich von der Wahl von h ab. Dieselbe Menge M kann also auf ganz unterschiedliche Art Modul ein und desselben Rings sein. (Das ist bei Vektorräumen nicht anders.)

9. Die Menge $M = C^\infty([-1, 1], \mathbb{R})$ der beliebig oft differenzierbaren Funktionen wird mit der Operation

$$(p_0 + p_1 X + \cdots + p_n X^n) \cdot f := p_0 f + p_1 f' + \cdots + p_n f^{(n)}$$

   zu einem $\mathbb{R}[X]$-Modul.

10. Umgekehrt: Ist $M = \mathbb{R}[X]$ und $R = C^\infty([-1,1],\mathbb{R})$ der Ring der beliebig oft differenzierbaren Funktionen (mit punktweiser Addition und Multiplikation als Ringoperationen), dann wird M z.B. mit der Operation

$$f \cdot (p_0 + p_1 X + \cdots + p_n X^n) := p_0 f(0) + p_1 f'(0) X + \cdots + p_n f^{(n)}(0) X^n$$

zu einem R-Modul.

11. $\mathbb{K}^n$ wird mit der üblichen Matrix-Vektor-Multiplikation zu einem $\mathbb{K}^{n\times n}$-Modul.

12. Für jedes fest gewählte $A \in \mathbb{K}^{n\times n}$ wird $\mathbb{K}^n$ zu einem $\mathbb{K}[X]$-Modul, wenn man definiert $p \cdot v := p(A)v$.

13. Sei $a \in \mathbb{K}[X]$ fest. Für $p = p_0 + p_1 X + \cdots + p_d X^d \in \mathbb{K}[X]$ definiert man $p(a) := p_0 + p_1 a + \cdots + p_d a^d \in \mathbb{K}[X]$. Der Ring $\mathbb{K}[X]$ wird zusammen mit der Operation

04-17

$$\cdot\colon \mathbb{K}[X] \times \mathbb{K}[X] \to \mathbb{K}[X], \quad p \cdot q := p(a)q$$

zu einem $\mathbb{K}[X]$-Modul.

Definition 74. Sei R ein Ring und M ein R-Modul.

1. $N \subseteq M$ heißt *Untermodul* (engl. *submodule*) von M, falls gilt $N \neq \emptyset$ und für alle $r_1, r_2 \in R$ und alle $n_1, n_2 \in N$ gilt $r_1 n_1 + r_2 n_2 \in N$.

2. Eine Menge $E \subseteq M$ heißt *Erzeugendensystem* (engl. *generating set*) von M, falls gilt: für jedes $m \in M$ existieren $e_1, \ldots, e_n \in E$ und $r_1, \ldots, r_n \in R$ mit $m = r_1 e_1 + \cdots + r_n e_n$. Ist E ein Erzeugendensystem von M, so schreibt man $M = \langle E \rangle$. Statt $\langle \{e_1, \ldots, e_m\} \rangle$ schreibt man auch $\langle e_1, \ldots, e_m \rangle$.

3. Der Modul M heißt *endlich erzeugt* (engl. *finitely generated*), falls es ein Erzeugendensystem von M mit nur endlich vielen Elementen gibt.

Beispiel.

1. $\mathbb{Z}^n$ ist ein Untermodul von $\mathbb{R}^n$.

2. Ideale sind Untermoduln des Rings R, aufgefasst als Modul über sich selbst.

3. $\mathbb{R}$ ist als $\mathbb{Z}$-Modul nicht endlich erzeugt.

4. Ein Erzeugendensystem von $R[X]$ als R-Modul ist $\{1, X, X^2, \ldots\}$. Ein Erzeugendensystem von $\mathbb{Z}^2$ als $\mathbb{Z}$-Modul ist $\{\binom{0}{1}, \binom{1}{0}\}$.

Dass die Theorie der Moduln wesentlich komplizierter ist als die Theorie der Vektorräume, liegt vor allem daran, dass die Theorie der Ringe wesentlich komplizierter ist als die der Körper. Ringe können sehr unterschiedliche Eigenschaften haben, und natürlich sind auch die zugehörigen Moduln von sehr unterschiedlicher Natur. Uns interessieren in den folgenden Abschnitten vor allem Moduln über Ringen, die noch eine gewisse Ähnlichkeit zu Körpern haben. Moduln über solchen Ringen haben dann eine gewisse Ähnlichkeit zu Vektorräumen.

Definition 75.

1. Ein Ring R heißt *nullteilerfrei,* falls gilt:

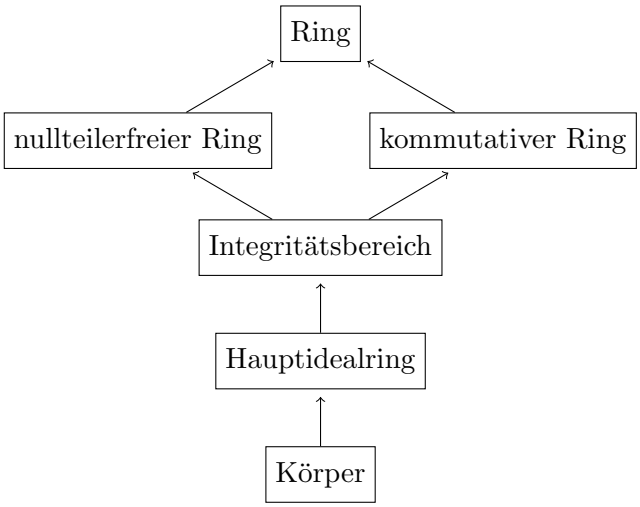
$$\forall\ a, b \in R : ab = 0 \Rightarrow a = 0 \vee b = 0.$$

(Ein Element $a \in R$ heißt *Nullteiler* (engl. *zero divisor*), falls $\exists\ b \in R \setminus \{0\} : ab = 0 \vee ba = 0$.) 04-17

2. Ein kommutativer nullteilerfreier Ring heißt *Integritätsbereich* (engl. *integral domain*).

3. Ein Integritätsbereich R heißt *Hauptidealring* (engl. *principal ideal domain*), falls es für jedes Ideal $A \subseteq R$ ein Element $a \in R$ gibt mit $A = \langle a \rangle$.

05-08



Beispiel.

1. Der Ring $\mathbb{K}^{n\times n}$ ist weder kommutativ noch nullteilerfrei, denn z.B. gilt

$$\underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{\neq 0} \underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{\neq 0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

2. Der Ring $\mathbb{Z}_6$ ist kommutativ, aber nicht nullteilerfrei, denn z.B. gilt

$$\underbrace{[2]_{\equiv_6}}_{\neq 0} \cdot \underbrace{[3]_{\equiv_6}}_{\neq 0} = [6]_{\equiv_6} = [0]_{\equiv_6}$$

3. Der Ring $\mathbb{Z}[X]$ ist ein Integritätsbereich, aber kein Hauptidealring. Zum Beispiel das Ideal $\langle 2, X \rangle$ lässt sich nicht von einem einzigen Element aus $\mathbb{Z}[X]$ erzeugen, denn so ein Erzeuger müsste sowohl ein Teiler von 2 als auch ein Teiler von X sein. Da kommt nur 1 in Frage, aber 1 liegt nicht in $\langle 2, X \rangle$. Man kann zeigen, dass der Polynomring $R[X]$ über einem Ring R genau dann ein Hauptidealring ist, wenn R ein Körper ist. 05-03 04-17

4. Der Ring $\mathbb{K}[X]$ ist ein Hauptidealring, denn in diesem Ring gilt 05-03

$$\langle p_1, \ldots, p_n \rangle = \langle \gcd(p_1, \ldots, p_n) \rangle$$

für jede Wahl $p_1, \ldots, p_n \in \mathbb{K}[X]$. Analog ist auch $\mathbb{Z}$ ein Hauptidealring.

5. Jeder Körper ist ein Hauptidealring. In einem Körper gibt es nämlich nur die beiden Ideale $\{0\}$ und $\langle 1 \rangle = \mathbb{K}$.

38 Konstruktionen

Satz 117. Sei R ein Ring und M ein R-Modul. Seien $N_1, N_2 \subseteq M$ Untermoduln von M. Dann gilt:

1. Der Schnitt $N_1 \cap N_2$ ist ein Untermodul von M.
2. Die Summe $N_1 + N_2 := \{ n_1 + n_2 : n_1 \in N_1, n_2 \in N_2 \} \subseteq M$ ist ein Untermodul von M.

Beweis. Genau wie für Satz 37. ∎

Wenn $N_1, N_2 \subseteq M$ so sind, dass $N_1 \cap N_2 = \{0\}$ ist, dann spricht man von einer *direkten Summe* und kann $N_1 \oplus N_2$ statt $N_1 + N_2$ schreiben.

Satz 118. Sei R ein Ring und M_1, M_2 seien R-Moduln. Dann ist auch $M_1 \times M_2$ zusammen mit der Operation

$$\cdot \colon R \times (M_1 \times M_2) \to M_1 \times M_2, \qquad r \cdot (m_1, m_2) := (r \cdot m_1, r \cdot m_2)$$

ein R-Modul.

Beweis. Genau wie für Satz 45. ∎

Satz 119. Sei R ein Ring, M ein R-Modul und N ein Untermodul von M. Auf M wird durch

$$m_1 \sim m_2 \quad :\Leftrightarrow \quad m_1 - m_2 \in N$$

eine Äquivalenzrelation erklärt. Die Menge $M/N := M/\!\sim$ der Äquivalenzklassen bildet zusammen mit der Operation

$$\cdot \colon R \times M/N \to M/N, \qquad r \cdot [m]_\sim := [rm]_\sim$$

einen R-Modul.

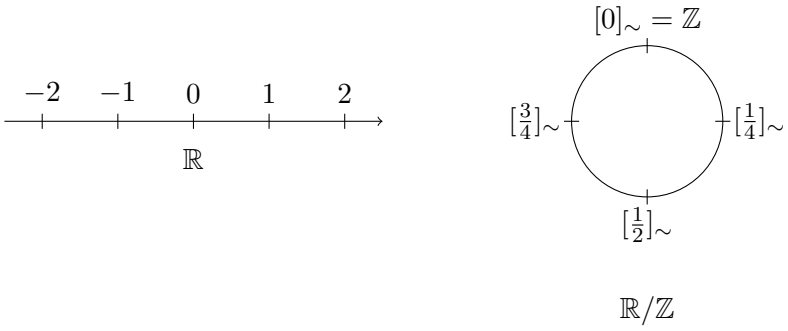
Beweis. Genau wie für Satz 46. ∎

Definition 76. Der Modul M/N aus dem vorigen Satz heißt der *Quotientenmodul* (engl. *quotient module*) von M nach N.

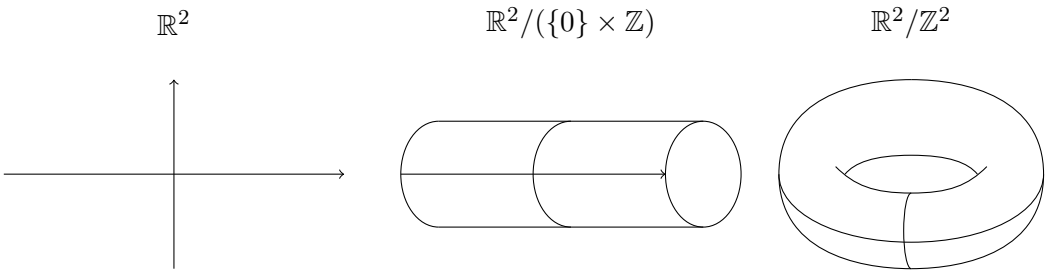
Beispiel.

1. $\mathbb{Z}_6 = \mathbb{Z}/6\mathbb{Z}$.

2. Betrachte die Menge $\mathbb{R}$ der reellen Zahlen als $\mathbb{Z}$-Modul. Die Elemente von $\mathbb{R}$ kann man sich bekanntlich als Punkte auf einer Geraden (der „Zahlengeraden“) veranschaulichen. Zur Veranschaulichung der Elemente von $\mathbb{R}/\mathbb{Z}$ bietet sich statt einer Gerade ein Kreis mit Umfang 1 an. Wenn man sich vorstellt, dass die Zahlengerade auf diesem Kreis aufgewickelt wird, dann kommen genau die zueinander äquivalenten Elemente von $\mathbb{R}$ auf denselben Kreispunkten zu liegen. Zum Beispiel ist $[\frac{3}{4}]_\sim = \frac{3}{4} + \mathbb{Z} = \{\dots, -\frac{1}{4}, \frac{3}{4}, \frac{7}{4}, \dots\}$.



Auch $\mathbb{R}^2$ lässt sich als $\mathbb{Z}$-Modul interpretieren. Die Elemente von $\mathbb{R}^2$ sind Punkte der Ebene. Die Menge $\{0\} \times \mathbb{Z}$ bildet einen Untermodul von $\mathbb{R}^2$, und die Elemente von $\mathbb{R}^2/(\{0\} \times \mathbb{Z})$ kann man sich vorstellen als die Punkte auf einem unendlich langen Zylinder, um den die Ebene $\mathbb{R}^2$ gewickelt wurde. Interessant ist auch der $\mathbb{Z}$-Modul $\mathbb{R}^2/\mathbb{Z}^2$: dieser entsteht, indem man die Ebene zunächst um eine der Achsen zu einem Zylinder aufwickelt, und den Zylinder dann um die andere Achse wickelt, so dass ein Torus-artiges Gebilde entsteht.



Definition 77. Sei R ein Ring und seien M_1, M_2 zwei R-Moduln. Eine Funktion $h\colon M_1 \to M_2$ heißt *(Modul-)Homomorphismus,* falls gilt

$$\forall\, r_1, r_2 \in R\ \forall\, m_1, m_2 \in M_1 : h(r_1 m_1 + r_2 m_2) = r_1 h(m_1) + r_2 h(m_2).$$

Ist $h\colon M_1 \to M_2$ ein Homomorphismus, so heißt

$$\ker h := \{\, m \in M_1 : h(m) = 0 \,\} \subseteq M_1$$

der *Kern* von h und

$$\operatorname{im} h := \{\, h(m) : m \in M_1 \,\} \subseteq M_2$$

das *Bild* von h.

Ein bijektiver Homomorphismus heißt *Isomorphismus,* und M_1, M_2 heißen (zueinander) *isomorph,* geschrieben $M_1 \cong M_2$, falls es einen Isomorphismus $h\colon M_1 \to M_2$ gibt.

Satz 120. Sei R ein Ring.

1. Die Verkettung zweier Homomorphismen ist ein Homomorphismus. Die Umkehrfunktion eines Isomorphismus ist ein Isomorphismus.

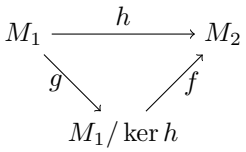
2. Für je drei R-Moduln M_1, M_2, M_3 mit gilt $M_1 \cong M_1$, $M_1 \cong M_2 \Rightarrow M_2 \cong M_1$, $M_1 \cong M_2, M_2 \cong M_3 \Rightarrow M_1 \cong M_3$.

3. Sind M_1, M_2 zwei R-Moduln und ist $h\colon M_1 \to M_2$ ein Homomorphismus, so ist $\ker h$ ein Untermodul von M_1 und $\operatorname{im} h$ ein Untermodul von M_2.

Beweis. Genau wie für Satz 48. ∎

Satz 121. (Homomorphiesatz für Moduln) Sei R ein Ring, M_1, M_2 seien R-Moduln, $h\colon M_1 \to M_2$ sei ein Homomorphismus. Dann gibt es einen surjektiven Homomorphismus $g\colon M_1 \to M_1/\ker h$ und einen injektiven Homomorphismus $f\colon M_1/\ker h \to M_2$, so dass $h = f \circ g$. Wenn h surjektiv ist, dann ist auch f surjektiv. Insbesondere gilt immer $M_1/\ker h \cong \operatorname{im} h$.

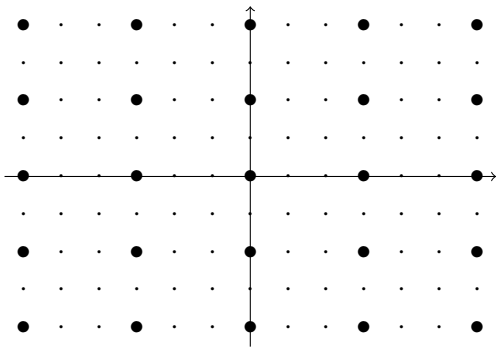
Beweis. Genau wie für Satz 54. ∎



Beispiel. Sei $R = \mathbb{Z}$, $M_1 = \mathbb{Z}^2$, $M_2 = \mathbb{Z}_6$ und $h\colon M_1 \to M_2$, $h(x,y) = [2x+3y]_{\equiv_6}$. Es ist klar, dass h ein $\mathbb{Z}$-Modulhomomorphismus ist. Wegen

$$\begin{aligned} h(0,0) &= [0]_{\equiv_6} & h(0,1) &= [3]_{\equiv_6} \\ h(-1,1) &= [1]_{\equiv_6} & h(2,0) &= [4]_{\equiv_6} \\ h(1,0) &= [2]_{\equiv_6} & h(1,1) &= [5]_{\equiv_6} \end{aligned}$$

ist $\operatorname{im} h = \mathbb{Z}_6$ (d.h. h ist surjektiv). Was ist der Kern? Wegen $h(3,0) = h(0,2) = [0]_{\equiv_6}$ gehören zumindest $(3,0)$ und $(0,2)$ zum Kern, und damit auch alle $\mathbb{Z}$-Linearkombinationen dieser beiden Vektoren:



Weitere Vektoren enthält der Kern nicht. Wenn nämlich $(x, y) \in \mathbb{Z}^2$ im Kern ist, dann auch $(x, y) + a(3, 0) + b(0, 2)$ für jede Wahl von $a, b \in \mathbb{Z}$. Für jedes gegebene $(x, y) \in \mathbb{Z}^2$ lassen sich $a, b \in \mathbb{Z}$ finden, so dass $(x, y) + a\,(3, 0) + b\,(0, 2) \in \{(0, 0), (1, 0), (2, 0), (0, 1), (1, 1), (2, 1)\}$ ist, und da von diesen Punkten $(0, 0)$ der einzige ist, an dem h den Wert 0 annimmt, kann $(x, y) \in \mathbb{Z}^2$ nur dann im Kern liegen, wenn es eine $\mathbb{Z}$-Linearkombination von $(3, 0)$ und $(0, 2)$ ist.

Es gilt also $\ker h = \langle \binom{3}{0}, \binom{0}{2} \rangle = \mathbb{Z}\binom{3}{0} + \mathbb{Z}\binom{0}{2}$, und aus dem Satz folgt, dass $\mathbb{Z}^2/\langle \binom{3}{0}, \binom{0}{2} \rangle$ und $\mathbb{Z}_6$ als $\mathbb{Z}$-Moduln isomorph sind.

Definition 78. Sei R ein Ring, M_1, M_2 seien zwei R-Moduln. Dann wird die Menge aller Homomorphismen $h\colon M_1 \to M_2$ mit $\mathrm{Hom}(M_1, M_2)$ bezeichnet.

Genau wie im Fall der Vektorräume wird $\mathrm{Hom}(M_1, M_2)$ zusammen mit den Operationen

$$h_1 + h_2 := x \mapsto h_1(x) + h_2(x)$$
$$r \cdot h := x \mapsto rh(x)$$

zu einem R-Modul.

Damit sind alle Konstruktionen, die wir für Vektorräume behandelt haben, auf den Fall von Moduln übertragen, mit Ausnahme des Tensorprodukts. Für dessen Definition haben wir im Abschnitt 15 verwendet, dass jeder Vektorraum eine Basis hat. Das geht bei Moduln nicht so einfach, denn wie wir bald sehen werden, hat nicht jeder Modul eine Basis.

Die Definition des Tensorprodukts $U \otimes V$ zweier $\mathbb{K}$-Vektorräume U, V war so gemacht, dass die Elemente von $U \otimes V$ Linearkombinationen von Paaren (u, v) sind, wobei u ein Basiselement von U und v ein Basiselement von V ist. Für die Paare (u, v) schreibt man $u \otimes v$. Die Notation $u \otimes v$ wurde dann erweitert auf beliebige Elemente $u \in U, v \in V$, wobei sich herausstellte, dass folgende Rechenregeln gelten: $(u_1 + u_2) \otimes v = (u_1 \otimes v) + (u_2 \otimes v)$, $u \otimes (v_1 + v_2) = (u \otimes v_1) + (u \otimes v_2)$ und $\alpha(u \otimes v) = (\alpha u) \otimes v = u \otimes (\alpha v)$ für alle $u, u_1, u_2 \in U$, $v, v_1, v_2 \in V$ und $\alpha \in \mathbb{K}$.

Für das Tensorprodukt von Moduln geht man von Linearkombinationen beliebiger Paare $(u, v) \in U \times V$ aus. Das ergibt zunächst einen viel zu großen Modul. Durch Quotientenbildung unterwirft man die formalen Linearkombinationen in diesem Modul den Rechenregeln, die für Tensoren erfüllt sein sollen. Das ist eine sehr typische Anwendung von Quotientenbildung: Man beginnt mit einem großen Raum aller „Ausdrücke“ und dividiert dann die „Relationen“ heraus, die zwischen den Ausdrücken gelten sollen.

Definition 79. Sei R ein kommutativer Ring.

1. Sei S eine Menge und $\mathbf{F}_R(S)$ die Menge aller Funktionen $f\colon S \to R$ mit $|\{x \in S : f(x) \neq 0\}| < \infty$. Dann heißt $\mathbf{F}_R(S)$ (zusammen mit den naheliegenden Operationen) der *freie R-Modul* über S.

Für $f \in \mathbf{F}_R(S)$ mit $f(s_1) = r_1, \ldots, f(s_n) = r_n$ und $f(s) = 0$ für alle $s \in S \setminus \{s_1, \ldots, s_n\}$ verwendet man die Notation

$$f = r_1 s_1 + \cdots + r_n s_n.$$

05-06 2. Seien M_1, M_2 zwei R-Moduln und

$$\begin{aligned} E = &\{\, \alpha(u,v) - (\alpha u, v) : \alpha \in R, u \in M_1, v \in M_2 \,\} \\ &\cup \{\, \alpha(u,v) - (u, \alpha v) : \alpha \in R, u \in M_1, v \in M_2 \,\} \\ &\cup \{\, (u_1, v) + (u_2, v) - (u_1 + u_2, v) : u_1, u_2 \in M_1, v \in M_2 \,\} \\ &\cup \{\, (u, v_1) + (u, v_2) - (u, v_1 + v_2) : u \in M_1, v_1, v_2 \in M_2 \,\} \subseteq \mathbf{F}_R(M_1 \times M_2). \end{aligned}$$

Dann heißt

$$M_1 \otimes_R M_2 := \mathbf{F}_R(M_1 \times M_2)/\langle E \rangle$$

das *Tensorprodukt* von M_1 und M_2.

Statt $[(m_1, m_2)]_\sim$ schreibt man $m_1 \otimes m_2$.

Beispiel. Es gilt $R[X] \otimes_R R[Y] \cong R[X][Y]$.
Jedes Element von $\mathbf{F}_R(R[X] \times R[X])$ hat zunächst die Form

$$r_1 \cdot (p_1, q_1) + \cdots + r_n \cdot (p_n, q_n)$$

für gewisse $r_1, \ldots, r_n \in R$, $p_1, \ldots, p_n \in R[X]$, und $q_1, \ldots, q_n \in R[Y]$. Dabei hat jedes p_i die Form

$$p_i = p_{i,0} + p_{i,1} X + \cdots + p_{i,d} X^d$$

für gewisse $p_{i,j} \in R$ und jedes q_i hat die Form

$$q_i = q_{i,0} + q_{i,1} Y + \cdots + q_{i,d} Y^d$$

für gewisse $q_{i,j} \in R$.
Modulo $\langle E \rangle$ ist jedes Paar $(p, q) \in R[X] \times R[Y]$ äquivalent zu einer Linearkombination

$$\begin{aligned} &s_{0,0} \cdot (1,1) + s_{0,1} \cdot (1, Y) + \cdots + s_{0,d} \cdot (1, Y^d) \\ &+ s_{1,0} \cdot (X, 1) + s_{1,1} \cdot (X, Y) + \cdots + s_{1,d} \cdot (X, Y^d) \\ &+ \cdots\cdots \\ &+ s_{d,0} \cdot (X^d, 1) + s_{d,1} \cdot (X^d, Y) + \cdots + s_{d,d} \cdot (X^d, Y^d) \end{aligned}$$

für gewisse $s_{0,0}, \ldots, s_{d,d} \in R$, z.B.

$$\begin{aligned} &(3 + 4X - X^2, 1 - Y) \\ &\sim (3 + 4X - X^2, 1) + (3 + 4X - X^2, -Y) \\ &\sim (3 + 4X - X^2, 1) - (3 + 4X - X^2, Y) \\ &\sim (3,1) + (4X, 1) + (-X^2, 1) - (3, Y) - (4X, Y) - (-X^2, Y) \\ &\sim 3(1,1) + 4(X,1) - (X^2, 1) - 3(1, Y) - 4(X, Y) + (X^2, Y). \end{aligned}$$

Damit ist jede R-Linearkombination von Paaren $(p, q) \in R[X] \times R[Y]$ modulo $\langle E \rangle$ äquivalent zu einer R-Linearkombination von Paaren (X^i, Y^j) mit $i, j \in \mathbb{N}$. Die Elemente von

$$R[X] \otimes_R R[Y] = \mathbf{F}_R(R[X] \times R[Y])/\langle E \rangle$$

lassen sich deshalb interpretieren als Polynome in zwei Variablen X, Y mit Koeffizienten in R.

39 Freiheit und Torsion

Definition 80. Sei R ein Ring und M ein R-Modul.

1. Eine Menge $E \subseteq M$ heißt *linear unabhängig,* falls für jede Wahl von paarweise verschiedenen Elementen $e_1, \ldots, e_n \in E$ gilt: 04-29

$$\forall\, r_1, \ldots, r_n \in R : r_1 e_1 + \cdots + r_n e_n = 0 \Rightarrow r_1 = \cdots = r_n = 0.$$

2. Ein linear unabhängiges Erzeugendensystem von M heißt *Basis* von M.

3. M heißt *frei* (engl. *free*), falls M eine Basis hat.

Beispiel.

1. Jeder $\mathbb{K}$-Vektorraum ist wegen Satz 43 ein freier $\mathbb{K}$-Modul. 04-17, 04-29

2. $\{\binom{1}{0}, \binom{0}{1}\}$ ist eine Basis des $\mathbb{Z}$-Moduls $\mathbb{Z}^2$.

3. Der $\mathbb{Z}$-Modul $\mathbb{Z}_6 = \mathbb{Z}/6\mathbb{Z}$ hat keine Basis, denn für jedes $e \in \mathbb{Z}_6$ gilt $6 \cdot e = [0]_{\equiv_6}$, und als Element von $\mathbb{Z}$ ist $r = 6$ von Null verschieden. Die einzige linear unabhängige Teilmenge von $\mathbb{Z}_6$ ist deshalb die leere Menge, und diese ist offensichtlich kein Erzeugendensystem von $\mathbb{Z}_6$.

   Erzeugendensysteme von $\mathbb{Z}_6$ sind $\{[1]_{\equiv_6}\}$ und $\{[5]_{\equiv_6}\}$. Die Mengen $\{[2]_{\equiv_6}\}$, $\{[3]_{\equiv_6}\}$, $\{[4]_{\equiv_6}\}$ sind keine Erzeugendensysteme von $\mathbb{Z}_6$, sondern erzeugen echte Untermoduln. Aber z. B. $\{[2]_{\equiv_6}, [3]_{\equiv_6}\}$ ist auch ein Erzeugendensystem von $\mathbb{Z}_6$. Bei Moduln kann es also sein, dass ein Erzeugendensystem minimal ist in dem Sinn, dass keine echte Teilmenge ebenfalls ein Erzeugendensystem ist (wie z.B. $\{[2]_{\equiv_6}, [3]_{\equiv_6}\}$), es aber trotzdem auch ein Erzeugendensystem mit weniger Elementen gibt (wie z.B. $\{[1]_{\equiv_6}\}$). Bei Vektorräumen geht das nicht.

4. Das Ideal $I = \langle 2, X\rangle \subseteq \mathbb{Z}[X]$ wird von $E = \{2, X\}$ erzeugt. Es handelt sich dabei aber nicht um eine Basis, weil $X \cdot 2 - 2 \cdot X = 0$, d.h. 2 und X sind linear abhängig über $\mathbb{Z}[X]$. Die Teilmengen $\{2\}$ und $\{X\}$ von E sind zwar linear unabhängig, bilden aber keine Basen von I, weil $2 \notin \langle X\rangle$ und $X \notin \langle 2\rangle$. Tatsächlich hat I gar keine Basis.

5. $\mathbb{Q}$ ist als $\mathbb{Z}$-Modul nicht frei. Für jede Wahl von $\frac{u}{v}, \frac{p}{q} \in \mathbb{Q}$ gilt $(vp)\frac{u}{v} - (uq)\frac{p}{q} = 0$, d.h. je zwei Elemente von $\mathbb{Q}$ sind über $\mathbb{Z}$ linear abhängig. Linear unabhängige Teilmengen von $\mathbb{Q}$ können deshalb höchstens ein Element enthalten. Es gibt aber keine Erzeugendensysteme von $\mathbb{Q}$ mit nur einem Element, denn für jede Wahl von $e \in \mathbb{Q} \setminus \{0\}$ gilt $\frac{1}{2}e \notin \langle e\rangle = \mathbb{Z}e$, und $\{0\}$ ist natürlich auch kein Erzeugendensystem. 05-08

6. Ein Modul M ist genau dann frei, wenn es eine Menge S gibt, so dass $M \cong \mathbf{F}_R(S)$ gilt.

Satz 122. Sei R ein Hauptidealring und M ein endlich erzeugter freier Modul. Dann ist jeder Untermodul N von M auch endlich erzeugt und frei.

Beweis. Sei $E = \{e_1, \ldots, e_n\}$ eine Basis von M. Wir zeigen durch Induktion nach i, dass $N_i := N \cap \langle e_1, \ldots, e_i \rangle$ für jedes $i = 0, \ldots, n$ frei und endlich erzeugt ist.

Für $i = 0$ ist $\{e_1, \ldots, e_i\} = \emptyset$ und $\langle e_1, \ldots, e_i \rangle = \{0\}$ und $N_0 = \{0\}$, was sicher frei und endlich erzeugt ist.

Induktionsschritt $i - 1 \to i$. Jedes $m \in N_i = N \cap \langle e_1, \ldots, e_i \rangle$ lässt sich schreiben als

$$m = r_1 e_1 + \cdots + r_i e_i$$

für gewisse $r_1, \ldots, r_i \in R$. Die Menge aller $r_i \in R$, die in diesen Darstellungen auftreten
05-08 können, bildet ein Ideal von R, und nach Annahme über R ist dieses Ideal von einem Element erzeugt. Sei $r \in R$ so ein Erzeuger.

Falls $r = 0$ ist, ist $N_i = N_{i-1}$ und das ist nach Induktionsvoraussetzung frei und endlich erzeugt. Wenn $r \neq 0$ ist, wähle ein $m \in N_i$, das sich schreiben lässt als

$$m = r_1 e_1 + \cdots + r_{i-1} e_{i-1} + r e_i$$

für gewisse $r_1, \ldots, r_{i-1} \in R$. Nach Induktionsvoraussetzung hat N_{i-1} eine endliche Basis $\{b_1, \ldots, b_k\}$. Wir zeigen, dass $B := \{b_1, \ldots, b_k, m\}$ eine Basis von N_i ist.

B ist linear unabhängig: Seien $s_1, \ldots, s_k, s \in R$ so, dass $s_1 b_1 + \cdots + s_k b_k + sm = 0$ ist. Wenn $s \neq 0$ ist, dann ist sm eine Linearkombination von $e_1, \ldots, e_i$, in der e_i explizit vorkommt, weil $r \neq 0$ ist und R als Hauptidealring keine Nullteiler hat. Da die $b_1, \ldots, b_k$ nur Linearkombinationen von $e_1, \ldots, e_{i-1}$ sind, wäre auch $s_1 b_1 + \cdots + s_k b_k + sm$ eine Linearkombination von $e_1, \ldots, e_i$, in der e_i explizit vorkommt. Eine solche kann es nicht geben, weil $\{e_1, \ldots, e_n\}$ linear unabhängig ist. Es gilt also auf jeden Fall $s = 0$. Daraus folgt aber sofort auch $s_1 = \cdots = s_k = 0$, weil $\{b_1, \ldots, b_k\}$ als Basis von N_{i-1} linear unabhängig ist.

B ist ein Erzeugendensystem, weil nach Wahl von r jedes $\tilde{m} \in N_i$ von der Form $\tilde{m} = \tilde{r}m + n$ für ein $\tilde{r} \in R$ und ein $n \in N_{i-1}$ ist. ∎

Beispiel. $M = \mathbb{Z}^3$ ist ein endlich erzeugter freier Modul. $N = \langle \begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}4\\5\\6\end{pmatrix}, \begin{pmatrix}7\\8\\9\end{pmatrix} \rangle$ ist ein Untermodul. Das angegebene Erzeugendensystem ist nicht linear unabhängig, aber nach dem Satz muss es auch ein linear unabhängiges Erzeugendensystem von N geben. Ein solches findet man im allgemeinen **nicht** dadurch, dass man eine Basis des $\mathbb{Q}$-Untervektorraums $\langle \begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}4\\5\\6\end{pmatrix}, \begin{pmatrix}7\\8\\9\end{pmatrix} \rangle \subseteq \mathbb{Q}^3$ berechnet, deren Einträge in $\mathbb{Z}$ liegen. Dieser Vektorraum enthält nämlich zum Beispiel das Element

$$\begin{pmatrix}0\\1\\2\end{pmatrix} = \frac{4}{3}\begin{pmatrix}1\\2\\3\end{pmatrix} - \frac{1}{3}\begin{pmatrix}4\\5\\6\end{pmatrix},$$

das in $\mathbb{Z}^3$ liegt, aber nicht in N. (Das sieht man z.B. daran, dass N nur Vektoren enthält, deren dritte Koordinate durch drei teilbar ist, weil das schon für alle drei Elemente des Erzeugendensystems so ist.) Eine Basis von N ist $\{\begin{pmatrix}2\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\1\\3\end{pmatrix}\}$. Wie man eine solche Basis findet, werden wir im folgenden Abschnitt sehen.

Definition 81. Sei R ein Integritätsbereich und M ein R-Modul.

1. Ein Element $m \in M$ heißt *Torsionselement* (engl. *torsion element*), falls es ein $r \in R \setminus \{0\}$ gibt mit $rm = 0$.
2. Die Menge M_{tor} aller Torsionselemente von M heißt der *Torsionsuntermodul* (engl. *torsion submodule*) von M.
3. M heißt *Torsionsmodul* (engl. *torsion module*), falls $M = M_{\text{tor}}$ ist, und *torsionsfrei* (engl. *torsion-free*), falls $M_{\text{tor}} = \{0\}$ ist.

Beispiel.

1. $\mathbb{Z}_6$ ist ein Torsionsmodul, denn für jedes $x \in \mathbb{Z}_6$ gilt $6 \cdot x = 0$.
2. $\mathbb{Z}^6$ ist torsionsfrei. Allgemeiner: jeder freie Modul ist torsionsfrei. Insbesondere ist jeder $\mathbb{K}$-Vektorraum, aufgefasst als $\mathbb{K}$-Modul, torsionsfrei, d.h. für jeden $\mathbb{K}$-Vektorraum V gilt $V_{\text{tor}} = \{0\}$. 05-08
3. $M = \mathbb{K}^{n\times n}$ ist als $\mathbb{K}[X]$-Modul im Sinn von Definition 57 ein Torsionsmodul, weil nach Satz 80 für jedes $A \in M$ ein $p \in \mathbb{K}[X] \setminus \{0\}$ existiert mit $p(A) = 0$.
4. $\mathbb{Q}$ ist als $\mathbb{Z}$-Modul torsionsfrei, aber nicht frei. $\mathbb{Q}/\mathbb{Z}$ ist ein Torsionsmodul, weil für jedes $\frac{p}{q} \in \mathbb{Q}$ gilt $q\frac{p}{q} = p \in \mathbb{Z} = [0]_\sim$, also $q \cdot [\frac{p}{q}]_\sim = [0]_\sim$.
   $\mathbb{R}/\mathbb{Z}$ ist weder frei noch torsionsfrei. Es gilt $(\mathbb{R}/\mathbb{Z})_{\text{tor}} = \mathbb{Q}/\mathbb{Z}$.

Satz 123. Sei R ein Integritätsbereich und M ein R-Modul. Dann ist M_{tor} ein Untermodul von M.

Beweis. Auf jeden Fall gilt $1 \cdot 0 = 0$, also $0 \in M_{\text{tor}}$, also $M_{\text{tor}} \neq \emptyset$.

Seien nun $r_1, r_2 \in R$ und $m_1, m_2 \in M_{\text{tor}}$. Zu zeigen: $r_1m_1 + r_2m_2 \in M_{\text{tor}}$, d.h. es gibt ein $s \in R \setminus \{0\}$ mit $s(r_1m_1 + r_2m_2) = 0$. 05-12

Wegen $m_1, m_2 \in M_{\text{tor}}$ gibt es $s_1, s_2 \in R \setminus \{0\}$ mit $s_1m_1 = s_2m_2 = 0$. Wähle $s = s_1s_2$. Dann gilt $s \neq 0$, weil R ein Integritätsbereich ist, und 05-12

$$\begin{aligned} s(r_1m_1 + r_2m_2) &= s_1s_2(r_1m_1 + r_2m_2) \\ &= s_1s_2r_1m_1 + s_1s_2r_2m_2 \\ &= s_2r_1\underbrace{s_1m_1}_{=0} + s_1r_2\underbrace{s_2m_2}_{=0} = 0, \end{aligned}$$

wie gewünscht. ∎

Satz 124. Sei R ein Integritätsbereich und M ein R-Modul. Dann ist M/M_{tor} torsionsfrei.

Beweis. Sei $m \in M$ beliebig. Zu zeigen: Wenn $r \in R \setminus \{0\}$ so ist, dass $r \cdot [m]_\sim = [0]_\sim$ ist, dann ist $[m]_\sim = [0]_\sim$, also $m \in M_{\text{tor}}$.

Sei also $r \in R \setminus \{0\}$ so, dass $r \cdot [m]_\sim = [0]_\sim$. Dann gilt $rm \in M_{\text{tor}}$, dann existiert $s \in R \setminus \{0\}$ mit $s(rm) = 0$, dann $(sr)m = 0$. Weil R Integritätsbereich ist, ist $sr \neq 0$. Also gilt $m \in M_{\text{tor}}$.
∎

Satz 125. Sei R ein Integritätsbereich und M ein endlich erzeugter R-Modul. Dann gibt es einen freien Untermodul M_0 von M, so dass M/M_0 ein Torsionsmodul ist.

Beweis. Sei E ein endliches Erzeugendensystem von M und $E_0 \subseteq E$ eine linear unabhängige Teilmenge von E, die maximal ist in dem Sinn, dass $E_0 \cup \{e\}$ für jedes $e \in E \setminus E_0$ linear abhängig ist. Dann ist $M_0 := \langle E_0 \rangle$ ein freier Untermodul von M. Wir zeigen, dass M/M_0 ein Torsionsmodul ist.

Nach Wahl von E_0 ist $E_0 \cup \{e\}$ für jedes $e \in E \setminus E_0$ linear abhängig, etwa

$$ae + r_1 e_1 + \cdots + r_m e_m = 0$$

für gewisse $a, r_1, \ldots, r_m \in R$ und $e_1, \ldots, e_m \in E_0$. Da E_0 linear unabhängig ist, muss $a \neq 0$ sein. Andererseits ist $ae \in \langle E_0 \rangle = M_0$, d.h. $a \cdot [e]_\sim = [0]_\sim$ in M/M_0, d.h. e ist ein Torsionselement.

Da $\{[e]_\sim : e \in E \setminus E_0\}$ ein Erzeugendensystem von M/M_0 ist und Linearkombinationen von Torsionselementen wieder Torsionselemente sind (Satz 123), folgt die Behauptung. ∎

Satz 126. Sei R ein Hauptidealring und M ein R-Modul. Wenn M endlich erzeugt und torsionsfrei ist, dann ist M auch frei.

05-12 **Beweis.** Sei E ein endliches Erzeugendensystem von M und sei $E_0 = \{e_1, \ldots, e_n\} \subseteq E$ eine linear unabhängige Teilmenge, die maximal ist in dem Sinn, dass $E_0 \cup \{e\}$ für jedes $e \in E \setminus E_0$ linear abhängig ist. Schreibe $E \setminus E_0 = \{u_1, \ldots, u_k\}$. Dann gibt es für jedes $i = 1, \ldots, k$ ein $a_i \in R$ sowie $r_1, \ldots, r_n \in R$, so dass $a_i, r_1, \ldots, r_n$ nicht alle Null sind und

$$a_i u_i = r_1 e_1 + \cdots + r_n e_n$$

gilt, d.h. es gilt $a_i u_i \in \langle E_0 \rangle$. Wegen der linearen Unabhängigkeit von E_0 muss $a_i \neq 0$ sein.

05-12 Mit $a = a_1 a_2 \cdots a_k$ gilt dann $au_i \in \langle E_0 \rangle$ für alle i. Für den Homomorphismus $h\colon M \to M$, $h(x) = ax$ gilt deshalb $\operatorname{im} h \subseteq \langle E_0 \rangle$.

Weil $a \neq 0$ ist und M torsionsfrei, gilt $\ker h = \{0\}$. Aus Satz 121 folgt deshalb $M \cong M/\ker h \cong \operatorname{im} h \subseteq \langle E_0 \rangle$.

Als Untermodul des freien und endlich erzeugten Moduls $\langle E_0 \rangle$ ist $\operatorname{im} h$ nach Satz 122 ebenfalls frei und endlich erzeugt (hier geht ein, dass R ein Hauptidealring ist). Daraus folgt wegen

05-03 $M \cong \operatorname{im} h$ schließlich, dass M frei ist. ∎

Satz 127. Sei R ein Hauptidealring, M ein endlich erzeugter R-Modul. Dann existiert ein freier Untermodul F von M, so dass $M = F \oplus M_{\text{tor}}$.

Beweis. Betrachte den Homomorphismus $h\colon M \to M/M_{\text{tor}}$ mit $h(m) = [m]_\sim$. Es ist klar, dass $\ker h = M_{\text{tor}}$ ist und dass h surjektiv ist. Nach Satz 124 ist M/M_{tor} torsionsfrei. Weil M endlich erzeugt ist, ist auch M/M_{tor} endlich erzeugt. Weil R ein Hauptidealring ist, folgt nach Satz 126 aus „torsionsfrei“ und „endlich erzeugt“, dass M/M_{tor} frei ist.

Sei $\{b_1, \ldots, b_k\}$ eine Basis von M/M_{tor}. Da h surjektiv ist, gibt es $m_1, \ldots, m_k \in M$ mit $h(m_1) = b_1, \ldots, h(m_k) = b_k$. Die Elemente $m_1, \ldots, m_k \in M$ sind linear unabhängig, weil $b_1, \ldots, b_k \in M/M_{\text{tor}}$ linear unabhängig sind. Also ist $F := \langle m_1, \ldots, m_k \rangle$ ein freier Untermodul von M.

Es gilt $F \cap M_{\text{tor}} = \{0\}$, weil $F \cap M_{\text{tor}}$ als Untermodul des freien Moduls F auch frei ist und freie Moduln keine anderen Torsionselemente als 0 enthalten können.

Es gilt auch $M = F + M_{\text{tor}}$, denn ist $x \in M$ beliebig, so ist $h(x) = r_1 b_1 + \cdots + r_k b_k$ für gewisse $r_1, \ldots, r_k \in R$. Dann gilt

$$x \sim r_1 m_1 + \cdots + r_k m_k,$$

also

$$x = \underbrace{r_1 m_1 + \cdots + r_k m_k}_{\in F} + m$$

für ein gewisses $m \in M_{\text{tor}}$. ∎

40 Lineare Gleichungssysteme über Ringen

Sei R ein Ring und $A \in R^{n \times m}$ eine Matrix mit Einträgen in R. Die Multiplikation von Matrizen über Ringen ist definiert wie bei Körpern (Def. 19). Für gegebenes $A \in R^{n \times m}$ ist die Menge $\ker_R A := \{\, x \in R^m : Ax = 0 \,\}$ ein Untermodul von R^m, und man kann sich fragen, wie man ein Erzeugendensystem für diesen Untermodul konstruiert. Die Frage ist also, wie man ein lineares Gleichungssystem

$$\begin{aligned} a_{1,1} x_1 + \cdots + a_{1,m} x_m &= 0 \\ &\vdots \\ a_{n,1} x_1 + \cdots + a_{n,m} x_m &= 0 \end{aligned}$$

löst, wenn die $a_{i,j}$ gegebene Ringelemente sind und die Variablen x_j für unbekannte Elemente des Rings stehen.

Es ist kein Lösungsverfahren bekannt, das für beliebige Ringe funktioniert. Andererseits wissen wir aus Abschnitt 9, dass es ein allgemeines Lösungsverfahren für den Fall gibt, dass der Ring ein Körper ist. Wir betrachten jetzt den Fall, wo der Ring ein Hauptidealring ist. In diesem Fall ist R^m ein freier Modul und $\ker_R A$ als Untermodul von R^m nach Satz 122 auch frei. Wir können also nach einer Basis von $\ker_R A$ fragen.

Wir können ebenso nach einer Basis des R-Untermoduls von R^m fragen, der von den Zeilen von A erzeugt wird. Klar ist, dass sich dieser Untermodul nicht ändert, wenn man

- eine Zeile von A mit einem Element $c \in R$ multipliziert, das in R ein multiplikatives Inverses besitzt,
- das c-fache einer Zeile zu einer anderen dazuaddiert, für ein beliebiges $c \in R$,
- zwei Zeilen vertauscht.

