Diskrete Strukturen

und

Lineare Algebra I für Informatiker

Skript zur Vorlesung

Dr. Timo Hanke
Prof. Dr. Gerhard Hiß
Lehrstuhl D für Mathematik
RWTH Aachen

Letzte Aktualisierung:
15. Oktober 2012

Unter der freundlichen Mithilfe von:
Wolf-Daniel Andres, Grischa Studzinski und Florian Weingarten.

Inhaltsverzeichnis

Grundlagen

Kapitel 1

Mathematische Grundbegriffe

1.1 Aussagen

1.1.1 Definition und Beispiele

Definition. *Mathematische Aussagen* oder kurz *Aussagen* sind sprachliche Ausdrücke, die auch Formeln und Symbole enthalten können, und die einen eindeutigen *Wahrheitswert* besitzen, der entweder *wahr* oder *falsch* lautet.

Beispiel. Mathematische Aussagen sind:

(i) '$2 + 3 = 5$' (wahr)

(ii) 'Alle Punkte auf einem Kreis haben den gleichen Abstand zum Mittelpunkt' (wahr)

(iii) 'Jede ganze Zahl größer als 2 ist Summe zweier Primzahlen' (unbekannt)

(iv) 'Jede reelle Zahl ist ein Quadrat einer reellen Zahl' (falsch)

(v) 'Es gibt eine ganze Zahl, deren Quadrat gleich ihrem Doppelten ist' (wahr)

Die Aussage (iii) ist eine mathematische Aussage, denn sie besitzt einen Wahrheitswert, auch wenn uns dieser nicht bekannt ist. Die *Goldbach'sche Vermutung* besagt, dass der Wahrheitswert von (iii) wahr lautet. Keine mathematischen Aussagen sind dagegen 'Aachen ist schön' und 'Die Mensapreise sind zu hoch'.

1.1.2 Zusammensetzung und Verneinung

Definition. Für beliebige Aussagen A und B definieren wir die Wahrheitswerte für folgende *zusammengesetzte Aussagen*:

(i) Die *Verneinung* oder *Negation* $\neg A$ ist genau dann wahr, wenn A falsch ist.

(ii) Die *Konjunktion* $A \wedge B$ ist genau dann wahr, wenn sowohl A als auch B wahr ist.

(iii) Die *Disjunktion* $A \vee B$ ist genau dann wahr, wenn A oder B wahr ist oder beide wahr sind.

(iv) Das *exklusive oder* $A \,\mathrm{XOR}\, B$ ist genau dann wahr, wenn A oder B wahr ist, aber nicht beide wahr sind.

(v) Die *Implikation* $A \Rightarrow B$ ist genau dann falsch, wenn A wahr ist und B falsch ist.

(vi) Die *Äquivalenz* $A \Leftrightarrow B$ ist genau dann wahr, wenn A und B den gleichen Wahrheitswert besitzen.

Sprechweise. Zu $\neg A$ sagt man „nicht A“, zu $A \wedge B$ „A und B“, zu $A \vee B$ „A oder B“, zu $A \,\mathrm{XOR}\, B$ „A x-or B“ oder „entweder A oder B“, zu $A \Rightarrow B$ „A impliziert B“ oder „aus A folgt B“ oder „wenn A dann B“, zu $A \Leftrightarrow B$ „A ist äquivalent zu B“ oder „A gilt genau dann, wenn B gilt“.

Die Wahrheitswerte der eingeführten zusammengesetzten Aussagen können in folgender Tabelle zusammengefasst werden:

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \,\mathrm{XOR}\, B$	$A \Rightarrow B$	$A \Leftrightarrow B$
w	w	f	w	w	f	w	w
w	f	f	f	w	w	f	f
f	w	w	f	w	w	w	f
f	f	w	f	f	f	w	w

Beispiel.

(i) Die Verneinung von ‘2+3 = 5’ lässt sich als ‘Es gilt nicht, dass 2+3 = 5 ist’ oder kürzer als ‘2 + 3 ist ungleich 5’ formulieren.

(ii) Die Verneinung von ‘Das Glas ist voll’ lässt sich als ‘Das Glas ist nicht voll’ formulieren, nicht aber als ‘Das Glas ist leer’.

(iii) Die Verneinung von ‘Alle Gläser sind voll’ lässt sich als ‘Nicht alle Gläser sind voll’, oder als ‘Es gibt ein Glas, das nicht voll ist’ formulieren.

(iv) ‘Wenn $2 + 3 = 6$, dann ist $2 + 3 = 7$’ ist wahr.

1.1.3 Aussageformen

Definition. Eine *Aussageform* ist eine sprachlicher Ausdruck, der Variablen enthält, und der für jede Belegung alle vorkommenden Variablen mit konkreten Objekten zu einer Aussage wird.

Bemerkung. Eine Aussageform ist selbst keine Aussage. Die Zusammensetzung von Aussageformen mittels $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$, etc. ist wieder eine Aussageform.

Beispiel.

(i) ‘Wenn $x > 0$, dann ist x ein Quadrat.’ ist eine Aussageform. Wird die Variable x mit einer beliebigen reellen Zahl belegt, so erhalten wir eine Aussage (einen eindeutigen Wahrheitswert).

(ii) ‘Student x hat mindestens 50% der Klausur-Punkte erzielt’ ist eine Aussageform. Wird die Variable x mit einem beliebigen Studenten belegt, so erhalten wir eine Aussage (einen eindeutigen Wahrheitswert).

(iii) Es sei $A(x)$ die Aussageform ‘Student x hat in der Klausur volle Punktzahl erzielt’ und $B(x)$ die Aussageform ‘Student x hat das Modul bestanden’. Dann ist auch $A(x) \Rightarrow B(x)$ eine Aussageform. Für jede Belegung der Variable x mit einem Studenten ist $A(x) \Rightarrow B(x)$ eine wahre Aussage. Das liegt daran, dass der Fall $A(x)$ wahr und $B(x)$ falsch (der einzige Fall in dem $A(x) \Rightarrow B(x)$ falsch ist) nicht vorkommt.

(iv) Es sei $A(t)$ die Aussageform ‘Der Projektor im Hörsaal ist zum Zeitpunkt t aus’ und $B(t)$ die Aussageform ‘Der Hörsaal ist zum Zeitpunkt t leer’. Für jede Belegung der Variable t mit einem Zeitpunkt ist $A(t) \Rightarrow B(t)$ eine Aussage. Deren Wahrheitswert hängt allerdings von t ab. Wann ist sie falsch?

Bemerkung. Wenn $A(x) \Rightarrow B(x)$ unabhängig von x stets wahr ist (wie in Beispiel (iii)), dann drückt $\Rightarrow$ offensichtlich einen kausalen Zusammenhang aus.

1.1.4 Tautologien

Beispiel.

(i) $A \,\mathrm{XOR}\, B$ hat stets den gleichen Wahrheitswert wie $(A \wedge \neg B) \vee (\neg A \wedge B)$, egal um welche Aussagen es sich bei A und B handelt. Wir sagen daher, dass XOR durch $\neg, \wedge, \vee$ augedrückt werden kann.

(ii) $A \Rightarrow B$ hat stets den gleichen Wahrheitswert wie $\neg(A \wedge \neg B)$.

(iii) $A \Leftrightarrow B$ hat stets den gleichen Wahrheitswert wie $\neg(A \,\mathrm{XOR}\, B)$.

Übung. Man zeige, dass XOR durch $\neg, \vee$ ausgedrückt werden kann.

Definition. Ein *logischer Term* ist ein Ausdruck bestehend aus Variablen, z.B. $A, B, \ldots$, die verknüpft sind mit den Symbolen $\neg, \wedge, \vee, \mathrm{XOR}, \Rightarrow, \Leftrightarrow$. Eine *Tautologie* ist ein logischer Term, der für jede Belegung seiner Variablen mit Wahrheitswerten „wahr“ ergibt.

Beispiel. $A \wedge \neg B$ und $A \Rightarrow ((B \Rightarrow \neg C) \vee D)$ sind logische Terme, aber keine Tautologien. $(A \Rightarrow B) \Leftrightarrow \neg(A \wedge \neg B)$ ist eine Tautologie. Bedeutsame Tautologien sind:

(i) Modus Ponens:
$$(A \wedge (A \Rightarrow B)) \Rightarrow B$$

(ii) Tertium non datur (Gesetz des ausgeschlossenen Dritten):
$$A \vee \neg A$$

(iii) de Morgan-Gesetze:
$$\neg(A \wedge B) \Leftrightarrow (\neg A \vee \neg B),$$
$$\neg(A \vee B) \Leftrightarrow (\neg A \wedge \neg B)$$

(iv) Kontrapositionsgesetz:
$$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$$

Übung.

(i) Man schreibe die Tautologien auf, die von Beispiel (i),(ii) und (iii) geliefert werden.

(ii) Ist $(A \Leftrightarrow B) \Leftrightarrow ((A \Rightarrow B) \wedge (B \Rightarrow A))$ eine Tautologie?

(iii) Man folgere aus den Tautologien des Beispiels durch Einsetzen, dass auch $\neg(A \wedge \neg A)$ eine Tautologie ist.

Bemerkung. Tautologien helfen bei Beweisen: Zeigt man etwa, dass sowohl A als auch $A \Rightarrow B$ wahr sind, so folgt nach Modus Ponens, dass auch B wahr ist. Möchte man $A \Rightarrow B$ zeigen, so kann man nach dem Kontrapositionsgesetz anstelle dessen auch $\neg B \Rightarrow \neg A$ zeigen (z.B. statt 'Wenn x kein Quadrat ist, dann $x \leq 0$' zeigt man 'Wenn $x > 0$, dann x ein Quadrat').

1.1.5 Sprachliche Konventionen

Wir einigen uns auf folgende Konventionen

(i) *Ein* bedeutet stets „mindestens ein" und ist von „genau ein" zu unterscheiden.

(ii) In einer Auzählung von Objekten $x_1, \ldots, x_n$ heißen $x_1, \ldots, x_n$ *paarweise verschieden*, wenn keine zwei Objekte der Aufzählung gleich sind (d.h. wenn in der Aufzählung keine Wiederholungen vorkommen). Davon zu unterscheiden ist „verschieden" im Sinne von „nicht alle gleich". Wenn wir von „n verschiedenen Objekten $x_1, \ldots, x_n$" sprechen, impliziert das, dass $x_1, \ldots, x_n$ paarweise verschieden sind.

1.2 Mengen

1.2.1 Definition und Beispiele

> "Unter einer *Menge* verstehen wir jede Zusammenfassung M von bestimmten wohlunterscheidbaren Objekten unserer Anschauung oder unseres Denkens [welche die *Elemente* von M genannt werden] zu einem Ganzen."
>
> Georg Cantor, 1895

Bei der Auslegung von Cantor's Begriff einer „Zusammenfassung" ist allerdings Vorsicht geboten. Das wusste schon Cantor selbst und zeigte, dass die Betrachtung der „Menge aller Mengen" zu einem Widerspruch führt: nach der *Zweiten Cantor'schen Antinomie* wäre sie „größer" als sie selbst. Man kann auch ohne Betrachtung der „Größe" einer Menge einen rein logischen Widerspruch aus der „Menge aller Mengen" ableiten, die *Russel'sche Antinomie* (siehe Übung b unten). Wir einigen uns auf die folgende Definition des Mengenbegriffs.

Definition a. Eine *Menge* M ist etwas, zu dem jedes beliebige Objekt x entweder *Element* der Menge ist, geschr. $x \in M$, oder nicht, geschr. $x \notin M$.

Mengen sind also gerade dadurch gekennzeichnet, dass '$x \in M$' für jedes Objekt x eine Aussage ist (einen eindeutigen Wahrheitswert hat), also gerade dadurch, dass '$x \in M$' eine Aussageform ist. Umgekehrt ist für jede Aussageform $A(x)$ die Zusammenfassung aller x, für die $A(x)$ wahr ist, eine Menge (vgl. Schreibweise (iii) unten).

Bemerkung. Mengen, die sich selbst enthalten führen nicht per se zu einem Widerspruch. In der weitverbreitetsten Mengenlehre (der *Zermelo-Fraenkel-Mengenlehre*), der wir uns anschliessen wollen, sind Mengen, die sich selbst als Elemente enthalten, allerdings nicht erlaubt.

Definition b. Sind M, N zwei Mengen, so heißt N eine *Teilmenge* von M und M eine *Obermenge* von N, geschr. $N \subseteq M$, wenn für alle $x \in N$ gilt: $x \in M$. Das Zeichen $\subseteq$ bzw. die Aussage $N \subseteq M$ heißt *Inklusion*.

Zwei Mengen M und N heißen *gleich*, geschr. $M = N$, wenn $M \subseteq N$ und $N \subseteq M$.

Eine Menge M heißt *endlich*, wenn M nur endlich viele Elemente besitzt. Man schreibt in diesem Fall $|M|$ für die Anzahl der Elemente von M. Anderenfalls heißt M *unendlich* und man schreibt $|M| = \infty$. Die Zahl $|M|$ heißt die *Mächtigkeit* von M.

Schreibweise.

(i) *Aufzählen.* Die Elemente werden aufgelistet und mit Mengenklammern eingeschlossen. Reihenfolge und Wiederholungen spielen bei der Mengenaufzählung keine Rolle, z.B.

$$\{1, 3, 17\} = \{3, 1, 17\} = \{1, 3, 17, 1, 3\}.$$

(ii) *Beschreiben.* Mengen können durch Worte beschrieben werden, etwa:

$$\text{Menge der natürlichen Zahlen} = \{1, 2, 3, 4, 5, \ldots\}$$
$$\text{Menge der ganzen Zahlen} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$$

(iii) *Aussondern.* Es sei M eine Menge. Ist $A(x)$ eine Aussageform, so bezeichnet

$$\{x \in M \mid A(x)\}$$

diejenige Teilmenge von M, die aus allen Elementen besteht, für die $A(x)$ wahr ist (gespr. „Menge aller x aus M mit $A(x)$"). Benennen wir beispielsweise die Menge der natürlichen Zahlen mit $\mathbb{N}$, so ist $\{n \in \mathbb{N} \mid n \text{ ist ungerade}\}$ die Menge der ungeraden natürlichen Zahlen, also $\{1, 3, 5, 7, \ldots\}$.

(iv) *Abbilden.* Ist M eine Menge und $e(x)$ für jedes $x \in M$ ein *Ausdruck* (z.B. $e(x) = x^2, |x|$, etc.), so bezeichnet

$$\{e(x) : x \in M\}$$

die Menge aller Ausdrücke $e(x)$, wobei x alle Elemente der Menge M durchläuft. Beispielsweise ist $\{n^2 : n \in \mathbb{N}\}$ die Menge der Quadratzahlen. Abbilden und Aussondern können kombiniert werden, sodass z.B. $\{n^2 : n \in \mathbb{N} \,|\, n \text{ ungerade}\}$ die Menge aller Quadrate von ungeraden natürlichen Zahlen bezeichnet, also $\{1, 9, 25, 49, \ldots\}$. In der Regel wird letztere Menge auch kürzer als $\{n^2 \,|\, n \in \mathbb{N} \text{ ungerade}\}$ geschrieben.