Ist nämlich $\tilde{A}$ eine Matrix, die auf diese Weise aus A entsteht, dann lässt sich jede R-Linearkombination von Zeilen von $\tilde{A}$ auch als R-Linearkombination von Zeilen von A schreiben, und weil jede der genannten Operationen durch eine Operation des gleichen Typs rückgängig gemacht werden kann, ist auch jede R-Linearkombination von Zeilen von A eine R-Linearkombination von Zeilen von $\tilde{A}$.

Die Operationen entsprechen der Multiplikation der Matrix von links mit Elementarmatrizen, die die Eigenschaft haben, dass ihre Einträge und auch die Einträge in ihren Inversen Ringelemente sind. Sie sind also als Elemente von $R^{n \times n}$ invertierbar. Wenn R ein Teilring eines Körpers $\mathbb{K}$ ist, kann es sein, dass eine Matrix $A \in R^{n \times n}$ als Element von $\mathbb{K}^{n \times n}$ invertierbar ist, aber als Element von $R^{n \times n}$ nicht. Um Missverständnisse zu vermeiden, benutzt man deshalb statt „invertierbar“ ein anderes Wort, wenn man Invertierbarkeit über dem Ring meint.

Definition 82. Sei $\mathbb{K}$ ein Körper, $R \subseteq \mathbb{K}$ ein Unterring von $\mathbb{K}$ (d.h. eine Teilmenge von $\mathbb{K}$, die 0 und 1 enthält und bezüglich der Addition, Subtraktion und Multiplikation von $\mathbb{K}$ abgeschlossen ist). Sei $A \in R^{n\times n}$. 04-17

1. A heißt *invertierbar* (engl. *invertible*), falls $\exists\ B \in \mathbb{K}^{n\times n} : AB = BA = I_n$.
2. A heißt *unimodular,* falls $\exists\ B \in R^{n\times n} : AB = BA = I_n$.

Die Menge aller unimodularen $(n \times n)$-Matrizen über R wird mit $\mathrm{GL}(n, R)$ bezeichnet.

Beispiel.

1. $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathbb{Z}^{2\times 2} \subseteq \mathbb{Q}^{2\times 2}$ ist invertierbar, aber nicht unimodular, denn $A^{-1} = \frac{1}{2}\begin{pmatrix} -4 & 2 \\ 3 & -1 \end{pmatrix}$ liegt nicht in $\mathbb{Z}^{2\times 2}$.
2. $A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in \mathbb{Z}^{3\times 3} \subseteq \mathbb{Q}^{3\times 3}$ ist unimodular, denn $A^{-1} = \begin{pmatrix} 1 & 2 & -2 \\ 0 & -1 & 1 \\ -1 & -2 & -3 \end{pmatrix}$ liegt in $\mathbb{Z}^{3\times 3}$.
3. $A = \begin{pmatrix} 1 & 0 \\ 0 & X \end{pmatrix} \in \mathbb{Q}[X]^{2\times 2} \subseteq \mathbb{Q}(X)^{2\times 2}$ ist invertierbar, aber nicht unimodular, denn $A^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1/X \end{pmatrix}$ liegt nicht in $\mathbb{Q}[X]^{2\times 2}$.
4. $A = \begin{pmatrix} 1+2X+X^2 & 2+3X+2X^2 \\ 3-X^2 & 4+X-2X^2 \end{pmatrix} \in \mathbb{Q}[X]^{2\times 2} \subseteq \mathbb{Q}(X)^{2\times 2}$ ist unimodular, denn $A^{-1} = \frac{1}{2}\begin{pmatrix} -4-X+2X^2 & 2+3X+2X^2 \\ 3-X^2 & -1-2X-X^2 \end{pmatrix}$ liegt in $\mathbb{Q}[X]^{2\times 2}$.

   Trotz der Unterscheidung zwischen „unimodular“ und „invertierbar“ kann es zu Missverständnissen kommen. Zum Beispiel ist die Matrix A zwar unimodular als Element von $\mathbb{Q}[X]^{2\times 2}$, aber nicht als Element von $\mathbb{Z}[X]^{2\times 2}$.
5. Eine Matrix, die nicht invertierbar ist, ist erstrecht nicht unimodular.

Satz 128. (Verallgemeinerung von Satz 30) Sei $\mathbb{K}$ ein Körper und R ein Unterring von $\mathbb{K}$. Sei $A \in R^{n\times n}$. Dann gilt: A ist genau dann unimodular, wenn $\det(A)$ ein multiplikatives Inverses in R hat.

05-12 **Beweis.** „$\Rightarrow$“ Wenn A unimodular ist, existiert ein $B \in R^{n\times n}$ mit $AB = I_n$. Dann gilt $\det(AB) = \det(A)\det(B) = \det(I_n) = 1$. Also ist $\frac{1}{\det(A)} = \det(B) \in R$.

„$\Leftarrow$“ $\det(A)$ hat ein multiplikatives Inverses in R. Damit ist insbesondere $\det(A) \neq 0$ und A ist in $\mathbb{K}^{n\times n}$ invertierbar. Sei $B = A^{-1}$. Die j-te Spalte von B ist die Lösung $x \in \mathbb{K}^n$ des inhomogenen linearen Gleichungssystems $Ax = e_j$, wobei e_j der j-te Einheitsvektor ist. Nach Satz 34 (Cramers Regel) hat die i-te Komponente von x die Form $\frac{\det(\tilde{A})}{\det(A)}$, wobei $\tilde{A}$ die Matrix ist, die aus A entsteht, wenn man die i-te Spalte durch e_j ersetzt. Wegen $\tilde{A} \in R^{n\times n}$ gilt $\det(\tilde{A}) \in R$, und nach Annahme gilt $\frac{1}{\det(A)} \in R$. Da i und j beliebig waren, folgt, dass alle Komponenten von B in R liegen. ∎

Beispiel.

1. Eine Elementarmatrix $\mathrm{diag}(1,\dots,1,c,1,\dots,1)$ ist genau dann unimodular, wenn $c \in R$ in R ein multiplikatives Inverses hat.

2. Die Elementarmatrizen, die das Vertauschen zweier Zeilen oder das Addieren des c-fachen (mit $c \in R$) einer Zeile zu einer anderen ausdrücken, sind immer unimodular.

In Abschnitt 9 haben wir gezeigt, wie man eine beliebige Matrix über einem Körper durch elementare Zeilenumformungen in eine Treppennormalform überführen kann. Wenn man nicht dividieren kann, sind die Bedingungen der Treppennormalform in Def. 23 im allgemeinen nicht zu erfüllen. Wir werden jetzt eine abgeschwächte Variante der Treppenormalform einführen, die auch für Hauptidealringe funktioniert. Der Einfachheit halber betrachten wir ab jetzt nur noch den Ring $R = \mathbb{Z}$.

Definition 83. Eine Matrix $A = ((a_{i,j}))_{i,j=1}^{n,m} \in \mathbb{Z}^{n\times m}$ heißt *Hermite-Normalform* (HNF), falls es $j_1 < j_2 < \cdots < j_k$ gibt mit

1. $a_{1,j_1},\dots,a_{k,j_k}$ sind positiv.

2. $a_{i,1} = \cdots = a_{i,j_i-1} = 0$ für alle $i = 1,\dots,k$, und $a_{i,j} = 0$ für alle $i > k$ und alle $j = 1,\dots,m$.

3. Für alle $i = 1,\dots,k$ gilt $0 \le a_{1,j_k},\dots,a_{k-1,j_k} < a_{k,j_k}$.

Informal gesagt hat eine HNF die Gestalt einer Treppenform im Sinn von Def. 23, wobei auf den Treppenstufen statt 1 beliebige positive Zahlen stehen dürfen und oberhalb jeder Treppenstufe nur nichtnegative Zahlen, die kleiner sind als die Zahl auf der zugehörigen Stufe.

Beispiel. Die Matrix

$$\begin{pmatrix} 2 & * & * & 2 & 0 & * & 1 & * & * \\ 0 & 0 & 0 & 3 & 0 & * & 3 & * & * \\ 0 & 0 & 0 & 0 & 1 & * & 2 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

ist eine Hermite-Normalform. Dabei stehen die Symbole $*$ für beliebige Elemente von $\mathbb{Z}$.

Satz 129. Für jedes $A \in \mathbb{Z}^{n\times m}$ existiert eine unimodulare Matrix $U \in \mathbb{Z}^{n\times n}$ und genau eine Hermite-Normalform $H \in \mathbb{Z}^{n\times m}$ mit $UA = H$.

Statt eines formalen Beweises geben wir nur einen Algorithmus an, der die Hermite-Normalform zu einer beliebigen gegebenen Matrix berechnet. Wie im Abschnitt 9 ist der Algorithmus so notiert, dass er die Einträge der Eingabematrix A im Verlauf der Rechnung durch andere Einträge ersetzt. Die Notation $A[i,j]$ bezeichnet jeweils den Eintrag, der zum aktuellen Zeitpunkt an der Stelle (i,j) von A steht.

Algorithmus 7. Eingabe: $A = ((a_{i,j}))_{i,j=1}^{n,m} \in \mathbb{Z}^{n\times m}$
Ausgabe: Die Hermite-Normalform $H \in \mathbb{Z}^{n\times m}$ von A.

1 $r = 1$
2 für $c = 1, \dots, m$:
3 wenn $A[r,c], \dots, A[n,c]$ nicht alle Null sind:
4 wähle ein $p \in \{r, \dots, n\}$, so dass $|A[p,c]| = \min(\{|A[r,c]|, \dots, |A[n,c]|\} \setminus \{0\})$ ist.
5 wenn $p \neq r$, dann vertausche die p-te und die r-te Zeile von A
05-12 6 wenn $A[r,c] < 0$, dann multipliziere die r-te Zeile von A mit (-1)
7 für $i = r+1, \dots, m$:
05-12 8 bestimme ein $q \in \mathbb{Z}$, für das $|A[i,c] - qA[r,c]|$ minimal ist.
05-12 9 addiere das $(-q)$-fache der r-ten Zeile zur i-ten Zeile von A.
10 wiederhole die Schritte 4–9, bis $A[r+1,c] = \dots = A[n,c] = 0$ gilt.
11 für $i = 1, \dots, r-1$:
05-12 12 bestimme ein $q \in \mathbb{Z}$, für das $0 \leq A[i,c] - qA[r,c] < A[r,c]$ gilt.
05-12 13 addiere das $(-q)$-fache der r-ten Zeile zur i-ten Zeile von A.
14 $r = r + 1$
15 gib A als Ergebnis zurück.

Der Algorithmus geht von links nach rechts durch die Matrix und bearbeitet jede Spalte so lange, bis sie die Bedingungen aus der Definition erfüllt. Dabei werden nur elementare Zeilenoperationen auf die Matrix angewendet, die unimodularen Matrizen entsprechen, und die die schon bearbeiteten
05-12 Spalten links von der aktuellen Spalte nicht verändern. Damit ist klar, dass die Ausgabematrix H des Algorithmus eine Hermite-Normalform ist, und dass es eine unimodulare Matrix U gibt mit $UA = H$.

Es ist weniger klar, dass der Algorithmus für jede Eingabematrix zum Ende kommt. Es wäre denkbar, dass eine Situation entsteht, in der die Schritte 4–9 unendlich oft wiederholt werden, ohne dass die Schleifenabbruchbedingung jemals erfüllt wird. Dass das nicht passieren kann, folgt daraus, dass nach jedem Durchgang durch diese Schritte die Zahlen $A[i,c]$ für $i = r+1, \dots, n$ echt kleiner sind als der Treppenstufeneintrag $A[r,c]$. Dieser wird also bei jeder Wiederholung dieser Schritte durch ein kleineres Element von $\mathbb{N}$ ersetzt. Da jede streng monoton fallende Folge in $\mathbb{N}$ endlich ist, muss auch der Algorithmus nach endlich vielen Schritten fertig werden.

Beispiel.

$$\begin{pmatrix} \boxed{2} & 3 & 5 & 0 & -1 \\ 6 & 2 & 8 & -2 & 1 \\ 7 & -4 & 3 & 3 & 2 \\ 3 & -4 & -1 & 3 & 1 \end{pmatrix} \begin{matrix} -3 \quad -3 \quad -1 \\ + \\ + \\ + \end{matrix} \leftrightarrow \begin{pmatrix} \boxed{2} & 3 & 5 & 0 & -1 \\ 0 & -7 & -7 & -2 & 4 \\ 1 & -13 & -12 & 3 & 5 \\ 1 & -7 & -6 & 3 & 2 \end{pmatrix}$$

$$\leftrightarrow \begin{pmatrix} \boxed{1} & -7 & -6 & 3 & 2 \\ 0 & -7 & -7 & -2 & 4 \\ 1 & -13 & -12 & 3 & 5 \\ 2 & 3 & 5 & 0 & -1 \end{pmatrix} \begin{matrix} -1 \quad -2 \\ \\ + \\ + \end{matrix} \leftrightarrow \begin{pmatrix} 1 & \boxed{-7} & -6 & 3 & 2 \\ 0 & -7 & -7 & -2 & 4 \\ 0 & -6 & -6 & 0 & 3 \\ 0 & 17 & 17 & -6 & -5 \end{pmatrix} \begin{matrix} \\ | \cdot (-1) \\ \\ \\ \end{matrix}$$

$$
\begin{aligned}
&\leftrightarrow \begin{pmatrix} 1 & \boxed{-7} & -6 & 3 & 2 \\ 0 & 6 & 6 & 0 & -3 \\ 0 & -7 & -7 & -2 & 4 \\ 0 & 17 & 17 & -6 & -5 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \end{matrix} \\
&\leftrightarrow \begin{pmatrix} 1 & \boxed{-7} & -6 & 3 & 2 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & 6 & 6 & 0 & -3 \\ 0 & -1 & -1 & -6 & 4 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \end{matrix} \\
&\leftrightarrow \begin{pmatrix} 1 & 0 & 1 & \boxed{17} & -5 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 0 & 4 & -3 \\ 0 & 0 & 0 & 0 & -6 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \end{matrix} \\
&\leftrightarrow \begin{pmatrix} 1 & 0 & 1 & 1 & \boxed{1} \\ 0 & 1 & 1 & 2 & 5 \\ 0 & 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & 0 & 6 \end{pmatrix} .
\end{aligned}
\qquad
\begin{aligned}
&\leftrightarrow \begin{pmatrix} 1 & \boxed{-7} & -6 & 3 & 2 \\ 0 & 6 & 6 & 0 & -3 \\ 0 & -1 & -1 & -2 & 1 \\ 0 & -1 & -1 & -6 & 4 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \end{matrix} \\
&\leftrightarrow \begin{pmatrix} 1 & 0 & 1 & \boxed{17} & -5 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 0 & -12 & 3 \\ 0 & 0 & 0 & -4 & 3 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \end{matrix} \\
&\leftrightarrow \begin{pmatrix} 1 & 0 & 1 & 1 & \boxed{7} \\ 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 0 & 4 & -3 \\ 0 & 0 & 0 & 0 & -6 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \end{matrix}
\end{aligned}
$$

Satz 130. Sei $A \in \mathbb{Z}^{n\times m}$ und sei $H \in \mathbb{Z}^{m\times(n+m)}$ die Hermite-Normalform von $(A^\top | I_m)$. Es seien $B \in \mathbb{Z}^{k\times n}$, $U \in \mathbb{Z}^{k\times m}$ und $N \in \mathbb{Z}^{(m-k)\times m}$ so, dass 05-20 05-16

$$H = \left(\begin{array}{c|c} B & U \\ \hline 0 & N \end{array}\right),$$

wobei k so gewählt ist, dass B keine Nullzeilen hat. Dann gilt:

1. $\binom{B}{0}$ ist die Hermite-Normalform von $A^\top$, und $UA^\top = B$. 05-19
2. Die Zeilen von B bilden eine Basis des $\mathbb{Z}$-Untermoduls von $\mathbb{Z}^n$, der von den Spalten von A erzeugt wird.
3. Die Zeilen von N bilden eine Basis von $\ker_{\mathbb{Z}} A$.

Beweis.

1. Da H eine Hermite-Normalform ist, ist auch B eine Hermite-Normalform. Ist $M \in \mathbb{Z}^{m\times m}$ eine unimodulare Matrix mit $M(A^\top|I_m) = H$, so gilt $MA^\top = \binom{B}{0}$ und $MI_m = M = \binom{U}{N}$, also $\binom{U}{N}A^\top = \binom{B}{0}$, also $UA^\top = B$. 05-16
2. Aus Teil 1 folgt, dass es sich um ein Erzeugendensystem handelt. Dass die von Null verschiedenen Zeilen einer Hermite-Normalform linear unabhängig sind, zeigt man wie in Satz 22.
3. Da H eine Hermite-Normalform ist, ist auch N eine Hermite-Normalform. N kann keine Nullzeilen enthalten, weil die Zeilen von I_m linear unabhängig sind. Es folgt, dass die Zeilen von N linear unabhängig sind. 05-16

Aus der Rechnung im Beweis zu Teil 1 folgt $NA^\top = 0$. Damit sind die Zeilen von N in $\ker_{\mathbb{Z}} A$ enthalten. Sei umgekehrt $x \in \ker_{\mathbb{Z}} A$. Zu zeigen: x ist eine $\mathbb{Z}$-Linearkombination der Zeilen von N.

Sei $M \in \mathbb{Z}^{m\times m}$ wie im Beweis zu Teil 1. Da M unimodular ist, ist auch $M^\top$ unimodular (wegen Satz 28 und Satz 30), und wir können $x = M^\top y$ für ein $y \in \mathbb{Z}^m$ schreiben. Es gilt dann $0 = Ax = AM^\top y$, also $y(MA^\top) = y\binom{B}{0} = 0$. Da die Zeilen von B linear unabhängig sind, muss y die Form $y = (0,\dots,0,*,\dots,*)$ haben. Daraus folgt, dass $x = M^\top y = (U^\top | N^\top)y$ eine Linearkombination der Zeilen von N ist. ∎

Beispiel.

1. Betrachte die Matrix $A = \begin{pmatrix} 2 & 4 & 3 & 2 \\ 3 & 2 & 3 & 0 \end{pmatrix} \in \mathbb{Z}^{2\times 4}$. Die Hermite-Normalform von

$$\left(\begin{array}{cc|cccc} 2 & 3 & 1 & & & \\ 4 & 2 & & 1 & & \\ 3 & 3 & & & 1 & \\ 2 & 0 & & & & 1 \end{array}\right) \quad \text{ist} \quad \left(\begin{array}{cc|cccc} 1 & 0 & 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 2 & -2 & -2 \\ \hline & & 2 & 0 & -2 & 1 \\ & & 0 & 3 & -2 & -3 \end{array}\right).$$

Daraus folgt, dass die Spalten von A den ganzen $\mathbb{Z}^2$ aufspannen, und dass

$$\{\begin{pmatrix} 2 \\ 0 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ -2 \\ -3 \end{pmatrix}\}$$

eine Basis des $\mathbb{Z}$-Moduls $K := \ker_{\mathbb{Z}} A = \{x \in \mathbb{Z}^4 : Ax = 0\} \subseteq \mathbb{Z}^4$ ist.

Mit dem Gauß-Algorithmus findet man

$$A \leftrightarrow \begin{pmatrix} 1 & 0 & 3/4 & -1/2 \\ 0 & 1 & 3/8 & 3/4 \end{pmatrix},$$

und daraus folgt, dass der $\mathbb{Q}$-Vektorraum $U := \{x \in \mathbb{Q}^4 : Ax = 0\} \subseteq \mathbb{Q}$ die Basis

$$\{\begin{pmatrix} 3/4 \\ 3/8 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1/2 \\ 3/4 \\ 0 \\ -1 \end{pmatrix}\}$$

hat. Durch Skalierung der beiden Basiselemente mit 8 bzw. 4 bekommt man die alternative Basis

$$B = \{\begin{pmatrix} 6 \\ 3 \\ -8 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \\ 0 \\ -4 \end{pmatrix}\}.$$

Obwohl die Elemente dieser Basis in $\mathbb{Z}^4$ liegen, handelt es sich **nicht** um eine $\mathbb{Z}$-Modul-Basis von M, denn z.B. der Vektor $\begin{pmatrix} 2 \\ 0 \\ -2 \\ 1 \end{pmatrix} \in K$ lässt sich nicht als $\mathbb{Z}$-Linearkombination der beiden Vektoren in B schreiben.

Umgekehrt gilt aber, dass jede $\mathbb{Z}$-Modulbasis von $\ker_{\mathbb{Z}} A$ auch eine $\mathbb{Q}$-Vektorraumbasis von U ist.

2. Betrachte die $\mathbb{Z}$-Moduln $A = \langle \binom{2}{3}, \binom{1}{-2} \rangle \subseteq \mathbb{Z}^2$ und $B = \langle \binom{1}{1}, \binom{3}{-2} \rangle \subseteq \mathbb{Z}^2$. Gesucht ist eine Basis von $A \cap B$. 05-16

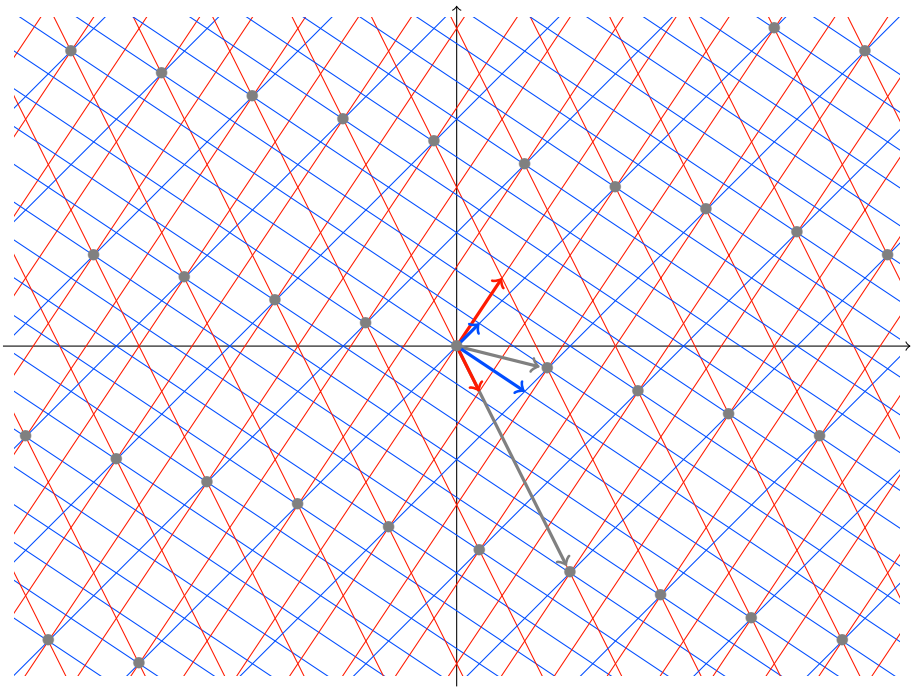
   Das Problem lässt sich genau wie bei Vektorräumen auf die Lösung eines linearen Gleichungssystems zurückführen: Eine $\mathbb{Z}$-Modul-Basis für den Kern von

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

   ist $\{ \begin{pmatrix} 1 \\ 2 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 4 \\ -3 \end{pmatrix} \}$.

   Daraus folgt $A \cap B = \langle 1\binom{2}{3} + 2\binom{1}{-2}, 0\binom{2}{3} + 5\binom{1}{-2} \rangle = \langle \binom{4}{-1}, \binom{5}{-10} \rangle$.

   In der folgenden Abbildung ist das Gitter A rot und das Gitter B blau dargestellt. Die Gitterpunkte sind jeweils die Punkte, wo sich zwei gleichfarbige Geraden kreuzen. Das Gitter $A \cap B$ besteht aus den Punkten, wo sich sowohl zwei rote als auch zwei blaue Geraden kreuzen. Diese Punkte sind als ausgefüllte Kreise markiert.



41 Die Smith-Normalform

Nach Satz 122 ist jeder $\mathbb{Z}$-Untermodul von $\mathbb{Z}^n$ frei und endlich erzeugt. Mit der Hermite-Normalform kann man aus einem gegebenen Erzeugendensystem eine Basis berechnen. Quotientenmoduln $\mathbb{Z}^n/M$

für Untermoduln $M \subseteq \mathbb{Z}^n$ sind im allgemeinen nicht frei. Nach Satz 127 gibt es aber für jeden solchen Modul eine Zerlegung in einen freien Modul und einen Torsionsmodul. Allerdings gibt der Satz keinen Hinweis darauf, wie man diese Zerlegung explizit berechnen kann. Manchmal lässt sich aus dem Beweis eines Satzes ein Berechnungsverfahren extrahieren. Man spricht dann von einem konstruktiven Beweis. Der Beweis von Satz 127 ist nicht konstruktiv. Wie findet man also für einen gegebenen Untermodul $M \subseteq \mathbb{Z}^n$ die Zerlegung von $\mathbb{Z}^n/M$? Die Frage lässt sich einfach beantworten, wenn M von $\mathbb{Z}$-Vielfachen von Einheitsvektoren erzeugt wird.

Beispiel. Für $M = \langle \begin{pmatrix}5\\0\\0\end{pmatrix}, \begin{pmatrix}0\\2\\0\end{pmatrix} \rangle \subseteq \mathbb{Z}^3$ ist

$$\mathbb{Z}^3/M = \langle [\begin{pmatrix}1\\0\\0\end{pmatrix}]_\sim, [\begin{pmatrix}0\\1\\0\end{pmatrix}]_\sim, [\begin{pmatrix}0\\0\\1\end{pmatrix}]_\sim \rangle = \underbrace{\underbrace{\langle [\begin{pmatrix}1\\0\\0\end{pmatrix}]_\sim \rangle}_{\cong \mathbb{Z}_5} + \underbrace{\langle [\begin{pmatrix}0\\1\\0\end{pmatrix}]_\sim \rangle}_{\cong \mathbb{Z}_2}}_{\text{Torsionsanteil}} + \underbrace{\langle [\begin{pmatrix}0\\0\\1\end{pmatrix}]_\sim \rangle}_{\cong \mathbb{Z}}_{\text{freier Anteil}}.$$

Wenn M keine solche Basis hat, dann kann man noch die Freiheit ausnutzen, für $\mathbb{Z}^n$ eine andere Basis als die Standardbasis zu wählen. Es würde nämlich genügen, zum gegebenen Modul $M \subseteq \mathbb{Z}^n$ eine Basis B_1 von $\mathbb{Z}^n$ zu finden, so dass M eine Basis B_2 hat, so dass die Koordinatendarstellungen der Elemente von B_2 bezüglich der Basis B_1 ganzzahlige Vielfache von Einheitsvektoren sind.

Beispiel. $M = \langle \begin{pmatrix}5\\0\\0\end{pmatrix}, \begin{pmatrix}2\\2\\0\end{pmatrix} \rangle \subseteq \mathbb{Z}^3$ hat keine Basis bestehend aus $\mathbb{Z}$-Vielfachen von Ein-

05-19 heitsvektoren. Bezüglich der Basis $(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix})$ von $\mathbb{Z}^3$ haben die Erzeuger von M die Form $\begin{pmatrix}5\\0\\0\end{pmatrix}, \begin{pmatrix}0\\2\\0\end{pmatrix}$. Da ein Basiswechsel in $\mathbb{Z}^3$ einem Isomorphismus $\mathbb{Z}^3 \to \mathbb{Z}^3$ entspricht, folgt $\mathbb{Z}^3/M \cong \mathbb{Z}_5 \times \mathbb{Z}_2 \times \mathbb{Z}$.

Andere Möglichkeit: Bezüglich der Basis $B = (\begin{pmatrix}3\\1\\0\end{pmatrix}, \begin{pmatrix}7\\2\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix})$ von $\mathbb{Z}^3$ haben die Vektoren $\begin{pmatrix}5\\0\\0\end{pmatrix}, \begin{pmatrix}2\\2\\0\end{pmatrix}$ die Koordinatendarstellungen $b_1 = \begin{pmatrix}-10\\5\\0\end{pmatrix}$, $b_2 = \begin{pmatrix}10\\-4\\0\end{pmatrix}$, denn es gilt

$$\begin{pmatrix}5\\0\\0\end{pmatrix} = -10\begin{pmatrix}3\\1\\0\end{pmatrix} + 5\begin{pmatrix}7\\2\\0\end{pmatrix} + 0\begin{pmatrix}0\\0\\1\end{pmatrix}$$
$$\begin{pmatrix}2\\2\\0\end{pmatrix} = 10\begin{pmatrix}3\\1\\0\end{pmatrix} - 4\begin{pmatrix}7\\2\\0\end{pmatrix} + 0\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

Auch

$$b_1 + b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \text{und} \qquad 4b_1 + 5b_2 = \begin{pmatrix} 10 \\ 0 \\ 0 \end{pmatrix}$$

bilden eine Basis von M, wobei die Koordianten der Vektoren auf den rechten Seiten wieder bezüglich B zu verstehen sind. Es gilt also auch $\mathbb{Z}^3/M \cong \mathbb{Z}_1 \times \mathbb{Z}_{10} \times \mathbb{Z}$. (Beachte $\mathbb{Z}_1 = \mathbb{Z}/\mathbb{Z} = \{0\}$.)

Der folgende Satz besagt, dass eine analoge Rechnung für jeden Modul $M \subseteq \mathbb{Z}^n$ möglich ist.

Satz 131. (Smith-Normalform) Für alle $A \in \mathbb{Z}^{n\times m}$ gibt es unimodulare Matrizen $U \in \mathbb{Z}^{n\times n}$ und $V \in \mathbb{Z}^{m\times m}$, so dass

$$UAV = \begin{pmatrix} s_1 & & \\ & s_2 & \\ & & \ddots \end{pmatrix} \in \mathbb{Z}^{n\times m}$$

mit Diagonaleinträgen $s_1, \ldots, s_n \in \mathbb{N}$, für die gilt $s_1 \mid s_2 \mid s_3 \mid \cdots$. Diese $s_1, s_2, \ldots$ sind eindeutig durch A bestimmt.

Die Diagonalmatrix aus dem Satz wird als die Smith-Normalform von A bezeichnet. Sie verallgemeinert eine Beobachtung aus Abschnitt 11. Dort hatten wir gesehen, dass sich jede Matrix $A \in \mathbb{K}^{n\times m}$ durch Anwendung elementarer Zeilen- und Spaltenoperationen auf eine Diagonalform bringen lässt, wobei auf der Diagonalen zunächst eine Reihe von Einsen und dann eine Reihe von Nullen steht.

Auch die Smith-Normalform bekommt man, indem man systematisch Zeilen- und Spaltenoperationen auf die Matrix anwendet. Man eliminiert zuerst die Einträge, die nicht auf der Diagonalen stehen. Das geht so ähnlich wie bei der Berechnung der Hermite-Normalform. Anschließend muss man noch die Teilbarkeitseigenschaft der Diagonalelemente herstellen. Dazu benutzt man die $\mathbb{Z}$-Version von Satz 72, die besagt, dass es zu je zwei ganzen Zahlen s_i, s_j zwei ganze Zahlen a, b gibt, so dass $g := \gcd(s_i, s_j) = as_i + bs_j$ gilt. Wenn s_i, s_j, zwei Diagonaleinträge mit $s_i \nmid s_j$ sind, dann kann man wegen 05-20

$$\begin{pmatrix} s_i & \\ & s_j \end{pmatrix} \overset{\substack{+ \quad b \\ \downarrow \quad \lceil \\ \lfloor \quad \downarrow \\ a \quad +}}{\rightsquigarrow} \begin{pmatrix} s_i & \\ g & s_j \end{pmatrix} \overset{\substack{\leftarrow \quad + \\ \rfloor \quad \leftarrow \\ -s_i/g \quad \leftarrow}}{\rightsquigarrow} \begin{pmatrix} g & s_j \\ 0 & -s_i s_j/g \end{pmatrix} \overset{\substack{-s_j/g \quad + \\ \lceil \quad \downarrow \\ \lfloor \quad \downarrow}}{\rightsquigarrow} \begin{pmatrix} g & 0 \\ 0 & s_i s_j/g \end{pmatrix} \overset{\mid \cdot(-1)}{\rightsquigarrow}$$

die Einträge s_i und s_j durch g und $s_i s_j/g$ ersetzen. Die Operationen sind unimodular, weil s_i/g und s_j/g in $\mathbb{Z}$ liegen.

Da in dieser Rechnung die positive ganze Zahl s_i durch eine echt kleinere positive ganze Zahl g ersetzt wird (wegen $s_i \nmid s_j$ ist $s_i = g$ nicht möglich), muss nach endlich vielen Wiederholungen eine Diagonale entstehen, in der jeder Diagonaleintrag den folgenden teilt.

Algorithmus 8. Eingabe: $A \in \mathbb{Z}^{n\times m}$

Ausgabe: $s_1, s_2, \cdots \in \mathbb{N}$, so dass $\begin{pmatrix} s_1 & & \\ & s_2 & \\ & & \ddots \end{pmatrix} \in \mathbb{Z}^{n\times m}$ die Smith-Normalform von A ist.

1 für $k = 1, \ldots, \min(n, m)$:

2 wenn $A[i, j] = 0$ für alle $i, j \geq k$, dann weiter mit Schritt 14.

3 wähle eine Position (i, j) mit $i, j \geq k$, für die $|A[i, j]|$ minimal und nicht Null ist.

4 falls $i \neq k$, vertausche die i-te Zeile und die k-te Zeile von A.

5 falls $j \neq k$, vertausche die j-te Spalte und die k-te Spalte von A.

6 falls $A[k,k] < 0$, multipliziere die k-te Zeile (oder Spalte) mit (-1).

7 für $i = k+1, \ldots, n$:

8 bestimme ein $q \in \mathbb{Z}$, für das $|A[i,k] - qA[k,k]|$ minimal ist.

9 addiere das $(-q)$-fache der k-ten Zeile zur i-ten Zeile von A.

10 für $j = k+1, \ldots, m$:

11 bestimme ein $q \in \mathbb{Z}$, für das $|A[k,j] - qA[k,k]|$ minimal ist.

12 addiere das $(-q)$-fache der k-ten Spalte zur j-ten Spalte von A.

13 falls $A[i,k] \neq 0$ für ein $i > k$ oder $A[k,j] \neq 0$ für ein $j > k$, weiter mit Schritt 3

14 solange es Paare (i,j) mit $i < j$ und $A[i,i] \nmid A[j,j]$ gibt:

15 wähle so ein Paar und berechne $g = \gcd(A[i,i], A[j,j])$.

16 ersetze $A[j,j]$ durch $A[i,i]A[j,j]/g$ und dann $A[i,i]$ durch g.

Beispiel.

05-19,05-20 1. Es gilt $\mathbb{Z}^2/\langle \binom{3}{5}, \binom{4}{6} \rangle \cong \mathbb{Z}_1 \times \mathbb{Z}_2 \cong \mathbb{Z}_2$, weil

$$\begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{matrix} -1 \\ + \end{matrix} \rightsquigarrow \begin{pmatrix} 3 & 4 \\ 2 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 3 & 1 \\ 2 & 0 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} \begin{matrix} + \\ -1 \end{matrix} \rightsquigarrow \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} .$$

Aus der Smith-Normalform lässt sich ablesen, dass $\mathbb{Z}^n/M$ zu einem Produkt $\mathbb{Z}^k \times \mathbb{Z}_{s_1} \times \cdots \times \mathbb{Z}_{s_m}$ isomorph ist: k ist die Zahl der Nullen auf der Diagonalen und $s_1, \ldots, s_m$ sind die Diagonaleinträge, die ≥ 2 sind. Etwaige 1-Einträge kann man weglassen, weil $\mathbb{Z}_1 = \mathbb{Z}/\mathbb{Z} = \{0\}$ ist und für jeden Modul U gilt $U \times \{0\} \cong U$.

2. Beim Modul $\mathbb{Z}^k \times \mathbb{Z}_{s_1} \times \cdots \times \mathbb{Z}_{s_m}$ ist $\mathbb{Z}^k \times \{0\}^m$ der freie Untermodul F aus Satz 127 und $\{0\} \times \mathbb{Z}_{s_1} \times \cdots \times \mathbb{Z}_{s_m}$ der Torsionsuntermodul. Diese Moduln entsprechen gewissen Untermoduln von $\mathbb{Z}^n/M$. Wenn wir diese identifizieren wollen, brauchen wir neben der Smith-Normalform S auch die unimodularen Matrizen U und V aus Satz 131. Diese Matrizen codieren den Isomorphismus zwischen $\mathbb{Z}^n/M$ und $\mathbb{Z}^k \times \mathbb{Z}_{s_1} \times \cdots \times \mathbb{Z}_{s_m}$.

   Geeignete Matrizen U und V findet man, indem man den Algorithmus statt auf A auf die erweiterte Matrix $\begin{pmatrix} A & I_n \\ I_m & 0 \end{pmatrix}$ anwendet, wobei man Zeilenoperationen nur auf die ersten n Zeilen und Spaltenoperationen nur auf die ersten m Spalten anwendet. Zum

Beispiel so:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 2 & 1 & 4 & 0 & 1 & 0 \\ 4 & 5 & 2 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & & & \\ 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & & \end{array}\right) \overset{-2,\,-4}{\underset{+,\,+}{\rightsquigarrow}} \left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & -3 & 6 & -2 & 1 & 0 \\ 0 & -3 & 6 & -4 & 0 & 1 \\ \hline 1 & 0 & 0 & & & \\ 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & & \end{array}\right) \overset{|\cdot(-1)}{\underset{+}{\rightsquigarrow}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 3 & -6 & 2 & -1 & 0 \\ 0 & 0 & 0 & -2 & -1 & 1 \\ \hline 1 & -2 & 1 & & & \\ 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & & \end{array}\right) \overset{2,\,+}{\rightsquigarrow}$$

$$\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & -2 & -1 & 1 \\ \hline 1 & -2 & -3 & & & \\ 0 & 1 & 2 & & & \\ 0 & 0 & 1 & & & \end{array}\right)$$

Im zweiten Schritt haben wir zugleich Zeilen- und Spaltenoperationen verwendet. Welche dieser beiden Operationen man zuerst durchführt, ist egal, weil die Zeilenoperation eine Multiplikation von Links und die Spaltenoperation eine Multiplikation von rechts ist und Matrixmultiplikation assoziativ ist. Trotzdem sollte man mit solchen Abkürzungen vorsichtig sein, weil man dabei leicht vermeidbare Fehler macht. 06-02

Für

$$M = \langle \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}, \begin{pmatrix} -1 \\ 4 \\ 2 \end{pmatrix} \rangle \subseteq \mathbb{Z}^3$$

ergibt sich aus der Smith-Normalform im Block oben links $\mathbb{Z}^3/M \cong \mathbb{Z} \times \mathbb{Z}_3$. Mit den Basiswechselmatrizen in den beiden anderen Blöcken kann man die Untermoduln von $\mathbb{Z}^3/M$ identifizieren, die zu $\mathbb{Z}$ bzw. $\mathbb{Z}_3$ isomorph sind. Zunächst gilt

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

Die zweite Multiplikation entspricht einem Basiswechsel von M, d.h. es gilt

$$M = \langle \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ -3 \end{pmatrix} \rangle.$$

Die erste Multiplikation besagt, dass $\begin{pmatrix}1\\0\\0\end{pmatrix}$ und $\begin{pmatrix}0\\3\\0\end{pmatrix}$ die Koordinatendarstellungen von $\begin{pmatrix}1\\2\\4\end{pmatrix}$ und $\begin{pmatrix}0\\-3\\-3\end{pmatrix}$ bezüglich der Basis $(\begin{pmatrix}1\\2\\4\end{pmatrix}, \begin{pmatrix}0\\-1\\-1\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix})$ sind.

Für den Quotientenmodul $\mathbb{Z}^3/\langle \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\3\\0\end{pmatrix} \rangle$ ist klar, dass der Untermodul $\langle [\begin{pmatrix}0\\0\\1\end{pmatrix}]_\sim \rangle$ isomorph zu $\mathbb{Z}$ und der Untermodul $\langle [\begin{pmatrix}0\\1\\0\end{pmatrix}]_\sim \rangle$ isomorph zu $\mathbb{Z}_3$ ist. Übersetzt man diese Erkenntnis zurück in die ursprüngliche Basis, so findet man, dass die Zerlegung von $\mathbb{Z}^3/M$ in seinen freien und seinen Torsionsanteil

$$\mathbb{Z}^3/M = \underbrace{\langle [\begin{pmatrix}0\\0\\1\end{pmatrix}]_\sim \rangle}_{\cong \mathbb{Z}} \oplus \underbrace{\langle [\begin{pmatrix}0\\-1\\-1\end{pmatrix}]_\sim \rangle}_{= (\mathbb{Z}^3/M)_{\mathrm{tor}} \cong \mathbb{Z}_3}$$

lautet. In der Tat gilt $3 \cdot [\begin{pmatrix}0\\-1\\-1\end{pmatrix}]_\sim = [\begin{pmatrix}0\\-3\\-3\end{pmatrix}]_\sim = [0]_\sim$, da $\begin{pmatrix}0\\-3\\-3\end{pmatrix} \in M$. Dagegen gibt es kein $a \in \mathbb{Z} \setminus \{0\}$ mit $a \cdot [\begin{pmatrix}0\\0\\1\end{pmatrix}]_\sim = [0]_\sim$.

Satz 132. (Elementarteilersatz) Sei M ein endlich erzeugter $\mathbb{Z}$-Modul. Dann gibt es ein $k \in \mathbb{N}$ und ganze Zahlen $s_1, \ldots, s_m \geq 2$ mit $s_1 \mid \cdots \mid s_m$, so dass $M \cong \mathbb{Z}^k \times \mathbb{Z}_{s_1} \times \cdots \times \mathbb{Z}_{s_m}$. Die Zahlen $k, s_1, \ldots, s_m$ sind eindeutig durch M bestimmt.

Beweis. Sei $E = \{e_1, \ldots, e_n\} \subseteq M$ ein Erzeugendensystem von M. Für den Homomorphismus $h\colon \mathbb{Z}^n \to M$, $h(a_1, \ldots, a_n) := a_1 e_1 + \cdots + a_n e_n$ gilt dann $M \cong \mathbb{Z}^n / \ker h$. Es genügt also, die Aussage für $\mathbb{Z}$-Moduln der Form $\mathbb{Z}^n/U$ mit Untermoduln $U \subseteq \mathbb{Z}^n$ zu zeigen. Für solche $\mathbb{Z}$-Moduln folgt sie aber unmittelbar aus Satz 131. ∎

42 Kurze Vektoren in Gittern

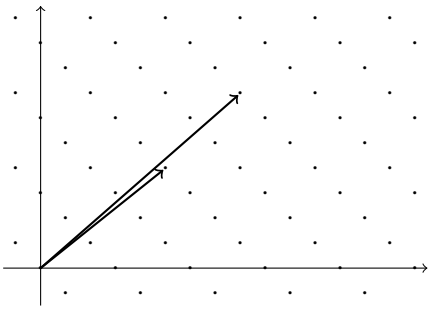
Wenn man wissen möchte, wie engmaschig ein Gitter $L \subseteq \mathbb{Z}^n$ ist, darf man sich nicht an der Länge der Basisvektoren (z.B. bezüglich der Standardnorm) orientieren. Zwar ist klar, dass für jeden Basisvektor b die Vielfachen αb für $\alpha \in \mathbb{Z} \setminus \{0\}$ stets länger sind als b, aber Linearkombinationen mehrerer Basisvektoren können durchaus kürzere Vektoren ergeben.

Beispiel. $L = \langle \binom{5}{4}, \binom{8}{7} \rangle \subseteq \mathbb{Z}^2$ enthält den Vektor

$$\binom{1}{2} = (-3)\binom{5}{4} + 2\binom{8}{7}$$

und es gilt

$$\|\binom{1}{2}\| = \sqrt{5} < \sqrt{41} = \min\{\|\binom{5}{4}\|, \|\binom{8}{7}\|\}$$



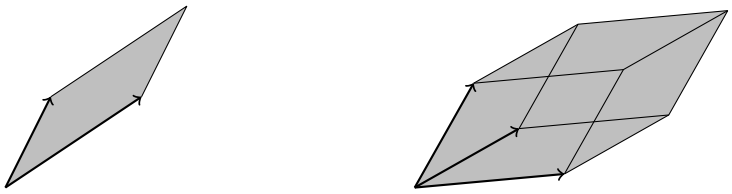
Definition 84. Seien $x_1, \ldots, x_k \in \mathbb{R}^n$ linear unabhängig über $\mathbb{R}$. Dann heißt

$$\Pi(x_1, \ldots, x_k) := \{\, \alpha_1 x_1 + \cdots + \alpha_k x_k : \alpha_1, \ldots, \alpha_k \in [0,1) \,\} \subseteq \mathbb{R}^n$$

der von $x_1, \ldots, x_k$ aufgespannte *Spat* (engl. *parallelepiped*).
Dabei ist $[0,1) = \{\, x \in \mathbb{R} : 0 \leq x < 1 \,\}$ das rechts halb-offene Einheitsintervall.
Ist $P = \Pi(x_1, \ldots, x_k)$ der von $x_1, \ldots, x_k$ aufgespannte Spat, so heißt

$$V(P) := \begin{cases} 0 & \text{falls } k < n \\ |\det(x_1, \ldots, x_n)| & \text{falls } k = n \end{cases}$$

das (n-dimensionale) *Volumen* von P.



Satz 133. Ist $L \subseteq \mathbb{Z}^n$ ein Gitter und sind $\{b_1, \ldots, b_k\}$ und $\{b'_1, \ldots, b'_k\}$ zwei Basen von L, so gilt $V(\Pi(b_1, \ldots, b_k)) = V(\Pi(b'_1, \ldots, b'_k))$.