Beispiel. Häufig auftretende Mengen sind:

Symbol	Beschreibung	Definition
$\emptyset$	leere Menge	$\{\}$
$\underline{n}$	n-elementige Menge	$\{1, 2, \ldots, n\}$
$\mathbb{N}$	natürliche Zahlen	$\{1, 2, 3, \ldots\}$
$\mathbb{N}_0$	natürliche Zahlen einschl. 0	$\{0, 1, 2, 3, \ldots\}$
$\mathbb{P}$	Primzahlen	$\{2, 3, 5, 7, 11, 13, \ldots\}$
$\mathbb{Z}$	ganze Zahlen	$\{\ldots, -2, -1, 0, 1, 2, \ldots\}$
$\mathbb{Q}$	rationale Zahlen	$\{\frac{a}{b} : a \in \mathbb{Z}, b \in \mathbb{N}\}$
$\mathbb{R}$	reelle Zahlen	$\{a_1 a_2 \ldots a_r, b_1 b_2 \ldots : a_i, b_i \in \{0, 1, \ldots, 9\}\}$
$\mathbb{R}_{>0}$	positive reelle Zahlen	$\{x \in \mathbb{R} \,|\, x > 0\}$
$\mathbb{R}_{\geq 0}$	nicht-negative reelle Zahlen	$\{x \in \mathbb{R} \,|\, x \geq 0\}$
$\mathbb{C}$	komplexe Zahlen	$\{a + bi : a, b \in \mathbb{R}\}$

Nur die ersten beiden Mengen der Tabelle sind endlich, nämlich $|\emptyset| = 0$ und $|\underline{n}| = n$ für alle $n \in \mathbb{N}_0$. Es gilt:

$$\emptyset = \underline{0} \subseteq \underline{1} \subseteq \underline{2} \subseteq \ldots \subseteq \mathbb{N} \subseteq \mathbb{N}_0 \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

Übung a. Was gilt für eine Menge M: $x \in M \,\text{XOR}\, x \notin M$ für alle x? $x \in M \Leftrightarrow \neg(x \notin M)$? $\neg(x \in M) \Leftrightarrow x \notin M$?

Übung b (Russel's Antinomie). Die „Menge aller Mengen" würde als Teilmenge enthalten die „Menge" $\mathcal{M}$ aller Mengen, die sich nicht selbst als Element enthalten. Ist dann $\mathcal{M} \in \mathcal{M}$ oder $\mathcal{M} \notin \mathcal{M}$?

1.2.2 Quantifizierte Aussagen

Es sei $A(x)$ eine Aussageform. Nach Definition (1.1.3) ist $A(x)$ für jedes x eine Aussage. Setzt man in $A(x)$ für x in ein konkretes Objekt ein, so sagt

man, x wird *spezifiziert*. Zwei weitere Möglichkeiten, aus $A(x)$ eine Aussage zu machen, bestehen darin, x zu *quantifizieren*:

'Für alle $x \in M$ gilt $A(x)$' und 'Es gibt ein $x \in M$, für das $A(x)$ gilt'.

Diese sprachlichen Ausdrücke sind Aussagen, denn x ist keine (freie) Variable mehr!

Beispiel.

(i) Sei $A(x)$ die Aussageform '$x > 5$'. Dann ist 'Es existiert ein $x \in \mathbb{N}$ mit $A(x)$' wahr, weil z.B. $A(7)$ wahr ist. Dagegen ist 'Für alle $x \in \mathbb{N}$ gilt $A(x)$' falsch, weil z.B. $A(2)$ falsch ist.

(ii) Sei $A(t)$ die Aussageform 'Zum Zeitpunkt t gilt: Projektor ist aus $\Rightarrow$ Hörsaal ist leer'. Ist t ein konkreter Zeitpunkt, an dem der Projektor an ist oder der Hörsaal leer, so ist die Aussage $A(t)$ wahr. Da es solche Zeitpunkte gibt, ist 'Es gibt eine Zeit t mit $A(t)$' wahr. Ist t dagegen ein konkreter Zeitpunkt, an dem der Projektor aus ist und der Hörsaal nicht leer, so ist die Aussage $A(t)$ falsch. Da es auch solche Zeitpunkte gibt, ist auch 'Es gibt eine Zeit t mit $\neg A(t)$' wahr und 'Für alle Zeiten t gilt $A(t)$' falsch.

(iii) Die Verneinung von 'Für alle $x \in M$ gilt $A(x)$' lässt sich als 'Es existiert $x \in M$ mit $\neg A$' bzw. 'Es existiert $x \in M$ für das A nicht gilt' formulieren. Die Verneinung von 'Für alle $x \in \mathbb{R}$ gilt $x^2 > 0$' lässt sich als 'Es existiert ein $x \in \mathbb{R}$ mit $x^2 \leq 0$' formulieren.

(iv) Die Verneinung von 'Es existiert ein $x \in M$ mit $A(x)$' lässt sich als 'Für alle $x \in M$ gilt $\neg A$' formulieren. Die Verneinung von 'Es gibt eine Person im Hörsaal, die ihr Handy aus hat' lässt sich als 'Alle Personen im Hörsaal haben ihr Handy an' formulieren.

Bemerkung. Gelegentlich schreibt man (missbräuchlich) nur eine Aussageform $A(x)$ auf, meint damit aber die Aussage 'Für alle $x \in M$ gilt $A(x)$'. Das geht nur, wenn die Menge M aus dem Zusammenhang klar ist.

Übung. Wie lautet der Wahrheitswert der Aussagen 'Für alle $x \in \emptyset$ gilt $A(x)$' und 'Es gibt $x \in \emptyset$ mit $A(x)$'?

1.2.3 Konstruktion von Mengen

Definition (Mengenoperationen)**.** Es seien M, N beliebige Mengen.

(i) $M \cap N := \{x \in M \mid x \in N\}$ heißt *Durchschnitt* von M und N.

(ii) $M \cup N := \{x \mid x \in M \text{ oder } x \in N\}$ heißt *Vereinigung* von M und N.

(iii) $M \backslash N := \{x \in M \mid x \notin N\}$ heißt die *Differenzmenge*, gespr. „M ohne N“.

(iv) $M \times N := \{(x, y) \mid x \in M \text{ und } y \in N\}$ heißt *kartesisches Produkt* von M und N.
Hierbei ist (x, y) ein *geordnetes Paar*. Zwei geordnete Paare (x, y) und (x', y') sind genau dann gleich, wenn $x = x'$ und $y = y'$.

(v) $M^n := \{(x_1, \ldots, x_n) \mid x_1, \ldots, x_n \in M\}$ heißt *n-faches kartesisches Produkt* von M ($n \in \mathbb{N}$).
Hierbei ist $(x_1, \ldots, x_n)$ ein *n-Tupel* über M. Zwei n-Tupel $(x_1, \ldots, x_n)$ und $(y_1, \ldots, y_n)$ über M sind genau dann gleich, wenn $x_i = y_i$ für jedes $i \in \{1, \ldots, n\}$. Die $x_1, \ldots, x_n$ heißen die *Einträge* des Tupels.

(vi) $\mathrm{Pot}(M) := \{S \mid S \subseteq M\}$ heißt die *Potenzmenge* von M.

Beispiel.

(i) Die leere Menge ist Teilmenge jeder beliebigen Menge (auch von sich selbst).

(ii) Es gilt:

$$\begin{aligned} \mathrm{Pot}(\emptyset) &= \{\emptyset\}, \\ \mathrm{Pot}(\{1\}) &= \{\emptyset, \{1\}\}, \\ \mathrm{Pot}(\{1, 2\}) &= \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}, \\ &\vdots \end{aligned}$$

(iii) Für jede Menge M gilt $M^2 = M \times M$.

(iv) Für jede Menge M gilt $M^n = \underbrace{M \times \cdots \times M}_{n\text{-mal}}$.

(v) Ein Element von $\mathbb{Z}^5$ ist z.B. $(1, -3, 0, 0, 27)$.

Übung.

(i) Wieviele Elemente hat $\mathrm{Pot}(\underline{n})$ für $n \in \mathbb{N}_0$?

(ii) Wieviele Elemente hat M^n für $n \in \mathbb{N}$?

1.2.4 Mengenpartitionen

Definition.

(i) Zwei Mengen A, B heißen *disjunkt*, wenn $A \cap B = \emptyset$.

(ii) Endlich viele Mengen $M_1, \ldots, M_n$ heißen *paarweise disjunkt*, wenn je zwei davon disjunkt sind, d.h. wenn für alle $i, j \in \{1, \ldots, n\}$ mit $i \neq j$ gilt: $M_i \cap M_j = \emptyset$.

(iii) Es sei $\mathcal{M}$ eine Menge von Mengen ($\mathcal{M}$ darf hier unendlich sein). Die Elemente von $\mathcal{M}$ heißen *paarweise disjunkt*, wenn je zwei davon disjunkt sind, d.h. wenn für alle $M, M' \in \mathcal{M}$ mit $M \neq M'$ gilt: $M \cap M' = \emptyset$.