Beweis. Im Fall $k < n$ gilt $V(\Pi(b_1, \ldots, b_k)) = V(\Pi(b'_1, \ldots, b'_k)) = 0$. Betrachte den Fall $n = k$. Jedes b_i ist eine $\mathbb{Z}$-Linearkombination von $b'_1, \ldots, b'_k$ und umgekehrt. Es gibt also Matrizen $T, T' \in \mathbb{Z}^{n \times n}$ mit

$$(b_1, \ldots, b_k) = (b'_1, \ldots, b'_k)T' \quad \text{und} \quad (b'_1, \ldots, b'_k) = (b_1, \ldots, b_k)T.$$

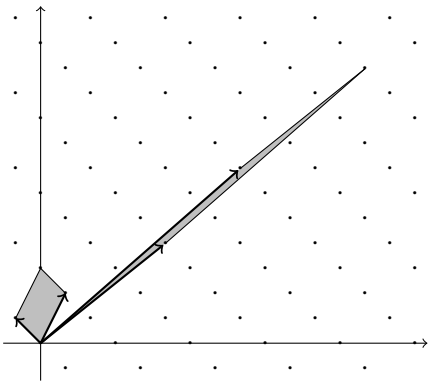
Aus $n = k$ und der linearen Unabhängigkeit von $b_1, \ldots, b_k$ und von $b'_1, \ldots, b'_k$ folgt $T'T = TT' = I_n$. Es handelt sich bei T und T' also um unimodulare Matrizen. Deshalb gilt $|\det(T)| = |\det(T')| = 1$, und deshalb

$$\begin{aligned} V(\Pi(b_1, \ldots, b_n)) &= |\det(b_1, \ldots, b_n)| \\ &= |\det((b'_1, \ldots, b'_n)T')| \\ &= |\det(b'_1, \ldots, b'_n)\det(T')| \\ &= |\det(b'_1, \ldots, b'_n)| \\ &= V(\Pi(b'_1, \ldots, b'_n)), \end{aligned}$$

wie behauptet. ∎

Beispiel. Eine alternative Basis für das Gitter $L = \langle \binom{5}{4}, \binom{8}{7} \rangle \subseteq \mathbb{Z}^2$ ist $\{\binom{1}{2}, \binom{-1}{1}\}$. Die Flächeninhalte der zugehörigen Parallelogramme sind

$$\left|\begin{vmatrix} 5 & 8 \\ 4 & 7 \end{vmatrix}\right| = |5 \cdot 7 - 4 \cdot 8| = 3$$
$$\left|\begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix}\right| = |1 \cdot 1 - 2 \cdot (-1)| = 3.$$



Die Umkehrung des Satzes ist im allgemeinen falsch. Zum Beispiel ist auch $\langle \binom{3}{0}, \binom{0}{1} \rangle$ ein Gitter, dessen Basis einen Spat mit Volumen 3 aufspannt, aber es ist offensichtlich ein anderes Gitter als L, denn $\binom{0}{1} \notin L$.

Das Volumen des Spats, der von den Basisvektoren eines Gitters aufgespannt wird, hängt wegen des vorherigen Satzes nicht von der Wahl der Basis ab, sondern nur vom Gitter selbst.

Die geometrische Anschauung legt die Vermutung nahe, dass eine Basis bestehend aus kurzen Vektoren relativ nah an einer Orthogonalbasis sein muss. Eine solche fast orthogonale Basis gibt einen adäquaten Eindruck davon, wie engmaschig das Gitter ist.

Wir wissen aus Satz 100, wie man aus einer beliebigen Basis eines Unterraums von $\mathbb{R}^n$ eine ONB berechnet. Das Volumen des Spats einer ONB ist immer 1, aber man kann den Algorithmus aus

Satz 100 leicht so abändern, dass er eine gegebene geordnete Basis $(b_1, \dots, b_k)$ eines Unterraums von $\mathbb{R}^n$ in eine orthogonale Basis $(v_1, \dots, v_k)$ mit $V(\Pi(b_1, \dots, b_k)) = V(\Pi(v_1, \dots, v_k))$ überführt. Dazu setzt man $v_1 = b_1$ und

$$v_i = b_i - \sum_{j=1}^{i-1} \frac{\langle b_i | v_j \rangle}{\langle v_j | v_j \rangle} v_j$$

für $j = 2, \dots, k$. Wie im Beweis von Satz 100 rechnet man nach, dass dann $\langle v_i | v_j \rangle = 0$ für $i \neq j$ gilt. Nach Konstruktion gilt dann (im Fall $n = k$)

$$(v_1, \dots, v_n) = (b_1, \dots, b_n) \begin{pmatrix} 1 & * & \cdots & * \\ & 1 & \ddots & \vdots \\ & & \ddots & * \\ & & & 1 \end{pmatrix},$$

und da die Matrix rechts die Determinante 1 hat, folgt $V(\Pi(v_1, \dots, v_n)) = V(\Pi(b_1, \dots, b_n))$.

Satz 134. Sei $\langle \cdot | \cdot \rangle$ ein Skalarprodukt auf $\mathbb{R}^n$ und $\| \cdot \|$ die zugehörige Norm. Sei $L \subseteq \mathbb{Z}^n$ ein Gitter und $\{b_1, \dots, b_k\}$ eine Basis von L. Es sei $v_1 = b_1$ und

$$v_i = b_i - \sum_{j=1}^{i-1} \frac{\langle b_i | v_j \rangle}{\langle v_j | v_j \rangle} v_j$$

für $i = 2, \dots, k$. Dann gilt

$$\forall\, x \in L \setminus \{0\} : \|x\| \geq \min\{\|v_1\|, \dots, \|v_k\|\}.$$

Beweis. Sei $x \in L \setminus \{0\}$, etwa $x = \alpha_1 b_1 + \cdots + \alpha_k b_k$ für gewisse $\alpha_1, \dots, \alpha_k \in \mathbb{Z}$. Wegen $x \neq 0$ muss mindestens ein α_i von Null verschieden sein. Sei $\ell \in \{1, \dots, k\}$ maximal mit $\alpha_\ell \neq 0$. Dann gilt also $x = \alpha_1 b_1 + \cdots + \alpha_\ell b_\ell$.

Wegen $b_i = v_i + \sum_{j=1}^{i-1} \frac{\langle b_i | v_j \rangle}{\langle v_j | v_j \rangle} v_j$ $(i = 1, \dots, \ell)$ gilt 05-24

$$\begin{aligned} x &= \alpha_1 b_1 + \alpha_2 b_2 + \cdots + \alpha_{\ell-1} b_{\ell-1} + \alpha_\ell b_\ell \\ &= \underbrace{\alpha_1 v_1 + \alpha_2 \Big(v_2 + \frac{\langle b_1 | v_1 \rangle}{\langle v_1 | v_1 \rangle} v_1 \Big) + \cdots + \alpha_{\ell-1} \Big(v_{\ell-1} + \sum_{j=1}^{\ell-2} \frac{\langle b_i | v_j \rangle}{\langle v_j | v_j \rangle} v_j \Big)}_{\in \langle v_1, \dots, v_{\ell-1} \rangle} + \alpha_\ell \Big(v_\ell + \underbrace{\sum_{j=1}^{\ell-1} \frac{\langle b_i | v_j \rangle}{\langle v_j | v_j \rangle} v_j}_{\in \langle v_1, \dots, v_{\ell-1} \rangle} \Big) \\ &= \nu_1 v_1 + \cdots + \nu_{\ell-1} v_{\ell-1} + \alpha_\ell v_\ell \end{aligned}$$

für gewisse $\nu_1, \dots, \nu_{\ell-1} \in \mathbb{Q}$. Wegen der Orthogonalität der v_i folgt

$$\begin{aligned} \|x\|^2 = \langle x | x \rangle &= \langle \nu_1 v_1 + \cdots + \nu_{\ell-1} v_{\ell-1} + \alpha_\ell v_\ell | \nu_1 v_1 + \cdots + \nu_{\ell-1} v_{\ell-1} + \alpha_\ell v_\ell \rangle \\ &= \underbrace{\nu_1^2 \langle v_1 | v_1 \rangle + \cdots + \nu_{\ell-1}^2 \langle v_{\ell-1} | v_{\ell-1} \rangle}_{\geq 0} + \alpha_\ell^2 \langle v_\ell | v_\ell \rangle \\ &\geq \alpha_\ell^2 \|v_\ell\|^2 \\ &\geq \min\{\|v_1\|^2, \dots, \|v_k\|^2\}, \end{aligned}$$

wobei im letzten Schritt ausgenutzt wurde, dass für $\alpha_\ell \in \mathbb{Z} \setminus \{0\}$ jedenfalls $\alpha_\ell^2 \geq 1$ gelten muss. ∎

Der Satz bestätigt die geometrische Intuition, dass die Elemente von Orthogonalbasen relativ kurz sind. Allerdings handelt es sich bei $v_1, \ldots, v_k$ im allgemeinen nicht um Elemente von L, weil die Koeffizienten $\frac{\langle b_i | v_j \rangle}{\langle v_j | v_j \rangle}$ in der Definition der v_i im allgemeinen nicht in $\mathbb{Z}$ sondern nur in $\mathbb{Q}$ liegen. Um kurze Vektoren zu finden, die auch selbst in L liegen, betrachtet man Gitter-Basen, bei denen sich nur wenig ändert, wenn man sie als Untervektorraumbasen auffasst und aus ihnen eine Orthogonalbasis berechnet. Dass sich wenig ändert, wird durch die beiden Bedingungen ausgedrückt, dass die Koeffizienten $\frac{\langle b_i | v_j \rangle}{\langle v_j | v_j \rangle}$ bei der Berechnung der Orthogonalbasis klein sein sollen, und dass jedes v_i nicht viel länger sein soll als v_{i-1}. Das führt auf die folgende Definition.

Definition 85. Sei $L \subseteq \mathbb{Z}^n$ ein Gitter und $(b_1, \ldots, b_k)$ eine geordnete Basis von L. Seien $v_1, \ldots, v_k \in \mathbb{Q}^k$ wie in Satz 134. Die Basis $(b_1, \ldots, b_k)$ heißt *reduziert* (engl. *reduced*), falls gilt:

1. Für alle i, j mit $1 \leq j < i \leq k$ gilt $|\frac{\langle b_i | v_j \rangle}{\langle v_j | v_j \rangle}| \leq \frac{1}{2}$
2. Für alle $i = 1, \ldots, k-1$ gilt $\|v_i\|^2 \leq 2\|v_{i+1}\|^2$.

Beispiel. Die Basis $(\binom{5}{4}, \binom{8}{7})$ ist nicht reduziert, denn für die beiden Vektoren

$$v_1 = \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 8 \\ 7 \end{pmatrix} - \frac{\langle \binom{8}{7} | \binom{5}{4} \rangle}{\langle \binom{5}{4} | \binom{5}{4} \rangle} \begin{pmatrix} 5 \\ 4 \end{pmatrix} = \begin{pmatrix} -12/41 \\ 15/41 \end{pmatrix}$$

05-20 des zugehörigen Orthogonalsystems gilt $\|v_1\|^2 = 41 > \frac{18}{41^2} = 2\|v_2\|^2$. Die zweite Bedingung ist
05-24 also verletzt. Die erste ist auch schon verletzt, denn $\frac{\langle \binom{8}{7} | \binom{5}{4} \rangle}{\langle \binom{5}{4} | \binom{5}{4} \rangle} = \frac{68}{41} > \frac{1}{2}$.

Die Basis $(\binom{-1}{1}, \binom{1}{2})$, die dasselbe Gitter erzeugt, ist reduziert, denn für

$$v_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \frac{\langle \binom{1}{2} | \binom{-1}{1} \rangle}{\langle \binom{-1}{1} | \binom{-1}{1} \rangle} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 3/2 \end{pmatrix}$$

gilt $\|v_1\|^2 = 2 \leq 9 = 2\|v_2\|^2$, und $|\frac{\langle \binom{1}{2} | \binom{-1}{1} \rangle}{\langle \binom{-1}{1} | \binom{-1}{1} \rangle}| = \frac{1}{2} \leq \frac{1}{2}$.

Satz 135. Sei $L \subseteq \mathbb{Z}^n$ ein Gitter und $(b_1, \ldots, b_k)$ eine reduzierte Basis von L. Weiter seien $v_1, \ldots, v_k \in \mathbb{Q}^k$ wie in Satz 134. Dann gilt $\|b_i\|^2 \leq 2^{i-1}\|v_i\|^2$ für $i = 1, \ldots, k$.

Beweis. Schreibe $\mu_{i,j} = \frac{\langle b_i | v_j \rangle}{\langle v_j | v_j \rangle}$ für $1 \leq j < i \leq k$. Dann gilt $b_i = v_i + \sum_{j<i} \mu_{i,j} v_j$ für $i = 1, \ldots, k$. Wegen der Orthogonalität der v_j untereinander gilt

$$\begin{aligned} \|b_i\|^2 = \langle b_i | b_i \rangle &= \langle v_i + \sum_{j<i} \mu_{i,j} v_j | v_i + \sum_{j<i} \mu_{i,j} v_j \rangle \\ &= \langle v_i | v_i \rangle + \sum_{j<i} \underbrace{\mu_{i,j}^2}_{\leq \frac{1}{4}} \underbrace{\langle v_j | v_j \rangle}_{\leq 2^{i-j}\|v_i\|^2} \\ &\leq \underbrace{\left(1 + \frac{1}{4} \sum_{j<i} 2^{i-j}\right)}_{= \frac{1}{2} + \frac{1}{4} 2^i \leq 2^{i-1}} \|v_i\|^2 \leq 2^{i-1}\|v_i\|^2. \end{aligned}$$

Im letzten Schritt wurde die Summenformel $\sum_{j=1}^{i-1} q^j = \frac{q^i - q}{q-1}$ verwendet, die für alle $q \in \mathbb{R} \setminus \{1\}$ gilt (geometrische Reihe). ∎

Satz 136. Sei $L \subseteq \mathbb{Z}^n$ ein Gitter und $(b_1, \ldots, b_k)$ eine reduzierte Basis von L. Weiter seien $x_1, \ldots, x_m \in L$ linear unabhängig. Dann gilt

$$\max\{\|b_1\|^2, \ldots, \|b_m\|^2\} \leq 2^{k-1} \max\{\|x_1\|^2, \ldots, \|x_m\|^2\}.$$

Insbesondere gilt $\|b_1\|^2 \leq 2^{k-1}\|x\|^2$ für alle $x \in L \setminus \{0\}$.

Beweis. Wir beginnen ähnlich wie im Beweis von Satz 134. Wegen $x_1, \ldots, x_m \in L$ gibt es für jedes $i = 1, \ldots, m$ eine Wahl von $\alpha_{i,\ell} \in \mathbb{Z}$ $(\ell = 1, \ldots, k)$, so dass $x_i = \sum_{\ell=1}^{k} \alpha_{i,\ell} b_\ell$. Da die $x_1, \ldots, x_m$ nach Annahme linear unabhängig sind, kann es keine Indices $\ell_1, \ldots, \ell_{m-1} \in \{1, \ldots, k\}$ geben, so dass $x_1, \ldots, x_m$ im von $b_{\ell_1}, \ldots, b_{\ell_{m-1}}$ erzeugten $\mathbb{Q}$-Vektorraum liegen. Es muss also mindestens m verschiedene Indices $\ell_1, \ldots, \ell_m \in \{1, \ldots, k\}$ geben, so dass für jedes dieser ℓ ein passendes $i \in \{1, \ldots, k\}$ existiert, für das $\alpha_{i,\ell} \neq 0$ ist. Mit anderen Worten: jedes der $b_{\ell_1}, \ldots, b_{\ell_m}$ muss in der Linearkombination von wenigstens einem der x_i mit einem von Null verschiedenen Koeffizienten vorkommen. 05-31

Wir können annehmen, dass $1 \leq \ell_1 < \ell_2 < \cdots < \ell_m \leq k$ gilt. Dann ist insbesondere $q \leq \ell_q$ für alle q. Wähle ein beliebiges $q \in \{1, \ldots, m\}$ und ein passendes $p \in \{1, \ldots, m\}$ mit $\alpha_{p,\ell_q} \neq 0$. Setze $\mu_{i,i} = 1$ für $i = 1, \ldots, k$ und $\mu_{i,j} = \frac{\langle b_i | v_j \rangle}{\langle v_j | v_j \rangle}$ für $1 \leq j < i \leq k$. Dann gilt $b_i = \sum_{j \leq i} \mu_{i,j} v_j$ für $i = 1, \ldots, m$ und 05-24

$$\begin{aligned}
\|x_p\|^2 &= \|\sum_{i=1}^{k} \alpha_{p,i} b_i\|^2 = \|\sum_{i=1}^{k} \alpha_{p,i} \sum_{j \leq i} \mu_{i,j} v_j\|^2 \\
&= \langle \sum_{i=1}^{k} \alpha_{p,i} \sum_{j \leq i} \mu_{i,j} v_j \,|\, \sum_{i=1}^{k} \alpha_{p,i} \sum_{j \leq i} \mu_{i,j} v_j \rangle \\
&= \sum_{i=1}^{k} \sum_{j \leq i} \underbrace{\alpha_{p,i}^2 \mu_{i,j}^2 \langle v_j | v_j \rangle}_{\geq 0} \geq \underbrace{\alpha_{p,\ell_q}^2}_{\geq 1} \underbrace{\mu_{\ell_q,\ell_q}^2}_{=1} \|v_{\ell_q}\|^2 \\
&\geq \|v_{\ell_q}\|^2 \underset{\substack{\uparrow \\ \ell_q \geq q \\ \text{und Def. 85}}}{\geq} 2^{q-\ell_q} \|v_q\|^2 \underset{\substack{\uparrow \\ \text{Satz 135}}}{\geq} 2^{q-\ell_q-(q-1)} \|b_q\|^2 \underset{\substack{\uparrow \\ \ell_q \leq k}}{\geq} 2^{1-k} \|b_q\|^2.
\end{aligned}$$

Damit ist gezeigt, dass es für jedes $q \in \{1, \ldots, m\}$ ein $p \in \{1, \ldots, m\}$ gibt mit $\|x_p\|^2 \geq 2^{1-k}\|b_q\|^2$. Daraus folgt $\max\{\|x_1\|^2, \ldots, \|x_m\|^2\} \geq 2^{1-k} \max\{\|b_1\|^2, \ldots, \|b_m\|^2\}$, und daraus die Behauptung. ∎

Der Satz sagt, dass die Vektoren in einer reduzierten Basis nur um einen bestimmten Faktor größer sein können als die kürzesten von Null verschiedenen Vektoren des Gitters. Zwar kann es sein, dass es im Gitter noch kürzere Elemente gibt, aber diese zu finden ist nach derzeitigem Kenntnisstand nur mit unvertretbar hohem Rechenaufwand möglich. Zur Berechnung einer reduzierten Basis gibt es dagegen einen effizienten Algorithmus, den sogenannten *LLL-Algorithmus* (benannt nach seinen Entdeckern Lenstra, Lenstra und Lovász). In gewisser Weise erkauft man sich die Effizienz des Algorithmus dadurch, dass man nicht die kürzesten Vektoren von $L \setminus \{0\}$ verlangt, sondern sich mit Vektoren begnügt, für die zumindest garantiert ist, dass sie nicht wesentlich länger sind als die kürzesten. 06-06

Wir werden die Einzelheiten des LLL-Algorithmus zur Berechnung einer reduzierten Basis hier nicht besprechen, sondern nur darauf verweisen, dass der Algorithmus in den meisten Computeralgebrasystemen implementiert ist, z.B. in Maple (im Befehl `IntegerRelations[LLL]`), in Mathematica (im Befehl `LatticeReduce`), und in Sage (in der Methode `LLL` der Klasse für Matrizen über $\mathbb{Z}$).

Obwohl der LLL-Algorithmus nicht unbedingt die kürzesten Vektoren eines Gitters findet, ist er ein extrem nützliches Werkzeug. Man wendet ihn an, indem man ein Gitter konstruiert, in dem die Vektoren, die man kennt, sehr lang sind, und die, die man finden will, sehr kurz.

Beispiel.

1. Sei $A \in \mathbb{Z}^{n\times m}$. Wir suchen eine Basis von $\ker_{\mathbb{Z}} A$. Da $\mathbb{Z}$ ein Integritätsbereich ist, gilt $\ker_{\mathbb{Z}} A = \ker_{\mathbb{Z}} wA$ für jedes $w \in \mathbb{Z} \setminus \{0\}$. Wenn wir w hinreichend groß wählen, werden alle $\mathbb{Z}$-Linearkombinationen von Spalten von A lang sein im Vergleich zu den Vektoren in einer reduzierten Basis von $\ker_{\mathbb{Z}} A$.

05-24 Sei $L \subseteq \mathbb{Z}^{n+m}$ das Gitter, das von den Zeilen von $(wA^\top | I_m) \in \mathbb{Z}^{m\times(m+n)}$ erzeugt wird. Dann gilt

$$(0,\dots,0,x_1,\dots,x_m) \in L \quad \Longleftrightarrow \quad (x_1,\dots,x_m) \in \ker_{\mathbb{Z}}(A).$$

Wir können also hoffen, dass sich solche Vektoren in einer reduzierten Basis von L befinden.

Betrachte zum Beispiel $A = \begin{pmatrix} 2 & 9 & 10 & 5 & 6 \\ 10 & 3 & 2 & 7 & 7 \\ 9 & 5 & 10 & 6 & 1 \end{pmatrix}$. Eine reduzierte Basis des Gitters, das von den Zeilen von $(A^\top | I_5)$ erzeugt wird, lautet

	Norm
$\{(-1,1,-2,-1,-1,0,2,0),$	≈ 3.4641
$(2,-1,0,0,-1,1,-1,1),$	≈ 3.0000
$(0,0,3,-1,-2,1,2,0),$	≈ 4.3589
$(2,-3,-1,0,2,-1,0,-1),$	≈ 4.4721
$(6,5,0,0,3,-2,1,-1)\}.$	≈ 8.7178

Diese Basis enthält keine Vektoren der Form $(0,0,0,x_1,x_2,x_3,x_4,x_5)$, weil $w = 1$ zu klein gewählt war. (Je nach verwendeter Software erhält man unter Umständen ein anderes Ergebnis. Die reduzierte Basis eines Gitters ist nicht eindeutig bestimmt.)

Für $(10A^\top | I_5)$ bekommt man

	Norm
$\{(-10,10,10,-2,-3,1,4,0),$	≈ 18.1659
$(0,0,0,-7,-9,3,13,0),$	≈ 17.5499
$(0,10,-10,-3,-5,2,5,1),$	≈ 16.2481
$(10,0,10,-2,-4,2,3,1),$	≈ 15.2971
$(10,0,0,-3,12,-10,13,-11)\}.$	≈ 25.3574

Das liefert die Lösung $(-7,-9,3,13,0) \in \ker_{\mathbb{Z}} A$. Es müsste aber noch mindestens einen weiteren linear unabhängigen Vektor in $\ker_{\mathbb{Z}} A$ geben. Offenbar war $w = 10$ auch

zu klein. Man könnte als nächstes $w = 100$ testen und würde tatsächlich zwei linear unabhängige Elemente von $\ker_{\mathbb{Z}} A$ finden. Es spricht aber auch nichts dagegen, ein noch viel größeres w zu wählen, z.B. $w = 10^{10}$. Für diese Wahl erhält man die reduzierte Basis

$$\begin{array}{ll} & \text{Norm} \\ \{(0,0,0,-7,-9,3,13,0), & \approx 17.5499 \\ \ (0,0,0,-4,46,-35,32,-36), & \approx 75.3459 \\ \ (0,0,-10^{10},-1,16,-12,10,-12), & \approx 10^{10} \\ \ (0,10^{10},0,5,-12,11,-18,13), & \approx 10^{10} \\ \ (10^{10},0,0,-3,12,-10,13,-11)\}. & \approx 10^{10} \end{array}$$

Die Vektoren, die den Elementen von $\ker_{\mathbb{Z}} A$ entsprechen, sind jetzt deutlich kürzer als die übrigen Vektoren. In der Tat gilt $\ker_{\mathbb{Z}} A = \langle \begin{pmatrix} -7 \\ -9 \\ 3 \\ 13 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 46 \\ -35 \\ 32 \\ -36 \end{pmatrix} \rangle$.

2. Den gleichen Trick kann man verwenden, um Beziehungen zwischen reellen Zahlen aufzuspüren, die man nur bis zu einer gewissen (endlichen) Genauigkeit kennt.

   Seien $x_1, \ldots, x_n \in \mathbb{R}$ und seien $\xi_1, \ldots, \xi_n \in \mathbb{Q}$ so, dass $|x_i - \xi_i| < \varepsilon$ für ein gewisses (kleines) $\varepsilon > 0$ und $i = 1, \ldots, n$ gilt.

   Wir wollen wissen, ob $x_1, \ldots, x_n$ linear abhängig über $\mathbb{Z}$ sind, d.h. ob es einen Vektor $(e_1, \ldots, e_n) \in \mathbb{Z}^n \setminus \{0\}$ gibt mit

   $$e_1 x_1 + \cdots + e_n x_n = 0.$$

   Betrachte dazu das Gitter

   $$L = \langle \begin{pmatrix} \lfloor w\xi_1 \rfloor \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} \lfloor w\xi_n \rfloor \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \rangle \subseteq \mathbb{Z}^{n+1},$$

   für ein (großes) $w \in \mathbb{Z}$. Dabei steht $\lfloor u \rfloor := \max\{z \in \mathbb{Z} : z \leq u\}$ für die kleinste ganze Zahl $z \in \mathbb{Z}$, die höchstens so groß ist wie $u \in \mathbb{R}$.

   Sind $e_1, \ldots, e_n \in \mathbb{Z}$ wie vorher, so gilt

   $$\begin{aligned} & |e_1 w\xi_1 + \cdots + e_n w\xi_n| \\ & = |e_1 w(x_1 - \xi_1) + \cdots + e_n w(x - \xi_n)| \\ & \leq (|e_1| + \cdots + |e_n|) w\varepsilon. \end{aligned}$$

   Deshalb werden, wenn ε hinreichend klein und w hinreichend groß ist, die Vektoren $(e_1 w\lfloor\xi_1\rfloor + \cdots + e_n w\lfloor\xi_n\rfloor, e_1, \ldots, e_n)$ in L vergleichsweise kurz sein, und man kann hoffen, dass LLL sie findet.

Eine typische Anwendung ist „runden rückwärts“. Es ist meist einfach, von einer exakten Darstellung einer Zahl, wie z.B. $x = \frac{1}{3}$, zu einer gerundeten wie z.B. $\xi = 0.3333$ zu kommen. Aber wie kommt man wieder zurück? Wenn man weiß (oder vermutet), dass 0.3333 ein Näherungswert für eine rationale Zahl $\frac{p}{q}$ ist, dann kann man Kandidaten für p und q finden, indem man LLL z.B. auf das Gitter

$$\langle \begin{pmatrix} 100 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \lfloor 100 \cdot 0.3333 \rfloor \\ 0 \\ 1 \end{pmatrix} \rangle = \langle \begin{pmatrix} 100 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 33 \\ 0 \\ 1 \end{pmatrix} \rangle$$

05-23 anwendet. Ein möglicher Output ist $(1, 1, -3), (30, -3, 10)$. Das bedeutet, dass $1 \cdot 1 + (-3) \cdot 0.3333 \approx 0$, also $0.3333 \approx \frac{1}{3}$ ist.

Welche rationale Zahl ist 0.548387? Eine reduzierte Basis für

$$\langle \begin{pmatrix} 1000000 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 548387 \\ 0 \\ 1 \end{pmatrix} \rangle$$

ist $(-3, 17, -31), (32026, -1315, 2398)$. Es gilt also $0.548387 \approx \frac{17}{31}$.

Die Zahl $x \approx \xi = 0.9342585459$ ist eine (irrationale) Lösung einer quadratischen Gleichung. Welcher? Eine quadratische Gleichung ist eine lineare Abhängigkeit zwischen 1, x, und x^2. Wir können also Kandidaten finden, indem wir LLL z.B. auf das Gitter

$$\langle \begin{pmatrix} 10000000000 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 9342585459 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 8728390305 \\ 0 \\ 0 \\ 1 \end{pmatrix} \rangle$$

anwenden. Eine reduzierte Basis dieses Gitters ist

	Norm
$\{(12, 3, 8, -12),$	$\approx$ 19.0
$(20425, -1643, -10440, 13057),$	$\approx$ 26445.4
$(1622, 24545, -20647, -6021)\}$	$\approx$ 32674.7

Das zeigt, dass x nah bei einer Nullstelle des Polynoms $3 + 8X - 12X^2$ liegt. Es beweist **nicht,** dass die reelle Zahl x, von der wir nur die ersten 10 Nachkommastellen kennen, tatsächlich eine Nullstelle dieses Polynoms ist. Die Rechnung erlaubt es nur, 0.934259 zu $\frac{2+\sqrt{13}}{6}$ runden.

3. Rechnungen in einem endlichen Körper $\mathbb{Z}_p$ mit p prim kann man auch als eine Art Approximation auffassen. Zwar sagt $[\frac{6}{11}]_{\equiv 1091} = [794]_{\equiv 1091}$ nicht, dass $\frac{6}{11}$ und 794 als reelle Zahlen nah beieinander liegen, aber diese Zahlen sind sich in gewisser Weise arithmetisch ähnlich. Um von einer „Approximation“ 794 modulo 1091 zurück zur rationalen Zahl $\frac{6}{11}$ zu kommen, wendet man LLL auf das Gitter

$$\langle \begin{pmatrix} 794 \\ 1 \end{pmatrix}, \begin{pmatrix} 1091 \\ 0 \end{pmatrix} \rangle$$

an. Für alle Vektoren (p, q) in diesem Gitter mit $q \neq 0$ gilt $[\frac{p}{q}]_{\equiv 1091} = [794]_{\equiv 1091}$. Plausible Kandidaten entsprechen kurzen Vektoren. Tatsächlich ist $(6, 11), (-79, 37)$ im vorliegenden Beispiel eine reduzierte Basis.

Teil VIII

Algorithmische Lineare Algebra

43 Komplexität

Ein *Algorithmus* (engl. *algorithm*) ist ein systematisches Verfahren zur Lösung eines Problems. In den vorangegangenen Abschnitten haben wir schon mehrere Beispiele für solche Verfahren gesehen. All diese Algorithmen haben gemeinsam, dass die Probleme, die sie lösen, prinzipiell unendlich viele verschiedene mögliche Eingaben haben (z.B. Matrizen $A \in \mathbb{K}^{n\times m}$ für beliebige $n, m \in \mathbb{N}$). Wenn man den möglichen Eingaben eine Größe zuordnet, wird man akzeptieren (müssen), dass der Algorithmus für größere Eingaben mehr Rechenzeit braucht als für weniger große. Zugleich will man nicht mehr Rechenzeit als nötig investieren. Um die *Effizienz* eines Algorithmus zu beurteilen, betrachtet man Funktionen T, die die Rechenzeit des Algorithmus in Abhängigkeit von der Problemgröße messen. Dabei ist zu berücksichtigen, dass die Rechenzeit für verschiedene Eingaben der gleichen Größe nicht unbedingt gleich sein muss. Man unterschiedet zwischen der durchschnittlichen Rechenzeit (*average case*) und der maximal möglichen Rechenzeit (*worst case*) für eine Eingabegröße.

Darüber hinaus gibt es verschiedene Optionen, wie man die Laufzeit eines Algorithmus messen kann. Eine naheliegende Möglichkeit besteht darin, den Algorithmus zu programmieren und die Laufzeit für konkrete Eingaben experimentell zu bestimmen. Eine Messung der Laufzeit, z.B. in Sekunden, erfasst aber nicht nur die allgemeinen Eigenschaften des Algorithmus, sondern hängt immer auch ab von spezielle Eigenschaften des verwendeten Computers, von der verwendeten Programmiersprache, von den Details der Programmierung, und von der Wahl der Testbeispiele. Im Übrigen wüsste man gerne vorher, wie lange eine (große) Rechnung dauert. Bei einer experimentellen Bestimmung weiß man es erst hinterher. Um die Analyse der Laufzeit eines Algorithmus von den Eigenschaften einer konkreten Implementierung zu trennen, misst man die Laufzeit deshalb nicht in echter Zeit (z.B. in Sekunden), sondern zählt stattdessen, wie viele „Grundoperationen“ für die Rechnung durchgeführt werden.

Für Algorithmen in der linearen Algebra bietet es sich an, die Verknüpfungen $+, \cdot, -, /$ aus $\mathbb{K}$ als Grundoperationen zu betrachten. Man sagt, ein Algorithmus hat die *arithmetische Komplexität* $T(n)$, wenn er für jede Eingabe der Größe n höchstens $T(n)$ Operationen in $\mathbb{K}$ durchführt. Nach Bedarf kann die Größe der Eingabe auch durch mehr als einen Parameter beschrieben werden, z.B. $T(n, m)$ für Algorithmen, die als Eingabe eine Matrix $A \in \mathbb{K}^{n\times m}$ nehmen.

Beispiel.

1. Betrachte den folgenden Algorithmus zur Polynommultiplikation.

   Algorithmus 9. Eingabe: $p, q \in \mathbb{K}[X]$ mit $\deg(p) = n$, $\deg(q) = m$
   Ausgabe: pq

   1 schreibe $p = p_0 + p_1X + \cdots + p_nX^n$, $q = q_0 + q_1X + \cdots + q_mX^m$

   2 setze $r_0 = r_1 = \cdots = r_{n+m} = 0$

   3 für $i = 0, \ldots, n$:

   4 für $j = 0, \ldots, m$:

   5 $r_{i+j} = r_{i+j} + p_iq_j$

   6 gib $r_0 + r_1X + \cdots + r_{n+m}X^{n+m}$ als Ergebnis zurück.

   In den Schritten 1, 2, 6 finden keine arithmetischen Operationen statt. Schritt 5 besteht aus zwei Operationen in $\mathbb{K}$, einer Addition und einer Multiplikation. Schritt 4 bewirkt, dass Schritt 5 insgesamt $m+1$ mal ausgeführt wird, und Schritt 3 bewirkt, dass Schritt 4 insgesamt $n + 1$ mal ausgeführt wird, so dass Schritt 5 insgesamt $(n + 1)(m + 1)$ mal ausgeführt wird. Die Komplexität des Algorithmus ist also $T(n) = 2(n + 1)(m + 1)$.

2. Betrachte den folgenden Algorithmus zur Polynomendivision.

Algorithmus 10. Eingabe: $p, q \in \mathbb{K}[X]$ mit $\deg(p) = n \geq 0$, $\deg(q) = m \leq n$
Ausgabe: $\mathrm{quo}(p, q)$ und $\mathrm{rem}(p, q)$

1 setze $u = 0$, $v = p$

2 solange $\deg(v) \geq \deg(q)$:

3 setze $t = \frac{\mathrm{lc}(v)}{\mathrm{lc}(q)} X^{\deg(v)-\deg(q)}$

4 setze $u = u + t$ und $v = v - t\,q$

5 gib u und v als Ergebnis zurück.

Der Grad von v ist zu Beginn n und wird in jedem Schleifendurchlauf um mindestens 1 verringert, bis er unter m fällt. Die Schleife kann also höchstens $(n - m + 1)$ mal durchlaufen werden. Die Berechnung von t kostet eine Operation in $\mathbb{K}$. Da t ein Polynom mit nur einem Term ist, kostet die Addition $u = u + t$ höchstens eine Addition, die Berechnung von $t\,q$ höchstens $(m+1)$ Operationen, und die Subtraktion $v - t\,q$ höchstens $m + 1$ Operationen, weil $t\,q$ höchstes $m + 1$ Terme hat. Die Gesamtzahl $T(n, m)$ der Körperoperationen ist deshalb beschränkt durch $(n-m+1)(1+1+(m+1)+(m+1)) = 2(n - m + 1)(m + 2)$.

3. Betrachte den folgenden Algorithmus zur Multiplikation zweier $n \times n$-Matrizen.

Algorithmus 11. Eingabe: $A, B \in \mathbb{K}^{n\times n}$
Ausgabe: AB

1 schreibe $A = ((a_{i,j}))_{i,j=1}^{n}$, $B = ((b_{i,j}))_{i,j=1}^{n}$.

2 setze $c_{i,j} = 0$ für $i, j = 1, \ldots, n$.

3 für $i = 1, \ldots, n$:

4 für $j = 1, \ldots, n$:

5 für $k = 1, \ldots, n$:

6 $c_{i,j} = c_{i,j} + a_{i,k} b_{k,j}$

7 gib $((c_{i,j}))_{i,j=1}^{n}$ als Ergebnis zurück.

Die arithmetische Komplexität dieses Algorithmus ist $T(n) = 2n^3$.

Die „Komplexität“ eines Algorithmus sagt nichts darüber aus, wie „kompliziert“ ein Algorithmus ist, sondern nur darüber, wie lange er braucht. Tatsächlich ist es oft so, dass die Algorithmen mit der geringsten Komplexität zugleich die kompliziertesten sind. Die Bezeichnungsweise kommt vom Begriff der Komplexität eines Problems, die definiert ist als die Komplexität des effizientesten Algorithmus, der dieses Problem löst. Wenn man z.B. von der Komplexität der Polynommultiplikation spricht, dann meint man nicht die Anzahl der Operationen, die ein bestimmter Algorithmus für die Durchführung der Multiplikation braucht, sondern man meint die kleinstmögliche Anzahl von Operationen, die prinzipiell nötig ist, um die Polynommultiplikation durchzuführen. Die Komplexität eines Problems sagt also etwas über die „Kompliziertheit“ des Problems aus. Man weiss nur für sehr wenige Probleme die genaue Komplexität. Die Komplexität jedes Algorithmus ist aber immer eine obere Schranke für die Komplexität des Problems, das er löst.

Bei der Komplexität von Algorithmen interessiert man sich vor allem für das Verhalten bei großen Problemgrößen, weil es wichtiger scheint, z.B. eine 100-jährige Rechnung auf eine 1-jährige Rechnung

zu reduzieren als eine Rechnung von 100 Millisekunden auf eine Millisekunde zu beschleunigen. Bei großen Problemgrößen ist die Anzahl der Operationen so hoch, dass es auf einige Operationen mehr oder weniger nicht ankommt. Es ist deshalb meistens nicht nötig, die Komplexität eines Algorithmus exakt zu bestimmen, sondern es genügt zu wissen, wie sich die Komplexität asymptotisch verhält, wenn man die Problemgröße gegen unendlich gehen lässt. Um asymptotisches Verhalten auszudrücken, kann man folgende Notation verwenden.

06-06 **Definition 86.** Seien $f, g\colon \mathbb{N} \to \mathbb{R}$ zwei Funktionen. Man schreibt $f(n) = \mathrm{O}(g(n))$, falls gilt

$$\exists\, c > 0\ \exists\, n_0 \in \mathbb{N}\ \forall\, n \geq n_0 : |f(n)| \leq c\, g(n)$$

Beispiel.

1. Es gilt $4n^2 + 3n + 5 = \mathrm{O}(n^2)$ und $4n^2 + 3n + 5 = \mathrm{O}(n^3)$, aber **nicht** $n^3 = \mathrm{O}(100n^2)$.
2. Die Multiplikation zweier Polynome vom Grad n kostet $\mathrm{O}(n^2)$ Operationen in $\mathbb{K}$.
3. Die Polynomendivision zweier Polynome mit Graden n bzw. m kostet $\mathrm{O}(nm)$ Operationen in $\mathbb{K}$.
4. Die Multiplikation zweier $n \times n$-Matrizen kostet $\mathrm{O}(n^3)$ Operationen in $\mathbb{K}$.

Satz 137. Sei $A \in \mathbb{K}^{n\times m}$ eine Matrix mit $\mathrm{Rang}(A) = r$. Dann braucht die Berechnung einer Treppenform von A nicht mehr als

$$\frac{1}{6}r\big(3(2m-1)(2n-r) + 4r^2 - 6nr - 1\big) = \mathrm{O}(\max\{n,m\}^3)$$

Operationen in $\mathbb{K}$. Weitere $(r-1)(r-2)(m-r)$ Operationen in $\mathbb{K}$ genügen, um aus der Treppenform eine Treppennormalform zu berechnen.

Beweis. Wir orientieren uns an Algorithmus 4. Dieser Algorithmus geht von links nach rechts durch die Spalten von A.

Wenn in der ersten Spalte nur Nullen stehen, sind für diese Spalte keine Körperoperationen nötig und der Algorithmus fährt mit der Behandlung der $n \times (m-1)$-Matrix vom Rang r fort, die aus A durch Streichen der ersten Spalte entsteht.

Wenn in der ersten Spalte wenigstens ein von Null verschiedenes Element steht, wird dieses gegebenenfalls durch Vertauschen zweier Zeilen an die Stelle $(1,1)$ gebracht. Das kostet keine Körperoperationen. Danach wird $a_{1,j}$ für $j = 2, \ldots, m$ durch $a_{1,1}$ dividiert ($(m-1)$ Operationen) und anschließend $a_{1,1}$ auf 1 gesetzt (0 Operationen). Dann wird für jedes $i = 2, \ldots, n$ das $(-a_{i,1})$-fache der ersten Zeile zur $i$ten Zeile addiert. Der neue Eintrag für die Position $(i,1)$ ist nach Konstruktion 0 und muss nicht berechnet werden. Für jeden anderen Eintrag (i,j) mit $j = 2, \ldots, m$ berechnet sich der neue Eintrag durch $a_{i,j}^{\mathrm{neu}} = a_{i,j} - a_{i,1}a_{1,j}$. Die Berechnung von $a_{i,j}^{\mathrm{neu}}$ braucht 2 Operationen (eine Multiplikation und eine Subtraktion). Das sind $2(m-1)$ Operationen für jedes $i = 2, \ldots, n$ und $(m-1) + (n-1)2(m-1) = (m-1)(2n-1)$ Operationen insgesamt (incl. der Operationen für die Normierung der ersten Zeile). Danach bleibt eine Treppenform für eine $(n-1) \times (m-1)$-Matrix vom Rank $r-1$ zu berechnen.

Ist $T(n,m,r)$ die Anzahl der nötigen Operationen, so gilt also

$$T(n,m,r) \leq \max\big(T(n,m-1,r), T(n-1,m-1,r-1) + (m-1)(2n-1)\big),$$

für alle $n, m, r \geq 1$. Weiter gilt $T(n, m, 0) = 0$ für alle $n, m \geq 1$, denn die einzige Matrix, deren Rang Null ist, ist die Null-Matrix, und diese ist ohne weitere Rechnung bereits eine Treppenform.

Betrachte

$$f(n, m, r) := \frac{1}{6} r\big(3(2m-1)(2n-r) + 4r^2 - 6nr - 1\big).$$

Dann gilt:

1. $f(n, m, 0) = 0$ für alle $n, m \in \mathbb{N}$
2. $f(n, m, r) \geq 0$ für alle $n, m, r \in \mathbb{N}$ mit $0 \leq r \leq \min(n, m)$
3. $f(n-1, m-1, r-1) + (m-1)(2n-1) = f(n, m, r)$ für alle $n, m, r \in \mathbb{N}$
4. $f(n, m-1, r) = f(n, m, r) - (2n-r)r \leq f(n, m, r)$ für alle $n, m, r \in \mathbb{N}$ mit $r \leq n$.

Wir zeigen $T(n, m, r) \leq f(n, m, r)$ durch Induktion nach m.

Induktionsanfang: Wenn $m = 0$ ist, muss auch $r = 0$ sein, und es gilt $T(n, m, 0) = f(n, m, 0)$, wie behauptet.

Induktionsschluss $m-1 \to m$: Nach Induktionsannahme gilt $T(n, m-1, r) \leq f(n, m-1, r)$ für alle $n, r \in \mathbb{N}$ mit $r \leq m-1$. Daraus folgt zusammen mit den oben gemachten Bemerkungen

$$\begin{aligned} T(n, m, r) &\leq \max\big(T(n, m-1, r), T(n-1, m-1, r-1) + (m-1)(2n-1)\big) \\ &\leq \max\big(\underbrace{f(n, m-1, r)}_{\leq f(n,m,r)}, \underbrace{f(n-1, m-1, r-1) + (m-1)(2n-1)}_{= f(n,m,r)}\big) = f(n, m, r). \end{aligned}$$

∎

Satz 138. Für jedes $A \in \mathbb{K}^{n \times n}$ kann man $\operatorname{Rang}(A)$, $\det(A)$ sowie Basen von $\ker A$, $\operatorname{im} A$, $\operatorname{coker} A$, $\operatorname{coim} A$ mit $O(n^3)$ Operationen in $\mathbb{K}$ ausrechnen. Wenn A invertierbar ist, genügen $O(n^3)$ Operationen, um A^{-1} zu berechnen.

Beweis. Alle genannten Daten lassen sich aus einer Treppennormalform von A oder $(A|I_n) \in \mathbb{K}^{n \times 2n}$ ablesen. Die Behauptungen folgen deshalb sofort aus Satz 137.

Für $\det(A)$ muss man sich die Einträge merken, durch die die Zeilen geteilt werden. Wenn $\operatorname{Rang} A = n$ ist, ist deren Produkt die Determinante. Das Produkt auszurechnen kostet höchstes n Operationen und fällt deshalb nicht ins Gewicht. Wenn $\operatorname{Rang} A < n$ ist, ist $\det(A) = 0$ ohne weitere Rechnung. ∎

44 Dünn besetzte Matrizen

Man sagt, eine Matrix ist *dünn besetzt* (engl. *sparse matrix*), wenn viele ihrer Einträge Null sind. Wenn es kaum Nulleinträge gibt, spricht man von einer *dicht besetzten* Matrix (engl. *dense matrix*). Es gibt keine scharfe Abgrenzung sondern einen fließenden Übergang zwischen diesen Begriffen. Unter Umständen kann dieselbe Matrix abhängig vom Kontext mal als dünn besetzt und mal als dicht besetzt aufgefasst werden.

Beim Rechnen mit dünn besetzten Matrizen kann man ausnutzen, dass $0 \cdot x = 0$ und $0 + x = x$ für alle $x \in \mathbb{K}$ gilt, und dadurch Operationen einsparen. Wenn man im Vorhinein weiß, wo die Nicht-Null-Einträge der Matrix stehen, ist das meist nicht allzu schwer.

Beispiel. Sind $A, B \in \mathbb{K}^{n\times n}$ Diagonalmatrizen, so lässt sich AB mit $\mathrm{O}(n)$ Operationen berechnen.