(iv) Es sei M eine Menge. Eine *Partition* von M ist eine Menge $\mathcal{P}$ nichtleerer, paarweise disjunkter Teilmengen von M mit $M = \bigcup_{C \in \mathcal{P}} C$. Die Elemente $C \in \mathcal{P}$ heißen *Teile* der Partition.

Bemerkung. Für jede Partition $\mathcal{P}$ von M ist $\mathcal{P} \subseteq \mathrm{Pot}(M) \setminus \{\emptyset\}$.

Beispiel.

(i) $\mathcal{P} = \{\{n \in \mathbb{N} \mid n \text{ gerade}\}, \{n \in \mathbb{N} \mid n \text{ ungerade}\}\}$ stellt eine Partition von $\mathbb{N}$ mit zwei Teilen dar.

(ii) $\mathcal{P} = \{\{n \in \mathbb{N} \mid n \text{ hat } k \text{ Dezimalstellen}\} \mid k \in \mathbb{N}\}$ stellt eine Partition von $\mathbb{N}$ mit unendlich vielen Teilen dar.

(iii) Die einzige Partition von $\emptyset$ ist $\mathcal{P} = \emptyset$.

Übung. Man mache sich klar:

(i) Sind A, B endliche, disjunkte Mengen, so gilt $|A \cup B| = |A| + |B|$.

(ii) Sind $M_1, \ldots, M_n$ endliche, paarweise disjunkte Mengen, so gilt

$$|\bigcup_{i=1}^{n} M_i| = \sum_{i=1}^{n} |M_i|.$$

1.3 Beweisprinzipien

1.3.1 Direkter Beweis

Prinzip. Ziel: $A \Rightarrow B$ ist wahr. Um das Ziel zu zeigen, nehmen wir an, dass A wahr ist und folgern daraus mittels logischer Schlüsse, dass B wahr ist. Wenn das gelungen ist, ist auch $A \Rightarrow B$ wahr.

Beispiel. Für alle $n \in \mathbb{N}$ gilt: n ungerade $\Rightarrow n^2$ ungerade.

Beweis. Sei $n \in \mathbb{N}$ beliebig, sei A die Aussage 'n ist ungerade' und B die Aussage 'n^2 ist ungerade'. Wir nehmen an, A ist wahr, d.h. n ist ungerade. Wir folgern, dass B wahr ist: Da n ungerade ist, existiert ein $k \in \mathbb{N}$ mit $n = 2k - 1$. Dann ist $n^2 = (2k-1)^2 = 4k^2 - 4k + 1 = 2(2k^2 - 2k) + 1$, eine ungerade Zahl. Damit ist gefolgert, dass B wahr ist. Nach dem Beweisprinzip des direkten Beweises ist also $A \Rightarrow B$ wahr. Da $n \in \mathbb{N}$ beliebig gewählt war, gilt dies für alle $n \in \mathbb{N}$. □

Übung. Was passiert, wenn sich aus A ein Widerspruch folgern lässt, A also falsch ist?

1.3.2 Beweis durch Kontraposition

Prinzip. Ziel: $A \Rightarrow B$ ist wahr. Stattdessen zeigen wir, dass $\neg B \Rightarrow \neg A$ wahr ist. Wenn das gelungen ist, ist auch $A \Rightarrow B$ wahr.

Beweis des Prinzips. Dieses Prinzip beruht auf der bekannten Tautologie $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$ aus Beispiel (1.1.4). □

Beispiel. Für alle $n \in \mathbb{N}$ gilt: n^2 gerade $\Rightarrow n$ gerade.

Beweis des Prinzips. Sei $n \in \mathbb{N}$ beliebig, sei A die Aussage 'n^2 ist gerade' und B die Aussage 'n ist gerade'. Wir zeigen, dass $\neg B \Rightarrow \neg A$ wahr ist: Diese Aussage lautet 'n ist ungerade $\Rightarrow n^2$ ist ungerade' und wurde schon in (1.3.1) gezeigt. Damit ist nach dem Beweisprinzip der Kontraposition auch $A \Rightarrow B$ wahr. Da $n \in \mathbb{N}$ beliebig gewählt war, gilt dies für alle $n \in \mathbb{N}$. □

1.3.3 Beweis durch Widerspruch

Prinzip. Ziel: A ist wahr. tir zeigen, dass $\neg A \Rightarrow (B \wedge \neg B)$ wahr ist. Wenn das gelungen ist, ist auch A wahr. ($B \wedge \neg B$ ist hier der Widerspruch und die Aussage B kann frei gewählt werden.)

Beweis des Prinzips. $B \wedge \neg B$ ist stets falsch (vgl. Übung 1.1.4). Wenn $\neg A \Rightarrow (B \wedge \neg B)$ wahr ist, ist also $\neg A$ falsch (vgl. Definition von $\Rightarrow$), d.h. A ist wahr. □

Beispiel. Es sei A die Aussage $\sqrt{2} \notin \mathbb{Q}$.

Beweis. Wir nehmen an, $\neg A$ ist wahr, d.h. $\sqrt{2} \in \mathbb{Q}$. Dann gibt es $n, m \in \mathbb{N}$, die nicht beide gerade sind und $\sqrt{2} = \frac{m}{n}$ erfüllen ($\sqrt{2}$ wird als Bruch geschrieben und dieser gekürzt). Seien solche n, m gewählt. Durch Quadrieren

folgt $2n^2 = m^2$, d.h. m^2 ist gerade. Also ist m gerade nach Beispiel (1.3.2). Sei $k \in \mathbb{N}$ mit $m = 2k$. Dann gilt $2n^2 = m^2 = 4k^2$, also $n^2 = 2k^2$, d.h. n^2 ist gerade. Also ist n gerade nach Beispiel (1.3.2). Insgesamt wurde gezeigt, dass sowohl n als auch m gerade sind. Das ist ein Widerspruch (die Aussage B kann hier 'n und m sind nicht beide gerade' gewählt werden). Also ist die Annahme $\sqrt{2} \in \mathbb{Q}$ falsch, und damit ist die Behauptung $\sqrt{2} \notin \mathbb{Q}$ wahr. □

1.3.4 Vollständige Induktion

Prinzip. Ziel: Für alle $n \in \mathbb{N}$ gilt $A(n)$.
Wir zeigen als *Induktionsanfang*, dass $A(1)$ wahr ist, und als *Induktionsschritt*, dass $A(n) \Rightarrow A(n+1)$ für alle $n \in \mathbb{N}$ wahr. Dann ist $A(n)$ für alle $n \in \mathbb{N}$ wahr. Man spricht präziser von einer vollständigen Induktion *über* n. Im Induktionsschritt nennt man die Aussage $A(n)$ die *Induktionsvoraussetzung*.

Beweis des Prinzips. Das Prinzip beruht auf der folgenden Eigenschaft von $\mathbb{N}$, die wir als gegeben annehmen:

> Für jede Teilmenge $A \subseteq \mathbb{N}$ gilt: Ist $1 \in A$ und ist für jedes $n \in A$ auch $n+1 \in A$, dann ist $A = \mathbb{N}$.

Bei der vollständigen Induktion zeigen wir gerade, dass die Menge $A := \{n \in \mathbb{N} \mid A(n) \text{ ist wahr}\}$ diese Bedingung erfüllt, also gleich $\mathbb{N}$ ist. □

Bemerkung. Eine alternative Möglichkeit, die Aussage 'Für alle $n \in \mathbb{N}$ gilt $A(n)$' zu zeigen, wäre, ein *beliebiges* $n \in \mathbb{N}$ zu wählen und dann $A(n)$ mit einem der Prinzipien (1.3.1)–(1.3.3) zu beweisen. (Genau so wurde in Beispiel (1.3.1) und (1.3.2) vorgegangen.) Da vollständige Induktion nur für $\mathbb{N}$ möglich ist, ist diese Alternative sogar der einzige Weg, um Aussagen 'Für alle $x \in M$ gilt $A(x)$' zu zeigen, bei denen die Menge M „größer" als $\mathbb{N}$ ist.

Beispiel. Für alle $n \in \mathbb{N}$ gilt $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.