05-31 Eine dicht besetzte $n \times m$-Matrix wird im Rechner durch eine Liste von Listen von Einträgen dargestellt. Dieses Datenformat entspricht der gewöhnlichen Schreibweise für Matrizen auf Papier. Es hat einen Speicherbedarf von $\mathrm{O}(nm)$. Für dünn besetzte Matrizen sollte man ein anderes Datenformat verwenden, damit nicht ein großer Teil des Speichers mit lauter Nullen belegt wird. Für eine Diagonalmatrix genügt es zum Beispiel, die Einträge auf der Diagonale abzuspeichern (sowie die Tatsache, dass es sich um eine Diagonalmatrix handelt). Dieses Datenformat entspricht der Schreibweise $\mathrm{diag}(a_1, \ldots, a_n)$. Für eine Diagonalmatrix der Größe $n \times n$ braucht man dann nur $\mathrm{O}(n)$ viel Speicher.

In diesem Fall nutzt man neben der dünnen Besetzung auch die besondere *Struktur* der Matrix aus. Für unstrukturierte dünn besetzte Matrizen, also solche, bei denen die Lage der Nullen keinem klaren Schema folgt, gibt es mehrere brauchbare Datenformate. Bei einem davon codiert man die Matrix durch eine Menge von Tripeln $(i, j, a_{i,j})$, die für jeden Nicht-Null-Eintrag der Matrix dessen Position und den Wert enthalten. Die Codierung einer $n \times m$-Matrix mit s Nicht-Null-Einträgen braucht dann $\mathrm{O}(s)$ viel Speicher.

Beispiel. Eine Codierung der Matrix

$$\begin{pmatrix}
0&0&5&0&0&0&0&0&0&0&0&4&0&0&0&0&0&0&0&0&0&0&0&0&8\\
0&0&0&0&0&0&0&0&0&0&0&0&0&5&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&7&0&0&0&0&8&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&4&0&0&0&0&5&0&0&0&0&0&0&0&0&2&0&0&0&0&0&0\\
0&0&0&0&0&0&7&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&4&0&0&0&0&0&0&0&9&0&0&0&0\\
0&0&0&0&0&0&0&2&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&2&0\\
0&0&0&5&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&4&0&3&0&0&0&0&0&0&4&0&0&0&0&0&0&1&0&7&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&5&0&0&4&0&0&6&0&0&0&0&0\\
0&0&0&0&3&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&7&0&0&0&3&0&0&0&0&0&0&9&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&3&0&0&0&0&0&0&0&0&0&8&0&0&4\\
7&8&0&0&0&0&0&0&0&0&0&0&0&0&0&7&4&0&0&0&0&0&0&0&0\\
0&0&3&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&6&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&4&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&7&0&0&0&0&0&0&0&0&0&0&6&0&0&0&0&0&0&0&0&0&0&0&0\\
3&0&0&0&0&0&0&0&0&0&0&0&0&0&0&8&0&2&0&0&7&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&7&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&3&0&0&0&0&3&0&0&0&0
\end{pmatrix}$$

im beschriebenen Format ist

$$\big\{(16,17,4),\ (22,13,6),\ (23,1,3),\ (22,2,7),\ (18,12,4),\ (8,8,2),\ (9,22,1),\ (1,12,4),\\ (15,22,8),\ (3,9,7),\ (9,4,5),\ (11,5,4),\ (11,23,9),\ (3,14,8),\ (13,5,3),\ (14,9,7),\\ (12,17,4),\ (4,19,2),\ (11,7,7),\ (11,23,7),\ (11,21,1),\ (12,20,6),\ (4,5,4),\ (2,14,5),$$

$(16,16,7),\ (15,12,3),\ (8,24,2),\ (11,7,3),\ (23,18,2),\ (1,3,5),\ (16,2,8),\ (24,21,7),$
$(14,20,9),\ (4,10,5),\ (14,13,3),\ (25,21,3),\ (17,3,3),\ (7,13,4),\ (12,14,5),\ (1,25,8),$
$(7,21,9),\ (17,19,6),\ (16,1,7),\ (15,25,4),\ (23,21,7),\ (18,6,1),\ (25,16,3),\ (11,14,4),$
$(5,7,7),\ (23,16,8)\}$

Algorithmus 12. Eingabe: $A \in \mathbb{K}^{n\times m}$, $B \in \mathbb{K}^{m\times k}$ im oben beschriebenen Datenformat. Ausgabe: $C = AB$ im gleichen Format

1 setze $C = \emptyset$
2 für alle Einträge $(i, j, a_{i,j})$ von A:
3 für alle Einträge $(p, q, b_{p,q})$ von B:
4 falls $j = p$:
5 durchsuche C nach einem Eintrag $(i, q, c_{i,q})$
6 wenn es einen solchen gibt:
7 ersetze darin $c_{i,q}$ durch $c_{i,q} + a_{i,j}b_{p,q}$
8 sonst:
9 $C = C \cup \{(i, q, a_{i,j}b_{p,q})\}$
10 gib C als Ergebnis zurück

Satz 139. Angewandt auf Matrizen A, B mit je höchstens s Nicht-Null-Einträgen braucht Algorithmus 12 nicht mehr als $2s^2$ Operationen in $\mathbb{K}$.

Beweis. Bei jeder Ausführung der Schritte 4–8 finden höchstens 2 Operationen statt, nämlich entweder eine Addition und eine Multiplikation in Schritt 7, oder nur eine Multiplikation in Schritt 9, oder gar keine Operationen. Schritt 3 bewirkt, dass die Schritte 4–8 höchstens s mal durchlaufen werden, und Schritt 2 bewirkt, dass Schritt 3 höchstens s mal durchlaufen wird, so dass die Schritte 4–8 insgesamt höchstens s^2 mal durchlaufen werden. ∎

Auch beim Lösen von Gleichungsystemen kann man Nullen ausnutzen, besonders dann, wenn man etwas über die Lage der Nullen weiss, etwa, dass alle Einträge unterhalb der Diagonalen Null sind.

Beispiel. Sei $A = ((a_{i,j}))_{i,j=1}^{n} \in \mathbb{K}^{n\times n}$ eine invertierbare obere Dreiecksmatrix, d.h. die Diagonaleinträge $a_{i,i}$ $(i = 1, \ldots, n)$ sind alle von Null verschieden und $a_{i,j} = 0$ für alle $1 \leq j < i \leq n$.

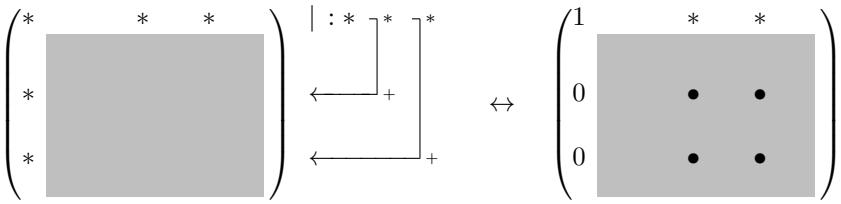
$$A = \begin{pmatrix} * & \cdots & \cdots & * \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix}$$

Dann kann man zu gegebenem $b \in \mathbb{K}^{n\times n}$ die Lösung $x \in \mathbb{K}^n$ des Gleichungssystems $Ax = b$ mit $\mathrm{O}(n^2)$ Operationen berechnen, und zwar mit folgendem Algorithmus:

Algorithmus 13. Eingabe: $A = ((a_{i,j}))_{i,j=1}^{n} \in \mathbb{K}^{n\times n}$ wie oben, $b \in \mathbb{K}^n$
Ausgabe: $x \in \mathbb{K}^n$ so, dass $Ax = b$

1 für $i = n, n-1, \ldots, 1$:
2 $x_i = b_i$
3 für $j = i+1, \ldots, n$: 06-02
4 $x_i = x_i - a_{i,j}x_j$
5 $x_i = \frac{x_i}{a_{i,i}}$
6 gib $(x_1, \ldots, x_n)$ als Ergebnis zurück.

Wenn man nicht vorher weiss, wo die Nullen stehen, verwendet man eine Variante des Gauss-Algorithmus. Dabei wählt man für die Elimination Zeilen aus, in denen möglichst viele Nullen stehen. Eine solche Wahl hat zwei Vorteile: erstens braucht man für die Spalten, in denen die ausgewählte Zeile eine Null hat, keine Rechnungen durchführen, und zweitens bleiben etwaige Nullen, die in diesen Spalten stehen, erhalten. Letzteres ist für den weiteren Verlauf der Rechnung von Vorteil.



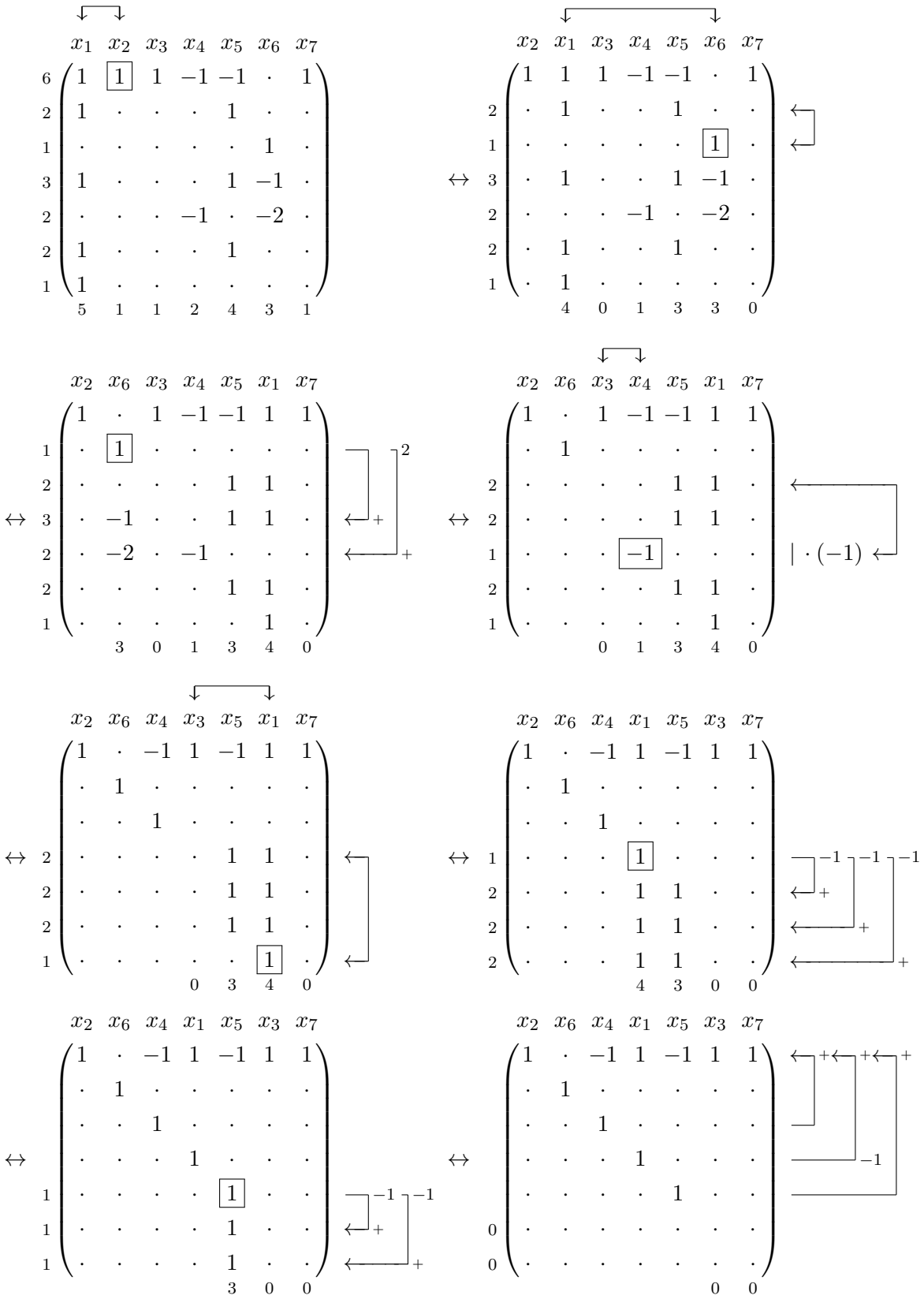
Wenn in der ausgewählten Zeile p Nicht-Null-Einträge stehen und die erste Spalte der aktuellen Teilmatrix q Nicht-Null-Einträge hat, dann muss man mit einem sogenannten *fill-in* von bis zu $(p-1)(q-1)$ Einträgen in die Restmatrix rechnen. Man kann die Zahl der Optionen erhöhen, indem man neben Vertauschungen von Zeilen auch Vertauschungen von Spalten zulässt. Eine Zeilenoperation entspricht ja einer Multiplikation mit einer Elementarmatrix von links und ändert den Kern der Matrix nicht ($Ax = 0 \iff UAx = 0$ für alle $U \in \mathrm{GL}(n, \mathbb{K})$). Eine Spaltenoperation entspricht einer Multiplikation von rechts, und eine solche verändert zwar im allgemeinen den Kern, aber es gilt natürlich $Ax = 0 \iff AVV^{-1}x = 0$ für alle $V \in \mathrm{GL}(n, \mathbb{K})$. Vertauschungen von Spalten kann man deshalb gestatten, sofern man jede solche Vertauschung auch auf die Koordinaten des Lösungsvektors anwendet. Für die Handrechnung empfiehlt es sich, die Koordinaten des gesuchten Vektors (also die Variablen des Gleichungssystems) über die entsprechenden Spalten der Matrix zu schreiben.

Beispiel. Gesucht ist $\ker A$ für

$$A = \begin{pmatrix} 1 & 1 & 1 & -1 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & -2 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{Q}^{7\times 7}.$$

In der folgenden Rechnung ist für jede Zeile und jede Spalte notiert, wie viele Nicht-Null-Einträge sie hat. Zur Elimination wird in jedem Schritt ein Nicht-Null-Eintrag (i, j) mit

minimalem fill-in verwendet. Zur besseren Lesbarkeit schreiben wir · statt 0.



$$\leftrightarrow \begin{array}{c} \begin{array}{ccccccc} x_2 & x_6 & x_4 & x_1 & x_5 & x_3 & x_7 \end{array} \\ \begin{array}{c} \\ \\ \\ \\ \\ 0 \\ 0 \end{array} \begin{pmatrix} 1 & \cdot & \cdot & \cdot & \cdot & 1 & 1 \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & & 0 & 0 \end{pmatrix} \end{array} \quad \Longrightarrow \quad \ker A = \left\langle \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} \right\rangle \quad \begin{array}{cc} 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & -1 \end{array} \quad \begin{array}{c} x_2 \\ x_6 \\ x_4 \\ x_1 \\ x_5 \\ x_3 \\ x_7 \end{array}$$

Die Vertauschungen der Spalten werden durch die entsprechenden Vertauschungen der Koordinaten der Lösungsvektors kompensiert.

45 Strukturierte Matrizen

Unter strukturierten Matrizen versteht man Matrixen, bei denen zwischen den Einträgen gewisse
05-24 bekannte Beziehungen bestehen, die man in Algorithmen ausnutzen kann. Ein Beispiel für strukturierte dünn besetzte Matrizen sind Diagonalmatrizen. Eine strukturierte Matrix muss aber nicht unbedingt dünn besetzt sein.

06-02 **Definition 87.** Sei $A = ((a_{i,j}))_{i,j=1}^{n} \in \mathbb{K}^{n \times n}$.

1. A heißt *Hankel-Matrix,* falls für alle $i, j < n$ gilt $a_{i,j+1} = a_{i+1,j}$.
2. A heißt *Töplitz-Matrix,* falls für alle $i, j < n$ gilt $a_{i+1,j+1} = a_{i,j}$.

Beispiel. $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$ ist eine Hankel-Matrix und $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 5 & 4 & 1 \end{pmatrix}$ ist eine Töplitz-Matrix.

Satz 140.

1. Das Produkt zweier Hankel-Matrizen der Größe $n \times n$ kann mit $\mathrm{O}(n^2)$ Körperoperationen berechnet werden.
2. Das Produkt zweier Töplitz-Matrizen der Größe $n \times n$ kann mit $\mathrm{O}(n^2)$ Körperoperationen berechnet werden.

Beweis.

1. Seien $A = ((a_{i,j}))_{i,j=1}^{n}$, $B = ((b_{i,j}))_{i,j=1}^{n}$, $C = AB = ((c_{i,j}))_{i,j=1}^{n}$. Dann gilt $c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$ für alle i, j. Die Berechnung dieser Summe für ein Paar (i, j) kostet $2n - 1$ Operationen (n Multiplikationen und $n - 1$ Additionen). Wir können zunächst $c_{1,1}, c_{1,2}, \ldots, c_{1,n}$ und $c_{2,1}, \ldots, c_{n,1}$ auf diese Weise berechnen. Das kostet insgesamt $(2n - 1)^2$ Operationen.

   Für jedes $i, j < n$ gilt

$$c_{i+1,j+1} = \sum_{k=1}^{n} a_{i+1,k} b_{k,j+1}$$

$$
\begin{aligned}
&= \sum_{k=1}^{n-1} a_{i+1,k} b_{k,j+1} + a_{i+1,n} b_{n,j+1} \\
&= \sum_{k=1}^{n-1} a_{i,k+1} b_{k+1,j} + a_{i+1,n} b_{n,j+1} \\
&= \sum_{k=2}^{n} a_{i,k} b_{k,j} + a_{i+1,n} b_{n,j+1} \\
&= c_{i,j} - a_{i,1} b_{1,j} + a_{i+1,n} b_{n,j+1}.
\end{aligned}
$$

Jeder der Einträge $c_{i+1,j+1}$ lässt sich also mit nur vier Operationen aus dem vorherigen Eintrag $c_{i,j}$ berechnen.

Da $(n-1)^2$ Einträge zu berechnen sind, kostet die Berechnung der $c_{i,j}$ für $i,j > 1$ nicht mehr als $4(n-1)^2$ Operationen. Zusammen mit den Kosten für die Berechnung von $c_{1,j}$ und $c_{i,1}$ $(i,j = 1,\ldots,n)$ ergibt sich ein Aufwand von $(2n-1)^2 + 4(n-1)^2 = 8n^2 - 12n + 5 = \mathrm{O}(n^2)$, wie behauptet.

2. Analog. ∎

Beispiel. Betrachte $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$ und $B = \begin{pmatrix} 6 & 7 & 8 \\ 7 & 8 & 9 \\ 8 & 9 & 0 \end{pmatrix}$. Sei $C = ((c_{i,j}))_{i,j=1}^{3} = AB$.

Wir berechnen zuerst

$$
\begin{aligned}
c_{1,1} &= 1 \cdot 6 + 2 \cdot 7 + 3 \cdot 8 = 44 \\
c_{1,2} &= 1 \cdot 7 + 2 \cdot 8 + 3 \cdot 9 = 50 \\
c_{1,3} &= 1 \cdot 8 + 2 \cdot 9 + 3 \cdot 0 = 26 \\
c_{2,1} &= 2 \cdot 6 + 3 \cdot 7 + 4 \cdot 8 = 65 \\
c_{3,1} &= 3 \cdot 6 + 4 \cdot 7 + 5 \cdot 8 = 86
\end{aligned}
$$

Die restlichen Einträge ergebnen sich zu

$$
\begin{aligned}
c_{2,2} &= c_{1,1} - a_{1,1} b_{1,1} + a_{2,3} b_{3,2} = 44 - 1 \cdot 6 + 4 \cdot 9 = 74 \\
c_{3,3} &= c_{2,2} - a_{2,1} b_{1,2} + a_{3,3} b_{3,3} = 74 - 2 \cdot 7 + 5 \cdot 0 = 60 \\
c_{2,3} &= c_{1,2} - a_{1,1} b_{1,2} + a_{2,3} b_{3,3} = 50 - 1 \cdot 7 + 4 \cdot 0 = 43 \\
c_{3,2} &= c_{2,1} - a_{2,1} b_{1,1} + a_{3,3} b_{3,2} = 65 - 2 \cdot 6 + 5 \cdot 9 = 98
\end{aligned}
$$

Beachten Sie, dass das Produkt zweier Hankel-Matrizen im allgemeinen keine Hankel-Matrix mehr ist:

$$
C = AB = \begin{pmatrix} 44 & 50 & 26 \\ 65 & 74 & 43 \\ 86 & 98 & 60 \end{pmatrix}
$$

Für Hankel- und Töplitz-Matrizen lassen sich auch die Inversen (sofern sie existieren) und die Determinanten mit nur $\mathrm{O}(n^2)$ Operationen berechnen. Wir wollen hier nicht ausführen, wie das geht, sondern uns stattdessen einer anderen Art von strukturierten Matrizen zuwenden.

Definition 88. Seien $u_1, \dots, u_n \in \mathbb{K}$ paarweise verschieden. Dann heißt

$$V(u_1, \dots, u_n) := \begin{pmatrix} 1 & u_1 & \cdots & u_1^{n-1} \\ 1 & u_2 & \cdots & u_2^{n-1} \\ \vdots & & & \vdots \\ 1 & u_n & \cdots & u_n^{n-1} \end{pmatrix} \in \mathbb{K}^{n \times n}$$

die *Vandermonde Matrix* zu $u_1, \dots, u_n$.

Die Vandermonde-Matrix tritt auf im Zusammenhang mit der Auswertung von Polynomen. Ist $p = p_0 + p_1 X + \cdots + p_{n-1} X^{n-1} \in \mathbb{K}[X]$ und $u_1, \dots, u_n \in \mathbb{K}$, so gilt

$$\begin{pmatrix} p(u_1) \\ \vdots \\ \vdots \\ p(u_n) \end{pmatrix} = \begin{pmatrix} 1 & u_1 & \cdots & u_1^{n-1} \\ 1 & u_2 & \cdots & u_2^{n-1} \\ \vdots & & & \vdots \\ 1 & u_n & \cdots & u_n^{n-1} \end{pmatrix} \begin{pmatrix} p_0 \\ \vdots \\ \vdots \\ p_{n-1} \end{pmatrix}$$

In Abschnitt 12 haben wir gesehen, dass

$$\det V(u_1, \dots, u_n) = \prod_{i=2}^{n} \prod_{j=1}^{i-1} (u_i - u_j) \neq 0$$

ist, wenn die u_i paarweise verschieden sind. Die Matrix $V(u_1, \dots, u_n)$ ist dann also invertierbar. Daraus folgt, dass es für jede Wahl von paarweise verschiedenen $u_1, \dots, u_n \in \mathbb{K}$ und jede Wahl von (nicht notwendigerweise paarweise verschiedenen) $y_1, \dots, y_n \in \mathbb{K}$ genau ein Polynom $p = p_0 + p_1 X + \cdots + p_{n-1} X^{n-1} \in \mathbb{K}[X]$ gibt mit $p(u_i) = y_i$ für $i = 1, \dots, n$. Man nennt dieses p das *Interpolationspolynom* für $(u_1, y_1), \dots, (u_n, y_n)$.

Die Berechnung des Interpolationspolynoms durch Lösen des Gleichungssystems mit dem Gauß-Algorithmus würde $\mathrm{O}(n^3)$ Operationen kosten. Wir wollen zeigen, dass es schon mit $\mathrm{O}(n^2)$ Operationen gelingt. Dazu erinnere man sich zunächst, dass die Auswertung eines Polynoms dasselbe ist wie die Division mit Rest durch ein lineares Polynom: $\mathrm{rem}(p, X - u) = p(u)$ für $p \in \mathbb{K}[X]$ und $u \in \mathbb{K}$. Die Menge aller Polynome, die bei einem vorgegebenen $u \in \mathbb{K}$ einen vorgegebenen Wert y annehmen, lässt sich schreiben als

$$\{\, y + (X - u)p : p \in \mathbb{K}[X] \,\} = y + (X - u)\mathbb{K}[X] = [y]_{\sim_{X-u}} \in \mathbb{K}[X]/(X - u)\mathbb{K}[X].$$

Dabei gilt, wie üblich, die Definition $p \sim_{X-u} q :\iff X - u \mid p - q$. Die Menge der Polynome, die bei zwei vorgegebenen $u_1, u_2 \in \mathbb{K}$ $(u_1 \neq u_2)$ zwei vorgegebene Werte $y_1, y_2 \in \mathbb{K}$ annehmen, ist demnach

$$[y_1]_{\sim_{X-u_1}} \cap [y_2]_{\sim_{X-u_2}}.$$

Zwei Polynome haben an zwei Punkten $u_1, u_2 \in \mathbb{K}$ mit $u_1 \neq u_2$ genau dann denselben Wert, wenn ihre Differenz an diesen Stellen den Wert 0 hat, und dies ist genau dann der Fall, wenn die Differenz ein Vielfaches von $(X - u_1)(X - u_2)$ ist. Es ist deshalb nicht überraschend, dass der Schnitt einer
06-02 Äquivalenzklasse bezüglich $\sim_{X-u_1}$ mit einer Äquivalenzklasse bezüglich $\sim_{X-u_2}$ eine Äquivalenzklasse bezüglich $\sim_{(X-u_1)(X-u_2)}$ ist.

Der folgende Satz sagt, dass das nicht nur für Polynome $X - u_1, X - u_2$ sondern für beliebige teilerfremde Polynome gilt. Es gibt sogar eine analoge Aussage für $\mathbb{Z}$ statt $\mathbb{K}[X]$.

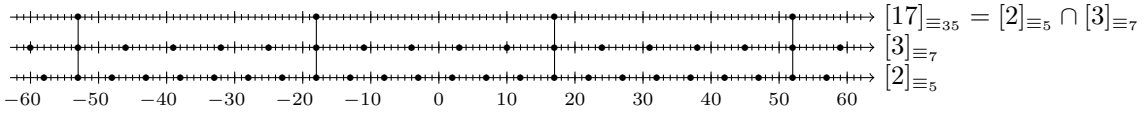
Satz 141.

1. (Newton-Interpolation) Seien $p_1, p_2 \in \mathbb{K}[X]$ mit $\gcd(p_1, p_2) = 1$, und seien $u_1, u_2 \in \mathbb{K}[X]$. Weiter seien $q_1, q_2 \in \mathbb{K}[X]$ so, dass $q_1 p_1 + q_2 p_2 = 1$. Dann gilt

$$[u_1]_{\sim_{p_1}} \cap [u_2]_{\sim_{p_2}} = [u_1 + p_1 q_1 \operatorname{rem}(u_2 - u_1, p_2)]_{\sim_{p_1 p_2}}.$$

2. (Chinesischer Restsatz) Seien $p_1, p_2 \in \mathbb{Z}$ mit $\gcd(p_1, p_2) = 1$, und seien $u_1, u_2 \in \mathbb{Z}$. Weiter seien $q_1, q_2 \in \mathbb{Z}$ so, dass $q_1 p_1 + q_2 p_2 = 1$. Dann gilt

$$[u_1]_{\equiv_{p_1}} \cap [u_2]_{\equiv_{p_2}} = [u_1 + p_1 q_1 \operatorname{rem}(u_2 - u_1, p_2)]_{\equiv_{p_1 p_2}}.$$



Beweis.

1. „$\subseteq$“ Sei $m \in [u_1]_{\sim_{p_1}} \cap [u_2]_{\sim_{p_2}}$. Zu zeigen: $m \sim_{p_1 p_2} u_1 + p_1 q_1 \operatorname{rem}(u_2 - u_1, p_2)$.

   Wegen $m \in [u_1]_{\sim_{p_1}}$ gilt $m = u_1 + v p_1$ für ein $v \in \mathbb{K}[X]$, und wegen $m \in [u_2]_{\sim_{p_2}}$ gilt $m = u_2 + w p_2$ für ein $w \in \mathbb{K}[X]$.

   Daraus folgt:

$$\begin{aligned}
m &= (q_1 p_1 + q_2 p_2) m \\
&= q_1 p_1 (u_2 + w p_2) + q_2 p_2 (u_1 + v p_1) \\
&= q_1 p_1 u_2 + q_2 p_2 u_1 + p_1 p_2 (q_1 w + q_2 v) \\
&\sim_{p_1 p_2} q_1 p_1 u_2 + q_2 p_2 u_1 \\
&= q_1 p_1 u_2 + q_2 p_2 u_1 + q_1 p_1 u_1 - q_1 p_1 u_1 \\
&= u_1 + q_1 p_1 (u_2 - u_1) \\
&= u_1 + q_1 p_1 \big(p_2 \operatorname{quo}(u_2 - u_1, p_2) + \operatorname{rem}(u_2 - u_1, p_2)\big) \\
&\sim_{p_1 p_2} u_1 + q_1 p_1 \operatorname{rem}(u_2 - u_1, p_2),
\end{aligned}$$

   wie behauptet.

   „$\supseteq$“ Sei $m \in [u_1 + p_1 q_1 \operatorname{rem}(u_2 - u_1, p_2)]_{\sim p_1 p_2}$, etwa $m = u_1 + p_1 q_1 \operatorname{rem}(u_2 - u_1, p_2) + p_1 p_2 w$ für ein $w \in \mathbb{K}[X]$. Mit $w = v + \operatorname{quo}(u_2 - u_1, p_2) q_1$ gilt dann $m = u_1 + p_1 q_1 (u_2 - u_1) + p_1 p_2 v$. 06-06

   Daraus folgt $m \sim_{p_1} = u_1$, also $m \in [u_1]_{\sim_{p_1}}$, und wegen $m = u_1 + (1 - q_2 p_2)(u_2 - u_1) + p_1 p_2 v = u_2 + p_2 (q_2 (u_1 - u_2) + p_1 v)$ auch $m \sim_{p_2} = u_2$. 06-06

2. Genauso. ∎

Beispiel.

1. Gesucht ist ein Polynom $u \in \mathbb{K}[X]$ mit $u(1) = 2$, $u(2) = -1$, $u(3) = 2$ und $u(4) = 5$. Man findet ein solches durch wiederholte Anwendung des Satzes.

u_1	u_2	p_1	p_2	q_1	q_2	$u_1+p_1q_1\,\mathrm{rem}(u_2-u_1,p_2)$
2	-1	$X-1$	$X-2$	1	-1	$5-3X$
$5-3X$	2	$(X-1)(X-2)$	$X-3$	$\frac{1}{2}$	$-\frac{1}{2}X$	$11-12X+3X^2$
$11-12X+3X^2$	5	$(X-1)(X-2)(X-3)$	$X-4$	$\frac{1}{6}$	$-\frac{1}{2}+\frac{1}{3}X-\frac{1}{6}X^2$	$\underline{17-23X+9X^2-X^3}$

2. Gesucht ist jetzt eine Zahl $u \in \mathbb{Z}$ mit $u \equiv_3 1$, $u \equiv_5 3$, $u \equiv_7 4$, und $u \equiv_{13} 5$. Man findet
06-06 eine solche Zahl durch eine analoge Rechnung:

u_1	u_2	p_1	p_2	q_1	q_2	$u_1+p_1q_1\,\mathrm{rem}(u_2-u_1,p_2)$
1	3	3	5	2	-1	13
13	4	15	7	1	-2	88
88	5	105	13	1	-8	$\underline{928}$

Algorithmus 14. (Newton-Interpolation)
Eingabe: $u_1, \ldots, u_n \in \mathbb{K}$ paarweise verschieden, und $y_1, \ldots, y_n \in \mathbb{K}$.
Ausgabe: das Polynom $f = f_0 + f_1X + \cdots + f_{n-1}X^{n-1} \in \mathbb{K}[X]$ mit $f(u_i) = y_i$ für alle i.

1 setze $f = y_1$, $p = 1$
2 für $i = 2, \ldots, n$:
3 setze $p = (X - u_{i-1})p$
4 setze $f = f + p\frac{1}{p(u_i)}\,\mathrm{rem}(y_i - f, X - u_i)$
5 gib f als Ergebnis zurück.

Satz 142. Algorithmus 14 ist korrekt und braucht nicht mehr als $O(n^2)$ Operationen in $\mathbb{K}$.

Beweis. Korrektheit: Wir zeigen durch Induktion nach i, dass am Beginn des $i$ten Schleifendurchlaufs gilt $p = (X - u_1) \cdots (X - u_{i-1})$ und $f(u_j) = y_j$ für alle $j \le i$.

Für $i = 2$ ist das durch Schritt 1 gewährleistet. Angenommen es gilt für $i - 1$. Nach Schritt 3 gilt dann $p = (X - u_1) \cdots (X - u_{i-2})(X - u_{i-1})$. Es gilt $\gcd(p, X - u_i) = 1$, weil $u_1, \ldots, u_n$ nach Voraussetzung paarweise verschieden sind. Ferner folgt aus $p = \mathrm{quo}(p, X - u_i)(X - u_i) + \mathrm{rem}(p, X - u_i)$ und $\mathrm{rem}(p, X - u_i) = p(u_i)$ die Gleichung

$$\frac{1}{p(u_i)}p - \frac{\mathrm{quo}(p, X - u_i)}{p(u_i)}(X - u_i) = 1,$$

06-06 und deshalb folgt aus Satz 141, dass nach Schritt 4 $f(u_i) = y_i$ gilt, und dass $\mathrm{rem}(f, p)$ der vorherige Wert von f ist. Da für diesen nach Induktionsannahme gilt $f(u_j) = y_j$ für alle $j \le i - 1$ und weil $p(u_j) = 0$ für alle $j \le i - 1$ gilt, folgt $f(u_j) = y_j$ für alle $j \le i$.

Komplexität: Bei jedem Durchlauf durch die Schritte 3 und 4 sind f und p Polynome vom Grad $\le n$. Da $X - u_{i-1}$ und $X - u_i$ Polynome vom Grad 1 sind, kostet die Berechnung von $(X - u_i)p$ jeweils nur $O(n)$ Operationen, ebenso jene von $p(u_i) = \mathrm{rem}(p, X - u_i)$ und die von $\mathrm{rem}(y_i - f, X - u_i) = (y_i - f)(u_i)$. Die Berechnung von $\mathrm{rem}(y_i - f, X - u_i)/p(u_i) \in \mathbb{K}$ kostet eine Operation, die Multiplikation dieses Körperelements mit p kostet $O(n)$ Operationen, und

die Subtraktion des Resultats von f noch einmal $O(n)$. Insgesamt kostet die Durchführung der Schritte 3 und 4 also $O(n)$ Operationen in $\mathbb{K}$. Wegen Schritt 2 werden die Schritte 3 und 4 insgesamt $n-1$ mal durchlaufen. Daraus ergibt sich die behauptete Komplexität. ∎

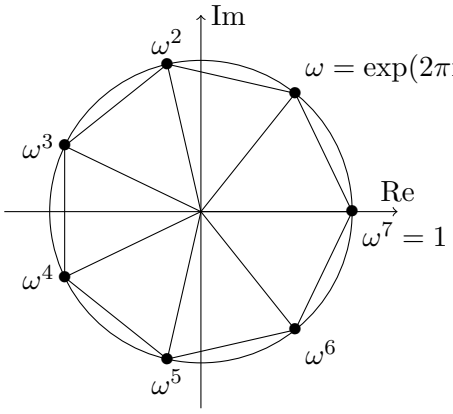
Wenn es noch schneller gehen soll, kann man sich durch eine geeigente Wahl von $u_1, \ldots, u_n$ eine noch strukturiertere Matrix verschaffen. Für diese spezielle Vandermonde Matrix lässt sich die Matrix-Vektor-Multiplikation mit einem beliebigen (unstrukturierten) Vektor mit nur $O(n \log(n))$ Operationen durchgeführen. Auch die Interpolation ist dann mit diesem Aufwand möglich.

Definition 89. Sei $n \in \mathbb{N} \setminus \{0\}$.

1. $\omega \in \mathbb{K}$ heißt *n-te Einheitswurzel* (engl. *root of unity*), falls $\omega^n = 1$ ist.
2. Eine n-te Einheitswurzel heißt *primitiv,* falls $\omega^i \neq 1$ für $i = 1, \ldots, n-1$.
3. Ist ω eine primitive n-te Einheitswurzel, so heißt die Matrix
$$\mathrm{DFT}_n^{(\omega)} := V(1, \omega, \ldots, \omega^{n-1}) = ((\omega^{ij}))_{i,j=0}^{n-1} \in \mathbb{K}^{n \times n}$$
die *diskrete Fouriertransformation* der Größe n zur Einheitswurzel ω. Man beachte, dass die Einträge der Matrix von 0 bis $n-1$ statt wie üblich von 1 bis n indiziert werden.

Beispiel.

1. 1 ist eine n-te Einheitswurzel für jedes $n \in \mathbb{N}$, aber nur für $n = 1$ eine primitive.
2. -1 ist eine primitive zweite Einheitswurzel. In $\mathbb{Q}$ und $\mathbb{R}$ gibt es außer 1 und -1 keine weiteren Einheitswurzeln.
3. $\mathrm{i} \in \mathbb{C}$ ist eine primitive vierte Einheitswurzel, denn $\mathrm{i}^4 = 1$ und $\mathrm{i}^1 = \mathrm{i} \neq 1$, $\mathrm{i}^2 = -1 \neq 1$ und $\mathrm{i}^3 = -\mathrm{i} \neq 1$. Es gilt
$$\mathrm{DFT}_4^{(\mathrm{i})} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \end{pmatrix}$$
Allgemeiner: zu jedem $n \in \mathbb{N} \setminus \{0\}$ ist $\omega = \exp(2\pi\mathrm{i}/n)$ eine primitive n-te Einheitswurzel in $\mathbb{C}$. Die Zahlen $\omega, \omega^2, \ldots, \omega^n = 1$ bilden in der komplexen Zahlenebene die Ecken eines regelmäßigen n-Ecks, dessen Ecken auf dem Einheitskreis liegen.



Die n-ten Einheitswurzeln sind genau die Nullstellen des Polynoms $X^n - 1$:

$$X^n - 1 = (X-1)(X-\omega)(X-\omega^2)\cdots(X-\omega^{n-1}).$$

Wegen $\omega^i\omega^j = \omega^{i+j} = \omega^{i+j+kn}$ für alle $i, j, k \in \mathbb{Z}$ bilden die n-ten Einheitswurzeln zusammen mit der Multiplikation eine abelsche Gruppe, die isomorph zu $(\mathbb{Z}_n, +)$ ist.

4. In $\mathbb{Z}_{17} = \mathbb{Z}/17\mathbb{Z}$ ist 9 eine primitive 8-te Einheitswurzel:

i	1	2	3	4	5	6	7	8
9^i	9	13	15	16	8	4	2	1

Jedes der Elemente in der zweiten Zeile ist eine 8-te Einheitswurzel, die Elemente $9, 15, 8, 2$ sind primitiv. Die Elemente 13 und 4 sind primitive vierte Einheitswurzeln, und $16 = -1$ ist eine primitive zweite Einheitswurzel.

In einem Körper $\mathbb{Z}_p$ für $p \in \mathbb{Z}$ prim gilt $x^{p-1} = 1$ für jedes Element $x \in \mathbb{Z} \setminus \{0\}$ (kleiner Satz von Fermat). Jedes Element von $\mathbb{Z}_p \setminus \{0\}$ ist also eine Einheitswurzel.

06-07 **Satz 143.** (Cooley-Tuckey) Sei $n \in \mathbb{N} \setminus \{0\}$ und sei ω eine primitive $2n$-te Einheitswurzel. Weiter sei

$$\Delta = \operatorname{diag}(1, \omega, \omega^2, \ldots, \omega^{n-1}) \in \mathbb{K}^{n\times n}$$

und $P \in \mathbb{K}^{2n\times 2n}$ sei die Permutationsmatrix, für die gilt

$$P \cdot (x_0, x_1, x_2, x_3, \ldots) = (x_0, x_2, x_4, \ldots, x_1, x_3, x_5, \ldots).$$

Dann gilt

$$\mathrm{DFT}_{2n}^{(\omega)} = \begin{pmatrix} I_n & I_n \\ I_n & -I_n \end{pmatrix} \begin{pmatrix} I_n & 0 \\ 0 & \Delta \end{pmatrix} \begin{pmatrix} \mathrm{DFT}_n^{(\omega^2)} & 0 \\ 0 & \mathrm{DFT}_n^{(\omega^2)} \end{pmatrix} P.$$

Beweis. Sei $x = (x_0, x_1, \ldots, x_{2n-1}) \in \mathbb{K}^{2n}$ beliebig und $y = (y_0, \ldots, y_{2n-1}) = \mathrm{DFT}_{2n}^{(\omega)} \cdot x$. Für $i = 0, \ldots, 2n-1$ gilt dann

$$y_i = \sum_{j=0}^{2n-1} \omega^{ij} x_j = \sum_{j=0}^{n-1} \omega^{i(2j)} x_{2j} + \sum_{j=0}^{n-1} \omega^{i(2j+1)} x_{2j+1} = \sum_{j=0}^{n-1} (\omega^2)^{ij} x_{2j} + \omega^i \sum_{j=0}^{n-1} (\omega^2)^{ij} x_{2j+1}.$$

Da ω eine $2n$-te Einheitswurzel ist, gilt $\omega^{2n} = 1$. Daraus folgt $\omega^n \in \{-1, 1\}$ (unabhängig vom
06-07 Körper $\mathbb{K}$), und weil ω eine primitive $2n$-te Einheitswurzel ist, folgt $\omega^n = -1$. Für $i < n$ gilt deshalb $\omega^{i+n} = -\omega^i$ und $(\omega^2)^{i+n} = (\omega^2)^i \omega^{2n} = (\omega^2)^i$, und daraus folgt

$$y_{i+n} = \sum_{j=0}^{n-1} (\omega^2)^{ij} x_{2j} - \omega^i \sum_{j=0}^{n-1} (\omega^2)^{ij} x_{2j+1}.$$

Insgesamt erhält man

$$\begin{pmatrix} (y_i)_{i=0}^{n-1} \\ (y_{i+n})_{i=0}^{n-1} \end{pmatrix} = \begin{pmatrix} \mathrm{DFT}_n^{(\omega^2)} & \Delta\,\mathrm{DFT}_n^{(\omega^2)} \\ \mathrm{DFT}_n^{(\omega^2)} & -\Delta\,\mathrm{DFT}_n^{(\omega^2)} \end{pmatrix} \begin{pmatrix} (x_{2i})_{i=0}^{n-1} \\ (x_{2i+1})_{i=0}^{n-1} \end{pmatrix},$$

wie behauptet. ∎

Beispiel. Für $n = 2$ und $\omega = \mathrm{i} \in \mathbb{C}$ gilt

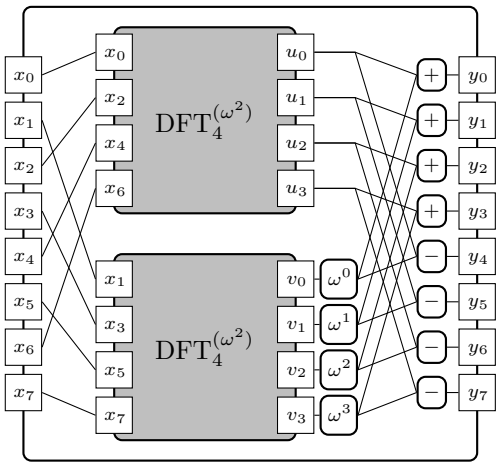
$$\mathrm{DFT}_4^{(\mathrm{i})} = \underbrace{\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & & \\ 0 & 1 & & \\ & & 1 & 0 \\ & & 0 & \mathrm{i} \end{pmatrix}}_{= \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & \mathrm{i} \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -\mathrm{i} \end{pmatrix}} \underbrace{\begin{pmatrix} 1 & 1 & & \\ 1 & -1 & & \\ & & 1 & 1 \\ & & 1 & -1 \end{pmatrix}}_{= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & \mathrm{i} & -\mathrm{i} \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -\mathrm{i} & \mathrm{i} \end{pmatrix}} \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \end{pmatrix}}$$

Die Formel aus Satz 143 führt die Anwendung einer DFT auf die Anwendung zweier DFTs der halben Größe zurück. Diese kleineren DFTs lassen sich mit der gleichen Formel in noch kleinere zerlegen, und so weiter, bis man bei einer ungeraden Problemgröße angekommen ist. Wenn die ursprüngliche Problemgröße eine Zweierpotenz ist, also $n = 2^k$ für ein $k \in \mathbb{N}$ gilt, dann erreicht man nach $k = \log_2(n)$ vielen Zerlegungsschritten die DFT der Größe 1×1. Für diese ist wegen $\mathrm{DFT}_1^{(1)} = I_1$ nichts zu rechnen.

Algorithmus 15. (FFT; Fast Fourier Transform)
Eingabe: Ein Vektor $x = (x_0, \ldots, x_{n-1}) \in \mathbb{K}^n$ mit $n = 2^k$ für ein $k \in \mathbb{N}$, und eine n-te Einheitswurzel $\omega \in \mathbb{K}$.
Ausgabe: $\mathrm{DFT}_n^{(\omega)}\, x$.

wenn $n = 1$ ist, gib x als Ergebnis zurück und stop.
wende den Algo. auf $(x_0, x_2, \ldots, x_{n-2})$ und ω^2 an. Das Ergebnis sei $u = (u_0, \ldots, u_{n/2-1})$.
wende den Algo. auf $(x_1, x_3, \ldots, x_{n-1})$ und ω^2 an. Das Ergebnis sei $v = (v_0, \ldots, v_{n/2-1})$.
$q = 1$
für $i = 0, \ldots, n/2 - 1$:
    $y_i = u_i + q v_i$
    $y_{i+n/2} = u_i - q v_i$
    $q = \omega q$
gib $(y_0, y_1, \ldots, y_{n-1})$ als Ergebnis zurück.

06-17



Satz 144. Algorithmus 15 braucht nicht mehr als $O(n \log(n))$ Operationen in $\mathbb{K}$.

Beweis. Es sei $T(n)$ die Zahl der Operationen für die Problemgröße $n = 2^k$. Die beiden
06-06 rekursiven Aufrufe in den Zeilen 2 und 3 brauchen dann je $T(n/2)$ Operationen. Darüber hinaus ist eine Operation für die Berechnung von ω^2 nötig. In den Zeilen 5–7 werden insgesamt fünf Operationen durchgeführt, und Zeile 4 bewirkt, dass diese drei Zeilen $n/2$ mal ausgeführt werden.