Beweis. Wir führen eine vollständige Induktion über n. Sei also $A(n)$ die Aussageform $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.

Induktionsanfang: Es ist $\sum_{i=1}^{1} i = 1 = \frac{1 \cdot 2}{2}$, d.h. $A(1)$ ist wahr.

Induktionsschritt: Sei jetzt $n \in \mathbb{N}$ beliebig. Wir zeigen $A(n) \Rightarrow A(n+1)$ mittels eines direkten Beweises. Wir nehmen an, dass $A(n)$ wahr ist, d.h. dass $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$ gilt. Dieses ist die Induktionsvoraussetzung (kurz IV).

Dann ist

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

Der Induktionsschritt ist damit erledigt, weil dies genau die Aussage $A(n+1)$ ist. □

1.4 Abbildungen

1.4.1 Definition und Beispiele

Definition. Seien M, N Mengen. Eine *Abbildung f von N nach M* ist eine „Vorschrift“ (z.B. eine Formel), die jedem $x \in N$ genau ein Element $f(x) \in M$ zuordnet, geschr.

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

Es heißen: N der *Definitionsbereich* von f, M der *Zielbereich* oder *Wertebereich* von f, $f(x)$ das *Bild* von x unter f, x ein *Urbild* von $f(x)$ unter f.

Zur Angabe einer Abbildung gehört die Angabe von Definitions- und Zielbereich dazu, d.h. zwei Abbildungen $f : N \to M$ und $g : N' \to M'$ sind nur dann gleich, wenn $N = N', M = M'$ und $f(x) = g(x)$ für alle $x \in N$.

Die Menge aller Abbildungen von N nach M wird mit $\mathrm{Abb}(N, M)$ oder mit M^N bezeichnet.

Beispiel.

(i) $f : \mathbb{N} \to \mathbb{R}, i \mapsto i^2$.

(ii) Es sei M eine Menge von Glasperlen , und sei F die Menge aller Farben. Dann gibt es die Abbildung $f : M \to F, x \mapsto$ Farbe von x.

(iii) Für jede Menge A von Personen gibt es die Abbildung $J : A \to \mathbb{Z}, p \mapsto$ Geburtsjahr von p.

(iv) Die Addition in $\mathbb{Z}$ kann als die Abbildung

$$\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, \quad (x, y) \mapsto x + y$$

aufgefasst werden.

(v) Für jede Menge M gibt es die *Identitätsabbildung* $\mathrm{id}_M : M \to M, x \mapsto x$.

(vi) Betrachten wir die Abbildungen

$$\begin{aligned} f &: \mathbb{R} \to \mathbb{R}, x \mapsto \sqrt{x^2}, \\ g &: \mathbb{R} \to \mathbb{R}, x \mapsto |x|, \\ h &: \mathbb{R} \to \mathbb{R}_{\geq 0}, x \mapsto |x|, \end{aligned}$$

so ist $f = g \neq h$.

(vii) $\mathrm{Abb}(\mathbb{R}, \mathbb{R}) = \{\mathbb{R} \to \mathbb{R}\} =$ Menge aller reellen Funktionen.

(viii) Für jede Menge M existiert genau eine Abbildung $\emptyset \to M$.

(ix) Für jede nicht-leere Menge N existiert keine Abbildung $N \to \emptyset$.

Bemerkung.

(i) Eine Abbildung $f : \mathbb{N} \to M$ wird auch *Folge in M* genannt. Oft benutzt man für Folgen die Schreibweise $a_1, a_2, a_3, ...$ oder $(a_i)_{i \in \mathbb{N}}$, wobei a_i für das Bild $f(i) \in M$ steht. Die Folge aus Beispiel (i) würde also auch geschrieben als $1, 4, 9, 16, ...$ oder als $(i^2)_{i \in \mathbb{N}}$.

Die Menge aller Folgen in M wird daher auch als $\mathrm{Abb}(\mathbb{N}, M)$ oder $M^{\mathbb{N}}$ geschrieben. Beispielsweise ist $2^{\mathbb{N}}$ die Menge der Binärfolgen, $\mathbb{R}^{\mathbb{N}}$ die Menge der reellen Folgen, usw.

(ii) Ein n-Tupel $(x_1, \ldots, x_n)$ über M kann als eine Abbildung $t : \underline{n} \to M$ aufgefasst werden durch die Zurordnung $i \mapsto x_i$. Das 5-Tupel $(1, -3, 0, 0, 27)$ über $\mathbb{Z}$ ist z.B. die Abbildung $t : \underline{5} \to \mathbb{Z}$ mit $t(1) = 1, t(2) = -3, t(3) = t(4) = 0, t(5) = 27$.

Übung. Bestimmen Sie $|\mathrm{Abb}(N, M)|$ für endliche Mengen N und M.

1.4.2 Bild und Urbild

Definition. Es sei $f : N \to M$ eine Abbildung.

(i) Für jede Teilmenge $X \subseteq N$ heißt $f(X) := \{f(x) \mid x \in X\}$ das *Bild von* X unter f.

(ii) Das Bild $f(N)$ von N unter f wird schlicht das *Bild* oder die *Bildmenge* von f genannt.

(iii) Für jede Teilmenge $Y \subseteq M$ heißt $f^{-1}(Y) := \{x \in N \mid f(x) \in Y\}$ das *Urbild von* Y unter f.

(iv) Die Mengen $f^{-1}(\{y\})$ mit $y \in M$ heißen die *Fasern von* f.

Die Schreibweise f^{-1} für das Urbild hat im Allgemeinen nichts mit Umkehrabbildungen zu tun.

Beispiel. Die Faser der Abbildung J aus Beispiel (1.4.1) zu 1900 ist die Menge aller Personen, die im Jahr 1900 geboren sind.

Bemerkung a. Für eine Abbildung $f : N \to M$ zwischen endlichen Mengen N und M gilt stets $|f(N)| \leq |N|$ und $|f(N)| \leq |M|$.

Bemerkung b. Die nicht-leeren Fasern einer Abbildung bilden eine Partition des Definitionsbereichs.

1.4.3 Injektive und surjektive Abbildungen

Definition. Es sei $f : N \to M$ eine Abbildung.

(i) f heißt *surjektiv*, falls $f(N) = M$.

(ii) f heißt *injektiv*, falls für alle $x, x' \in N$ gilt: $f(x) = f(x') \Rightarrow x = x'$.

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

Bemerkung a. Eine Abbildung $f : N \to M$ ist per Definition injektiv, surjektiv bzw. bijektiv, wenn jedes Element $y \in M$ höchstens ein, mindestens ein bzw. genau ein Urbild hat. Das ist genau dann der Fall, wenn alle Fasern von f höchstens ein, mindestens ein bzw. genau ein Element besitzen, also genau dann, wenn für jedes $y \in M$ die Gleichung $f(x) = y$ höchtens eine, mindestens eine bzw. genau eine Lösung $x \in N$ hat.

Beispiel.

(i) $f : \mathbb{Z} \to \mathbb{Z}, z \mapsto 2z$ ist injektiv, aber nicht surjektiv.

(ii) $f : \mathbb{R} \to \mathbb{R}, x \mapsto 2x$ ist bijektiv.

(iii) $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ ist weder injektiv noch surjektiv. In der Tat ist $f(\mathbb{R}) = \mathbb{R}_{\geq 0}$, also f nicht surjektiv. Weiter ist z.B. $f(2) = 4 = f(-2)$ aber $2 \neq -2$, folglich ist f nicht injektv.

(iv) Es sei $f : M \to F$ die Abbildung aus Beispiel (1.4.1)(ii). Die Faser $f^{-1}(\{\text{rot}\})$ ist die Menge der roten Perlen in M. Es ist f genau dann injektv, wenn von jeder Farbe höchstens eine Perle in M vorkommt, wenn also keine zwei Perlen aus M die gleiche Farbe haben. Weiter ist f genau dann surjektv, wenn von jeder Farbe (mindestens) eine Perle in M vorkommt.

(v) Die Abbildung $\emptyset \to M$ ist injektiv. Sie ist genau dann surjektiv, wenn $M = \emptyset$.