Daraus folgt die Abschätzung $T(n) \leq 2T(n/2) + \frac{5}{2}n + 1$. Wir zeigen induktiv $T(n) \leq 3nk$ für $k \geq 0$. Für $k = 0$ sind keine Operationen nötig, $T(1) = 5 \cdot 1 \cdot 0 = 0$ ist also richtig. Betrachte nun ein $k \in \mathbb{N}$, für das $T(2^k) \leq 3nk$ gilt. Dann gilt $T(2^{k+1}) = T(2n) = 2T(n) + 5n + 1 \leq 2(3nk) + 5n + 1 \leq 6n(k+1) = 3\,2^{k+1}(k+1)$, wie behauptet. ∎

Mit dem FFT-Algorithmus kann man zu einem gegebenen Polynom $p \in \mathbb{K}[X]$ vom Grad $n \leq 2^k$ den Wertevektor $(p(1), p(\omega), \ldots, p(\omega^{2^k-1})) \in \mathbb{K}^{2^k}$ für eine 2^k-te Einheitswurzel ω berechnen. Aus dem folgenden Satz folgt, dass man den FFT-Algorithmus auch zur Interpolation verwenden kann, wenn ω eine primitive 2^k-te Einheitswurzel ist.

Satz 145. Sei $\omega \in \mathbb{K}$ eine primitive n-te Einheitswurzel in $\mathbb{K}$.

1. Für $j = 1, \ldots, n-1$ gilt $\sum_{i=0}^{n-1} \omega^{ij} = 0$.
2. $1/\omega$ ist auch eine primitive n-te Einheitswurzel und $(\mathrm{DFT}_n^{(\omega)})^{-1} = \frac{1}{n}\,\mathrm{DFT}_n^{(1/\omega)}$.

Beweis.

1. Für alle $q \neq 1$ gilt $\sum_{i=0}^{n-1} q^i = \frac{q^n-1}{q-1}$ (geometrische Reihe). Daraus folgt

$$(q^n - 1) = (q-1)\sum_{i=0}^{n-1} q^i$$

für alle $q \neq 1$. Es gilt $\omega^n = 1$, weil ω eine n-te Einheitswurzel ist. Damit gilt auch $(\omega^j)^n = \omega^{jn} = (\omega^n)^j = 1$. Weil ω nach Annahme eine primitive n-te Einheitswurzel ist, gilt $\omega^j \neq 1$ für $j = 1, \ldots, n-1$. Für $q = \omega^j$ folgt also

$$0 = \underbrace{(\omega^j - 1)}_{\neq 0} \sum_{i=0}^{n-1} \omega^{ij},$$

wie behauptet.

2. Dass $1/\omega$ auch eine primitive n-te Einheitswurzel ist, folgt aus $(1/\omega^j) = 1 \iff \omega^j = 1$ für alle $j \in \mathbb{N}$.

   Sei $(i,j) \in \{0,\dots,n-1\}^2$ beliebig und $c_{i,j}$ der (i,j)-te Eintrag von $\mathrm{DFT}_n^{(\omega)} \mathrm{DFT}_n^{(1/\omega)}$ (wobei die Einträge dieser Matrix von 0 bis $n-1$ indiziert werden). Dann gilt gemäß Teil 1:

$$c_{i,j} = \sum_{k=0}^{n-1} \omega^{ik} (\frac{1}{\omega})^{kj} = \sum_{k=0}^{n-1} \omega^{(i-j)k} = \begin{cases} n & \text{falls } i = j \\ 0 & \text{sonst} \end{cases}$$ ∎

46 Schnelle Multiplikation

Für den Standardmultiplikationsalgorithmus für $n \times n$-Matrizen (Alg. 11) haben wir eine Komplexität von $\mathrm{O}(n^3)$ festgestellt. Im vorherigen Abschnitt haben wir gesehen, dass sich bestimmte strukturierte Matrizen, z.B. Töplitz-Matrizen, mit einer geringeren Komplexität multiplizieren lassen. In diesem Abschnitt werden wir sehen, dass sich $n \times n$-Matrizen auch ohne besondere Annahmen über die Struktur mit weniger als $\mathrm{O}(n^3)$ Operationen multiplizieren lassen.

Auch für die Polynommultiplikation lässt sich der Standardalgorithmus (Alg. 9) verbessern. Betrachten wir dieses Problem als erstes. Das Prinzip ist ähnlich wie beim FFT-Algorithmus (Alg. 15): man zerlegt das Problem in mehrere kleine Probleme des gleichen Typs, löst diese rekursiv mit demselben Algorithmus und setzt aus den Ergebnissen der kleinen Probleme das Ergebnis des ursprünglichen Problems zusammen.

Betrachte $a, b \in \mathbb{K}[X]$ mit $\deg a, \deg b < n = 2^k$. Schreibe $a = a_0 + a_1 X^{n/2}$, $b = b_0 + b_1 X^{n/2}$ für $a_0, a_1, b_0, b_1 \in \mathbb{K}[X]$ vom Grad $< n/2$. Dann gilt

$$ab = (a_0 + a_1 X^{n/2})(b_0 + b_1 X^{n/2}) = a_0 b_0 + (a_1 b_0 + a_0 b_1) X^{n/2} + a_1 b_1 X^n.$$

Man kann also die Multiplikation zweier Polynome a, b auf vier Multiplikationen von Polynomen halber Größe und $\mathrm{O}(n)$ Additionen zurückführen. Es zeigt sich allerdings, dass das gegenüber dem Standardalgorithmus noch keine Einsparung bringt. Die wesentliche Beobachtung ist, dass man mit einer Multiplikation weniger auskommt. Berechnet man nämlich nur die drei Produkte

$$u = a_0 b_0, \quad v = a_1 b_1, \quad w = (a_0 + a_1)(b_0 + b_1),$$

so gilt $ab = u + (w - u - v) X^{n/2} + v X^{n/2}$, d.h. die Multiplikation von ab ist auf nur drei Multiplikationen halber Größe und $\mathrm{O}(n)$ Additionen in $\mathbb{K}$ zurückgeführt.

Algorithmus 16. (Karatsuba)
Eingabe: $a, b \in \mathbb{K}[X]$ mit $\deg a, \deg b < n = 2^k$.
Ausgabe: ab

1 Falls $n \leq 1$, berechne ab direkt, gibt das Ergebnis zurück und stop.

2 Schreibe $a = a_0 + a_1 X^{n/2}$, $b = b_0 + b_1 X^{n/2}$ für $a_0, a_1, b_0, b_1 \in \mathbb{K}[X]$ vom Grad $< n/2$.

3 Berechne $u = a_0 b_0$ mit diesem Algorithmus.

4 Berechne $v = a_1 b_1$ mit diesem Algorithmus.

5 Berechne $w = (a_0 + a_1)(b_0 + b_1)$ mit diesem Algorithmus.

6 Berechne $u + (w - u - v) X^{n/2} + v X^n$ und gib dies als Ergebnis zurück.

Satz 146. Algorithmus 16 braucht nicht mehr als $O(n^{\log_2 3}) = O(n^{1.585})$ Operationen in $\mathbb{K}$.

Beweis. Sei $T(n)$ die Zahl der Operationen. Dann gilt $T(1) \leq 1$ und es gibt ein $c \geq 1$ so dass für alle $n = 2^k \in \mathbb{N}$ gilt $T(n) \leq 3T(n/2) + cn$.

Wir zeigen $T(2^k) \leq 2c3^k - c2^k = O(3^k) = O(n^{\log_2 3})$ für alle $k \in \mathbb{N}$ durch Induktion nach k.

Für $k = 0$ ist $2c3^0 - c2^0 = 2c - c = c \geq 1 \geq T(1)$. Ist $k \in \mathbb{N}$ so, dass die Abschätzung für $k-1$ gilt, so folgt

$$T(2^k) \leq 3T(2^{k-1}) + c2^k \leq 3(2c3^k - c2^k) + c2^k \leq 2c3^{k+1} - 3c2^k + c2^k = 2c3^{k+1} - c2^{k+1},$$

wie behauptet. ∎

Es geht noch schneller. Nehmen wir an, wir wollen zwei Polynome $a, b \in \mathbb{K}[X]$ vom Grad $< n$ multiplizieren. Dann gilt $\deg(ab) = \deg(a) + \deg(b) < 2n$. Das Produkt ab ist also durch seine Werte an $2n$ paarweise verschiedenen Punkten eindeutig festgelegt. Sind $u_1, \ldots, u_{2n} \in \mathbb{K}$ paarweise verschieden, so lassen sich die Werte $(ab)(u_i) = a(u_i)b(u_i)$ für $i = 1, \ldots, 2n$ mit geringem Aufwand aus den Werten $a(u_i)$ und $b(u_i)$ berechnen. Das motiviert einen Algorithmus nach folgendem Schema:

1 Wähle paarweise verschiedene $u_1, \ldots, u_{2n} \in \mathbb{K}$

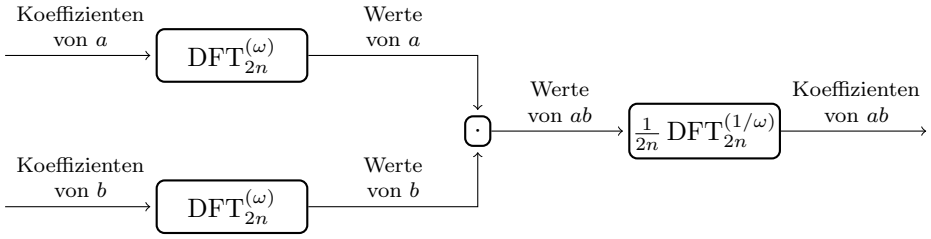
2 Berechne $a(u_1), \ldots, a(u_{2n})$ aus den bekannten Koeffizienten von a

3 Berechne $b(u_1), \ldots, b(u_{2n})$ aus den bekannten Koeffizienten von b

4 Berechne $a(u_1)b(u_1), \ldots, a(u_{2n})b(u_{2n})$

5 Berechne aus diesen Daten die Koeffizienten von ab und gib diese als Ergebnis zurück.

Die Multiplikationen in Schritt 4 kosten dann $O(n)$ Operationen. Wenn man als u_i die Potenzen einer primitiven $2n$-ten Einheitswurzel nimmt, kann man für die Evaluationen und die Interpolation den FFT-Algorithmus einsetzen, so dass diese Schritte nicht mehr als $O(n \log n)$ Operationen brauchen.



06-17

Satz 147. Sei $k \in \mathbb{N}$ und $n = 2^k$. Wenn es in $\mathbb{K}$ eine primitive $2n$-te Einheitswurzel gibt, dann kann man Polynome $a, b \in \mathbb{K}[X]$ mit $\deg(a), \deg(b) < n$ mit $O(n \log n)$ Operationen in $\mathbb{K}$ multiplizieren.

Beweis. Die Aussage folgt direkt aus der vorangegangenen Diskussion. ∎

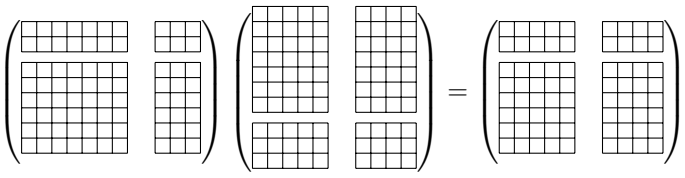
Nun zur Multiplikation von Matrizen. Die Grundidee ist ähnlich wie beim Algorithmus von Karatsuba. Die Matrix wird in Blöcke der halben Größe aufgeteilt, diese werden nach einem bestimmten Schema miteinander multipliziert, und die Ergebnisse der Teilprobleme werden zum Ergebnis des ursprünglichen Problems zusammengesetzt. Dass man mit Matrix-Blöcken rechnen kann wie mit skalaren Matrixeinträgen, klärt der folgende Satz.

Satz 148. Seien $A \in \mathbb{K}^{(n_1+n_2)\times(k_1+k_2)}$, $B \in \mathbb{K}^{(k_1+k_2)\times(m_1+m_2)}$, $C = AB \in \mathbb{K}^{(n_1+n_2)\times(m_1+m_2)}$. Schreibe

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}, \quad B = \begin{pmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{pmatrix}, \quad C = \begin{pmatrix} C_{1,1} & C_{1,2} \\ C_{2,1} & C_{2,2} \end{pmatrix}$$

für $A_{i,j} \in \mathbb{K}^{n_i \times k_j}$ $B_{i,j} \in \mathbb{K}^{k_i \times m_j}$ und $C_{i,j} \in \mathbb{K}^{n_i \times m_j}$ für $i, j \in \{1, 2\}$. Dann gilt

$$\begin{aligned} C_{1,1} &= A_{1,1}B_{1,1} + A_{1,2}B_{2,1}, & C_{1,2} &= A_{1,1}B_{1,2} + A_{1,2}B_{2,2}, \\ C_{2,1} &= A_{2,1}B_{1,1} + A_{2,2}B_{2,1}, & C_{2,2} &= A_{2,1}B_{1,2} + A_{2,2}B_{2,2}. \end{aligned}$$



Beweis. Wir beweisen den Spezialfall $n_1 = n_2 = k_1 = k_2 = m_1 = m_2 =: n$. Den allgemeinen Fall kann man sich zur Übung selbst überlegen.

Schreibe $A = ((a_{i,j}))_{i,j=1}^{2n}$, $B = ((b_{i,j}))_{i,j=1}^{2n}$, $C = ((c_{i,j}))_{i,j=1}^{2n}$. Dann gilt $A_{1,1} = ((a_{i,j}))_{i,j=1}^{n}$, $A_{1,2} = ((a_{i,j+n}))_{i,j=1}^{n}$, $A_{2,1} = ((a_{i+n,j}))_{i,j=1}^{n}$, $A_{2,2} = ((a_{i+n,j+n}))_{i,j=1}^{n}$, und entsprechend für die Blöcke in B und C.

Sei $(i, j) \in \{1, \ldots, 2n\}^2$ beliebig. Dann gilt

$$c_{i,j} = \sum_{k=1}^{2n} a_{i,k}b_{k,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j} + \sum_{k=n+1}^{2n} a_{i,k}b_{k,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j} + \sum_{k=1}^{n} a_{i,k+n}b_{k+n,j}.$$

Daraus folgt die Behauptung. ∎

Das Schema aus dem Satz führt eine Matrixmultiplikation der Größe $2n \times 2n$ auf acht Matrixmultiplikationen der Größe $n \times n$ und vier Additionen der Größe $n \times n$ zurück. Eine Einsparung gegenüber dem Standardalgorithmus ergibt sich damit noch nicht. Wie beim Karatsuba-Algorithmus kommt der Effizienzgewinn aus der Beobachtung, dass man mit einer Multiplikation weniger auskommt.

Algorithmus 17. (Strassen)
Eingabe: $A, B \in \mathbb{K}^{n\times n}$, $n = 2^k$
Ausgabe: $C = AB$

1 Falls $n = 1$ ist, berechne AB direkt, gib das Ergebnis zurück und stop.

2 Schreibe $A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$ und $B = \begin{pmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{pmatrix}$.

3 Berechne die folgenden Matrizen durch rekursive Anwendung dieses Algorithmus:

$$\begin{aligned} M_1 &= (A_{1,1} + A_{2,2})(B_{1,1} + B_{2,2}), & M_2 &= (A_{2,1} + A_{2,2})B_{1,1}, \\ M_3 &= A_{1,1}(B_{1,2} - B_{2,2}), & M_4 &= A_{2,2}(B_{2,1} - B_{1,1}), \\ M_5 &= (A_{1,1} + A_{1,2})B_{2,2}, & M_6 &= (A_{2,1} - A_{1,1})(B_{1,1} + B_{1,2}), \\ M_7 &= (A_{1,2} - A_{2,2})(B_{2,1} + B_{2,2}) \end{aligned}$$

4 Berechne die folgenden Matrizen:

$$C_{1,1} = M_1 + M_4 - M_5 + M_7, \qquad C_{1,2} = M_3 + M_5,$$
$$C_{2,1} = M_2 + M_4, \qquad C_{2,2} = M_1 - M_2 + M_3 + M_6$$

5 Gib $C = \begin{pmatrix} C_{1,1} & C_{1,2} \\ C_{2,1} & C_{2,2} \end{pmatrix}$ als Ergebnis zurück.

Satz 149. Algorithmus 17 ist korrekt und braucht nicht mehr als $O(n^{\log_2 7}) = \mathrm{O}(n^{2.81})$ Operationen in $\mathbb{K}$.

Beweis. Die Korrektheit ergibt sich durch direktes Nachrechnen mit Satz 148, z.B.

$$\begin{aligned} C_{1,2} = M_3 + M_5 &= A_{1,1}(B_{1,2} - B_{2,2}) + (A_{1,1} + A_{1,2})B_{2,2} \\ &= A_{1,1}B_{1,2} - A_{1,1}B_{2,2} + A_{1,1}B_{2,2} + A_{1,2}B_{2,2} \\ &= A_{1,1}B_{1,2} + A_{1,2}B_{2,2}. \end{aligned}$$

Für die Behauptung über die Komplexität sei $T(n)$ die Zahl der Operationen, die der Algorithmus für Matrizen der Größe $n \times n$ aufwendet. Dann gilt $T(1) = 1$ und es gibt ein $c \geq 1$ so dass für alle $n = 2^k \in \mathbb{N}$ gilt $T(n) \leq 7T(n/2) + cn^2$.
Wir zeigen $T(2^k) \leq 2c7^k - c4^{k+1} = \mathrm{O}(7^k) = \mathrm{O}(n^{\log_2 7})$ durch Induktion nach k.
Für $k = 0$ ist $2c7^0 - c4^1 = 7c - 4c = 3c \geq 1 = T(1)$. Ist $k \in \mathbb{N}$ so, dass die Abschätzung für $k - 1$ gilt, so folgt

$$T(2^k) \leq 7T(2^{k-1}) + c4^k \leq 7(2c7^{k-1} - c4^k) + c4^k = 2c7^k - 6c4^k \leq 2c7^k - c4^{k+1},$$

wie behauptet. ∎

Man kennt auch Algorithmen, die für sehr große n noch effizienter sind als der Algorithmus von Strassen. Der derzeitige Weltrekord ist aus dem Jahr 2014 und liegt bei einer Komplexität von $\mathrm{O}(n^\omega)$ für $\omega \approx 2.3728639$. Diese Algorithmen bringen allerdings erst bei unrealistisch großen Matrizen einen Gewinn und sind deshalb bis auf weiteres nur von theoretischem Interesse. Der Strassen-Algorithmus ist dagegen durchaus praxistauglich. Abhängig von der Implemetierung und vom Grundkörper schlägt er den Standardalgorithmus ab einer Problemgröße zwischen 100 und 1000.

Nach Satz 138 lassen sich die meisten Probleme der linearen Algebra mit einer Komplexität von $\mathrm{O}(n^3)$ lösen. Nachdem die Multiplikation von Matrizen schneller möglich ist, kann man sich fragen, ob sich auch die anderen Probleme mit weniger als $\mathrm{O}(n^3)$ Operationen lösen lassen. Tatsächlich ist das der Fall. Wenn $\omega > 2$ so ist, dass die Multiplikation zweier $n \times n$-Matrizen mit $\mathrm{O}(n^\omega)$ Operationen möglich ist, dann kann man $\det(A)$, A^{-1} (wenn A invertierbar ist), eine Treppenform von A, sowie
06-20 Basen von $\ker A$, $\operatorname{im} A$, $\operatorname{coker} A$ und $\operatorname{coim} A$ mit $\mathrm{O}(n^\omega)$ Operationen berechnen. Jede Verbesserung der Komplexität von Matrix-Multiplikation führt sofort auch zu einer Verbesserung der Komplexität für die anderen genannten Probleme. Diese Zusammenhänge der Komplexität erhält man, indem man die Berechnung so umorganisiert, dass der wesentliche Teil der Rechnung in einer Matrix-Multiplikation steckt.

Wir illustrieren dieses Prinzip hier exemplarisch für die Berechnung von A^{-1} für eine invertierbare Matrix $A \in \mathbb{K}^{n \times n}$ mit $n = 2^k$ für ein $k \in \mathbb{N}$. Teile A in Blöcke auf und wende den üblichen Algorithmus zum Invertieren von Matrizen auf die Blöcke an:

$$\left(\begin{array}{cc|cc} A_{1,1} & A_{1,2} & I_{n/2} & 0 \\ A_{2,1} & A_{2,2} & 0 & I_{n/2} \end{array}\right) \begin{array}{l} \mid \cdot P := A_{1,1}^{-1} \\ \longleftarrow \end{array} \begin{array}{l} \urcorner^{-A_{2,1}} \\ {}_{+} \end{array}$$

$$\leftrightarrow \left(\begin{array}{cc|cc} I_{n/2} & PA_{1,2} & P & 0 \\ 0 & A_{2,2}-A_{2,1}PA_{1,2} & I_{n/2}-A_{2,1}P & I_{n/2} \end{array}\right) \begin{array}{l} \leftarrow \\ \mid \cdot \underbrace{(A_{2,2}-A_{2,1}PA_{1,2})^{-1}}_{=:Q} \end{array} \begin{array}{l} + \\ -PA_{1,2} \end{array}$$

$$\leftrightarrow \left(\begin{array}{cc|cc} I_{n/2} & 0 & P-PA_{1,2}Q(I_{n/2}-A_{2,1}P) & -PA_{1,2}Q \\ 0 & I_{n/2} & Q(I_{n/2}-A_{2,1}P) & Q \end{array}\right)$$

Die Invertierung der $n \times n$-Matrix A lässt sich also zurückführen auf die Invertierung zweier Matrizen halber Größe plus einer konstanten Anzahl von Multiplikationen und Additionen von $(n/2) \times (n/2)$-Matrizen. Dabei ist strenggenommen noch zu berücksichtigen, dass $A_{1,1}$ möglicherweise nicht invertierbar ist, selbst wenn A invertierbar ist. Wenn man diesen Fall der Einfachheit halber ignoriert, erhält man folgenden Algorithmus.

Algorithmus 18. Eingabe: $A \in \mathbb{K}^{n\times n}$ invertierbar mit $n = 2^k$ für ein $k \in \mathbb{N}$.
Ausgabe: A^{-1} oder eine Fehlermeldung

1 wenn $n = 1$, dann:
2 wenn $A = 0$ ist, gib eine Fehlermeldung aus und brich die gesamte Rechnung ab.
3 anderenfalls berechne A^{-1} direkt, gib das Ergebnis zurück und stop.
4 schreibe $A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$ für $A_{i,j} \in \mathbb{K}^{n/2\times n/2}$
5 berechne $P = A_{1,1}^{-1}$ mit diesem Algorithmus
6 berechne $C_{2,2} = (A_{2,2} - A_{2,1}PA_{1,2})^{-1}$ mit diesem Algorithmus
7 berechne $C_{2,1} = C_{2,2}(I_{n/2} - A_{2,1}P)$
8 berechne $C_{1,2} = -PA_{1,2}C_{2,2}$ 06-09
9 berechne $C_{1,1} = P - PA_{1,2}C_{2,1}$
10 gib $\begin{pmatrix} C_{1,1} & C_{1,2} \\ C_{2,1} & C_{2,2} \end{pmatrix}$ als Ergebnis zurück.

Satz 150. Sei $\omega \geq 2$ so, dass je zwei $n\times n$-Matrizen mit $O(n^\omega)$ Operationen in $\mathbb{K}$ multipliziert werden können. Dann braucht Algorithmus 18 nicht mehr als $O(n^\omega)$ Operationen.

Beweis. Sei $c \in \mathbb{R}$ und $k_0 \in \mathbb{N}$ so, dass für alle $k \geq k_0$ die Multiplikation zweier $2^k \times 2^k$-Matrizen nicht mehr als $c2^{k\omega}$ Operationen braucht. Wir können o.B.d.A. c durch eine beliebige größere Konstante (unabhängig von n) ersetzen. Insbesondere können wir deshalb $c \geq 1$ annehmen.
Sei $T(n)$ die Zahl der Operationen, die der Algorithmus für Eingabegröße $n = 2^k$ braucht. Die Invertierung der gegebenen $n \times n$-Matrix wird zurückgeführt auf zwei Invertierungen von Matrizen halber Größe zu je $T(n/2)$ Operationen, acht Multiplikationen zu je $c(n/2)^\omega$ Operationen, drei Additionen und eine Multiplikation mit -1 zu je $(n/2)^2$ Operationen. Wegen $c \geq 1$ und $\omega \geq 2$ ergibt sich $T(n) \leq 2T(n/2) + 12c(n/2)^\omega$.
Wir zeigen $T(2^k) \leq \frac{12c}{2^\omega-2}2^{k\omega}$ für alle $k \geq k_0$. Für $k = k_0$ können wir o.B.d.A. annehmen, dass $T(2^k) \leq \frac{12c}{2^\omega-2}2^{k_0\omega}$ gilt, indem wir c gegebenenfalls durch eine größere Konstante austauschen. Stimmt die Abschätzung für $k \in \mathbb{N}$, so gilt

$$T(2^{k+1}) \leq 2T(2^k) + 12c2^{k\omega} \leq \frac{24c}{2^\omega-2}2^{k\omega} + 12c2^{k\omega} = \frac{24c+(2^\omega-2)12c}{2^\omega-2}2^{k\omega} = \frac{12c}{2^\omega-2}2^{(k+1)\omega}$$

■

06-20 **Satz 151.** Ist $\omega > 2$ so, dass invertierbare $n \times n$-Matrizen mit nicht mehr als $O(n^\omega)$ Operationen invertiert werden können, dann genügen auch $O(n^\omega)$ Operationen, um je zwei $n \times n$-Matrizen miteinander zu multiplizieren.

06-20 **Beweis.** Seien $c > 0$ und $n_0 \in \mathbb{N}$ so, dass für alle $n \geq n_0$ die Invertierung von $n \times n$-Matrizen mit cn^ω oder weniger Operationen möglich ist. Wir zeigen, dass dann die Multiplikationen von $n \times n$-Matrizen für alle $n \geq n_0$ mit $3^\omega cn^\omega$ Operationen möglich ist. Dazu seien $A, B \in \mathbb{K}^{n\times n}$. Durch Nachrechnen und Satz 148 lässt sich bestätigen, dass

$$\begin{pmatrix} I_n & A & 0 \\ 0 & I_n & B \\ 0 & 0 & I_n \end{pmatrix}^{-1} = \begin{pmatrix} I_n & -A & AB \\ 0 & I_n & -B \\ 0 & 0 & I_n \end{pmatrix}$$

gilt. Die Berechnung dieser Inversen braucht nach Voraussetzung nicht mehr als $c(3n)^\omega$ Operationen, und das gewünschte Matrixprodukt lässt sich ohne weitere Rechnung aus dem Ergebnis ablesen. ∎

47 Numerische Algorithmen

Mit den Elementen eines Körpers $\mathbb{K}$ rechnen zu können bedeutet, dass es Algorithmen gibt, die für je zwei gegebene Elemente $a, b \in \mathbb{K}$ die Elemente $a + b$, $a - b$, $a \cdot b$ und $\frac{a}{b}$ (falls $b \neq 0$) berechnen, sowie einen Algorithmus, der für jedes gegebene $a \in \mathbb{K}$ entscheiden kann, ob $a = 0$ ist. Die Körperelemente müssen den Algorithmen in einem bestimmten vorher festgelegten Datenformat übergeben werden. Unter einem Datenformat kann man sich eine Abbildung vorstellen, die jedem Körperelement eine endlich lange Zeichenkette (Symbolfolge) zuordnet. Für ganze Zahlen kann man z.B. die gewöhnliche Darstellung als endliche Folge von Ziffern mit gegebenenfalls vorangestelltem Vorzeichen verwenden. Auch für rationale Zahlen und für Elemente von endlichen Körpern, oder auch für den Körper $\mathbb{Q}(\sqrt{2})$ ist es nicht schwer, sich geeignete Darstellungen der Körperelemente zu überlegen, mit denen man rechnen kann.

Für die Körper $\mathbb{R}$ und $\mathbb{C}$ dagegen ist das aus prinzipiellen Gründen nicht möglich. Da diese Körper überabzählbar viele Elemente haben und es nur abzählbar viele endlich lange Zeichenketten gibt, kann es kein Datenformat für reelle oder komplexe Zahlen geben. Man löst dieses Dilemma durch *Approximation.* Statt mit der Zahl $x \in \mathbb{R}$ selbst rechnet man z.B. mit einer Zahl $\bar{x} \in \mathbb{Q}$, für die $|x - \bar{x}|$ klein ist. Man nennt dann $\bar{x}$ eine *Näherung* für x. Die Zahl $|x - \bar{x}|$ nennt man den *absoluten Fehler* und die Zahl $|x - \bar{x}|/|x|$ (falls $x \neq 0$) den *relativen Fehler* der Näherung $\bar{x}$. Wenn $\bar{x}, \bar{y}$ zwei Näherungen an bestimmte Zahlen $x, y \in \mathbb{R}$ sind und wir wissen, dass $|x - \bar{x}| < \epsilon$ und $|y - \bar{y}| < \epsilon$ für ein gewisses $\epsilon > 0$ gilt, dann ist $\bar{x} + \bar{y}$ eine Näherung an $x + y$, deren absoluter Fehler höchstes 2ϵ beträgt. Der Fehler der Näherung kann während einer Rechnung also größer werden.

Algorithmen, die mit Näherungen arbeiten, nennt man *numerische Algorithmen.* Bei solchen Algorithmen spielt nicht nur die Komplexität eine Rolle, sondern darüber hinaus die Frage, um wieviel größer der Fehler im Ergebnis verglichen mit dem Fehler in der Eingabe ist. Dieses Fehlerwachstum möchte man, ebenso wie die Komplexität, möglichst klein halten. Der Minimierung des Fehlerwachstums sind dabei Grenzen gesetzt.

Beispiel. Betrachte ein lineares Gleichungssystems $Ax = b$, von dem wir annehmen, dass $A \in \mathbb{Q}^{n\times n} \subseteq \mathbb{R}^{n\times n}$ exakt gegeben ist, während wir von der rechten Seite $b \in \mathbb{R}^n \setminus \{0\}$ nur eine Näherung $\bar{b} \in \mathbb{Q}^n$ kennen. Wenn $\bar{x}$ die Lösung des Gleichungssytems $A\bar{x} = \bar{b}$ ist, wie nah wird dann $\bar{x}$ bei der Lösung x von $Ax = b$ liegen? Es gilt

$$b - \bar{b} = Ax - A\bar{x} = A(x - \bar{x}),$$

06-09 also $A^{-1}(b-\bar{b}) = x - \bar{x}$, also

$$\|x-\bar{x}\| = \|A^{-1}(b-\bar{b})\| \leq \|A^{-1}\|\|b-\bar{b}\|\frac{\|b\|}{\|b\|} = \|A^{-1}\|\|AA^{-1}b\|\frac{\|b-\bar{b}\|}{\|b\|} \leq \|A^{-1}\|\|A\|\|x\|\frac{\|b-\bar{b}\|}{\|b\|},$$

also $\frac{\|x-\bar{x}\|}{\|x\|} \leq \|A^{-1}\|\|A\|\frac{\|b-\bar{b}\|}{\|b\|}$.
Damit ist gezeigt, dass bei einem relativen Fehler von ϵ in der rechten Seite b mit einem $\|A^{-1}\|\|A\|$-mal so großen relativen Fehler in der Näherung $\bar{x}$ von x zu rechnen ist.

Definition 90. Sei $A \in \mathbb{R}^{n\times n}$ invertierbar, $\|\cdot\|$ die zur Standardnorm auf $\mathbb{R}^n$ gehörende Matrixnorm. Dann heißt $\kappa(A) := \|A^{-1}\|\|A\|$ die *Konditionszahl* (engl. *condition number*) der Matrix A.

Wegen $1 = \|I_n\| = \|A^{-1}A\| \leq \|A^{-1}\|\|A\| = \kappa(A)$ ist die Konditionszahl nie kleiner als 1. Ist sie nicht zu viel größer, spricht man von einer gut konditionierten Matrix, anderenfalls von einer schlecht konditionierten.

Satz 152. Sei $A \in \mathbb{R}^{n\times n}$ invertierbar.

1. Sind $\sigma_n \geq \sigma_{n-1} \geq \cdots \geq \sigma_1 > 0$ die Singulärwerte von A, so gilt $\kappa(A) = \frac{\sigma_n}{\sigma_1}$. 06-09
2. Ist A symmetrisch und positiv definit, und sind $\lambda_n \geq \lambda_{n-1} \geq \cdots \geq \lambda_1 > 0$ die Eigenwerte von A, so gilt außerdem $\kappa(A) = \frac{\lambda_n}{\lambda_1}$. 06-09

Beweis.

1. Zunächst gilt $\|A\| = \sigma_n$ nach Satz 112. Wenn (U, Σ, V) eine Singulärwertzerlegung von A ist, ist $(V^{-1}, \Sigma^{-1}, U^{-1})$ eine Singulärwertzerlegung von A^{-1}. Die Singulärwerte von A^{-1} sind deshalb $1/\sigma_n, \ldots, 1/\sigma_1$. Daraus folgt $\|A^{-1}\| = \max\{1/\sigma_n, \ldots, 1/\sigma_1\} = 1/\sigma_1$, wieder wegen Satz 112. Insgesamt folgt $\kappa(A) = \|A^{-1}\|\|A\| = \frac{1}{\sigma_1}\sigma_n$, wie behauptet.
2. Nach Satz 115 ist $\sigma \neq 0$ genau dann ein Singulärwert von A, wenn σ^2 ein Eigenwert von $AA^\top$ ist. Da A symmetrisch ist, ist $AA^\top = A^2$. Wenn λ ein Eigenwert von A ist, ist λ^2 ein Eigenwert von A^2, also $\lambda^2 = \sigma^2$ für einen Singulärwert σ von A. Da Singulärwerte stets positiv sind und die Eigenwerte von A, da A positiv-definit ist, ebenfalls positiv sind, folgt, dass die Eigenwerte von A mit den Singulärwerten von A übereinstimmen. Daraus folgt die Behauptung. ∎ 06-09

Dass man beim Rechnen mit Näherungen nie ein exaktes Ergebnis erwarten kann, eröffnet beim Entwurf von Algorithmen zusätzliche Freiheiten. Man kann nämlich auch Rechnungen erlauben, die selbst bei einer gedachten exakten Rechnung mit einem Fehler behaftet wären. Man unterscheidet dann zwischen dem Fehler, der durch die unvermeidliche Ungenauigkeit in den Daten verursacht wird (*Datenfehler*), und dem, der sich aus bewusst vorgenommenen Vereinfachungen während der Rechnung ergibt (*Verfahrensfehler*). Man wird bereit sein, einen Verfahrensfehler zu akzeptieren, der klein im Vergleich zum Datenfehler ist, zumal wenn man sich dadurch eine bessere Komplexität erkaufen kann.

Typischerweise gehen numerische Algorithmen iterativ vor. Beginnend mit einer schlechten Näherung einer Lösung von $Ax = b$ berechnen sie nach und nach immer bessere Näherungen, bis die gewünschte Genauigkeit erreicht ist.

Satz 153. (Richardson) Es bezeichne $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$ sowie die dazugehörige Matrixnorm. Sei $A \in \mathbb{R}^{n\times n}$ und $b \in \mathbb{R}^n$.

1. Wenn $\|I_n - A\| < 1$ ist, ist A invertierbar.
2. Es sei $x_0 \in \mathbb{R}^n$, und $x_1, x_2, \ldots \in \mathbb{R}^n$ seien definiert durch $x_{k+1} = x_k + (b - Ax_k)$ $(k \in \mathbb{N})$. Ist A invertierbar, so gilt für den eindeutig bestimmten Vektor $x \in \mathbb{R}^n$ mit $Ax = b$

$$\|x - x_k\| \leq \|I_n - A\|^k \|x - x_0\|$$

für alle $k \in \mathbb{N}$.

Beweis.

1. Wäre A nicht invertierbar, dann gäbe es ein $x \neq 0$ mit $Ax = 0$. Dann wäre

$$\|x\| = \|(I_n - A)x\| \leq \underbrace{\|I_n - A\|}_{<1} \underbrace{\|x\|}_{\neq 0} < \|x\|,$$

und das kann nicht sein.

2. Sei $x = A^{-1}b$. Wir zeigen die Ungleichung durch Induktion nach k. Für $k = 0$ ist die Ungleichung wegen $\|I_n - A\|^0 = 1$ erfüllt. Nehmen wir an, sie ist für ein $k \in \mathbb{N}$ erfüllt. Dann ist sie auch für $k + 1$ erfüllt, denn

$$\begin{aligned}
\|x - x_{k+1}\| &= \|x - (x_k + (b - Ax_k))\| \\
&= \|x - x_k - (Ax - Ax_k)\| \\
&= \|(I_n - A)(x - x_k)\| \\
&\leq \|I_n - A\| \|x - x_k\| \\
&\leq \|I_n - A\| \|I_n - A\|^k \|x - x_0\| \\
&= \|I_n - A\|^{k+1} \|x - x_0\|. \qquad \blacksquare
\end{aligned}$$

Beispiel.

1. $A \approx \begin{pmatrix} 0.456146 & 0.0814547 & 0.0531599 \\ 0.0805973 & 0.47501 & 0.067736 \\ 0.0600193 & 0.0257225 & 0.529027 \end{pmatrix}$, $b \approx \begin{pmatrix} 2.55882 \\ 0.555556 \\ 0.0875 \end{pmatrix}$, $x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

Wegen $\|I_3 - A\| \approx 0.617327$ erwarten wir, dass das Verfahren konvergiert. Für die Folge der Näherungen findet man:

$$\begin{aligned}
x_1 &\approx (2.55882, 0.555556, 0.0875), \\
x_2 &\approx (3.90054, 0.635056, -0.0391589), \\
x_3 &\approx (4.6305, 0.577233, -0.181386), \\
&\vdots \\
x_{11} &\approx (5.5986, 0.295709, -0.479811),
\end{aligned}$$

$$
\begin{aligned}
x_{12} &\approx (5.60506, 0.292069, -0.482109),\\
x_{13} &\approx (5.60899, 0.289792, -0.483485),\\
&\vdots\\
x_{21} &\approx (5.61504, 0.286135, -0.485525),\\
x_{22} &\approx (5.61509, 0.286104, -0.48554),\\
x_{23} &\approx (5.61511, 0.286085, -0.485549),\\
&\vdots\\
x_{31} &\approx (5.61516, 0.286056, -0.485563),\\
x_{32} &\approx (5.61516, 0.286055, -0.485563),\\
x_{33} &\approx (5.61516, 0.286055, -0.485563),\\
&\vdots
\end{aligned}
$$

Die Komplexität des Verfahrens ergibt sich aus der Anzahl der Iterationen multipliziert mit der Zahl der Operationen, die man für eine Matrix-Vektor-Multiplikation braucht. Das sind im allgemeinen $\mathrm{O}(n^2)$ Operationen, es können bei dünn besetzten oder strukturierten Matrizen aber auch deutlich weniger sein.

2. $A \approx \begin{pmatrix} 0.830391 & -0.821397 & 0.260858 \\ 0.214607 & 0.00167985 & 0.119583 \\ 0.335831 & -0.711475 & 0.628112 \end{pmatrix}$, $b \approx \begin{pmatrix} 2.55882 \\ 0.555556 \\ 0.0875 \end{pmatrix}$, $x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

In diesem Fall gilt $\|I_n - A\| \approx 1.49229 > 1$. Die Bedingung $\|I_n - A\| < 1$ ist für die Konvergenz des Verfahrens hinreichend, aber nicht notwendig. In der Tat scheint das Verfahren in diesem Beispiel trotzdem zu konvergieren:

$$
\begin{aligned}
x_1 &\approx (2.55882, 0.555556, 0.0875)\\
x_2 &\approx (3.42633, 0.550572, -0.344029)\\
x_3 &\approx (3.68194, 0.411028, -0.799391)\\
&\vdots\\
x_{31} &\approx (3.6171, -0.0444187, -1.84496)\\
x_{32} &\approx (3.6171, -0.0444190, -1.84496)\\
x_{33} &\approx (3.6171, -0.0444192, -1.84496)\\
&\vdots
\end{aligned}
$$

Der Grund ist, dass es eine andere Norm gibt, für die $\|I_n - A\| < 1$ gilt. Satz 153 gilt für beliebige Matrixnormen, und die Iterationsvorschrift hängt nicht von der Wahl der Norm ab. Das Verfahren konvergiert also schon, wenn $\|I_n - A\| < 1$ für irgendeine Norm gilt.

Im vorliegenden Beispiel ist dies der Fall für die Matrixnorm, die zur Vektorraumnorm

$$
\|x\| := \sqrt{x \begin{pmatrix} 4130 & 8034 & 3203 \\ 8034 & 19288 & 8136 \\ 3203 & 8136 & 3626 \end{pmatrix}^{-1} x}
$$

gehört. Für diese Norm gilt nämlich $\|I_n - A\| \approx 0.617327 < 1$. Man kann zeigen, dass es eine Norm mit $\|I_n - A\| < 1$ genau dann gibt, wenn für alle Eigenwerte $\lambda \in \mathbb{C}$ von $I_n - A$ gilt $|\lambda| < 1$.

3. $A \approx \begin{pmatrix} 1.82459 & 0.325819 & 0.21264 \\ 0.322389 & 1.90004 & 0.270944 \\ 0.240077 & 0.10289 & 2.11611 \end{pmatrix}$, $b \approx \begin{pmatrix} 0.758621 \\ 0.037037 \\ 0.863636 \end{pmatrix}$, $x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

   Da die Eigenwerte von $I_n - A$ in der Nähe von -1.43877, -0.860404, -0.541556 liegen, gilt $\|I_n - A\| > 1$ für jede Norm, und Satz 153 kann beim besten Willen keine Konvergenz garantieren. In solchen Fällen kann man versuchen, das gegebene Gleichungssystem $Ax = b$ in ein anderes System umzuwandeln, für das das Verfahren konvergiert, zum Beispiel indem man ausnutzt, dass $Ax = b \iff PAQ^{-1}Qx = Pb$ für beliebige invertierbare Matrizen P, Q gilt. Man nennt solche Matrizen P, Q, für die $\|I_n - PAQ^{-1}\| < 1$ ist, *Vorkonditionierer* (engl. *preconditioner*) für das System.

   Natürlich kommen als Vorkonditionierer nur Matrizen P und Q in Frage, für die man PAQ^{-1} in vernünftiger Zeit berechnen kann. Es ist z.B. wenig hilfreich $P = I_n$ und $Q = A$ zu wählen, denn diese Wahl transformiert das gegebene System nicht in ein einfacheres, sondern auf sich selbst. Unter Umständen kann es schon helfen, $P = \tau I_n$ und $Q = I_n$ für eine geeignete Konstante $\tau \in \mathbb{R}$ zu wählen. Im vorliegenden Beispiel konvergiert z.B. die Iterationsvorschrift $x_{k+1} = x_k + \frac{1}{5}(b - Ax_k)$ gegen eine Lösung:

$$\begin{aligned} x_1 &\approx (0.151724, 0.00740741, 0.172727) \\ x_2 &\approx (0.240253, -0.00714279, 0.264915) \\ x_3 &\approx (0.293504, -0.0268675, 0.314136) \\ &\vdots \\ x_{31} &\approx (0.390551, -0.0993408, 0.368647) \\ x_{32} &\approx (0.390552, -0.0993415, 0.368647) \\ x_{33} &\approx (0.390553, -0.0993421, 0.368646) \\ &\vdots \end{aligned}$$

4. Selbst wenn $\|I_n - A\| < 1$ gilt und das Iterationsverfahren also nach Satz 153 konvergiert, kann es sinnvoll sein, einen Vorkonditionierer einzusetzen, und zwar mit dem Ziel, die Konvergenz zu beschleunigen. Man möchte, dass $\|x - x_k\| \leq q^k \|x - x_0\|$ für ein möglichst kleines q gilt. Wählt man für die Matrix A aus dem ersten Beispiel oben zum Beispiel den Parameter $\tau = 2$, d.h. verwendet man die Iterationsvorschrift $x_{k+1} = x_k + 2(b - Ax_k)$, so erhält man eine bessere Konvergenz:

$$\begin{aligned} x_1 &\approx (5.11765, 1.11111, 0.175) \\ x_2 &\approx (5.36688, 0.318001, -0.506636) \\ x_3 &\approx (5.59042, 0.330527, -0.45618) \\ &\vdots \\ x_{11} &\approx (5.61516, 0.286055, -0.485563) \\ x_{12} &\approx (5.61516, 0.286055, -0.485563) \end{aligned}$$

$$x_{13} \approx (5.61516, 0.286055, -0.485563)$$
$$\vdots$$

Satz 154. Es sei $\|\cdot\|$ die Standardnorm des $\mathbb{R}^n$. Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch und positiv definit, $\lambda_{\min}$ der kleinste und $\lambda_{\max}$ der größte Eigenwert von A. Seien $b \in \mathbb{R}^n$, $\tau > 0$, $q = \max\{|1-\tau\lambda_{\min}|, |1-\tau\lambda_{\max}|\}$. Es sei $x_0 \in \mathbb{R}^n$ beliebig, und $x_1, x_2, \ldots \in \mathbb{R}^n$ seien definiert durch $x_{k+1} = x_k - \tau(b - Ax_k)$ für $k \in \mathbb{N}$. Für das eindeutig bestimmte $x \in \mathbb{R}^n$ mit $Ax = b$ gilt dann

$$\|x - x_k\| \le q^k \|x - x_0\|$$

für alle $k \in \mathbb{N}$.

Beweis. Es sei $\langle\cdot|\cdot\rangle$ das Standardskalarprodukt auf $\mathbb{R}^n$. Nach Satz 105 hat A eine ONB aus Eigenvektoren. Sei $\{e_1, \ldots, e_n\}$ eine solche Basis und seien $\lambda_1, \ldots, \lambda_n$ die zugehörigen Eigenwerte. Dann gilt $Ae_i = \lambda_i e_i$ für $i = 1, \ldots, n$ und $(I_n - \tau A)e_i = (1-\tau\lambda_i)e_i$ für $i = 1, \ldots, n$. Somit sind die Vektoren $e_1, \ldots, e_n$ auch Eigenvektoren von $I_n - \tau A$, und die zugehörigen Eigenwerte sind $1-\tau\lambda_1, \ldots, 1-\tau\lambda_n$. Da A symmetrisch ist, ist auch $I_n - \tau A$ symmetrisch. Deshalb folgt aus Satz 112 und Satz 115, dass

$$\|I - \lambda A\| = \max\{|1-\tau\lambda_1|, \ldots, |1-\tau\lambda_n|\} = \max\{|1-\tau\lambda_{\min}|, |1-\tau\lambda_{\max}|\}.$$

Damit folgt die Behauptung aus Satz 153 angewandt auf τA und τb anstelle von A und b. ∎

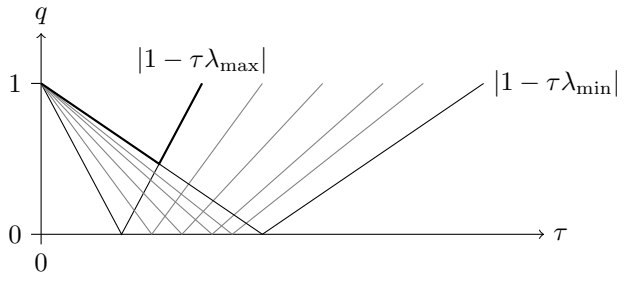
Um die Konvergenz des Iterationsverfahrens zu gewährleisten, wählt man τ zwischen 0 und $2/\lambda_{\max}$. Dann gilt 06-10

$$\left.\begin{array}{lll} -1 \le 1 - 2\dfrac{\lambda_{\min}}{\lambda_{\max}} < 1-\tau\lambda_{\min} < 1 & \Longrightarrow & |1-\tau\lambda_{\min}| < 1 \\ -1 = 1 - 2\dfrac{\lambda_{\max}}{\lambda_{\max}} < 1-\tau\lambda_{\max} < 1 & \Longrightarrow & |1-\tau\lambda_{\max}| < 1 \end{array}\right\} \Longrightarrow \quad q < 1.$$

Die beste Konvergenz erhält man, wenn q möglichst klein ist. Offenbar entspricht das kleinstmögliche q dem Wert τ, bei dem 06-14

$$|1-\tau\lambda_{\max}| = -(1-\tau\lambda_{\max}) = 1-\tau\lambda_{\min} = |1-\tau\lambda_{\min}|$$

gilt, also $\tau = \frac{2}{\lambda_{\min}+\lambda_{\max}}$.



Das zugehörige q ist dann

$$1 - \frac{2}{\lambda_{\min}+\lambda_{\max}}\lambda_{\min} = \frac{\lambda_{\min}+\lambda_{\max}-2\lambda_{\min}}{\lambda_{\min}+\lambda_{\max}} = \frac{\frac{\lambda_{\max}}{\lambda_{\min}}-1}{\frac{\lambda_{\max}}{\lambda_{\min}}+1} = \frac{\kappa(A)-1}{\kappa(A)+1}.$$

Wie man sieht, ist eine umso bessere Konvergenz zu erwarten, je näher die Konditionszahl bei 1 liegt.

48 Symbolische Algorithmen

In Abgrenzung von numerischen Algorithmen spricht man bei Algorithmen, die auf Körper spezialisiert sind, für die sich jedes Element mit endlich viel Information codieren läßt, von *symbolischen Algorithmen.* Zu diesen Körpern gehören alle endlichen Körper und $\mathbb{Q}$. Wenn $\mathbb{K}$ ein solcher Körper ist, dann auch der Körper $\mathbb{K}(X)$ der rationalen Funktionen über $\mathbb{K}$ (vgl. S. 37). Der Begriff „symbolisch“ bezieht sich auf das X von $\mathbb{K}(X)$: Anders als numerische Algorithmen können symbolische Algorithmen nämlich insbesondere mit Matrizen umgehen, in denen „Variablen“ vorkommen.

Wenn ein Körper unendlich groß ist und alle Elemente eine endlich lange Darstellung haben, dann können diese Darstellungen keine beschränkte Länge haben, d.h. für jedes $n \in \mathbb{N}$ muss es ein Körperelement geben, dessen Darstellung mehr als n Speicherzellen belegt. Es ist bei solchen Körpern deshalb nicht adäquat, für die Komplexität eines Algorithmus einfach die Anzahl der Körperoperationen anzugeben. Man muss darüber hinaus berücksichtigen, dass eine Körperoperation um so länger dauert, je länger die Codierung der Elemente ist, die addiert oder multipliziert werden sollen. Insbesondere ist zu beachten, dass die Körperelemente im Verlauf einer Rechnung immer länger werden können.

Beim Rechnen in endlichen Körpern besteht dieses Problem natürlich nicht, denn bei endlichen Körpern kann man für die Körperoperationen (zumindest prinzipiell) Wertetabellen abspeichern, so dass das Ergebnis jeder Operation mit dem gleichen Aufwand nachgeschlagen werden kann. Auch beim Rechnen mit Näherungen besteht das Problem nicht, weil man durch Rundung des Ergebnisses die Länge der Zahlen kurz halten kann:

$$\underbrace{6.91091}_{6\text{ Ziffern}} \cdot 10^{-3} \times \underbrace{8.23680}_{6\text{ Ziffern}} \cdot 10^{1} \approx \underbrace{5.61421}_{6\text{ Ziffern}} \cdot 10^{-1}$$

Vergleiche dazu eine exakte Rechnung in $\mathbb{Q}$:

$$\underbrace{\frac{23459769133495775}{3394597239563495799}}_{36\text{ Ziffern}} \times \underbrace{\frac{4589613856856600533}{56496729350456004}}_{36\text{ Ziffern}} = \underbrace{\frac{1076712814937489734012678290182480 75}{1917836414974238842155949129883271 96}}_{72\text{ Ziffern}}$$

Die Zahl rechts ist etwa so lang wie die beiden Zahlen links zusammen. Dasselbe Phänomen beobachtet man auch beim Rechnen mit rationalen Funktionen, z.B. in $\mathbb{Z}_5(X)$:

$$\underbrace{\frac{4X^5+2X^4+4X^3+X^2+X+2}{X^4+2X^3+X^2+X+3}}_{10\text{ Terme}} + \underbrace{\frac{2X^8+2X^7+4X^5+2X^3+3X}{X^8+3X^7+4X^5+3X^4+X^2+4}}_{11\text{ Terme}} = \underbrace{\frac{4X^{12}+3X^{11}+2X^8+4X^7+2X^5+4X^4+X^2+2X+4}{X^{11}+3X^{10}+X^9+X^8+4X^6+2X^4+4X^2+4X+1}}_{18\text{ Terme}}.$$

Die Behandlung von rationalen Zahlen (gemessen z.B. in der Gesamtzahl der Ziffern in Zähler und Nenner) und rationalen Funktionen über einem Körper (gemessen in der Summe der Polynomgrade von Zähler und Nenner) ist weitgehend analog. Der Fall der rationalen Funktionen ist dabei etwas einfacher. Ignoriert man in diesem Fall, dass auch die Elemente des Grundkörpers $\mathbb{K}$ möglicherweise anwachsen können, so kann man die Komplexität eines Problems über $\mathbb{K}(X)$ unter Berücksichtigung des Wachstums der Ausdrücke dadurch abschätzen, dass man statt der Zahl der Operationen in $\mathbb{K}(X)$ die Zahl der Operationen im Grundkörper $\mathbb{K}$ bestimmt.

Der folgende Satz illustriert, dass bei Rechnungen in der linearen Algebra ein gewisses Wachstum von Ausdrücken gar nicht zu vermeiden ist, weil schon die Ausdrücke im Ergebnis im allgemeinen größer sind als die in der Eingabe.

Satz 155.

1. Sei $A = ((a_{i,j}))_{i,j=1}^{n} \in \mathbb{K}[X]^{n\times n}$ mit $\deg(a_{i,j}) \leq d$ für alle $i, j \in \{1, \ldots, n\}$. Dann gilt $\det A \in \mathbb{K}[X]$ und $\deg(\det A) \leq nd$.

2. Sei $A = ((a_{i,j}))_{i,j=1}^n \in \mathbb{Z}^{n\times n}$ mit $|a_{i,j}| \le M$ für alle $i, j \in \{1, \dots, n\}$. Dann gilt $|\det(A)| \le n!M^n$.

   Ferner gilt die *Hadamardsche Ungleichung:*

$$|\det(A)| \le \prod_{i=1}^{n} \|(a_{i,1}, \dots, a_{i,n})\| \le n^{n/2} M^n$$

   wobei $\|\cdot\|$ die Standardnorm bezeichnet.

Beweis.

1. Für beliebige Polynome $p, q \in \mathbb{K}[X]$ gilt $\deg(pq) = \deg(p) + \deg(q)$ und $\deg(p+q) \le \max\{\deg(p), \deg(q)\}$. Deshalb folgt die Abschätzung unmittelbar aus Def. 28.

2. Die erste Abschätzung folgt ebenfalls direkt aus der Definition, denn

$$|\det(A)| = |\sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i,\pi(i)}| \le \sum_{\pi\in S_n} \prod_{i=1}^{n} |a_{i,\pi(i)}| \le \sum_{\pi\in S_n} M^n \le n!M^n.$$

   Für die zweite Abschätzung schreibe $a_i = (a_{i,1}, \dots, a_{i,n})$ für $i = 1, \dots, n$. Die Vektoren $b_1, \dots, b_n$ seien definiert durch $b_1 = a_1$ und $b_i = a_i - \sum_{j<i} \frac{\langle a_i|b_i\rangle}{\langle b_i|b_i\rangle} b_i$, wobei mit $\langle\cdot|\cdot\rangle$ das Standardskalarprodukt gemeint ist. Dann ist $\{b_1, \dots, b_n\}$ eine Orthogonalbasis von $\mathbb{R}^n$, es gibt eine Matrix

$$M = \begin{pmatrix} 1 & * & \cdots & * \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{Q}^{n\times n}$$

   mit $(a_1, \dots, a_n)^\top = M(b_1, \dots, b_n)^\top$, und es gilt $\|b_i\| \le \|a_i\|$ für alle i. Aus all dem folgt 06-10

$$\begin{aligned} |\det A| &= \sqrt{\det A \det A^\top} \\ &= \sqrt{\det(b_1, \dots, b_n) \underbrace{\det M^\top}_{=1} \underbrace{\det M}_{=1} \det(b_1, \dots, b_n)^\top} \\ &= \sqrt{\det\big((b_1, \dots, b_n)(b_1, \dots, b_n)^\top\big)} \\ &= \sqrt{\det \operatorname{diag}(\langle b_1|b_1\rangle, \dots, \langle b_n|b_n\rangle)} \\ &= \sqrt{\prod_{i=1}^{n} \langle b_i|b_i\rangle} = \prod_{i=1}^{n} \|b_i\| \le \prod_{i=1}^{n} \|a_i\|. \end{aligned}$$

   Der rechte Teil der Ungleichung folgt aus $\|a_i\| = \sqrt{a_{i,1}^2 + \cdots + a_{i,n}^2} \le \sqrt{nM^2} = n^{1/2}M$.

■

Wir betrachten als nächstes das Problem, die Lösung eines linearen Gleichungssystems $Ax = b$ für gegebene $A \in \mathbb{K}[X]^{n\times n}$ und $b \in \mathbb{K}[X]^n$ zu bestimmen. Zur Vereinfachung wollen wir annehmen, dass A invertierbar ist, und dass es darüber hinaus im Verlauf des Gauß-Algorithmus nicht nötig ist, jemals zwei Zeilen zu vertauschen. Die zweite Annahme ist nicht unbedingt nötig, aber durch sie wird die Formulierung von Sätzen, Beweisen und Algorithmen erheblich vereinfacht. Zum Beispiel vereinfacht sich der Algorithmus zur Berechnung einer Treppenform wie folgt.

Algorithmus 19. Eingabe: eine Matrix $A \in \mathbb{K}[X]^{n\times m}$, für die die Teilmatrix bestehend aus den ersten n Spalten die oben genannten Eigenschaften hat
Ausgabe: eine Treppenform für A

1 für $k = 1, \dots, n$:
2 dividiere die k-te Zeile durch $A[k,k]$.
3 für $i = k+1, \dots, n$:
4 addiere das $(-A[i,k])$-fache der k-ten Zeile zur i-ten Zeile
5 gib A als Ergebnis zurück.

Die Annahmen über A gewährleisten, dass $A[k,k]$ in Schritt 2 nie Null ist.

Wir nehmen an, dass die Einträge von A und b Polynome sind, betrachten sie aber als Elemente von $\mathbb{K}(X)^{n\times n}$ bzw. $\mathbb{K}(X)^n$, d.h. Divisionen sind erlaubt und wir sprechen über das klassische Problem der linearen Algebra und nicht etwa über die allgemeinere Modul-Variante. Trotzdem kann es sinnvoll sein, während der Rechnung die Einführung von Nennern zu vermeiden. Zum einen erspart man sich dadurch das Kürzen gemeinsamer Faktoren in Zähler und Nenner, für das jedesmal die Berechnung eines größten gemeinsamen Teilers nötig ist. Zum anderen ist es bei einer Komplexitätsabschätzung schwer vorherzusehen, wie groß diese gemeinsamen Teiler sein werden, und das erschwert die Abschätzung der Größe der Ausdrücke. Betrachten wir deshalb folgende Variante des Gauß-Algorithmus, bei der auf Divisionen verzichtet wird:

Algorithmus 20. Eingabe: eine Matrix $A \in \mathbb{K}[X]^{n\times m}$, für die die Teilmatrix bestehend aus den ersten n Spalten die oben genannten Eigenschaften hat
Ausgabe: Eine Treppenform von A, wobei auf den Treppenstufen statt 1 auch andere von Null verschiedene Einträge erlaubt sind.

1 für $k = 1, \dots, n$:
2 für $i = k+1, \dots, n$:
3 für $j = k, \dots, m$:
4 $A[i,j] = A[k,k]A[i,j] - A[i,k]A[k,j]$
5 gib A als Ergebnis zurück.

Beispiel.

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \end{pmatrix} \begin{array}{l} -a_{2,1} \quad -a_{3,1} \\ \mid \cdot a_{1,1} \quad + \\ \mid \cdot a_{1,1} \quad + \end{array} \leftrightarrow \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ 0 & b_{2,2} & b_{2,3} & b_{2,4} \\ 0 & b_{3,2} & b_{3,3} & b_{3,4} \end{pmatrix} \begin{array}{l} -b_{3,2} \\ \mid \cdot b_{2,2} \quad + \end{array}$$

$$\leftrightarrow \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ 0 & b_{2,2} & b_{2,3} & b_{2,4} \\ 0 & 0 & c_{3,3} & c_{3,4} \end{pmatrix}$$

06-10 Bei jedem Aufruf von Schritt 4 wird ein Eintrag $A[i,j]$ durch einen Eintrag ersetzt, dessen Grad potentiell doppelt so hoch ist. Wenn die Matrix A zu Beginn Polynome vom Grad $\leq d$ enthält, dann enthält sie am Ende Polynome vom Grad $2^n d$. Das ist viel. Die Funktion $n \mapsto 2^n$ wächst schneller als jedes Polynom, d.h. es gibt kein $p \in \mathbb{N}$, so dass $2^n = \mathrm{O}(n^p)$ ist. Man spricht von einer *exponentiellen Komplexität.* Algorithmen mit so hoher Komplexität sind meistens nicht praktikabel (Beispiel: wenn

man pro Operation eine Nanosekunde kalkuliert, braucht man für $2^{75} = 37778931862957161709568 \approx 3.8 \cdot 10^{22}$ Operationen mehr als eine Million Jahre).

Bei der folgenden Variante bleiben die Einträge auch polynomiell, wie wir in Satz 156 zeigen werden, und die Grade wachsen so viel langsamer, dass sich insgesamt eine polynomielle Komplexität ergibt (Satz 157).

Algorithmus 21. (Bareiss)
Eingabe und Ausgabe wie vorher.

1 setze $A[0,0] = 1$
2 für $k = 1, \dots, n$:
3 für $i = k+1, \dots, n$:
4 für $j = k, \dots, m$:
5 $A[i,j] = (A[k,k]A[i,j] - A[i,k]A[k,j])/A[k-1,k-1]$
6 gib A als Ergebnis zurück.

Satz 156. (Sylvester) Sei $A = ((a_{i,j}))_{i,j=1}^{n,m} \in \mathbb{K}[X]^{n\times m}$ so, dass $\det((a_{i,j}))_{i,j=1}^{k} \neq 0$ für alle $k = 1, \dots, n$ gilt.
Sei $a_{i,j}^{(0)} = a_{i,j}$ für alle i, j, und sei $a_{i,j}^{(k)}$ für $k = 1, \dots, n$ der Eintrag der Matrix A in Algorithmus 21 am Ende der k-ten Ausführung der Schritte 3–5 in Position (i, j). Dann gilt:

$$a_{i,j}^{(k)} = \begin{vmatrix} a_{1,1} & \cdots & a_{1,k} & a_{1,j} \\ \vdots & & \vdots & \vdots \\ a_{k,1} & \cdots & a_{k,k} & a_{k,j} \\ a_{i,1} & \cdots & a_{i,k} & a_{i,j} \end{vmatrix} \in \mathbb{K}[X]$$

für alle $k = 0, \dots, n$ und alle $i, j > k$. 06-10

Beweis. Für $k = 0, \dots, n$ und alle $i, j > k$ sei $b_{i,j}^{(k)}$ die Determinante auf der rechten Seite. 06-10
Zu zeigen: $a_{i,j}^{(k)} = b_{i,j}^{(k)}$ für alle i, j, k.
Induktion nach k. Für $k = 0$ ist nichts zu zeigen und für $k = 1$ lässt sich die Aussage direkt aus dem Algorithmus ablesen. Wir zeigen: wenn $k \in \{2, \dots, n\}$ so ist, dass die Behauptung für $k-1$ und $k-2$ gilt, dann gilt sie auch für k. Dazu genügt es zu zeigen, dass für alle $i, j > k$ gilt $b_{i,j}^{(k)} = (b_{k,k}^{(k-1)} b_{i,j}^{(k-1)} - b_{i,k}^{(k-1)} b_{k,j}^{(k-1)})/b_{k-1,k-1}^{(k-2)}$. Schreibe dazu

$$M_{k;i,j} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,k} & a_{1,j} \\ \vdots & & \vdots & \vdots \\ a_{k,1} & \cdots & a_{k,k} & a_{k,j} \\ a_{i,1} & \cdots & a_{i,k} & a_{i,j} \end{pmatrix} = \begin{pmatrix} A_k & B_{k;j} \\ C_{k;i} & D_{k;i,j} \end{pmatrix}$$

mit

$$A_k = \begin{pmatrix} a_{1,1} & \cdots & a_{1,k-1} \\ \vdots & & \vdots \\ a_{k-1,1} & \cdots & a_{k-1,k-1} \end{pmatrix} \in \mathbb{K}[X]^{(k-1)\times(k-1)}, \quad B_{k;j} = \begin{pmatrix} a_{1,k} & a_{1,j} \\ \vdots & \vdots \\ a_{k-1,k} & a_{k-1,j} \end{pmatrix} \in \mathbb{K}[X]^{(k-1)\times 2}$$

$$C_{k;i} = \begin{pmatrix} a_{k,1} & \cdots & a_{k,k-1} \\ a_{i,1} & \cdots & a_{i,k-1} \end{pmatrix} \in \mathbb{K}[X]^{2\times(k-1)}, \qquad D_{k;i,j} = \begin{pmatrix} a_{k,k} & a_{k,j} \\ a_{i,k} & a_{i,j} \end{pmatrix} \in \mathbb{K}[X]^{2\times 2}.$$

Dann gilt $A_k = M_{k-2;k-1,k-1}$, und daher $\det(A_k) = b^{(k-2)}_{k-1,k-1}$ nach Induktionsvoraussetzung. Nach Voraussetzung des Satzes ist A_k invertierbar. Mit Hilfe von Satz 148 überzeugt man sich, dass gilt:

$$M_{k;i,j} = \begin{pmatrix} A_k & B_{k;j} \\ C_{k;i} & D_{k;i,j} \end{pmatrix} = \begin{pmatrix} A_k & 0 \\ C_{k;i} & I_2 \end{pmatrix} \begin{pmatrix} I_{k-1} & A_k^{-1}B_{k;j} \\ 0 & D_{k;i,j} - C_{k;i}A_k^{-1}B_{k;j} \end{pmatrix}.$$

Daraus folgt unter Verwendung von Satz 85

$$\begin{aligned} \det(M_{k;i,j}) &= \det(A_k)\det(D_{k;i,j} - C_{k;i}A_k^{-1}B_{k;j}) \\ &= \det\bigl(\det(A_k)(D_{k;i,j} - C_{k;i}A_k^{-1}B_{k;j})\bigr)/\det(A_k). \end{aligned}$$

Die Behauptung folgt also, wenn wir zeigen können, dass

$$\det(A_k)(D_{k;i,j} - C_{k;i}A_k^{-1}B_{k;j}) = \begin{pmatrix} a^{(k-1)}_{k,k} & a^{(k-1)}_{k,j} \\ a^{(k-1)}_{i,k} & a^{(k-1)}_{i,j} \end{pmatrix}$$

gilt. Wir zeigen das für den Eintrag links unten. Das Argument für die drei anderen Einträge geht analog.

06-16 Ist c die zweite Zeile von $C_{k;i}$ und b die erste Spalte von $B_{k;j}$, so ist also zu zeigen

$$\det(A_k)a_{i,k} - \det(A_k)cA_k^{-1}b = a^{(k-1)}_{i,k}.$$

Für $x = (x_1, \ldots, x_{k-1}) := A_k^{-1}b$ gilt nach Satz 34 (Cramers Regel), dass $x_\ell = \frac{\det(A^{(\ell)}_k)}{\det(A_k)}$ für $\ell = 1, \ldots, k-1$, wobei $A^{(\ell)}_k$ für die Matrix steht, die aus A_k entsteht, wenn man die ℓ-te Spalte durch b ersetzt. Durch $k - \ell - 1$ Vertauschungen benachbarter Spalten lässt sich $A^{(\ell)}_k$ in eine Matrix $A^{[\ell]}_k$ überführen, bei der b ganz rechts steht. Wegen Satz 29 gilt $\det(A^{(\ell)}_k) = (-1)^{k-\ell-1}\det(A^{[\ell]}_k)$. Unter Verwendung von Satz 33 (Laplace-Entwicklung) und
06-16 der Induktionshypothese erhält man dann

$$\begin{aligned} &\det(A_k)a_{i,k} - \det(A_k)cA_k^{-1}b \\ &= \det(A_k)a_{i,k} - \det(A_k)\sum_{\ell=1}^{k-1} a_{i,\ell}\frac{(-1)^{k-\ell-1}\det(A^{[\ell]}_k)}{\det A_k} \\ &= \sum_{\ell=1}^{k-1}(-1)^{k+\ell}a_{i,\ell}\det(A^{[\ell]}_k) + (-1)^{k+k}a_{i,k}\det(A_k) \\ &= \det(M_{k-1;i,k}) = a^{(k-1)}_{i,k}, \end{aligned}$$

wie behauptet. ∎

Satz 157. Sei $A = ((a_{i,j}))^{n,m}_{i,j=1} \in \mathbb{K}[X]^{n\times m}$ so, dass $\det((a_{i,j}))^k_{i,j=1} \neq 0$ für alle $k = 1, \ldots, n$ gilt, und sei $d \in \mathbb{N}$ so, dass $\deg(a_{i,j}) \leq d$ für alle i, j. Dann genügen $O(d^2n^4m)$ Operationen in $\mathbb{K}$, um eine Treppenform von A zu berechnen.

Beweis. Wir orientieren uns an Algorithmus 21. Beim k-ten Durchlauf der Schleife aus Schritt 2 werden in Schritt 5 nach Satz 156 und Satz 155 zweimal zwei Polynome vom Grad höchstens kd multipliziert, deren Differenz gebildet und durch ein Polynom vom Grad höchstens $(k-1)d$ geteilt. Die beiden Multiplikationen kosten höchstens $2(kd+1)^2$ Operationen in $\mathbb{K}$, die Subtraktion höchstens $2kd+1$ weitere, und die Division noch einmal höchstens $2(kd-(k-1)d+1)((k-1)d+2) = 2(d+1)(kd-d+2)$. Die Anzahl der zu verarbeitenden Einträge im k-ten Durchlauf ist $(n-k)(m-k)$. Für die Gesamtkomplexität ergibt sich also

$$\sum_{k=1}^{n} \underbrace{(n-k)(m-k)}_{0\leq\ \cdot\ \leq nm} \underbrace{\big(2(kd+1)^2+2kd+1+2(d+1)(kd-d+2)\big)}_{0\leq\ \cdot\ \leq 4(kd+2)^2}$$
$$\leq 4nm\sum_{k=1}^{n}(kd+2)^2 = \frac{2}{3}mn^2(2d^2n^2+3d^2n+d^2+12dn+12d+24) = \mathrm{O}(d^2n^4m).$$

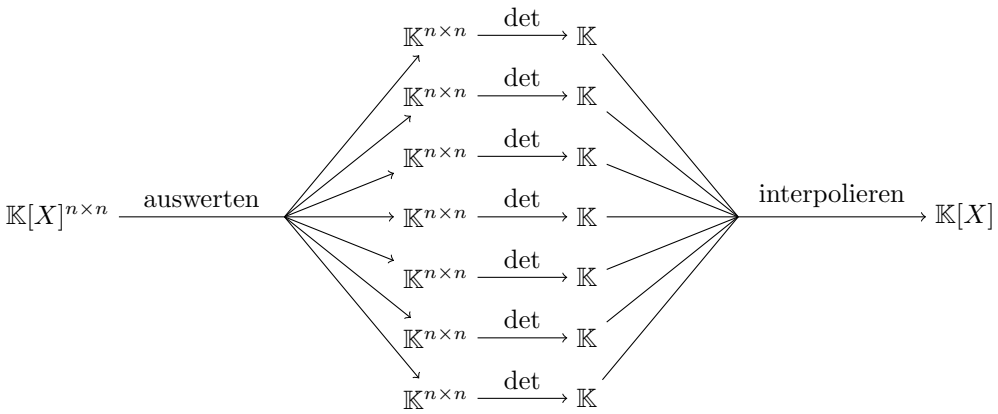
Von der verwendeten Summationsformel überzeugt man sich durch Induktion (oder mit Hilfe von Computeralgebra). ∎

Eine völlig andere Technik zur Eindämmung des Wachstums von Ausdrücken ist das Rechnen mit homomorphen Bildern. Die Idee ist dabei, eine Rechnung über $\mathbb{K}(X)$ auf eine oder mehrere Rechnungen in $\mathbb{K}$ zurückzuführen und aus den Ergebnissen der Rechnungen über dem kleineren Körper das Ergebnis der Rechnung über dem großen Körper zu rekonstruieren. Betrachten wir als Beispiel das Problem, zu einer gegebenen Matrix $A \in \mathbb{K}[X]^{n\times n}$ die Determinante zu berechnen. Per Definition entsteht die Determinante von A durch Additionen und Multiplikationen der Matrixeinträge. Da für Polynome $p, q \in \mathbb{K}[X]$ und Körperelemente $u \in \mathbb{K}$ gilt $(p+q)(u) = p(u)+q(u)$ und $(pq)(u) = p(u)q(u)$, folgt $\det(A(u)) = \det(A)(u)$ für alle $u \in \mathbb{K}$, wobei $A(u) \in \mathbb{K}^{n\times n}$ für die Matrix steht, die man erhält, wenn man jeden Matrixeintrag $p \in \mathbb{K}[X]$ von A durch $p(u) \in \mathbb{K}$ erstetzt, und $\det(A)(u) \in \mathbb{K}$ für die Auswertung des Polynoms $\det(A)$ an $u \in \mathbb{K}$. Es spielt also für das Ergebnis keine Rolle, ob man zuerst die polynomiellen Einträge von A bei u auswertet und dann die Determinante berechnet, oder zuerst die Determinante als Polynom berechnet und danach erst u einsetzt. Für die Komplexität macht das aber durchaus einen Unterschied, denn bei der Berechnung von $\det(A(u))$ ist kein X mehr im Spiel, so dass es keine anwachsenden Ausdrücke mehr gibt. Für eine Matrix $A \in \mathbb{K}[X]^{n\times n}$ mit Einträgen vom Grad höchstens d kann man deshalb wie folgt vorgehen:

1 wähle $nd+1$ paarweise verschiedene Körperelemente $u_0, \ldots, u_{nd} \in \mathbb{K}$ 06-19

2 berechne $\det(A(u_i))$ für $i = 0, \ldots, nd$

3 rekonstruiere $\det(A)$ aus diesen Daten und gib das Ergebnis zurück.



Die Berechnung jeder Matrix $A(u_i)$ kostet n^2 mal $O(d)$ Operationen, also insgesamt $O(d^2n^3)$. Die Berechnung jeder Determinante $\det(A(u_i))$ kostet dann nur $O(n^3)$ Operationen, also ingesamt $O(dn^4)$. Die Berechnung des Interpolationspolynoms kostet nur $O(d^2n^2)$ Operationen. Insgesamt braucht der Algorithmus also $O(d^2n^3 + dn^4)$ Operationen. Man hat zusätzlich den Vorteil, dass die Berechnung der $\det(A(u_i))$ für verschiedene u_i unabhängig voneinander parallel auf verschiedenen Rechnern durchgeführt werden kann.

Wenn man ein Gleichungssystem $Ax = b$ für gegebene $A \in \mathbb{K}[X]^{n\times n}$ und $b \in \mathbb{K}^n$ mit $\det(A) \neq 0$ zu lösen hat, werden die Koeffizienten des Lösungsvektores x im allgemeinen nicht polynomiell sein,
06-16 sondern rationale Funktionen mit nichttrivialen Nennern. Wir können deshalb diese Koeffizienten nicht einfach durch Interpolation berechnen. Aus Satz 34 wissen wir allerdings, dass $\det(A)x$ ein polynomieller Vektor ist. Man kann deshalb wie folgt vorgehen.

06-16 **Algorithmus 22.** Eingabe: $A \in \mathbb{K}[X]^{n\times n}$ mit $\det(A) \neq 0$, und $b \in \mathbb{K}[X]^n$. Der Maximalgrad unter den Einträgen von A und b sei $d \in \mathbb{N}$.
Ausgabe: $x \in \mathbb{K}(X)^n$ mit $Ax = b$.

1 berechne $\det(A) \in \mathbb{K}[X]$ wie oben beschrieben
2 wähle $u_0, \ldots, u_{dn} \in \mathbb{K}$ paarweise verschieden, so dass $\det(A)(u_i) \neq 0$ für alle i.
3 für $i = 0, \ldots, dn$:
4 berechne die Lösung $x^{(i)} = (x_1^{(i)}, \ldots, x_n^{(i)}) \in \mathbb{K}^n$ von $A(u_i)\, x^{(i)} = b(u_i)$
5 multipliziere alle Koeffizienten von $x^{(i)}$ mit $\det(A)(u_i)$
6 für $j = 1, \ldots, n$:
7 berechne $x_j \in \mathbb{K}[X]$ mit $x_j(u_i) = x_j^{(i)}$ für $i = 0, \ldots, dn$ und $\deg(x_j) \leq dn$.
8 gib $(\frac{x_1}{\det(A)}, \ldots, \frac{x_n}{\det(A)})$ als Ergebnis zurück.

In Schritt 2 werden Nullstellen von $\det(A)$ als Evaluationspunkte ausgeschlossen, weil für $u \in \mathbb{K}$ mit $\det(A)(u) = 0$ das System $A(u)x = b(u)$ mehrere Lösungen hat. Nur eine dieser Lösungen ist die Auswertung des Lösungsvektors von $Ax = b$ bei u, und es ist nicht klar, wie man diesen von den anderen Elementen der Lösungsmenge unterscheiden soll.

Beispiel. Sei $A = \begin{pmatrix} 3X+5 & 2X-7 \\ 8X+3 & 4X+1 \end{pmatrix} \in \mathbb{Q}[X]^{2\times 2}$ und $b = \begin{pmatrix} X-1 \\ 2X+1 \end{pmatrix}$.

Gesucht ist $x = (x_1, x_2) \in \mathbb{Q}(X)^2$, so dass $Ax = b$.

Zunächst berechnen wir die Determinante von A:

$$\left.\begin{aligned} \det A(0) &= \begin{vmatrix} 5 & -7 \\ 3 & 1 \end{vmatrix} = 26 \\ \det A(1) &= \begin{vmatrix} 8 & -5 \\ 11 & 5 \end{vmatrix} = 95 \\ \det A(2) &= \begin{vmatrix} 11 & -3 \\ 19 & 9 \end{vmatrix} = 156 \end{aligned}\right\} \quad \Longrightarrow \quad \det(A) = 26 + 73X - 4X^2.$$

Als nächstes berechnen wir Lösungen für spezielle Werte anstelle von X:

$$\begin{pmatrix} 5 & -7 \\ 3 & 1 \end{pmatrix} x(0) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad \Longrightarrow \quad x(0) = (\tfrac{3}{13}, \tfrac{4}{13})$$
$$\begin{pmatrix} 8 & -5 \\ 11 & 5 \end{pmatrix} x(1) = \begin{pmatrix} 0 \\ 3 \end{pmatrix} \quad \Longrightarrow \quad x(1) = (\tfrac{3}{19}, \tfrac{24}{95})$$

$$\begin{pmatrix} 11 & -3 \\ 19 & 9 \end{pmatrix} x(2) = \begin{pmatrix} 1 \\ 5 \end{pmatrix} \qquad \Longrightarrow \quad x(2) = (\tfrac{2}{13}, \tfrac{3}{13})$$

Multiplikation mit den Werten der Determinante bei $0, 1, 2$ ergibt

$$(\det(A)x)(0) = \begin{pmatrix} 6 \\ 8 \end{pmatrix}, \quad (\det(A)x)(1) = \begin{pmatrix} 15 \\ 24 \end{pmatrix}, \quad (\det(A)x)(2) = \begin{pmatrix} 24 \\ 36 \end{pmatrix}.$$

Interpolation liefert

$$\det(A)x = \begin{pmatrix} 6 + 9X \\ 8 + 18X - 2X^2 \end{pmatrix}$$

und schließlich

$$x = \begin{pmatrix} (6 + 9X)/(26 + 73X - 4X^2) \\ (8 + 18X - 2X^2)/(26 + 73X - 4X^2) \end{pmatrix}.$$

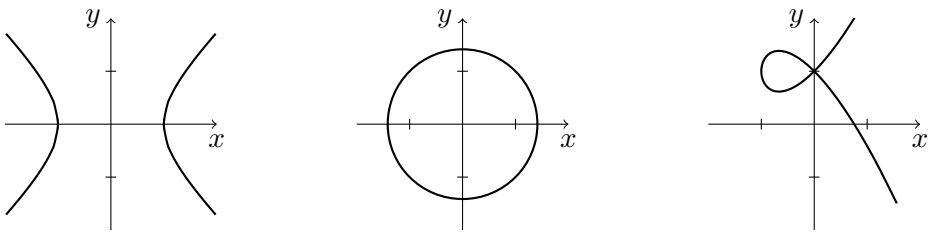
Teil IX

Anwendungen

49 Eliminationstheorie

Jedes Polynom $p \in \mathbb{R}[X]$ hat nur endlich viele Nullstellen. Die Zahl der Nullstellen ist durch den Grad von p beschränkt. Polynome in mehreren Variablen können aber durchaus unendlich viele Nullstellen haben. Zum Beispiel besteht die Nullstellenmenge des Polynoms $X^2 + Y^2 - 1 \in \mathbb{R}[X, Y]$ aus allen Punkten auf dem Einheitskreis. Das sind unendlich viele. Allgemein kann man sich die Nullstellenmenge eines Polynoms $p \in \mathbb{R}[X, Y]$ als eine Kurve im $\mathbb{R}^2$ vorstellen, die Nullstellenmenge eines Polynoms in $\mathbb{R}[X, Y, Z]$ als eine (i.a. gekrümmte) Fläche im $\mathbb{R}^3$, und so weiter.

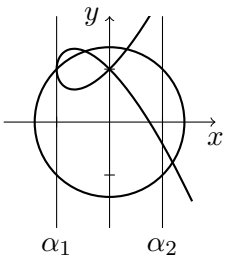
Beispiel. Links: die Kurve bestehend aus allen Punkten $(x, y) \in \mathbb{R}^2$ mit $(x-y)(x+y)-1 = 0$. Mitte: die Kurve bestehend aus allen Punkten $(x, y) \in \mathbb{R}^2$ mit $x^2 + y^2 - 2 = 0$. Rechts: die Kurve bestehend aus allen Punkten $(x, y) \in \mathbb{R}^2$ mit $(y - 1)^2 - x^2(x + 1) = 0$.



Es seien nun zwei Polynome $p, q \in \mathbb{R}[X, Y]$ gegeben. Wir wollen die Schnittpunkte der zugehörigen Kurven bestimmen, also die Punkte $(x, y) \in \mathbb{R}^2$, für die sowohl $p(x, y) = 0$ als auch $q(x, y) = 0$ gilt. Für zwei Polynome $p, q \in \mathbb{R}[X]$ wissen wir aus Abschnitt 25, dass die gemeinsamen Nullstellen von p, q genau die Nullstellen des größten gemeinsamen Teilers $\gcd(p, q) \in \mathbb{R}[X]$ sind. Das gilt für multivariate Polynome im allgemeinen so nicht.

Man kann die bivariaten Polynome $p, q \in \mathbb{R}[X, Y]$ auf univariate Polynome zurückführen, indem man für eine der beiden Variablen einen Wert α einsetzt, z. B. für X. Die Nullstellen des größten gemeinsamen Teilers der beiden resultierenden Polynome in $\mathbb{R}[Y]$ sind dann genau die Zahlen $\beta \in \mathbb{R}$, so dass $(\alpha, \beta) \in \mathbb{R}^2$ ein Schnittpunkt der beiden gegebenen Kurven ist. Mit anderen Worten: Wenn man eine Koordinate eines Schnittpunkts kennt, kann man die andere leicht berechnen.

Beispiel. Betrachte die beiden Kurven, die durch die Polynome $p = X^2 + Y^2 - 2$ und $q = (Y-1)^2 - X^2(X+1)$ definiert sind. Ersetzt man X in beiden Polynomen durch $\alpha_1 = -1$, so erhält man die beiden Polynome $Y^2 - 1$ und $(Y - 1)^2$, deren größter gemeinsame Teiler $Y - 1$ ist. Das Polynom $Y - 1$ hat nur die Nullstelle $\beta = 1$, und damit ist $(\alpha_1, \beta) = (-1, 1)$ der einzige Schnittpunkt der beiden Kurven, deren erste Koordinate -1 ist. Ersetzt man X in p und q durch $\alpha_2 = +1$, so erhält man $Y^2 - 1$ und $(Y - 1)^2 - 2$, und da diese beiden Polynome teilerfremd sind, gibt es keinen Schnittpunkt auf der vertikalen Geraden durch $(1, 0)$.



Es ist unwahrscheinlich, dass man eine Koordinate der Schnittpunkte durch Raten findet. Für fast jede Wahl von α wird man zwei teilerfremde univariate Polynome bekommen. Es stellt sich also die Frage, wie man zu gegebenen bivariaten Polynomen $p, q \in \mathbb{R}[X, Y]$ die Zahlen $\alpha \in \mathbb{R}$ berechnen kann, für die bei Ersetzung von X durch α zwei univariate Polynome in $\mathbb{R}[Y]$ mit einem nicht-trivialen größten gemeinsamen Teiler entstehen.

Zu diesem Zweck erinnere man sich zunächst daran, dass nach Satz 72 der größte gemeinsame Teiler g zweier Polynome $a, b \in \mathbb{R}[Y]$ als $\mathbb{R}[Y]$-Linearkombination geschrieben werden kann: $g = sa + tb$ für gewisse $s, t \in \mathbb{R}[Y]$. Aus dem Beweis von Satz 72 folgt auch, dass man immer geeignete s, t mit $\deg(s) < \deg(b) - \deg(g) \leq \deg(b)$ und $\deg(t) < \deg(a) - \deg(g) \leq \deg(a)$ finden kann.

Mit diesem Wissen kann man g, s, t statt mit dem erweiterten euklidischen Algorithmus auch direkt mit linearer Algebra berechnen. Denn $g = \gcd(a, b)$ ist dann einfach das Polynom von kleinstmöglichem Grad, das eine $\mathbb{R}$-Linearkombination der Polynome $a, Ya, \ldots, Y^{\deg(b)-1}a, b, Yb, \ldots, Y^{\deg(a)-1}b$ ist. Schreibt man deren Koeffizientenvektoren in die Zeilen einer Matrix, so muss der Koeffizientenvektor von g in der letzten von Null verschiedenen Zeile jeder Treppenform auftauchen.

Beispiel. Betrachte die Polynome $a = 2Y^5 + 3Y^4 + 3Y^3 + 4Y^2 - 5Y + 1$, $b = 2Y^4 + Y^3 - 2Y^2 + 4Y - 1$. Wir erwarten, dass sich $g = \gcd(a, b)$ als $\mathbb{R}$-Linearkombination von $a, Ya, Y^2a, Y^3a, b, Yb, Y^2b, Y^3b, Y^4b$ schreiben lässt. Wir schreiben die Koeffizienten dieser Polynome in die Zeilen einer Matrix und berechnen eine Treppenform. Beachte, dass die Spalten nach absteigenden Potenzen von Y angeordnet werden müssen. Aus

$$\begin{array}{r} \\ Y^3a \\ Y^2a \\ Ya \\ a \\ Y^4b \\ Y^3b \\ Y^2b \\ Yb \\ b \end{array}\!\!\begin{array}{c} \begin{array}{ccccccccc} Y^8 & Y^7 & Y^6 & Y^5 & Y^4 & Y^3 & Y^2 & Y & 1 \end{array} \\ \begin{pmatrix} 2 & 3 & 3 & 4 & -5 & 1 & 0 & 0 & 0 \\ 0 & 2 & 3 & 3 & 4 & -5 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & 3 & 4 & -5 & 1 & 0 \\ 0 & 0 & 0 & 2 & 3 & 3 & 4 & -5 & 1 \\ 2 & 1 & -2 & 4 & -1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & -2 & 4 & -1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & -2 & 4 & -1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & -2 & 4 & -1 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & -2 & 4 & -1 \end{pmatrix} \end{array} \leftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & * & * \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \frac{3}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

folgt $g = Y^2 + \frac{3}{2}Y - \frac{1}{2}$.

Die Matrix in dieser Rechnung hat $\deg(a)+\deg(b)$ Zeilen und Spalten. Dass a und b einen nicht-trivialen größten gemeinsamen Teiler haben, bedeutet, dass der Rang der Matrix kleiner als $\deg(a) + \deg(b)$ ist, und dies bedeutet wiederum, dass die Determinate der Matrix Null ist. Das motiviert folgende Definition.

Definition 91. Seien $a = a_0 + \cdots + a_n X^n \in \mathbb{K}[X]$, $b = b_0 + \cdots + b_m X^m \in \mathbb{K}[X]$ mit $a_n, b_m \neq 0$. Dann heißt

$$\mathrm{Syl}(a, b) = \begin{pmatrix} a_n & \cdots & \cdots & a_0 & 0 & \cdots & 0 \\ 0 & \ddots & & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & \ddots & 0 \\ 0 & \cdots & 0 & a_n & \cdots & \cdots & a_0 \\ b_m & \cdots & \cdots & b_0 & 0 & \cdots & 0 \\ 0 & \ddots & & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & \ddots & 0 \\ 0 & \cdots & 0 & b_m & \cdots & \cdots & b_0 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m \text{ Zeilen} \\ \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \text{ Zeilen} \end{matrix} \in \mathbb{K}^{(n+m)\times(n+m)}$$

die *Sylvester-Matrix* von a und b, und

$$\operatorname{res}(a,b) = \det \operatorname{Syl}(a,b)$$

heißt die *Resultante* (engl. *resultant*) von a und b.

Satz 158. Für zwei Polynome $a, b \in \mathbb{R}[X]$ gilt $\gcd(a,b) = 1 \iff \operatorname{res}(a,b) \neq 0$.

Beweis. Sei $n = \deg(a)$, $m = \deg(b)$.

„⇒“ Wir zeigen: wenn $\gcd(a,b) = 1$ ist, dann hat $\operatorname{Syl}(a,b)$ vollen Rang. Betrachte den Untervektorraum $V = \langle a, Xa, \ldots, X^{m-1}a, b, Xb, \ldots, X^{n-1}b\rangle \subseteq \mathbb{K}[X]$ von $\mathbb{K}[X]$. Nach Annahme $\gcd(a,b) = 1$ folgt aus Satz 72 (und den Gradschranken, die der Beweis von Satz 72 ebenfalls liefert), dass gilt $1 \in V$. Wir zeigen $X^i \in V$ für alle $i = 0, \ldots, n+m-1$. Dazu seien $s, t \in \mathbb{K}[X]$ mit $\deg(s) < m$, $\deg(t) < n$ und $1 = sa + tb$. Für $i \in \{0, \ldots, n+m-1\}$ gilt dann

$$\begin{aligned} X^i &= X^i sa + X^i tb \\ &= \big(\operatorname{quo}(X^i s, b)b + \operatorname{rem}(X^i s, b)\big)a + \big(\operatorname{quo}(X^i t, a)a + \operatorname{rem}(X^i t, a)\big)b \\ &= \operatorname{rem}(X^i s, b)a + \operatorname{rem}(X^i t, a)b + \operatorname{quo}(X^i sa, ba)ba + \operatorname{quo}(X^i tb, ab)ab \\ &= \underbrace{\operatorname{rem}(X^i s, b)a + \operatorname{rem}(X^i t, a)b}_{\in V} + \underbrace{\operatorname{quo}(X^i(1-tb) + X^i tb, ab)}_{=0,\ \text{da } i < n+m = \deg(ab)}\, ab \end{aligned}$$

„⇐“ Wenn $\operatorname{res}(a,b) \neq 0$ ist, ist $\operatorname{Syl}(a,b)$ invertierbar. Dann ist die Treppennormalform von $\operatorname{Syl}(a,b)$ die Einheitsmatrix I_{n+m}. Darum lässt sich jedes Monom X^i $(i = 0, \ldots, n+m-1)$ als Linearkombination von $a, Xa, \ldots, X^{m-1}a, b, Xb, \ldots, X^{n-1}b$ darstellen, insbesondere das Monom $1 = X^0$. Damit ist gezeigt, dass es Polynome $s, t \in \mathbb{K}[X]$ gibt mit $1 = sa + tb$. Jeder gemeinsame Teiler $d \in \mathbb{K}[X]$ von a und b ist dann auch ein Teiler von 1. Daraus folgt $\gcd(a,b) = 1$. ∎

Mit Hilfe der Resultante können wir nicht nur feststellen, *ob* zwei Polynome einen gemeinsamen Teiler haben, sondern auch *wann* dies der Fall ist. Wenn nämlich $p, q \in \mathbb{R}[X,Y]$ bivariate Polynome sind, die wir als Polynome in Y mit Koeffizienten in $\mathbb{R}[X]$ auffassen, dann ist die Resultante von p, q (bezüglich Y) ein Polynom in X. Die Nullstellen dieses Polynoms sind genau die X-Werte, für die p und q als Polynome in der verbleibenden Variablen Y einen gemeinsamen Teiler haben. Bei Resultanten von Polynomen in mehreren Variablen verwendet man statt $\operatorname{res}(p,q)$ auch die Notation $\operatorname{res}_X(p,q)$, die im Index die Variable angibt, bezüglich der die Resultante gebildet wird.