(vi) Hashfunktionen (bzw. „Checksummen“ oder „Fingerprints“), z.B. die bekannte md5 : $\{\text{Texte}\} \to \underline{2}^{128}$, die einen 128 bit–Hashwert produziert, sind nicht injektiv (da verschiedene Texte gleichen Hashwert haben können), sind surjektiv (um alle Hashwerte auszunutzen), und haben „gleich große“ Fasern (das macht gerade eine gute Hashfunktion aus!).

(vii) Verschlüsselungsfunktionen, etwa crypt : $\underline{2}^k \to \underline{2}^k$, sind injektiv, damit eine eindeutige Entschlüsselung möglich ist.

Bei Abbildungen zwischen endlichen Mengen bestehen weitere Zusammenhänge gemäß

Bemerkung b. Es sei $f : N \to M$ eine Abbildung zwischen endlichen Mengen N und M. Dann gilt $|f(N)| \leq |N|$ und $|f(N)| \leq |M|$. Weiter ist

(i) f genau dann injektiv, wenn $|f(N)| = |N|$,

(ii) f genau dann surjektiv, wenn $|f(N)| = |M|$.

In dem wichtigen Spezialfall $|N| = |M|$, z.B. wenn $N = M$, folgt aus injektiv also schon surjektiv und umgekehrt.

Übung. (i) Man mache sich klar, dass eine Abbildung $f : N \to M$ genau dann injektiv ist, wenn für alle $x_1, \ldots, x_r \in N$ gilt:

$x_1, \ldots, x_r$ paarweise verschieden $\Leftrightarrow f(x_1), \ldots, f(x_r)$ paarweise verschieden.

(ii) Man folgere aus Bemerkung b, dass für eine injektive Abbildung $f : N \to M$ stets $|N| \leq |M|$, und für eine surjektive Abbildung $f : N \to M$ stets $|N| \geq |M|$.

1.4.4 Einschränkung

Definition. Es sei $f : N \to M$ eine Abbildung und $N' \subseteq N$. Dann heißt die Abbildung

$$f|_{N'} : N' \to M, \quad x \mapsto f(x)$$

die *Einschränkung* von f auf N'.

Bemerkung. Jede Abbildung kann durch Einschränkung auf eine geeignete Teilmenge des Definitionsbereiches injektiv gemacht werden. Z.B. ist für f aus Beispiel (1.4.3)(iii) die Einschränkung $f|_{\mathbb{R}_{\geq 0}}$ injektiv, ebenso wie die Einschränkung $f|_{\mathbb{R}_{\leq 0}}$.

1.4.5 Kombinatorische Strukturen als Abbildungen

Tupel, Permutationen, Kombinationen und Multimengen (Definition erst in Kapitel 3) können mit Abbildungen bestimmter Art identifiziert werden.

Beispiel.

(i) Ein k-Tupel über A ist eine Abbildung $\underline{k} \to A$. Das Tupel $(a_1, \ldots, a_k)$ entspricht der Abbildung $f : \underline{k} \to A, i \mapsto a_i$.

(ii) Eine k-Permutation aus A ist eine injektive Abbildung $\underline{k} \to A$. Die Permutation $(a_1, \ldots, a_k)$ entspricht der Abbildung $f : \underline{k} \to A, i \mapsto a_i$.

(iii) Ist $|A| = n < \infty$, so ist eine Permutation aus A eine bijektive Abbildung $\underline{n} \to A$. Die Permutation $(a_1, \ldots, a_n)$ entspricht der Abbildung $f : \underline{n} \to A, i \mapsto a_i$.

(iv) Eine k-Kombination aus A ist eine Abbildung $A \to \{0, 1\}$ mit $|f^{-1}(\{1\})| = k$ (die Faser zu 1 hat k Elemente). Die Kombination $M \subseteq A$ entspricht der Abbildung $f : A \to \{0, 1\}$ mit $f(a) = 0$ falls $a \notin M$ und $f(a) = 1$ falls $a \in M$. Die Abbildung f bezeichnet man auch als *charakteristische Funktion* von M.

(v) Eine k-Multimenge ist eine Abbildung $A \to \mathbb{N}_0$ mit $\sum_{a \in A} f(a) = k$. Die Multimenge $M \subseteq A$ entspricht der Abbildung $f : A \to \mathbb{N}_0$, wobei $f(a)$ angibt, wie oft a in M vorkommt. Die Abbildung f wird als *Häufigkeitsfunktion* von M bezeichnet.

Das **Schubfachprinzip** kann mit Abbildungen so formuliert werden: Sei $f : N \to M$ eine Abbildung zwischen endlichen Mengen. Dann gilt

$$|N| > |M| \Rightarrow f \text{ nicht injektiv.}$$

Übung. Eine k-elementige Teilmenge M von A kann als k-Kombination oder als k-Multimenge aufgefasst werden. Vergleichen Sie die charakteristische Funktion von M mit der Häufigkeitsfunktion von M.

Übung. Bestimmen Sie die Anzahl injektiver Abbildungen von $\underline{n}$ nach $\underline{m}$.

Übung. Sei $f : N \to M$ eine Abbildung zwischen endlichen Mengen. Dann gilt

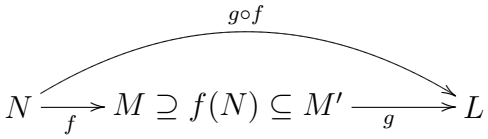
$$|N| < |M| \Rightarrow f \text{ nicht surjektiv.}$$

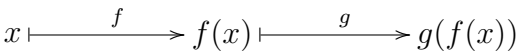
1.4.6 Komposition

Definition. Es seien N, M, M', L Mengen. Weiter seien $f : N \to M$ und $g : M' \to L$ zwei Abbildungen mit $f(N) \subseteq M'$. Dann heißt die Abbildung

$$g \circ f : N \to L, \quad x \mapsto (g \circ f)(x) := g(f(x))$$

die *Komposition von g mit f*. (Häufig ist $M = M'$.)





Beispiel. Für die Abbildungen

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto (x-3)^2,$$
$$g : \mathbb{R}_{\geq 0} \to \mathbb{R}, \quad x \mapsto \sqrt{x}$$

ergeben sich die Kompositionen

$$g \circ f : \mathbb{R} \to \mathbb{R}, x \mapsto \sqrt{(x-3)^2} = |x-3|,$$
$$f \circ g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \quad x \mapsto (\sqrt{x}-3)^2$$

Bemerkung. Es seien f, g, h Abbildungen.

(i) Es gilt $(h \circ g) \circ f = h \circ (g \circ f)$, sofern beide Seiten der Gleichung definiert sind. Daher kann die Komposition auch ohne Klammern kurz als $h \circ g \circ f$ geschrieben werden.

(ii) Wenn $g \circ f$ definiert ist, muß im Allgemeinen nicht $f \circ g$ definiert sein.

Übung.

(i) Man überlege sich ein Beispiel für drei Abbildungen f, g, h, in dem $h \circ (g \circ f)$ definiert ist, aber $(h \circ g) \circ f$ nicht.

(ii) Man überlege sich ein Beispiel, in dem $f \circ g = g \circ f$ bzw. $f \circ g \neq g \circ f$ gilt.

1.4.7 Umkehrabbildungen

Definition. Es seien $f : N \to M$ und $g : M \to N$ Abbildungen. Dann heißt g eine *linksseitige (rechtsseitige) Umkehrabbildung von* f, wenn $g \circ f = \mathrm{id}_N$ (wenn $f \circ g = \mathrm{id}_M$). Wir sprechen schlicht von einer *Umkehrabbildung von* f, wenn g sowohl links- als auch rechtsseitige Umkehrabbildung von f ist.

Satz a. *Sei $f : N \to M$ eine Abbildung und sei N nicht leer.*

(i) f besitzt genau dann eine linksseitige Umkehrabbildung, wenn f injektiv ist.

(ii) f besitzt genau dann eine rechtsseitige Umkehrabbildung, wenn f surjektiv ist.

(iii) f besitzt genau dann eine Umkehrabbildung, wenn f bijektiv ist.