Wenn man $p = q = 0$ als (i.a. nicht-lineares) Gleichungssystem in zwei Variablen X, Y auffasst, dann kann man mit der Resultante eine Variable aus dem Gleichungssytem eliminieren. Das Ergebnis ist eine neue Gleichung, die nur die andere Variable enthält. Für jede Lösung (α, β) des Gleichungssystems $p = q = 0$ ist α eine Nullstelle der Resultante bzgl. Y und β eine Nullstelle bzgl. X. Die Umkehrung gilt allerdings nicht: im allgemeinen hat die Resultante auch Nullstellen, die nicht der Koordinate einer Lösung des ursprünglichen Gleichungssystems entsprechen. Modernere Eliminationsverfahren, von denen Sie in weiterführenden Lehrveranstaltungen hören werden (insbesondere die sogenannten *Gröbnerbasen*) haben dieses Problem nicht.

Beispiel.

1. Für $p = X^2 + Y^2 - 2, q = (Y-1)^2 - X^2(X+1) \in \mathbb{R}[X,Y] = \mathbb{R}[X][Y]$

$$\operatorname{res}(p,q) = (1+X)(1 - X - 7X^2 + X^3 + 3X^4 + X^5).$$

Dieses Polynom hat vier reelle Nullstellen und zwei komplexe. Die Nullstelle -1 führt zu dem Schnittpunkt $(-1, 1)$, den wir bereits kennen. Die drei weiteren Schnittpunkte sind $(-.459062, 1.33763), (.326477, 1.37601), (1.1909, -.762731)$.

2. Man kann mit der Resultante auch die Punkte finden, an denen eine Kurve sich selbst schneidet. Solche Punkte zeichnen sich dadurch aus, dass die univariaten Polynome, die man erhält, wenn man eine der Koordinaten einsetzt, bei der anderen Koordinate eine mehrfache Nullstelle haben. Die mehrfachen Nullstellen eines Polynoms $p \in \mathbb{R}[Y]$ sind genau die gemeinsamen Nullstellen von p und der Ableitung p'. Die möglichen X-Koordinaten der Selbstüberschneidungspunkte sind also Nullstellen von $\operatorname{res}_Y(p, \frac{d}{dY}p) \in \mathbb{R}[X]$. Man nennt $\operatorname{disc}_Y(p) := \operatorname{res}_Y(p, \frac{d}{dY}p)$ die *Diskriminante* des Polynoms p.

   Beispiel: Für $p = Y^2 - X^2(X+1)$ ist $\operatorname{disc}_Y(p) = 4X^2(X+1)$. Die Nullstellen sind -1 und 0. Für diese Werte von X wird p zu Y^2, die doppelte Nullstelle ist also 0. Demnach finden wir die beiden speziellen Punkte $(-1, 0)$ und $(0, 0)$. Letzterer ist ein Punkt, an dem die Kurve sich selbst schneidet. Der Punkt $(-1, 0)$ ist dagegen ein Punkt, an dem die Kurve eine vertikale Tangente hat. Auch diese sogenannten *Verzweigungspunkte* (engl. *branch point*) findet man mit Hilfe der Diskriminante.

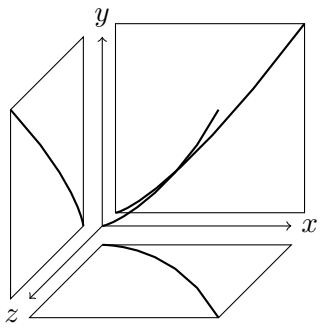
3. Zur Erinnerung: eine Zahl $\alpha \in \mathbb{C}$ heißt *algebraisch,* wenn es ein Polynom $p \in \mathbb{Q}[X] \setminus \{0\}$ gibt, das α als Nullstelle hat. Mit Hilfe der Resultante lässt sich einsehen, dass die Summe und das Produkt zweier algebraischer Zahlen wieder algebraische Zahlen sind. Sind nämlich $\alpha, \beta \in \mathbb{C}$ algebraisch und sind $p, q \in \mathbb{Q}[X] \setminus \{0\}$ so, dass $p(\alpha) = q(\beta) = 0$, dann ist $\alpha+\beta$ eine Nullstelle von $\operatorname{res}_Y(p(Y), q(X-Y)) \in \mathbb{Q}[X]$ und $\alpha\beta$ ist eine Nullstelle
06-22 von $\operatorname{res}_Y(p(Y), Y^{\deg q} q(X/Y)) \in \mathbb{Q}[X]$.

   Beispiel: $\sqrt{2}$ ist eine Nullstelle von $p = X^2 - 2$ und $\sqrt{3}$ ist eine Nullstelle von $q = X^2 - 3$. Die Summe $\sqrt{2} + \sqrt{3}$ ist eine Nullstelle des Polynoms $\operatorname{res}_Y(Y^2 - 2, (X-Y)^2 - 3) = X^4 - 10X^2 + 1$. Die anderen drei Nullstellen dieses Polynoms sind $\sqrt{2} - \sqrt{3}$, $-\sqrt{2} + \sqrt{3}$ und $-\sqrt{2} - \sqrt{3}$. Das Produkt $\sqrt{2}\sqrt{3}$ ist eine Nullstelle von $\operatorname{res}_Y(Y^2 - 2, Y^2((X/Y)^2 - 3)) = (X^2 - 6)^2$.

Um eine Kurve im dreidimensionalen Raum zu beschreiben, ist ein einziges Polynom nicht genug, denn die Nullstellenmenge eines Polynoms $p \in \mathbb{R}[X, Y, Z]$ in drei Variablen ist im allgemeinen eine Fläche. Man kann Kurven im $\mathbb{R}^3$ aber unter Umständen beschreiben als die Schnittmenge von zwei oder mehr solcher Flächen. Zu einer gegebenen Kurve $C \subseteq \mathbb{R}^3$ betrachtet man die Menge I aller Polynome $p \in \mathbb{R}[X, Y, Z]$, so dass $p(x, y, z) = 0$ für alle $(x, y, z) \in C$ gilt. Diese Menge I ist ein Ideal des Rings $\mathbb{R}[X, Y, Z]$, denn mit $p(x, y, z) = 0$ gilt auch $(rp)(x, y, z) = 0$ für jedes $r \in \mathbb{R}[X, Y, Z]$, und mit $p(x, y, z) = q(x, y, z) = 0$ gilt auch $(p + q)(x, y, z) = 0$. Wir können die Kurve C also durch ein Erzeugendensystem des Ideals I kodieren.

Einem Erzeugendensystem für Ideal $I \subseteq \mathbb{R}[X, Y, Z]$, das eine Kurve $C \subseteq \mathbb{R}^3$ beschreibt, sieht man nicht direkt an, wo im Raum die Kurve C sich befindet. Um sich einen Überblick zu verschaffen, kann man sich fragen, welche ebenen Kurven man erhält, wenn man C orthogonal auf eine der drei Koordinatenflächen projiziert. Diese Kurven werden durch die Polynome beschrieben, die in I liegen und nur die Variablen enthalten, die zur jeweiligen Koordinatenfläche gehören. Die Polynome $I \cap \mathbb{R}[X, Y]$ beschreiben zum Beispiel die Projektion von C auf die x/y-Ebene.

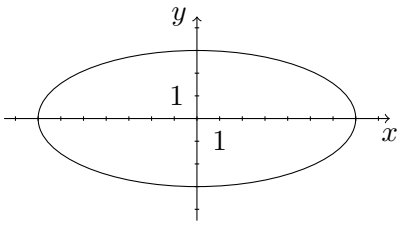
Beispiel. Das Ideal $I = \langle Y^2 - XZ, X^2Y - Z^2, X^3 - YZ \rangle \subseteq \mathbb{R}[X, Y, Z]$ beschreibt die unten dargestellte Kurve im $\mathbb{R}^3$. Gleichungen für die Kurven auf den Projektionsflächen bekommt man durch Berechnung von Resultanten, z.B. ist $\operatorname{res}_Z(Y^2 - XZ, X^3 - YZ) = Y^3 - X^4$.



50 Hauptachsentransformation

Beispiel.

1. Die Nullstellenmenge des bivariaten Polynoms $(X/7)^2 + (Y/3)^2 - 1 \in \mathbb{R}[X,Y]$ ist eine Ellipse. Die größte Ausdehnung hat diese Ellipse mit 14 Längeneinheiten entlang der horizontalen Achse. Die geringste Ausdehnung hat sie mit 6 Längeneinheiten entlang der vertikalen Achse. Man nennt diese beiden Achsen mit der größten bzw. niedrigsten Ausdehnung die *Hauptachsen* der Ellipse.



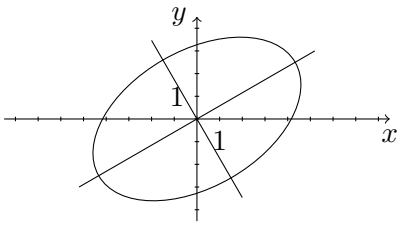
2. Auch die Nullstellenmenge des Polynoms $p = \frac{13}{225}X^2 - \frac{8}{75\sqrt{3}}XY + \frac{7}{75}Y^2 - 1 \in \mathbb{R}[X,Y]$ ist eine Ellipse. Wir wollen ihre Hauptachsen bestimmen. Im vorherigen Beispiel waren die Hauptachsen leicht aus dem Polynom abzulesen, weil es keinen gemischten Term XY gab. Die Idee ist deshalb, ein Koordinatensystem zu suchen, bezüglich dessen die jetzige Ellipse eine Gleichung der Form $(X/a)^2 + (Y/b)^2 - 1$ hat. Dabei soll der Koordinatenwechsel die Ellipse aber nicht verzerren. Wir brauchen also eine Isometrie. Dazu schreiben wir das Polynom p in der Form 06-22

$$p = \begin{pmatrix} X \\ Y \end{pmatrix} \underbrace{\begin{pmatrix} \frac{13}{225} & -\frac{4}{75\sqrt{3}} \\ -\frac{4}{75\sqrt{3}} & \frac{7}{75} \end{pmatrix}}_{=A} \begin{pmatrix} X \\ Y \end{pmatrix} - 1.$$

Die Matrix A ist symmetrisch und hat deshalb nach Satz 105 eine ONB aus Eigenvektoren. Wir haben die Zerlegung

$$A = \begin{pmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2}\sqrt{3} \end{pmatrix} \begin{pmatrix} \frac{1}{25} & 0 \\ 0 & \frac{1}{9} \end{pmatrix} \begin{pmatrix} \frac{1}{2}\sqrt{3} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2}\sqrt{3} \end{pmatrix}.$$

Daraus folgt, dass die beiden Hauptachsen der Ellipse den Richtungen $\begin{pmatrix}\sqrt{3}\\1\end{pmatrix}$ und $\begin{pmatrix}-1\\\sqrt{3}\end{pmatrix}$ entsprechen. Die größte Ausdehnung der Ellipse ist 10 (entlang der ersten Hauptachse), und die kleinste ist 6 Einheiten (entlang der zweiten Hauptachse).

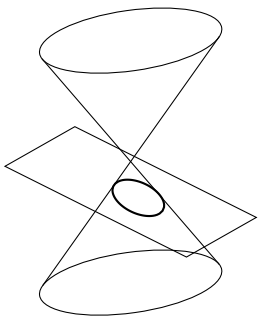


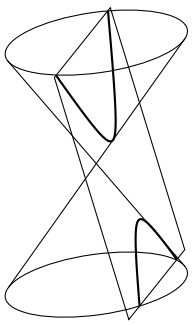
3. Bei der Nullstellenmenge des Polynoms $p = X^2 - XY - 5X + 2Y^2 + 6Y + 7 \in \mathbb{R}[X, Y]$ handelt es sich um eine Ellipse, deren Zentrum nicht im Ursprung liegt. Um das Zentrum zu bestimmen, ersetzt man X und Y durch $X + u$ und $Y + v$, wobei $u, v \in \mathbb{R}$ so gewählt sind, dass ein Polynom der Form $aX^2+bXY+cY^2+d$ entsteht. Im vorliegenden Beispiel ergibt sich

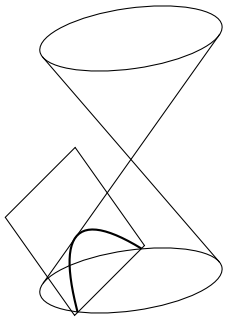
$$X^2 - XY + 2Y^2 + (2u - v - 5)X + (6 - u + 4v)Y + u^2 - uv - 5u + 2v^2 + 6v + 7 = 0.$$

Die geeignete Wahl für u, v erhält man durch Lösen des linearen Gleichungssystems $2u - v - 5 = 6 - u + 4v = 0$. Demnach liegt das Zentrum im Punkt $(u, v) = (2, -1)$. Durch Multiplikation mit einer geeigneten Konstanten bringt man die Gleichung als nächstes auf die Form $aX^2 + bXY + cY^2 - 1$. Danach kann man wie im vorigen Beispiel verfahren, um die beiden Hauptachsen zu bestimmen.

Allgemein bezeichnet man die Nullstellenmenge eines Polynoms $p = aX^2+bXY+cY^2+dX+eY+f \in \mathbb{R}[X, Y]$ mit $a, b, c, d, e, f \in \mathbb{R}$ als *Kegelschnitt* (engl. *conic*). Kegelschnitte sind genau die Kurven, die entstehen, wenn man einen (Doppel-)Kegel mit einer Ebene schneidet. Neben Ellipsen kommen als Kegelschnitte noch einige andere Arten von Kurven vor, insbesondere *Hyperbeln* (engl. *hyperbola*) und *Parabeln* (engl. *parabola*). Eine Hyperbel liegt vor, wenn sich die Gleichung durch eine geeigente Verschiebung und Rotation auf die Form $aX^2 - bY^2 - 1 = 0$ für $a, b > 0$ bringen lässt. Die Gleichung für eine Parabel hat nach geeigenter Verschiebung und Rotation die Form $aX^2 - 2Y = 0$. Die nötigen Rechnungen sind jeweils ähnlich wie oben für den Fall der Ellipse beschrieben.


Ellipse


Hyperbel


Parabel

Neben Ellipsen, Parabeln und Hyperbeln gibt noch einige degenerierte Fälle, z.B. entspricht $X^2 - Y^2$ zwei gekeuzten Geraden, $X^2 - 1$ zwei parallelen Geraden, $X^2 + Y^2$ einem einzelnen Punkt, und $X^2 + Y^2 + 1$ der leeren Menge.

Kegelschnitte werden seit dem Altertum untersucht. Nach Kepler treten sie zum Beispiel in der Himmelsmechanik als Bahnen von Himmelskörpern auf.

Die Nullstellenmenge eines Polynoms $p \in \mathbb{R}[X_1, \ldots, X_n]$ in n Variablen, das nur Terme $X_i^u X_j^v$ mit $u, v \in \{0, 1, 2\}$ und $u + v \leq 2$ enthält, heißt eine *Quadrik*. Im Fall $n = 2$ sind die Quadriken genau die Kegelschnitte. Durch eine geeignete Verschiebung und Drehung des Koordinatensystems lässt sich die definierende Gleichung jeder Quadrik auf eine der drei folgenden Formen bringen: 06-22

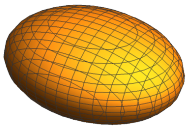
$$\left(\frac{X_1}{\alpha_1}\right)^2 + \cdots + \left(\frac{X_i}{\alpha_i}\right)^2 - \left(\frac{X_{i+1}}{\alpha_{i+1}}\right)^2 - \cdots - \left(\frac{X_j}{\alpha_j}\right)^2 = 1 \qquad \text{für gewisse } i, j \text{ mit } 1 \leq i \leq j \leq n,$$

$$\left(\frac{X_1}{\alpha_1}\right)^2 + \cdots + \left(\frac{X_i}{\alpha_i}\right)^2 - \left(\frac{X_{i+1}}{\alpha_{i+1}}\right)^2 - \cdots - \left(\frac{X_j}{\alpha_j}\right)^2 = 0 \qquad \text{für gewisse } i, j \text{ mit } 1 \leq i \leq j \leq n,$$

$$\left(\frac{X_1}{\alpha_1}\right)^2 + \cdots + \left(\frac{X_i}{\alpha_i}\right)^2 - \left(\frac{X_{i+1}}{\alpha_{i+1}}\right)^2 - \cdots - \left(\frac{X_j}{\alpha_j}\right)^2 = 2X_n \qquad \text{für gewisse } i, j \text{ mit } 1 \leq i \leq j < n.$$

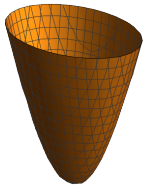
Dabei sind $\alpha_1, \ldots, \alpha_j$ jeweils positive reelle Zahlen.

Im Fall $n = 3$ sind Quadriken Flächen, von denen man ähnlich wie bei den Kegelschnitten verschiedene Typen unterscheiden kann. In der folgenden Übersicht sind degenerierte Fälle (z.B. Ebenen, Geraden, isolierte Punkte) nicht berücksichtigt.



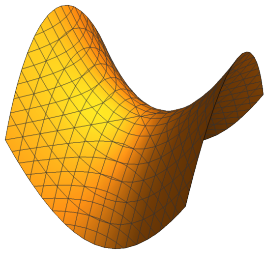
$\left(\frac{X}{\alpha}\right)^2 + \left(\frac{Y}{\beta}\right)^2 + \left(\frac{Z}{\gamma}\right)^2 = 1$

Ellipsoid



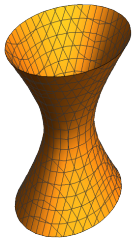
$\left(\frac{X}{\alpha}\right)^2 + \left(\frac{Y}{\beta}\right)^2 = 2Z$

elliptisches Paraboloid



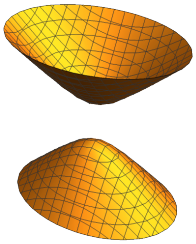
$\left(\frac{X}{\alpha}\right)^2 - \left(\frac{Y}{\beta}\right)^2 = 2Z$

hyperbolisches Paraboloid



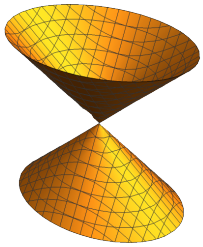
$\left(\frac{X}{\alpha}\right)^2 + \left(\frac{Y}{\beta}\right)^2 - \left(\frac{Z}{\gamma}\right)^2 = 1$

einschaliges Hyperboloid



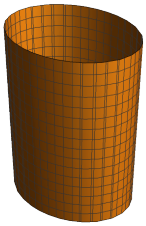
$-\left(\frac{X}{\alpha}\right)^2 - \left(\frac{Y}{\beta}\right)^2 + \left(\frac{Z}{\gamma}\right)^2 = 1$

zweischaliges Hyperboloid



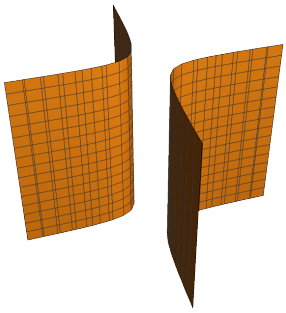
$\left(\frac{X}{\alpha}\right)^2 + \left(\frac{Y}{\beta}\right)^2 - \left(\frac{Z}{\gamma}\right)^2 = 0$

elliptischer Kegel



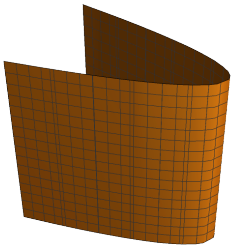
$\left(\frac{X}{\alpha}\right)^2 + \left(\frac{Y}{\beta}\right)^2 = 1$

elliptischer Zylinder



$\left(\frac{X}{\alpha}\right)^2 - \left(\frac{Y}{\beta}\right)^2 = 1$

hyperbolischer Zylinder



$\left(\frac{X}{\alpha}\right)^2 = 2Z$

parabolischer Zylinder

Mit der Hauptachsentransformation kann man für eine gegebene Gleichung feststellen, welchen Typ die zugehörige Fläche hat.

Beispiel. Sei $p = 3X^2 - 2XY + 2XZ + 8X + 5Y^2 - 2YZ - 20Y + 3Z^2 + 19 \in \mathbb{R}[X, Y, Z]$. Welchen Typ hat die Fläche $F = \{(x, y, z) \in \mathbb{R}^2 : p(x, y, z) = 0\} \subseteq \mathbb{R}^3$?

Erster Schritt: Bestimmung des Zentrums. Die Koeffizienten von X, Y, Z im Polynom $p(X + u, Y + v, Z + w)$ lauten $8 + 6u - 2v + 2w$, $-20 - 2u + 10v - 2w$ und $2u - 2v + 6w$. Sie werden genau dann Null, wenn $(u, v, w) = (-1, 2, 1)$ ist. Dieser Punkt ist das Zentrum.

Zeiter Schritt: Verschiebung und Normierung. Es gilt $p(u, v, w) = -5$. Wir setzen deshalb $q = \frac{1}{5}p(X + u, Y + v, Z + w) = \frac{3}{5}x^2 - \frac{2}{5}xy + \frac{2}{5}xz + y^2 - \frac{2}{5}yz + \frac{3}{5}z^2 - 1$. Damit sind die linearen Terme eliminiert und der konstante Term ist auf -1 normiert.

Dritter Schritt: Bestimmung der Hauptachsen. Es gilt $q = (X, Y, Z)A(X, Y, Z) - 1$ für die symmetrische Matrix

$$A = \frac{1}{5}\begin{pmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{pmatrix}.$$

Die Eigenwerte von A lauten $6/5, 3/5, 2/5$. Die zugehörigen Eigenräume sind $E_{6/5} = \langle \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \rangle$, $E_{3/5} = \langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \rangle$, $E_{2/5} = \langle \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \rangle$. Durch Normierung der Eigenvektoren und Multiplikation mit den Quadratwurzeln der Eigenwerte ergeben sich die Zeilen der orthogonalen Matrix B mit $A = B^\top B$. Die Matrix lautet

$$B = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & -2 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}.$$

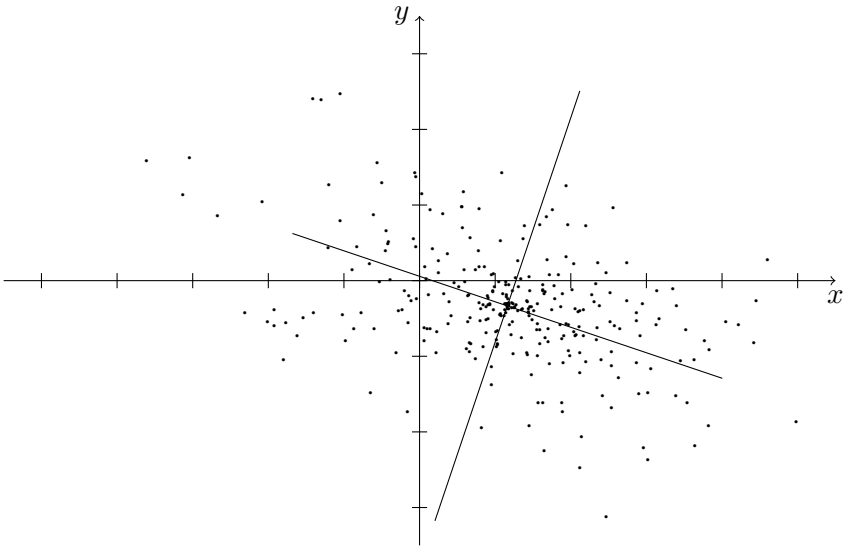
Daraus folgt die Darstellung

$$q = \left(\frac{x - 2y + z}{\sqrt{5}}\right)^2 + \left(\frac{x + y + z}{\sqrt{5}}\right)^2 + \left(\frac{x - z}{\sqrt{5}}\right)^2 - 1$$

Es handelt sich bei F also um ein Ellipsoid mit Zentrum $(-1, 2, 1)$ und den Hauptachsenrichtungen $a = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$, $b = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $c = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Die Ausdehnung entlang der Hauptachse a beträgt $2\sqrt{6/5}$ Längeneinheiten, entlang der Hauptachse b sind es $2\sqrt{3/5}$ und entlang von c sind es $2\sqrt{2/5}$ Einheiten.

Die Hauptachsentransformation wird auch eingesetzt, um statistische Zusammenhänge in großen Datenmengen aufzuspüren. Dabei hat man eine endliche Menge von Datenpunkten $W \subseteq \mathbb{R}^n$ gegeben und will wissen, in welcher Richtung die Daten am stärksten schwanken.

Beispiel. Gegeben sei die unten dargestellte Punktwolke $W = \{(x_1, y_1), \ldots, (x_{300}, y_{300})\} \subseteq \mathbb{R}^2$.



Das Zentrum der Punktwolke ergibt sich aus den Mittelwerten der Koordinaten: $(\mu_1, \mu_2) = \frac{1}{300} \sum_{i=1}^{300} (x_i, y_i)$. Man nennt diesen Punkt auch den *Schwerpunkt* des Datensatzes. Im obigen Beispiel liegt der Schwerpunkt bei $(1.16127, -0.333823)$.

Die Schwankung bezeichnet die durchschnittliche Abweichung vom Durchschnittswert. Bei einem eindimensionalen Datensatz $\{x_1, x_2, \ldots, x_n\}$ mit Mittelwert μ betrachtet man hierfür die *Varianz* $V = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2$ oder die *Streuung* $\sigma = \sqrt{V}$. Bei einem mehrdimensionalen Datensatz ist die *Kovarianz* $\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_1)(y_i - \mu_2)$ ein Maß dafür, wie stark die Schwankungen in die beiden Richtungen miteinander gekoppelt sind. Die Idee ist jetzt, das Koordinatensystem so drehen, dass die Kovarianz möglichst klein wird. Dazu wendet man den Spektralsatz auf die Matrix

$$A = \begin{pmatrix} \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_1)^2 & \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_1)(y_i - \mu_2) \\ \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_1)(y_i - \mu_2) & \frac{1}{n} \sum_{i=1}^{n} (y_i - \mu_2)^2 \end{pmatrix}$$

an. Der Eigenvektor zum größten Eigenwert ist dann die Richtung, bezüglich der die Varianz maximal ist. Die Vektoren der ONB der Beispielpunktwolke sind im Bild oben eingezeichnet.

Besonders nützlich ist die Hauptachsentransformation, wenn man die Lage einer Punktwolke in einem hochdimensionalen Raum veranschaulichen will. Zum Beispiel könnte es sein, dass man einer Gruppe

von 1000 Personen Blutproben entnommen hat und für jede Probe die Werte von zehn medizinisch interessanten Parametern gemessen wurden. Diese Daten bilden eine Punktwolke im $\mathbb{R}^{10}$. Um sich ein Bild von den Daten zu machen, kann man sich die Projektion der Punktwolke auf eine Ebene anschauen. Bei jeder solchen Projektion gehen Informationen verloren. Die Hauptachsentransformation dreht die Punktwolke in eine Position, bei der die Achsen nach absteigender Varianz angeordnet sind. Die erste Achse ist die Richtung der stärksten Schwankung, die zweite Achse ist die Richtung der stärksten Schwankung im orthogonalen Komplement der ersten Achse, die dritte Achse ist die Richtung der stärksten Schwankung innerhalb orthogonalen Komplements der ersten beiden Achsen, und so weiter. Es bietet sich an, die Daten auf die ersten beiden Hauptachsen zu projizieren und zu hoffen, dass die Information in den restlichen Dimensionen vernachlässigbar ist.

51 PageRank

Suchmaschinen durchsuchen das Internet und tabellieren für jede Webseite, welche Schlüsselbegriffe darin vorkommen. Wenn ein Benutzer nach einem Begriff sucht, gibt die Suchmaschine die Liste aller Webseiten aus, die den gesuchten Begriff enthält. Typischerweise sind das sehr viele. Dann kommt es darauf an, dass die Suchmaschine gut einschätzen kann, welche Suchresultate wohl die relevantesten sind, damit sie diese als erstes anzeigen kann. Der Bewertungsmechanismus der Relevanz von Seiten ist ein wesentliches Kriterium für den Erfolg der Suchmaschine.

In der Praxis wird eine Suchmaschine Webseiten nach mehreren ganz unterschiedlichen Kriterien bewerten und die verschiedenen Bewertungen auf eine bestimmte Weise zu einem Score verrechnen, nach dem die Suchergebnisse dann sortiert werden. Die Einzelheiten werden geheim gehalten, damit sie nicht von konkurrierenden Suchmaschinenanbietern imitiert werden, und damit Webseitenautoren es nicht so leicht haben, durch bestimmte Gestaltungstricks einen höheren Score zu erreichen.

Wir stellen hier den Bewertungsmechanismus vor, dessen Entdeckung zur Gründung von Google geführt hat, und der auch heute noch eine wesentliche Rolle im Bewertungsschema von Google spielt. Das Internet wird als ein Graph $G = (V, E)$ angesehen (vgl. Abschnitt 21), dessen Knoten die Webseiten sind, und bei dem es eine Kante von a nach b gibt, wenn die Webseite a einen Link zur Webseite b enthält. Die Relevanz einer Webseite w ergibt sich daraus, wie viele andere Webseiten einen Link auf w enthalten. Dabei sollen Links von relevanten Seiten stärker berücksichtigt werden als Links von weniger relevanten, und Links von Seiten, die viele Links enthalten, sollen weniger stark berücksichtigt werden als andere. Das motiviert die folgende Definition.

Definition 92. Sei $G = (V, E)$ ein Graph und $d \in (0, 1)$. Zu $v \in V$ sei $E_v = E \cap (\{v\} \times V)$ die Menge der von v weggehenden Kanten. Für $v, w \in V$ sei

$$p_{v,w} = \begin{cases} 1/|V| & \text{falls } |E_v| = 0 \\ 1/|E_v| & \text{falls } |E_v| \neq 0 \text{ und } (v,w) \in E \\ 0 & \text{sonst} \end{cases}$$

$$a_{v,w} = (1-d)\frac{1}{|V|} + d\, p_{v,w}$$

Eine Funktion $f\colon V \to \mathbb{R}$ heißt *PageRank* für G (zum Faktor d), falls gilt

$$\forall\ w \in V : f(w) = \sum_{v \in V} a_{v,w} f(v).$$

Die Zahlen $a_{v,w}$ lassen sich interpretieren als ein Votum der Seite v für die Seite w. Im Fall $d = 0$ verteilt jede Seite v ihr Votum gleichmäßig auf alle Seiten des Netzes. Im Fall $d = 1$ verteilt sie ihr Votum gleichmäßig auf alle Seiten, die von ihr aus verlinkt sind. Der allgemeine Fall d ist eine Überlagerung

dieser beiden Voten. Gemäß der Gleichung für f ergibt sich die Relevanz einer Seite w durch die Voten aller Seiten für w, wobei das Votum jeder Seite v für w mit der Relevanz von v gewichtet wird.

Aus Sicht der linearen Algebra besagt die Gleichung aus Def. 92, dass f ein Eigenvektor der Matrix $A^\top$ zum Eigenwert 1 ist, wobei $A = ((a_{v,w}))_{v,w \in V}$. Man kann zeigen, dass diese Matrix $A^\top$ den Eigenwert 1 haben muss, dass der zugehörige Eigenraum eindimensional ist, und dass er von einem Vektor mit lauter nichtnegativen Koordinaten aufgespannt wird. Daraus folgt, dass es für jeden Graphen $G = (V, E)$ genau einen PageRank f gibt mit $f(w) \geq 0$ für alle $w \in V$ und $\sum_{w \in V} f(w) = 1$. Diesen PageRank kann man auch als Wahrscheinlichkeitsverteilung interpretieren. Der Wert $f(w)$ besagt dann in etwa, mit welcher Wahrscheinlichkeit ein Surfer sich zu einem bestimmten Zeitpunkt auf der Seite w befindet, wenn er vor langer Zeit auf irgendeiner Seite gestartet ist und jeweils mit Wahrscheinlichkeit d einem zufälligen Link gefolgt ist und mit Wahrscheinlichkeit $1 - d$ auf eine zufällige beliebige Seite springt. (Wenn eine Seite keine Links enthält, springt der Surfer in jedem Fall auf eine zufällige beliebige Seite.)

Zur Berechnung des PageRanks verwendet man ein iteratives Verfahren. Experimente zeigen, dass man mit einer Wahl von $d \approx 0.85$ für einen Graphen mit $n = |V|$ Knoten schon nach $\mathrm{O}(\log n)$ Iterationsschritten eine brauchbare Näherung des PageRanks bekommt, wenn man mit der Näherungslösung f mit $f(w) = 1/n$ für alle $w \in V$ beginnt. Man wird annehmen dürfen, dass die durchschnittliche Zahl der Links pro Seite unabhängig von der Größe des Netzwerks ist, dass also der Graph nur $\mathrm{O}(n)$ viele Kanten hat. Das bedeutet, dass die Matrix $A^\top$ sich schreiben lässt als $A^\top = P + Q$, wobei P eine Matrix ist, bei der in jeder Spalte alle Einträge identisch sind, und Q eine dünn besetzte Matrix. Die Matrix-Vektor-Multiplikation, die im iterativen Verfahren durchzuführen ist, kostet deshalb nicht wesentlich mehr als $\mathrm{O}(n)$ Operationen, und die gesamte Berechnung des PageRanks nicht viel mehr als $\mathrm{O}(n \log n)$. Sie ist deshalb auch für große Graphen wie das WWW mit mehreren Milliarden Knoten noch machbar.

Beispiel. Betrachte den Graph $G = (\{1, 2, 3, 4, 5, 6, 7, 8, 9\}, E)$ mit

$$E = \{(1,2), (1,3), (1,4), (2,3), (2,1), (3,2), (4,1), (5,1), (6,1), (7,1), (8,9), (9,2)\}.$$

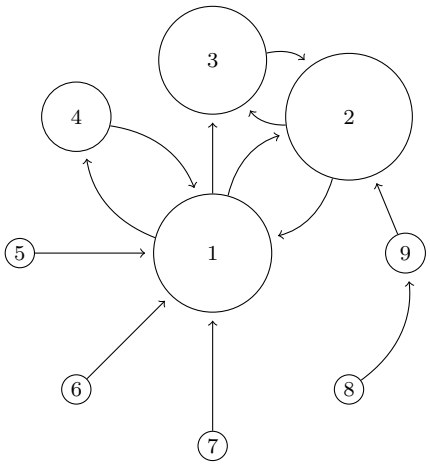
Für die Wahl $d = 0.85$ ist

$$A = 0.15 \underbrace{\frac{1}{9}\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}}_{=P} + 0.85 \underbrace{\begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_{=Q}$$

Der Eigenvektor von $A^\top$, dessen Koordinatensumme 1 ist, lautet

$$f = (0.271, 0.311, 0.226, 0.093, 0.016, 0.016, 0.016, 0.016, 0.03)$$

Im folgenden Bild sind diese Bewertungen durch die Größe der Knoten ausgedrückt. Der relevanteste Knoten ist demnach der Knoten 2, obwohl er viel weniger eingehende Kanten hat als der Knoten 1. Man sieht auch deutlich, dass Knoten 4 relevanter ist als Knoten 9, obwohl beide genau eine eingehende Kante haben. Der Grund ist jeweils, dass die Relevanz der Knoten, von dem die eingehenden Kanten ausgehen, in die Bewertung einfließt.



52 Rationale erzeugende Funktionen

Definition 93. Eine formale Potenzreihe $a = \sum_{n=0}^{\infty} a_n X^n \in \mathbb{K}[[X]]$ heißt *rational,* wenn es Polynome $p, q \in \mathbb{K}[X] \subseteq \mathbb{K}[[X]]$ gibt mit $q \neq 0$ und $qa = p$. In diesem Fall schreibt man $a = \frac{p}{q}$.

Beispiel.

1. Die Potenzreihe $a = 1 + X + X^2 + X^3 + \cdots \in \mathbb{K}[[X]]$ ist rational, denn es gilt $(1-X)a = a - Xa = 1$, also $a = \frac{1}{1-X}$.

2. Die Potenzreihe $a = \sum_{n=0}^{\infty} nX^n \in \mathbb{Q}[[X]]$ ist rational, es gilt $a = \frac{X}{(1-X)^2}$.

3. Wenn $a \in \mathbb{K}[[X]]$ so ist, dass $q_1 a = p_1$ und $p_2 a = q_2$ für gewisse Polynome $p_1, p_2, q_1, q_2 \in \mathbb{K}[X]$ gilt, dann gilt $q_2 p_1 a = q_1 q_2$ und $q_1 p_2 a = q_2 p_1$, und daher $(p_1 q_2 - p_2 q_1)a = 0$, und daher, weil $\mathbb{K}[[X]]$ ein Integritätsbereich ist, $p_1 q_2 = p_2 q_1$ also $\frac{p_1}{q_1} = \frac{p_2}{q_2}$. Wenn a also eine rationale Potenzreihe ist, dann gibt es nur eine rationale Funktion $\frac{p}{q} \in \mathbb{K}(X)$ mit $a = \frac{p}{q}$.

   Umgekehrt gibt es nicht zu jeder rationalen Funktion eine passende rationale Potenzreihe. Als Nenner kommen nur diejenigen Polynome $q \in \mathbb{K}[X]$ in Frage, die aufgefasst als Element von $\mathbb{K}[[X]]$ ein multiplikatives Inverses haben. Man kann zeigen, dass das genau jene $q \in \mathbb{K}[X]$ mit $q(0) \neq 0$ sind.

Zur Erinnerung: Gemäß Definition 47 heißt eine Folge $(a_n)_{n=0}^{\infty} \in \mathbb{K}^{\mathbb{N}}$ C-finit, wenn sie eine lineare Rekurrenz mit konstanten Koeffizienten erfüllt, d.h. wenn es $c_0, \ldots, c_r \in \mathbb{K}$ gibt, von denen wenigstens eines nicht Null ist, so dass gilt $c_0 a_n + \cdots + c_r a_{n+r} = 0$ für alle $n \in \mathbb{N}$.

Satz 159. Sei $(a_n)_{n=0}^{\infty} \in \mathbb{K}^{\mathbb{N}}$ eine Folge und $a = \sum_{n=0}^{\infty} a_n X^n \in \mathbb{K}[[X]]$ die formale Potenzreihe, deren Koeffizientenfolge $(a_n)_{n=0}^{\infty}$ ist. Dann gilt: $(a_n)_{n=0}^{\infty}$ ist genau dann C-finit, wenn a rational ist.

Beweis. „⇒“ $(a_n)_{n=0}^\infty$ ist C-finit. Dann gibt es $c_0, \dots, c_r \in \mathbb{K}$, nicht alle Null, dass für alle $n \in \mathbb{N}$ gilt $c_0 a_n + \cdots + c_r a_{n+r} = 0$. Sei $a = \sum_{n=0}^\infty a_n X^n$. Dann gilt $X^k a = \sum_{n=0}^\infty a_n X^{n+k}$ für jedes $k \in \mathbb{N}$, sowie $\sum_{n=0}^\infty a_{n+k} X^n = (a - a_0 - a_1 X - \cdots - a_{k-1} X^{k-1})/X^k$ für jedes $k \in \mathbb{N}$. Daraus folgt

$$\begin{aligned}
&c_0 a + c_1(a - a_0)/X + \cdots + c_r(a - a_0 - a_1 X - \cdots - a_{r-1}X^{r-1})/X^r \\
&= c_0 \sum_{n=0}^\infty a_n X^n + c_1 \sum_{n=0}^\infty a_{n+1} X^n + \cdots + c_r \sum_{n=0}^\infty a_{n+r} X^n \\
&= \sum_{n=0}^\infty \underbrace{(c_0 a_n + c_1 a_{n+1} + \cdots + c_r a_{n+r})}_{=0} X^n = 0,
\end{aligned}$$

und damit

$$\underbrace{(c_0 X^r + c_1 X^{r-1} + \cdots + c_r)}_{\in \mathbb{K}[X]} a = c_1 a_0 X^{r-1} + c_2(a_0 X^{r-2} + a_1 X^{r-1}) + \cdots\cdots + c_r(a_0 + \cdots + a_{r-1}X^{r-1}) \Big)}_{\in \mathbb{K}[X]}.$$

Das Polynom auf der linken Seite ist nicht Null, weil wenigstens ein c_i nicht Null ist. Damit ist gezeigt, dass a rational ist.

„⇐“ $a = \sum_{n=0}^\infty a_n X^n$ sei rational. Dann gibt es $p, q \in \mathbb{K}[X]$ mit $q \neq 0$ und $qa - p = 0$. Schreibe $q = q_0 + q_1 X + \cdots + q_r X^r$. Für jedes $k \in \mathbb{N}$ gilt $X^k a = \sum_{n=0}^\infty a_n X^{n+k}$. Für den Koeffizienten $[X^n]qa$ zum Term X^n in der Reihe qa gilt deshalb, wenn $n \geq r = \deg(q)$ ist,

$$\begin{aligned}
[X^n]qa &= [X^n]\big(q_0 a + q_1 X a + \cdots + q_r X^r a\big) \\
&= [X^n] \sum_{n=0}^\infty q_0 a_n X^n + [X^n] \sum_{n=0}^\infty q_1 a_n X^{n+1} + \cdots + [X^n] \sum_{n=0}^\infty q_r a_n X^{n+r} \\
&= q_0 a_n + q_1 a_{n-1} + \cdots + q_r a_{n-r}.
\end{aligned}$$

Für alle $n > \max(\deg(p), \deg(q)) =: s$ folgt damit aus $qa = p$, dass $q_0 a_n + q_1 a_{n-1} + \cdots + q_r a_{n-r} = 0$ ist. Für alle $n \in \mathbb{N}$ gilt deshalb die Rekurrenz $q_0 a_{n+s} + q_1 a_{n+s-1} + \cdots + q_r a_{n+s-r} = 0$, die von der geforderten Form ist. ∎

Beispiel.

1. Die Koeffizientefolge $(a_n)_{n=0}^\infty$ von $\frac{1}{1-X} = \sum_{n=0}^\infty X^n$ erfüllt die Rekurrenz $a_{n+1} - a_n = 0$, ist also C-finit.

2. Da die Folge $(F_n)_{n=0}^\infty$ der Fibonacci-Zahlen C-finit ist, muss die zugehörige erzeugende Funktion rational sein. In der Tat gilt $\sum_{n=0}^\infty F_n X^n = \frac{X}{1-X-X^2}$.

3. Die formale Potenzreihe $\sum_{n=0}^\infty n! X^n \in \mathbb{R}[[X]]$ kann nicht rational sein, da wir in Abschnitt 22 gezeigt haben, dass die Folge $(n!)_{n=0}^\infty$ nicht C-finit ist.

Eine formale Potenzreihe $a = \sum_{n=0}^\infty a_n X^n \in \mathbb{K}[X]$ bezeichnet man in der Kombinatorik als *erzeugende Funktion* (engl. *generating function*). Diese Bezeichnungsweise ist nicht ganz glücklich, denn es

handelt sich ebensowenig wie bei rationalen Funktionen um Funktionen im Sinn von Abschnitt 3. Man verwendet man erzeugende Funktionen $a = \sum_{n=0}^{\infty} a_n X^n \in \mathbb{Q}[[X]]$, um für eine Familie von Objekten anzugeben, dass es für jedes $n \in \mathbb{N}$ genau a_n Objekte der Größe n gibt. Was als Größe eines Objekts verstanden wird, hängt vom Problem ab. Bei einem Graphen (im Sinn von Abschnitt 21) kann z.B. die Zahl der Knoten oder die Zahl der Kanten als Maß für die Größe dienen. Für einen Pfad in einem Graphen ist die Länge (also die Anzahl der Segmente) ein naheliegendes Maß für die Größe. Nach Satz 65 lässt sich die Anzahl der Pfade der Länge $n \in \mathbb{N}$ zwischen je zwei Knoten eines Graphen aus der n-ten Potenz A^n der Adjazenzmatrix A des Graphen ablesen. Für das charakteristische Polynom $\chi = p_0 + p_1 X + \cdots + p_r X^r \in \mathbb{Q}[X]$ von A sagt der Satz von Cayley-Hamilton, dass $\chi(A) = 0$ gilt, also $p_0 I_n + p_1 A + \cdots + p_r A^r = 0$. Multiplikation dieser Gleichung mit A^n zeigt, dass alle Einträge von A^n, aufgefasst als Folge bezüglich n, C-finite sind. Nach Satz 159 sind die zugehörigen erzeugenden Funktionen also rational. Der folgende Satz präzisiert diesen Zusammenhang.

Satz 160. Für alle $A \in \mathbb{K}^{k\times k}$ und jede Wahl von $i, j \in \{1, \dots, k\}$ gilt

$$\sum_{n=0}^{\infty} (A^n)_{i,j} X^n = (-1)^{i+j} \frac{\det\bigl((I_k - AX)^{(j,i)}\bigr)}{\det(I_k - AX)}.$$

Dabei bezeichnet $(A^n)_{i,j}$ den Eintrag von A^n an der Position (i, j) und $(I_k - AX)^{(j,i)} \in \mathbb{K}[X]^{(k-1)\times(k-1)}$ bezeichnet die Matrix, die aus $I_k - AX \in \mathbb{K}[X]^{k\times k}$ durch Streichen der j-ten Zeile und der i-ten Spalte entsteht.