Bemerkung. Existiert eine Umkehrabbildung, so ist sie eindeutig bestimmt (Übung). Links- und rechtsseitige Umkehrabbildungen sind im Allgemeinen nicht eindeutig (Beispiel unten).

Schreibweise. Falls f bijektiv ist, so wird die eindeutige Umkehrabbildung mit f^{-1} bezeichnet. Die ist nicht zu verwechseln mit dem Urbild, dass ebenfalls mit f^{-1} bezeichnet wird. Was gemeint ist, ergibt sich aus dem Zusammenhang.

Beweis. (i) Es sind zwei Richtungen zu zeigen, wir zeigen zuerst den „wenn“-Teil. Dazu nehmen wir an, f sei injektiv und konstruieren eine linksseitige Umkehrabbildung g. Wähle $x_0 \in N$ beliebig ($N \neq \emptyset$) und definiere $g : M \to N$ durch

$$g(y) := \begin{cases} x & \text{falls } y = f(x) \text{ für ein } x \in N, \\ x_0 & \text{falls } y \notin f(N), \end{cases}$$

Das x in der ersten Zeile ist eindeutig, da f injektiv ist, also ist g wohldefiniert. Damit gilt $(g \circ f)(x) = g(f(x)) = x$ für alle $x \in N$, d.h. $g \circ f = \mathrm{id}_N$ wie gewünscht.

Wir zeigen jetzt die andere Richtung, den „genau dann“-Teil. Dazu nehmen wir an, $g : M \to N$ sei eine linksseitige Umkehrabbildung und folgern, dass f injektiv ist. Aus $g \circ f = \mathrm{id}_N$ folgt, dass für alle $x, x' \in N$ gilt:

$$f(x) = f(x') \Rightarrow g(f(x)) = g(f(x')) \Rightarrow \underbrace{(g \circ f)}_{=\mathrm{id}_N}(x) = \underbrace{(g \circ f)}_{=\mathrm{id}_N}(x') \Rightarrow x = x'.$$

Also ist f tatsächlich injektiv und der Beweis beendet.
(ii), (iii) siehe Vorlesung. □

Beispiel.

(i) $f : \mathbb{R} \to \mathbb{R}, x \mapsto 2x$ ist bijektiv mit der Umkehrabbildung

$$f^{-1} : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \frac{1}{2}x$$

(ii) $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 1}, x \mapsto x^2 + 1$ ist bijektiv mit der Umkehrabbildung

$$f^{-1} : \mathbb{R}_{\geq 1} \to \mathbb{R}_{\geq 0}, \quad x \mapsto \sqrt{x - 1}$$

(iii) $f : \mathbb{Z} \to \mathbb{Q}, a \mapsto a$ ist injektiv, aber nicht surjektiv.

$$\begin{aligned} g \circ f = \mathrm{id}_{\mathbb{Z}} &: \text{z.B. } g(x) := \lfloor x \rfloor \text{ oder } g(x) := \lceil x \rceil \\ f \circ g = \mathrm{id}_{\mathbb{Q}} &: \text{nicht möglich} \end{aligned}$$

(Hier bezeichnet $\lfloor x \rfloor$ die größte ganze Zahl $\leq x$, und $\lceil x \rceil$ die kleinste ganze Zahl $\geq x$.)

(iv) $f : \mathbb{R} \to \mathbb{R}_{\geq 0}, x \mapsto |x|$ ist surjektiv, aber nicht injektiv.

$$\begin{aligned} g \circ f = \mathrm{id}_{\mathbb{R}} &: \text{nicht möglich} \\ f \circ g = \mathrm{id}_{\mathbb{R}_{\geq 0}} &: \text{z.B. } g(x) := x \text{ oder } g(x) := -x \end{aligned}$$

Satz b. *Es seien $f : N \to M$ und $g : M \to L$ zwei bijektive Abbildungen. Wenn $g \circ f$ definiert ist, so ist $g \circ f$ ebenfalls bijektiv und es gilt:*

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

Beweis. als Übung. □

Übung.

(i) Zeigen Sie die restlichen Teile der Bemerkung.

(ii) Zeigen Sie den Satz.

(iii) Gilt der Satz auch, wenn man bijektiv durch injektiv ersetzt?

(iv) Gilt der Satz auch, wenn man bijektiv durch surjektiv ersetzt?

1.4.8 Mächtigkeit unendlicher Mengen

Definition. N und M heißen *gleichmächtig*, wenn eine bijektive Abbildung $N \to M$ existiert.

Übung a. $\mathbb{N}, \mathbb{Z}$ und $\mathbb{Q}$ sind gleichmächtig.

Satz (Cantor)**.** *Für jede Menge M sind M und* $\mathrm{Pot}(M)$ *nicht gleichmächtig.*

Beweis. Sei f eine beliebige Abbildung $f : M \to \mathrm{Pot}(M)$. Definiere $A_f := \{x \in M \mid x \notin f(x)\} \in \mathrm{Pot}(M)$. Angenommen, es gibt $m \in M$ mit $f(m) = A_f$. Falls $m \in A_f$, so folgt $m \notin f(m) = A_f$ (Widerspruch). Falls $m \notin A_f = f(m)$, so folgt $m \in A_f$ (Widerspruch). Also ist f nicht surjektiv. □

Übung b. Man folgere aus dem Satz:

(i) $\mathbb{N}$ und $\mathbb{R}$ sind nicht gleichmächtig.

(ii) Die Zusammenfassung aller Mengen ist keine Menge.

1.4.9 Abbildungen einer Menge in sich

Sind $f, g : M \to M$ zwei Abbildungen einer Menge M in sich, so kann man stets die Kompositionen $f \circ g$ und $g \circ f$ bilden.

Definition. Es sei $f : M \to M$ eine Abbildung und es sei $n \in \mathbb{N}$. Dann setzen wir

$$f^n := \underbrace{f \circ \ldots \circ f}_{n\text{-mal}}, \quad f^0 := \mathrm{id}_M.$$

Falls f bijektiv ist, so definieren wir auch $f^{-n} := (f^{-1})^n$.

Bemerkung.

(i) Es gilt $f^n(x) = f(f(\cdots f(x)))$.

(ii) Für bijektive Abbildungen einer Menge in sich selbst haben wir die üblichen Potenzrechenregeln:

$$f^{a+b} = f^a \circ f^b \quad \text{und} \quad f^{ab} = (f^a)^b \quad \text{für alle } a, b \in \mathbb{Z}.$$

1.5 Permutationen

1.5.1 Definition und Beispiele

Es sei A eine endliche Menge und $|A| = n$. Wir nummerieren die Elemente von A und schreiben $A = \{a_1, a_2, \ldots, a_n\}$.

Definition. Eine bijektive Abbildung $\pi : A \to A$ heißt *Permutation von A*. Wir verwenden für Permutationen die Schreibweise

$$\pi = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ \pi(a_1) & \pi(a_2) & \cdots & \pi(a_n) \end{pmatrix}.$$

Die Menge aller Permutationen von A wird mit S_A bezeichnet, also

$$S_A := \{\pi : A \to A \,|\, \pi \text{ bijektiv}\}.$$

In dem wichtigen Spezialfall $A = \underline{n}$ schreiben wir kurz S_n statt $S_{\underline{n}}$.

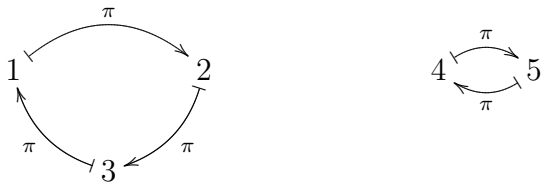
Bemerkung.

(i) Wenn $|A| = n$, dann $|S_A| = n!$. Das gilt auch für $n = 0$, denn $S_\emptyset$ hat genau ein Element (es existiert genau eine Abbildung $\emptyset \to \emptyset$ und die ist bijektiv).

(ii) Die Komposition von Permutationen von A ist wieder eine Permutation von A. Bei Permutationen sagt man statt Komposition auch *Produkt* und läßt das Zeichen $\circ$ einfach weg.

Beispiel.