Beweis. Für die Potenzreihe $a = \sum_{n=0}^{\infty} A^n X^n \in \mathbb{K}^{k\times k}[[X]]$ gilt

$$(I_k - AX)a = \sum_{n=0}^{\infty} A^n X^n - \sum_{n=1}^{\infty} A^n X^n = I_k,$$

also $a = (I_k - AX)^{-1}$.

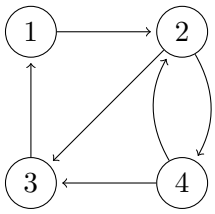
Für die Abbildung $h\colon \mathbb{K}^{k\times k}[[X]] \to \mathbb{K}[[X]]^{k\times k}$, die jede Potenzreihe $\sum_{n=0}^{\infty}((a_{i,j,n}))_{i,j=1}^{k} X^n \in \mathbb{K}^{k\times k}[[X]]$ auf die Matrix $((\sum_{n=0}^{\infty} a_{i,j,n} X^n))_{i,j=1}^{k} \in \mathbb{K}[[X]]^{k\times k}$ abbildet, lässt sich durch Nachrechnen überprüfen, dass $h(a + b) = h(a) + h(b)$ und $h(ab) = h(a)h(b)$ und $h(I_k) = I_k$ gilt. Daraus folgt, dass die Identität $a = (I_k - AX)^{-1}$ auch dann gilt, wenn man beide Seiten als Elemente von $\mathbb{K}[[X]]^{k\times k}$ auffasst.

Sei nun R ein kommutativer Ring und $U = ((u_{i,j}))_{i,j=1}^{k} \in R^{k\times k}$ eine beliebige unimodulare Matrix und $V = ((v_{i,j}))_{i,j=1}^{k} \in R^{k\times k}$ die Inverse von U. Dann gilt $UV = I_k$. Für jedes $j = 1, \dots, k$ ist die j-te Spalte von V also die Lösung des linearen Gleichungssystems $Ux = e_j$ (wobei e_j der j-te Einheitsvektor in $\mathbb{K}^k$ ist). Mit dem Satz von Cramer folgt $v_{i,j} = \det U^{(i)} / \det U$, wobei $U^{(i)}$ die Matrix ist, die aus U entsteht, wenn man die i-te Spalte durch e_j ersetzt. Durch Laplace-Entwicklung von nach der j-ten Spalte von $U^{(i)}$ bekommt man $\det U^{(i)} = (-1)^{i+j} \det U^{(j,i)}$, wobei $U^{(j,i)}$ aus $U^{(i)}$ (und damit auch aus U) durch Streichen der j-ten Zeile und der i-ten Spalte entsteht. Es gilt also $v_{i,j} = (-1)^{i+j} \det U^{(j,i)} / \det U$ für alle $i, j \in \{1, \dots, k\}^2$.

Wendet man dieses Argument auf $R = \mathbb{K}[X]$, $U = I_k - AX$ und $a = (I_k - AX)^{-1} = V \in \mathbb{K}[[X]]^{k\times k}$ an, so folgt die Behauptung des Satzes. ∎

Beispiel.

1. Sei $G = (\{1, \dots, 4\}, E)$ der folgende Graph:



Die Adjazenzmatrix von G ist $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} \in \mathbb{Q}^{4\times 4}$. Wenn a_n die Anzahl der Pfade der Länge n vom Knoten 1 zum Knoten 4 ist, dann gilt nach Satz 65 und Satz 160

$$\sum_{n=0}^{\infty} a_n X^n = -\frac{\begin{vmatrix} 0 & 1 & -X \\ -X & 0 & 1 \\ 0 & -X & -X \end{vmatrix}}{\begin{vmatrix} 1 & -X & 0 & 0 \\ 0 & 1 & -X & -X \\ -X & 0 & 1 & 0 \\ 0 & -X & -X & 1 \end{vmatrix}} = \frac{X^2(1+X)}{1-X^2-X^3-X^4}$$
$$= X^2 + X^3 + X^4 + 2X^5 + 3X^6 + 4X^7 + 6X^8 + 9X^9 + 13X^{10} + \cdots + 11351939356274689X^{100} + \cdots$$

2. Satz 160 ist nicht auf das Zählen von Pfaden in Graphen beschränkt. Er lässt sich immer dann anwenden, wenn sich die Anzahl von Objekten der Größe n mit einer Matrix-Vektor-Multiplikation aus der Anzahl der Objekte der Größe $n-1$ berechnen lässt. Man nennt die Matrix dann eine *Transfermatrix* für das Problem. Die Transfermatrix darf nicht von n abhängen.

   Betrachte zum Beispiel die Menge aller Wörter, die sich aus den Buchstaben a, b, c bilden lassen (vgl. Seite 28). Es ist leicht einzusehen, dass es insgesamt genau 3^n Wörter der Länge n gibt. Es sei L die Menge aller Wörter, in denen weder ein a auf ein b noch ein c auf ein a folgt. Wir wollen die Anzahl $f(n)$ aller Wörter in L der Länge n bestimmen.

   Dazu betrachtet man für jeden Buchstaben $u, v \in \{a, b, c\}$ die Anzahl $f_{u,v}(n)$ aller Wörter in L der Länge n, die mit u beginnen und mit v enden. Dann gilt $f(n) = \sum_{u,v} f_{u,v}(n)$ für $n \geq 1$ (und darüber hinaus $f(0) = 1$ oder $f(0) = 0$, je nach dem, ob wir das leere Wort mitzählen wollen oder nicht). Wörter der Länge $n+1$ lassen sich aus Wörtern der Länge n erzeugen, indem man einen Buchstaben anhängt. Wegen der beiden verbotenen Teilwörter ba und ac dürfen Wörter, die auf b enden, nicht um a erweitert werden, und Wörter, die auf a enden, nicht um c. Da es ansonsten keine Einschränkungen gibt, gilt für jedes $u \in \{a, b, c\}$ und jedes $n \geq 1$

$$\begin{pmatrix} f_{u,a}(n+1) \\ f_{u,b}(n+1) \\ f_{u,c}(n+1) \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}}_{=:A} \begin{pmatrix} f_{u,a}(n) \\ f_{u,b}(n) \\ f_{u,c}(n) \end{pmatrix}.$$

Die beiden Nullen in der Matrix entsprechen den beiden verbotenen Teilwörtern *ba* und *ac*.

Zusammen mit den Startwerten $f_{u,v}(1) = \delta_{u,v}$ für alle $u, v \in \{a, b, c\}$ folgt, dass $f_{u,v}(n)$ der Eintrag von A^{n-1} in der Zeile für v und der Spalte für u ist. Mit Hilfe von Satz 160 können wir jetzt für jedes Paar u, v die erzeugende Funktion bestimmen. Zum Beispiel gilt

$$\begin{aligned}\sum_{n=1}^{\infty} f_{a,b}(n)X^n = X\sum_{n=0}^{\infty}(A^n)_{a,b}X^n &= \frac{X^2(1-X)}{1-3X+2X^2-X^3}\\ &= X^2+2X^3+4X^4+9X^5+21X^6+\cdots.\end{aligned}$$

Demnach gibt es z.B. 9 Wörter in L, die mit a beginnen und mit b enden und die Länge 5 haben. Sie lauten: *aaaab*, *aaabb*, *aabbb*, *aabcb*, *abbbb*, *abbcb*, *abcab*, *abcbb*, *abccb*.

Für die Gesamtzahl $f(n)$ der Wörter in L der Länge n findet man

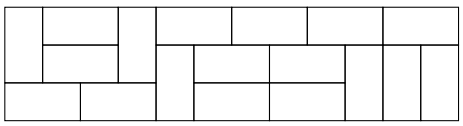
$$\sum_{n=0}^{\infty} f(n)X^n = \frac{1}{1-3X+2X^2-X^3} = 1+3X+7X^2+16X^3+37X^4+86X^5+\cdots$$

3. Wie viele Pfade mit n Zügen kann ein Läufer auf einem Schachbrett der Größe 8×8 fahren, wenn der Pfad in der Ecke A1 beginnt und dort auch wieder enden soll? (Außer dem Läufer befinden sich keine weiteren Figuren auf dem Schachbrett.)

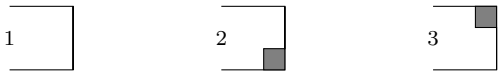
   Um diese Frage zu beantworten, numeriert man die Felder A1, ..., H8 des Schachbretts auf eine beliebige Weise mit $1, \ldots, 64$ durch und betrachtet die Matrix $A = ((a_{i,j}))_{i,j=1}^{64} \in \mathbb{Q}^{64\times 64}$ mit $a_{i,j} = 1$, wenn der Läufer in einem Zug vom Feld Nummer j zum Feld Nummer i ziehen kann, und mit $a_{i,j} = 0$ sonst. Satz 160 liefert dann die erzeugende Funktion

$$\begin{aligned}&\frac{(5X-1)\left(32X^6+32X^5-164X^4-52X^3+71X^2-16X+1\right)}{(X+1)\left(192X^7+184X^6-1066X^5-113X^4+468X^3-166X^2+22X-1\right)}\\ &= 1+7X^2+42X^3+325X^4+2564X^5+21203X^6+181070X^7+1583441X^8+\cdots.\end{aligned}$$

4. Unter einer Kachelung (engl. *tiling*) versteht man in der Kombinatorik eine Zerlegung einer großen (aber endlichen) Fläche in kleinere Flächen, die alle ein vorgegebenes Format haben. Zum Beispiel zeigt folgendes Bild eine Kachelung eines Rechtecks der Größe 3×12 mit kleinen Rechtecken vom Format 2×1 oder 1×2.



06-24 Wie viele solcher Kachelungen gibt es? Es ist klar, dass es für ein $3 \times n$-Rechteck mit ungeradem n gar keine Möglichkeit gibt, weil ein solches Rechteck eine ungerade Anzahl von Feldern hat und jede Kachel zwei Felder belegt, jede Kachelung also nur eine gerade Anzahl von Feldern belegen kann. Betrachten wir die Zahl der unvollständigen Kachelungen eines $3 \times (2n)$-Rechtecks, bei denen der rechte Rand eine der folgenden drei Formen hat:



Jede Kachelung mit einem rechten Rand vom Typ 1 ist entweder aus einer kleineren Kachelung des gleichen Typs entstanden, die man um  ,  oder  erweitert hat (drei Möglichkeiten), oder aus einer kleineren Kachelung vom Typ 2 durch Erweiterung um  (eine Möglichkeit), und so weiter. Die Matrix

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$$

enthält an Position (i, j) die Anzahl der Möglichkeiten, um von j nach i zu gelangen, wobei einige Möglichkeiten nicht berücksichtigt wurden, damit jede Kachelungen nur einmal gezählt wird. Zum Beispiel zählen wir für den Übergang von Typ 2 zu Typ 2 deshalb nur die Ergänzung mit  , aber nicht die mit , weil diese Erweiterung auch als eine Ergänzung zum Typ 1 gefolgt von einer Ergänzung zum Typ 2 aufgefasst werden kann und als solche in der Matrix schon berücksichtigt ist.

Mit Hilfe von Satz 160 lässt sich aus A die erzeugende Funktion für die Zahl a_n der Kachelungen eines $3 \times (2n)$-Rechtecks berechnen. Sie lautet:

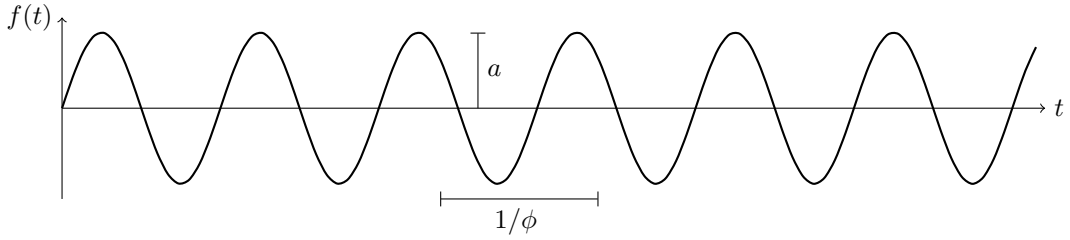
$$\sum_{n=0}^{\infty} a_n X^n = \frac{1 - X}{1 - 4X + X^2} = 1 + 3X + 11X^2 + 41X^3 + \cdots$$

53 Positive Polynome

TODO.

54 Signalverarbeitung

Der Eindruck von Schall entsteht durch Vibrationen der Luft, die vom Ohr aufgenommen werden. Die exakte Bewegung der Luftteilchen lässt sich nicht ohne weiteres mathematisch beschreiben, weil sie zu vielen Einflüssen unterliegt. Sie ist für den wahrgenommenen Ton aber auch unerheblich. Die Anatomie des Ohrs bewirkt, dass verschiedene Rezeptoren bei verschiedenen Luftbewegungsmustern ausschlagen. Jeder Rezeptor ist für eine bestimmte Frequenz ϕ zuständig. Einen Ton mit Frequenz ϕ kann man sich als Hin- und Herbewegung eines gedachten Luftteilchens gemäß einer Funktion der Form $f\colon \mathbb{R} \to \mathbb{R}$, $f(t) = a \sin(2\pi\phi t)$ vorstellen. Dabei ist a die Lautstärke (Amplitude).



Der akustische Gesamteindruck ergibt sich daraus, welche Frequenz-Rezeptoren bei den aktuell im Ohr eintreffenden Luftstößen wie stark angeregt werden. Bei der Überlagerung von Tönen ist zu beachten, dass sich Töne nicht nur in Amplitude und Frequenz unterscheiden können, sondern auch zeitlich gegeneinander versetzt sein können (Phasenverschiebung). Zwar sind zum Beispiel $t \mapsto \sin(2\pi t)$ und $t \mapsto \cos(2\pi t)$ als isolierte Töne nicht voneinander zu unterscheiden, aber $t \mapsto \sin(2\pi t) + \sin(2\pi t)$ ist ein anderer Ton als $t \mapsto \sin(2\pi t) + \cos(2\pi t)$.

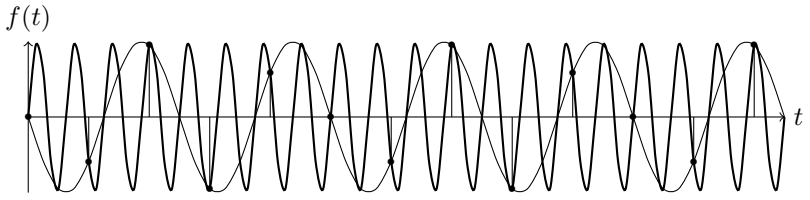
Es erweist sich als praktisch, statt trigonometrischer Funktionen die komplexe Exponentialfunktion zu verwenden. Man betrachtet dann Funktionen $f\colon \mathbb{R} \to \mathbb{C}$, $f(t) = a\exp(2\pi\mathrm{i}\phi t)$ mit reellem ϕ und komplexem a. Dabei ist $|a| \in \mathbb{R}$ die Amplitude und $\arg a \in [0, 2\pi)$ codiert die Phasenverschiebung. (Zur Erinnerung: Es gilt $a = |a|\exp(\mathrm{i}\arg a)$ und $\exp(2\pi\mathrm{i}\phi t) = \cos(2\pi\phi t) + \mathrm{i}\sin(2\pi\phi t)$.) Zur Beschreibung der Bewegung des gedachten Luftteilchens kann man dann wahlweise den Realteil oder den Imaginärteil der Funktion hernehmen.

Betrachten wir eine Audiosequenz von einer Sekunde Länge, die als Funktion $f\colon [0,1] \to \mathbb{C}$ gegeben ist, die die Bewegung eines gedachten Luftteilchens beschreibt. Der Raum $L^1([0,1],\mathbb{C})$ aller auf $[0,1]$ integrierbaren komplex-wertigen Funktionen bildet mit

$$\langle f|g\rangle = \frac{1}{2\pi}\int_0^1 f(t)\overline{g(t)}dt$$

einen Skalarproduktraum. Dabei ist $\overline{g(t)}$ die komplex-konjugierte Zahl von $g(t)$. Integrale von komplex-wertigen Funktionen sind genauso definiert wie Integrale von reellwertigen Funktionen. (Beachte: das Funktionsargument $t \in [0,1]$ ist reell.) Die Funktionen $t \mapsto \exp(2\pi\mathrm{i}nt)$ ($n \in \mathbb{N}$) bilden ein Orthonormalsystem in diesem Raum. Wegen der Ergebnisse aus Abschnitt 33 kann man $a = \langle f|t \mapsto \exp(2\pi\mathrm{i}nt)\rangle$ als den Anteil der Frequenz $2\pi n$ in der Audiosequenz f interpretieren (vgl. das Beispiel nach Def. 99). Die komplexe Zahl a codiert, welche Amplitude $|a|$ und welche Phasenverschiebung $\arg a$ dieser Anteil hat. Die Wirkung der Rezeptoren im Ohr entspricht also der Berechnung von Skalarprodukten mit den jeweiligen Frequenzfunktionen.

Eine Audiosequenz $f\colon [0,1] \to \mathbb{C}$ ist zunächst ein kontinuierliches Signal. Zur Bearbeitung wird das Signal digitalisiert, d.h. die Funktion f wird als Vektor $(f(0), f(\frac{1}{N}), f(\frac{2}{N}), \ldots, f(\frac{N-1}{N}))$ für ein $N \in \mathbb{N}$ codiert. Bei der Digitalisierung gehen Informationen verloren, weil Schwingungen mit zu hoher Frequenz nach der Digitalisierung nicht mehr von Schwingungen mit geringerer Frequenz zu unterscheiden sind.



Man kann zeigen, dass alle Frequenzanteile eines Signals dann aus der digitalisierten Version rekonstruierbar sind, wenn N mindestens doppelt so groß ist wie die größte im Signal vorkommende Frequenz (Abtasttheorem; engl. *sampling theorem*). Da akustische Signale Frequenzen im Bereich von ca. 15 bis ca. 20000 Schwingungen pro Sekunde enthalten, verwenden typische Audio-Dateiformate mindestens 40000 Funktionswerte pro Sekunde.

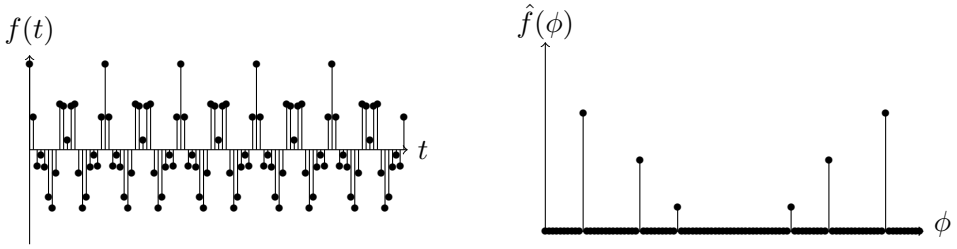
Wenn die Funktion f digitalisiert vorliegt, lassen sich die Integrale in den Skalarprodukten $\langle f|g\rangle$ nicht exakt berechnen. Man approximiert sie mit Riemannsummen. Der Anteil zur Frequenz n in der Sequenz f ist dann näherungsweise

$$\hat{f}(n) \approx \frac{1}{N}\sum_{k=0}^{N-1} f\big(\frac{k}{N}\big)\exp\big(-\frac{2\pi\mathrm{i}nk}{N}\big).$$

Die Berechnung des Vektors $\hat{f} = (\hat{f}(0), \hat{f}(1), \ldots, \hat{f}(N-1))$ aller Frequenzanteile des Signals f lässt sich also mit einer diskreten Fouriertransformation aus f berechnen. Nach Satz 144 kostet diese Berechnung nur $\mathrm{O}(N \log N)$ Operationen.

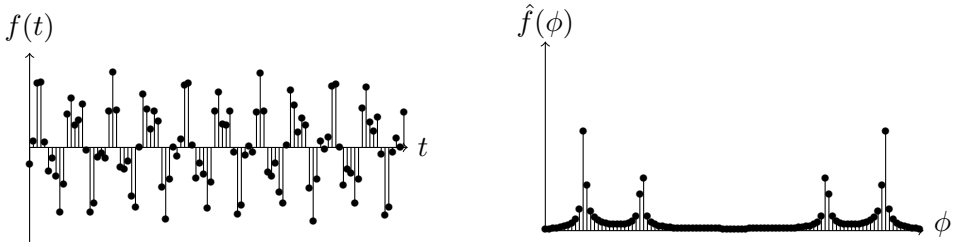
Beispiel.

1. Sei $f\colon [0,1] \to \mathbb{R} \subseteq \mathbb{C}$, $f(t) = 5\sin(2\pi 10(t+1)) + 3\sin(2\pi 25(t+2)) + \sin(2\pi 35t)$. Das Bild links zeigt den digitalisierten Signalvektor $x = (f(0), f(1/100), \ldots, f(99/100)) \in \mathbb{C}^{100}$. Rechts sind die Beträge der Koordinaten der Fouriertransformation $\hat{x} = \mathrm{DFT}_{100}^{(\omega)}\, x$ mit $\omega = \exp(-2\pi \mathrm{i}/100)$ dargestellt. Man erkennt die Spitzen bei den Frequenzen 10, 25 und 35. Die Spitzen in der rechten Bildhälfte sind Artifakte, die dadurch entstehen, dass wir nur den Realteil des Signals kennen.



2. Im obigen Beispiel sind die Spitzen markant, weil das Signal nur ganzzahlige Frequenzen enthält. Andere Frequenzen erkennt man auch an der Fouriertransformierten. Sie zeigen sich als Ausschläge der benachbarten ganzzahligen Frequenzen.

   Betrachte das Signal $f\colon [0,1] \to \mathbb{R} \subseteq \mathbb{C}$, $f(t) = 5\sin(2\pi 10.3t) + 3\sin(2\pi 25.6(t+1))$. Das Bild links zeigt wieder den Signalvektor $x = (f(0), f(1/100), \ldots, f(99/100)) \in \mathbb{C}^{100}$. Rechts sind die Beträge der Koordinaten der Fouriertransformierten von x dargestellt.



Die Fouriertransformation bietet zahlreiche Möglichkeiten zur Bearbeitung von Signalen. Um zum Beispiel in einem Musikstück die Bässe zu verstärken, berechnet man die Fouriertransformation, multipliziert die Amplituden der niedrigen Frequenzen mit einem Faktor $c > 1$, und transformiert das Ergebnis wieder zurück. Um das Rauschen in einem Signal zu verringern, kann man die Amplituden der hohen Frequenzen verkleinern (Tiefpassfilter). Um das Signal in einer höheren Frequenz zu hören ohne die Laufzeit zu verändern (eine schnellere Abspielgeschwindigkeit würde auch die Frequenz erhöhen), kann man die Daten der Fouriertransformierten etwas nach rechts verschieben und das Ergebnis zurücktransformieren. Ein entsprechend verlangsamtes Abspielen dieses Signals ist dann eine Zeitlupe des ursprünglichen Signals ohne Verringerung der Frequenzen.

Einige weitere Operationen, die man auf ein akustisches Signal anwenden will, lassen sich als Faltungen formulieren. Für die Definition der Faltung ist es zweckmäßig, den Definitionsbereich von Funktionen $f\colon [a,b] \to \mathbb{C}$ auf ganz $\mathbb{R}$ erweitern, indem man $f(t) = 0$ für $t \notin [a,b]$ setzt.

Definition 94. Seien $f, g\colon \mathbb{R} \to \mathbb{C}$ integrierbare Funktionen, für die es ein $T > 0$ gibt mit $f(t) = g(t) = 0$ für alle t mit $|t| > T$. Dann heißt

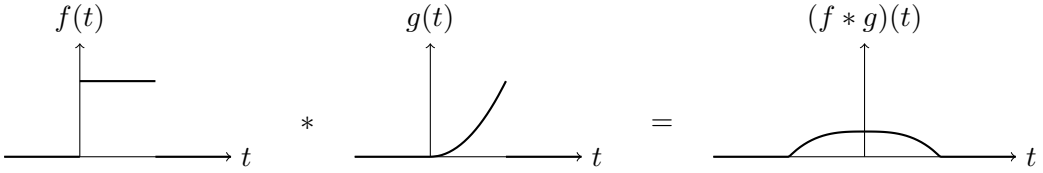
$$(f * g)\colon \mathbb{R} \to \mathbb{C}, \quad (f * g)(t) := \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau$$

die *Faltung* (engl. *convolution*) von f und g.

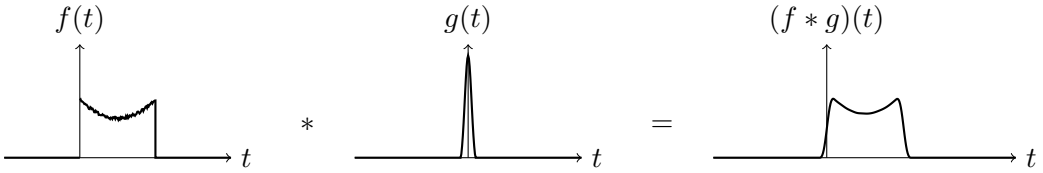
Beispiel.

1. Sei $f\colon \mathbb{R} \to \mathbb{R}$ mit $f(t) = 1$ für $t \in [0, 1]$ und $f(t) = 0$ sonst, und sei $g\colon \mathbb{R} \to \mathbb{R}$ mit $g(t) = t^2$ für $t \in [0, 1]$ und $g(t) = 0$ sonst. Dann gilt

$$\begin{aligned}(f * g)(t) &= \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau = \int_0^1 g(t - \tau) d\tau = \int_{\max(0,t)}^{\min(1,t+1)} (t - \tau)^2 d\tau \\ &= \begin{cases} 0 & \text{falls } t \leq -1 \text{ oder } t \geq 1 \\ \frac{1}{3}(1 + t^3) & \text{falls } t \in [-1, 0] \\ \frac{1}{3}(1 - t^3) & \text{falls } t \in [0, 1] \end{cases}\end{aligned}$$



2. Die Faltung einer Funktion f mit einer (schmalen) Glockenkurve g bewirkt eine Glättung des Funktionsgraphen von f. Der Funktionswert von $f * g$ an der Stelle t lässt sich dann interpretieren als Mittelwert der Funktionswerte von f in der Umgebung dieser Stelle t, wobei weiter weg liegende Stellen weniger stark gewichtet werden als nähere.



3. Mit Hilfe der Faltung kann man einem Audiosignal f ein Echo hinzufügen. Dazu wählt man als g eine Funktion, deren Graph aus einer hohen Spitze bei 0 und einigen niedrigeren Spitzen z.B. bei $-1, -2, -3$ besteht. Die Faltung $f * g$ ist dann eine Überlagerung aus dem Signal f und drei zeitversetzten leiseren Kopien von f.

06-23

4. Faltungen werden auch zur Spracherkennung eingesetzt. Wenn f ein langes Signal ist (z.B. eine Sekunde) und g ein kurzes (z.B. 0.1 Sekunden), dann hat die Faltung $f * g$ an einer Stelle t einen umso größeren Funktionswert, je größer der Verlauf von f in einer Umgebung von t dem umgekehrten Verlauf von g in einer Umgebung von 0 ähnelt. Wenn also f eine Audiosequenz ist, die gesprochene Sprache codiert, dann kann man durch Faltung mit kurzen Audiosequenzen, die jeweils nur einen rückwärts gesprochenen

06-23

isolierten Sprachlaut (ein sogenanntes *Phonem*, also z.B. ‘a’, ‘t’, ‘sch’,. . .) enthalten, erkennen, an welchen Stellen welche Laute gesprochen wurden. Die Treffgenauigkeit wird dabei nicht unbedingt besonders hoch sein, weil sich Phoneme untereinander u.U. ähnlicher sind als dasselbe Phonem, wenn es von verschiedenen Sprechern gesprochen ist. Aber zumindest kann man auf diese Weise für jeden Zeitpunkt t und jedes Phonem a so etwas wie eine Wahrscheinlichkeit $p(a,t)$ dafür schätzen, dass zum Zeitpunkt t das Phonem a gesprochen wurde. Aus diesen Rohdaten können andere Algorithmen plausible Hypothesen ableiten, welche Wörter wohl gesprochen wurden.

Die Faltung zweier digitalisierter Signale lässt sich mit dem FFT-Algorithmus effizient berechnen. Nehmen wir an, dass $f, g\colon \mathbb{R} \to \mathbb{C}$ außerhalb von $[\frac{1}{3}, \frac{2}{3}]$ Null sind. Das lässt sich durch Verschieben und Skalieren der Zeitachse leicht erreichen. Dann hat $f * g\colon \mathbb{R} \to \mathbb{C}$ außerhalb von $[0,1]$ den Wert Null. Als Digitalisierung von $f * g$ taugen also die Werte $(f * g)(n/N)$ für $n = 0, \ldots, N-1$. Es gilt

$$(f * g)\big(\frac{n}{N}\big) = \int_0^1 f(\tau) g\big(\frac{n}{N} - \tau\big) d\tau \approx \frac{1}{N} \sum_{k=0}^{N-1} f\big(\frac{k}{N}\big) g\big(\frac{n-k}{N}\big).$$

Die Werte von $f * g$ an den Stützstellen n/N ($n = 0, \ldots, N-1$) hängen demnach mit den Werten von f und g an diesen Stellen in der gleichen Weise zusammen wie die Koeffizienten des Produkts pq zweier Polynome $p, q \in \mathbb{K}[X]$ mit den Koeffizienten der Faktoren p und q. Man kann sie also mit Hilfe der FFT genau so berechnen, wie es in Abschnitt 46 für die Polynommultiplikation erklärt wurde.

Dass man die Faltung zweier digitalisierter Signale mit nur $\mathrm{O}(N \log N)$ Operationen berechnen kann, macht diese Operation (und damit den FFT-Algorithmus) für Anwendungen in der Signalverarbeitung besonders interessant. Die Anwendungen beschränken sich nicht auf die Verarbeitung akustischer Signale, sondern auch auf andere Daten, die man als Überlagerung von Schwingungen auffassen kann, z.B. beim Wechselstrom in der Elektrotechnik oder bei seismischen Messungen in der Geologie. Auch die Informationen in Bildern kann man als Überlagerungen von Schwingungen (in einem zweidimensionalen Raum) auffassen. Das erlaubt den Einsatz der FFT zur Mustererkennung oder zur Korrektur von Bildstörungen.

Index

Abbildung, 18
Abbildungsmatrix, 114
abelsche Gruppe, 25
Ableitung, 106, 308
absoluter Fehler, 290
Abstand, 193
Abtasttheorem, 322
Adjazenzmatrix, 318
affine subspace, 122
affiner Unterraum, 122
ähnlich, 167
Algebra, 170
algebraic closure, 162
algebraically closed, 161
algebraisch abgeschlossen, 161
algebraische Zahl, 162
algebraischer Abschluss, 162
algorithm, 268
Algorithmus, 268
Algorithmus von Bareiss, 299
Algorithmus von Euklid, 157
Algorithmus von Gauß, 55
Algorithmus von Karatsuba, 285
Algorithmus von Newton, 280
Algorithmus von Strassen, 287
Altertum, 311
Amplitude, 321
Anfangswert, 138
angeordneter Körper, 192
annihilierendes Polynom, 172
Ansatz, 171
antisymmetrisch, 12
Approximation, 211, 290
äquivalent, 54
Äquivalenzklasse, 14
Äquivalenzrelation, 13
Artifakt, 323
assoziativ, 23
assoziiert, 155
Audiosequenz, 322
aufspannen, 85, 89
Auge, 125
automatisches Beweisen, 142
Automorphismus, 104
average case, 268

Bahn, 311
Bareiss, 299
Basis, 89, 243
Basisergänzungssatz, 95
Basiswechselmatrix, 115
Bass, 323
Besselsche Ungleichung, 209
Bidualraum, 120
bijektiv, 18
Bild, 21, 29, 104, 230, 240
Bildstörung, 325
Binet, 170
Binomialkoeffizient, 178
Bit, 142
Blutprobe, 314
branch point, 308
Brennweite, 127
Buchstabe, 28, 142
Byte, 142

C-finit, 136
cardinality, 19
Cauchy-Schwarz-Ungleichung, 196
Cayley-Hamilton, 177
charakteristisches Polynom, 166
Chebyshev-Polynom, 208, 215
Chinesischer Restsatz, 279
Cholesky-Zerlegung, 204
Codewort, 143
column space, 85
common divisor, 156
Computeralgebra, 301

condition number, 291
conic, 310
convolution, 324
coprime, 158
Cramersche Regel, 83
cycle, 73, 133

Datenfehler, 291
Datenformat, 290
Datenkompression, 143
degree, 35
Dehnung, 46
dense matrix, 271
Determinante, 76, 116, 176
diagonalisierbar, 167
Diagonalmatrix, 167
dicht besetzt, 271
Differentialgleichung, 86
Digitalisierung, 322
Dimension, 93
direkte Summe, 98, 238
disjoint, 7
disjunkt, 7, 73
diskrete Fouriertransformation, 281
Diskriminante, 308
Distanz, 135
divisibility, 11
divisor, 155
Doppelkegel, 310
Drehspiegelung, 225
Drehung, 46, 225
Dreiecksmatrix, 273
duale Basis, 118
Dualraum, 117
dünn besetzt, 271

e, 162
Ebene, 122
echelon form, 53
Echo, 324
edge, 11, 129
Effizienz, 268
Eigenraum, 166
Eigenschaft, 6
eigenspace, 166
eigenvalue, 162
eigenvector, 162
Eigenvektor, 162
Eigenwert, 162
Einheitsintervall, 259
Einheitskreis, 305
Einheitskugel, 194
Einheitsmatrix, 47
Einheitsvektor, 45, 103
Einheitswurzel, 281
Eins, 25, 32, 233
einschalig, 311
Elektrotechnik, 325
Element, 6
Elementarmatrix, 54
Elementarteilersatz, 258
Elimination, 55
Eliminationstheorie, 305
Ellipse, 309
Ellipsoid, 311
elliptisch, 311
elliptische Kurve, 28
endlich erzeugt, 236
endliche Menge, 19
Endomorphismus, 104
engmaschig, 258
Entfernungstabelle, 135
equal, 7
erzeugen, 85, 89
erzeugende Funktion, 317
Erzeugendensystem, 89, 236
Erzeugermatrix, 143
euklidischer Algorithmus, 157
euklidischer Raum, 194
explizit, 60
exponentielle Komplexität, 298

Faktorraum, 101
Faltung, 324
Farbe, 127
Fast Fourier Transform, 283
Fehler, 290
Fehlererkennung, 143
FFT, 283
Fibonacci-Zahlen, 136, 138, 170, 317
field, 36
fill-in, 274
finite, 19
finitely generated, 236

Fixpunkt, 73
Fläche, 305
Folge, 136
formal power series, 34
formale Laurent-Reihe, 38
formale Potenzreihe, 34
Formel von Binet, 170
Fouriertransformation, 281
free, 243
frei, 241
freie Gruppe, 29
freier Vektorraum, 102
Frequenz, 321
Fundamentalsatz der Algebra, 161
Funktion, 18
Funktional, 117

ganze Zahlen, 25, 33, 36
Gauß-Algorithmus, 55
Gauß-Elimination, 55
gemeinsamer Teiler, 156
generating function, 317
generating set, 89, 236
Geologie, 325
geometrische Reihe, 263, 284
geordnete Basis, 113
Gerade, 122, 311
geschlossener Pfad, 133
Gewicht, 143
Gitter, 234, 258
Glättung, 324
Gleichheit, 7
Gleichungssystem, 51, 66
Glockenkurve, 324
Google, 314
größter gemeinsamer Teiler, 156
Gröbnerbasis, 307
Grad, 35
Gram-Schmidt, 212
Graph, 11, 129
Grauton, 129
Grauwert, 230
greatest common divisor, 156
group, 25
Grundfarbe, 127
Gruppe, 25
gut konditioniert, 291

Hadamardsche Ungleichung, 297
Halbgruppe, 25
Halbordnung, 12
Hamming-Code, 144
Hamming-Distanz, 143
Hankel-Matrix, 276
Hauptachse, 309
Hauptachsentransformation, 309
Hauptidealring, 237
Helligkeit, 129
Hermite-Normalform, 249
hermitesch, 196
Himmelsmechanik, 311
homogeneous, 66
homogenes Gleichungssystem, 66
Homomorphiesatz, 21, 32, 111, 240
Homomorphismus, 29, 104, 131, 239
hyper plane, 122
Hyperbel, 310
hyperbola, 310
hyperbolisch, 311
Hyperboloid, 311
Hyperebene, 122

Ideal, 233
ideal, 233
Identitätsfunktion, 18
image, 21, 29, 104
imaginäre Einheit, 37
Imaginärteil, 37, 105
implizit, 60
indefinit, 203
infinite, 19
Information, 142
inhomogeneous, 66
inhomogenes Gleichungssystem, 66
injektiv, 18
integral domain, 237
Integritätsbereich, 237
Interpolation, 279
Interpolationspolynom, 106, 278
intersection, 6
invariant, 179
Inverse, 21, 23, 47
invertible, 248
invertierbar, 248
Invertierbarkeit, 23

Isometrie, 223
isomorph, 29, 104, 131, 240
Isomorphiesatz, 112
Isomorphismus, 29, 104, 131, 240

Jordan-Normalform, 185

Kachelung, 320
Kante, 11, 129
Karatsuba, 285
kartesisches Produkt, 7
Kegel, 310, 311
Kegelschnitt, 310
Kepler, 311
Kern, 29, 51, 85, 104, 240
kernel, 29, 104
kleiner Satz von Fermat, 282
Knoten, 11, 129
Ko-Kern, 85
Kombinatorik, 317
kommutativ, 23, 32
Komplement, 7
Komplementärraum, 99
komplexe Zahlen, 37, 145
Komponente, 7
Komposition, 19
Konditionszahl, 291
Konjugation, 161
konjugiert, 196
konstruktiver Beweis, 254
Konvexkombination, 128
Koordinate, 113
Koordinatendarstellung, 113, 114
Körper, 36
Kovarianz, 313
Kryptographie, 28, 143
Kurve, 305

ℓ^2, 195
Länge, 193
Laplace-Entwicklung, 81
lattice, 234
Läufer, 320
leading coefficient, 154
leere Menge, 6
leeres Wort, 29, 319
Legendre-Polynom, 215
Leitkoeffizient, 154
Lemma von Zorn, 94
Lenstra, 263
Licht, 127
line, 122
linear, 159
linear abhängig, 42, 89
linear unabhängig, 42, 89, 243
lineare Abbildung, 104
lineare Gruppe, 47
linearer Code, 143
lineares Gleichungssystem, 51
Linearkombination, 42
Link, 314
linkseindeutig, 16
linkstotal, 16
LLL, 263
loop, 129
Lösung, 51
Lovász, 263
Luft, 321

Mächtigkeit, 19
Magma, 25
map, 18
Maple, 146, 264
Mathematica, 148, 264
Matrix, 44
Matrixmultiplikation, 269
Matrixprodukt, 44
Medizin, 314
Menge, 6
Messwert, 145
Metrik, 192
metrischer Raum, 192
minimal polynomial, 173
Minimalpolynom, 173
Modul, 233
module, 233
monic, 158
Monoid, 25
multiplicity, 155
Multiplikation, 268, 269
Musik, 323
Mustererkennung, 325

Näherung, 290
natürliche Zahlen, 6, 25, 33

negativ semidefinit, 203
Netzwerk, 129
Neutralelement, 23
Newton-Interpolation, 279, 280
normiert, 158
Null, 25, 32
Nullpolynom, 35
Nullstelle, 155
Nullstellenmenge, 305
Nullteiler, 237
nullteilerfrei, 237
Nullvektor, 85
numerisch, 145
numerischer Algorithmus, 290

obere Dreiecksmatrix, 273
Obermenge, 7
Objekt, 6
Ohr, 321
order(ing), 12
Ordnung, 12
orthogonal, 198
Orthogonalbasis, 206
orthogonale Gruppe, 226
orthogonale Matrix, 223
orthogonales Komplement, 198
Orthogonalprojektion, 221
Orthogonalsystem, 206
Orthonormalbasis, 206
Orthonormalsystem, 206

π, 162
p-adische Zahlen, 33, 235
PageRank, 314
Parabel, 310
parabola, 310
Paraboloid, 311
parallel, 124
parallelepiped, 259
Parallelogrammgleichung, 196
Parsevalsche Gleichung, 209
partielle Funktion, 18
path, 133
Permutation, 26
Permutation, Vorzeichen, 76
Permutationsmatrix, 49
perpendicular, 198
Pfad, 133, 318
Phasenverschiebung, 322
Phonem, 325
plane, 122
point, 122
Polynom, 34
Polynomendivision, 154, 268
Polynomfunktion, 105
Polynommultiplikation, 268
positiv definit, 203
Positivität, 321
Potenzmenge, 10
power set, 10
Prüfmatrix, 143
preconditioner, 294
preimage, 21
primitive Einheitswurzel, 281
principal ideal domain, 237
Projektion, 211, 220
projektiver Raum, 125
Punkt, 40, 122, 311
Punktwolke, 313
Pythagoras, 196
Python, 150

Quadrik, 311
Quotient, 155
quotient module, 238
quotient space, 101
Quotientenmodul, 238
Quotientenraum, 101

Rang, 62, 116
rank, 62
rationale Funktion, 37, 296
rationale Potenzreihe, 316
rationale Zahlen, 19, 25, 33, 36
Rauschen, 323
Realteil, 37, 105
rechnen, 290
rechtseindeutig, 16
rechtstotal, 16
reduced, 262
reduced echelon form, 53
reduzierte Basis, 262
Reed-Solomon-Code, 145
reelle Zahlen, 19, 25, 33, 36, 145

reflexiv, 12
Regel von Cramer, 83
Rekurrenz, 87
Relation, 10
relativer Fehler, 290
Relativitätstheorie, 27
Relevanz, 314
remainder, 155
repräsentantenunabhängig, 22
residue class ring, 33
Rest, 155
Restklassenring, 33, 36
resultant, 307
Resultante, 307
RGB, 128
Richardson, 292
Richtung, 40, 122
Riemannsumme, 322
Riesz, 217
Ring, 32, 233
root, 155, 192
root of unity, 281
Rot-Grün-Blindheit, 128
rotation, 225
row space, 85
runden, 230
runden rückwärts, 266

Sage, 150, 264
sampling theorem, 322
Satz über die Elementarteiler, 258
Satz von Binet, 170
Satz von Cauchy-Schwarz, 196
Satz von Cayley-Hamilton, 177, 318
Satz von Cooley-Tuckey, 282
Satz von Cramer, 83
Satz von Fermat, 282
Satz von Gram-Schmidt, 212
Satz von Hadamard, 297
Satz von Jordan, 185
Satz von Laplace, 81
Satz von Pythagoras, 196
Satz von Richardson, 292
Satz von Riesz, 217
Satz von Smith, 255
Satz von Sylvester, 299
scalar product, 194
Schachbrett, 320
Schachtelungstiefe, 10
Schall, 321
Scherung, 46
schlecht konditioniert, 291
Schleife, 129
Schnitt, 6, 88, 123, 238
Schnittpunkt, 305
Schwerpunkt, 313
seismisch, 325
selbstadjungiert, 219
semi group, 25
sequence, 136
set, 6
sign, 76
Signalverarbeitung, 321
signieren, 143
similar, 167
Singulärwert, 229
Singulärwertzerlegung, 227
singular value decomposition, 227
Skalarmultiplikation, 40, 85
Skalarprodukt, 194
Skalarproduktraum, 194
Smith-Normalform, 255
solution, 86
Spalte, 44
Spaltenraum, 85
span, 87
sparse matrix, 271
Spat, 259
Speicher, 230
Spektralsatz, 220
Spektrum, 127
spezielle lineare Gruppe, 80
spezielle orthogonale Gruppe, 226
Spracherkennung, 324
Spur, 166
Standardbasis, 89
Standardskalarprodukt, 195
Strassen, 287
Streuung, 313
Struktur, 272
submodule, 236
subset, 7
subspace, 87
Suchmaschine, 314

Summationsformel, 142, 301
Summe, 238
superset, 7
surjektiv, 18
Sylvester, 299
Sylvester-Matrix, 307
Symbolfolge, 290
symbolisch, 145
symbolischer Algorithmus, 296
Symmetrie, 26
Symmetriegruppe, 26
symmetrisch, 12, 203, 219
symmetrische Gruppe, 26
Syzygienmodul, 235

Teilbarkeit, 11
teilerfremd, 158
Teilmenge, 7
Tensor, 103
tensor product, 103
Tensorprodukt, 103, 242
Tiefpassfilter, 323
tiling, 320
Ton, 321
Töplitz-Matrix, 276
torisionsfrei, 245
torsion element, 245
torsion module, 245
torsion submodule, 245
torsion-free, 245
Torsionselement, 245
Torsionsmodul, 245
Torsionsuntermodul, 245
Torus, 239
total, 12
Totalordnung, 12
trace, 166
Trägermenge, 25
Transfermatrix, 319
transitiv, 12
Transposition, 48, 73, 118
transzendente Zahl, 162
Treppenform, 53
Treppennormalform, 53
Treppenstufe, 53
Tupel, 7

Umkehrfunktion, 21
unendlich, 126
unendliche Menge, 19
unimodular, 248
union, 6
unit sphere, 194
Untermodul, 236
Unterraum, 87
Unterring, 34, 248
Untervektorraum, 87
Urbild, 21

Vandermonde Matrix, 81, 278
Varianz, 313
vector space, 85
Vektor, 40, 85
Vektorraum, 85
Vereinigung, 6
Verfahrensfehler, 291
Verkettung, 19
Verknüpfung, 23
verschlüsseln, 143
vertex, 11, 129
Verzweigungspunkt, 308
Vibration, 321
Vielfachheit, 155
Vogel, 129
Volumen, 259
Vorkonditionierer, 294
Vorzeichen, 76
Votum, 314

Wahrscheinlichkeitsverteilung, 315
Wechselstrom, 325
wohldefiniert, 22
Wort, 28, 319
Wurzel, 192

Zahlengerade, 239
Zeichenkette, 290
Zeile, 44
Zeilenraum, 85
Zeitlupe, 323
Zentrum, 310
zero divisor, 237
Zoom, 127
Zornsches Lemma, 94
Zusammenhang, 6

Zusammenhangskomponente, 14
zweischalig, 311
Zyklus, 73, 133
Zylinder, 239, 312