(i) Die Permutation $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 5 & 4 \end{pmatrix} \in S_5$ läßt sich so veranschaulichen:



(ii) Ist $\psi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 1 & 2 \end{pmatrix}$ und π wie oben dann ergeben sich als Kompositionen

$$\pi \circ \psi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 2 & 3 \end{pmatrix} \quad \text{und} \quad \psi \circ \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 2 & 1 \end{pmatrix}.$$

1.5.2 Der Träger einer Permutation

Definition. Für $\pi \in S_A$ heißt

$$T_\pi := \{a \in A \,|\, \pi(a) \neq a\} \subseteq A$$

der *Träger* von π.

Beispiel.

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 8 & 3 & 6 & 2 & 7 & 4 & 1 & 9 & 11 & 10 \end{pmatrix}, \quad T_\pi = \{1, 2, 4, 5, 6, 7, 8, 10, 11\}.$$

Bemerkung. Es seien $\pi, \psi \in S_A$.

(i) $\pi(T_\pi) = T_\pi$.

(ii) Gilt $T_\pi \subseteq B$, so kann π auch als Element von S_B aufgefasst werden.

(iii) Haben π und ψ disjunkte Träger, so gilt $\pi \circ \psi = \psi \circ \pi$.

Beweis.

(i) Es reicht, die Inklusion $\pi(T_\pi) \subseteq T_\pi$ zu zeigen. Daraus folgt schon die Gleichheit, da es sich um endliche Mengen handelt und da $|\pi(T_\pi)| = |T_\pi|$ wegen der Injektivität von π gilt (vgl. Bem. 1.4.3b). Sei also a ein beliebiges Element aus T_π. Da $\pi(a) \neq a$ und π injektiv, folgt $\pi(\pi(a)) \neq \pi(a)$. Das bedeutet gerade $\pi(a) \in T_\pi$. Da $a \in T_\pi$ beliebig war, ist $\pi(T_\pi) \subseteq T_\pi$ gezeigt.

(ii) klar.

(iii) als Übung.

□

1.5.3 Zykel und Transpositionen

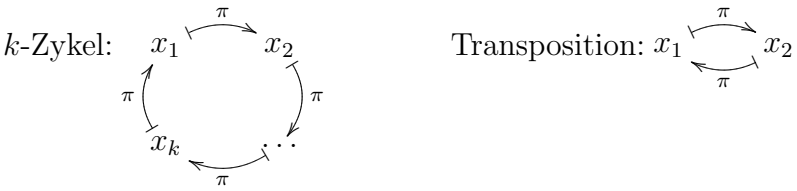
Definition. Es seien $x_1, x_2, \ldots, x_k \in A$ paarweise verschieden. Die Permutation $\sigma \in S_A$ mit

$$\sigma(x) = \begin{cases} x_{i+1} & \text{falls } x = x_i \text{ und } i < k, \\ x_1 & \text{falls } x = x_k, \\ x & \text{falls } x \neq x_1, x_2, \ldots, x_k, \end{cases}$$

heißt *Zykel der Länge k* oder kurz *k-Zykel* von S_A. Wir verwenden für σ die Schreibweise

$$\sigma = (x_1\ x_2\ \ldots\ x_k).$$

Die 2-Zykel heißen auch *Transpositionen* von S_A.



Bemerkung.

(i) Es gilt stets $(x_1\ x_2\ \ldots\ x_k)^k = \mathrm{id}$.

(ii) Es gilt stets $(x_1\ x_2\ \ldots\ x_k)^{-1} = (x_k\ x_{k-1}\ \ldots\ x_1)$.

(iii) Für Transpositionen τ gilt $\tau^{-1} = \tau$.

(iv) Jeder 1-Zykel ist die Identität.

(v) Jeder k-Zykel läßt sich als Produkt von $k-1$ Transpositionen schreiben:

$$(x_1\ x_2\ \ldots\ x_k) = (x_1\ x_2)(x_2\ x_3)\cdots(x_{k-1}\ x_k).$$

Eine solche Zerlegung ist im Allgemeinen nicht eindeutig (vgl. Beispiel a unten).

Beispiel a. Der 4-Zykel $\sigma := (1\ 5\ 2\ 4) \in S_5$ ist die Permutation

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

Es gilt

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 3 & 2 & 1 \end{pmatrix} = (4\ 2\ 5\ 1),$$
$$\sigma^2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 5 & 4 \end{pmatrix} = (1\ 2)(5\ 4),$$
$$\sigma^3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 3 & 1 & 5 \end{pmatrix} = (1\ 4\ 2\ 5),$$
$$\sigma^4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \mathrm{id}.$$

Es gilt

$$\sigma = (1\ 5)(5\ 2)(2\ 4) = (1\ 4)(1\ 2)(1\ 5).$$

Beispiel b. Möchte man eine Liste von n Elementen ordnen (z.B. eine Liste von Wörtern nach alphabetischer Reihenfolge), so ist eine Permutation $\pi \in S_n$ zu finden, die die (ungeordnete) Liste in ihre geordnete Reihenfolge überführt. Das i-te Wort der ungeordneten Liste steht in der geordneten Liste dann an $\pi(i)$-ter Stelle. Ein Sortieralgorithmus findet π im Allgemeinen nicht in einem Schritt, sondern nimmt nacheinander eine Reihe von Vertauschungen vor; er konstruiert somit π als ein Produkt $\pi_1 \circ \ldots \circ \pi_r$ einzelner (einfacherer) Umordnungen π_i. Der *Bubblesort*-Algorithmus kommt dabei z.B. mit Transpositionen π_i aus. Damit das immer funktioniert muss sich jede Permutation als Produkt von Transpositionen schreiben lassen. Davon überzeugen wir uns mit Hilfe von Satz (1.5.4) unten.

1.5.4 Zerlegung in Zykel

Satz. *Jede Permutation $\pi \in S_A$ läßt sich als Produkt von Zykeln schreiben, deren Träger paarweise disjunkt sind. Bis auf Reihenfolge und bis auf Erwähnung von 1-Zykeln ist diese Zerlegung eindeutig.*

Beweis. Siehe Beispiel. □

Man spricht kurz von einer Zerlegung von π in *paarweise disjunkte Zykeln*.

Beispiel a. Für π aus Beispiel (1.5.2) haben wir die Zerlegung

$$\begin{aligned}\pi &= (1\ 5\ 2\ 8)(3)(4\ 6\ 7)(9)(10\ 11)\\ &= (1\ 5\ 2\ 8)(4\ 6\ 7)(10\ 11).\end{aligned}$$

Die Träger der drei Zykeln lauten $\{1,5,2,8\}, \{4,6,7\}, \{10,11\}$ und sind paarweise disjunkt. Die einzelnen Zykeln zerlegen sich weiter in Transpositionen, z.B. $(1\ 5\ 2\ 8) = (1\ 5)(5\ 2)(2\ 8)$ und $(4\ 6\ 7) = (4\ 6)(6\ 7)$, also

$$\pi = (1\ 5)(5\ 2)(2\ 8)(4\ 6)(6\ 7)(10\ 11).$$

Die Zykelschreibweise lässt sich besonders leicht „potenzieren“:

$$\begin{aligned}\pi^{-1} &= (10\ 11)(7\ 6\ 4)(8\ 2\ 5\ 1),\\ \pi^2 &= (1\ 2)(5\ 8)(4\ 7\ 6),\\ \pi^3 &= (1\ 8\ 2\ 5)(10\ 11),\\ \pi^4 &= (4\ 6\ 7)\\ &\vdots\\ \pi^{11} &= (1\ 8\ 2\ 5)(4\ 7\ 6)(10\ 11) = \pi^{-1}\\ \pi^{12} &= \mathrm{id}.\end{aligned}$$

Definition. Es sei $\pi \in S_A$. Die *Zykelzahl* von $\pi \in S_A$ ist die Anzahl der Zykeln inklusive aller 1-Zykeln, die bei einer Zerlegung von π in paarweise disjunkte Zykeln auftreten.

Die Zykelzahl ist gemäß obigem Satz eindeutig bestimmt. Sie hängt allerdings nicht nur von π sondern auch von A ab!

Beispiel b. Die Zykelzahl von π aus Beispiel a ist 5. Da sich die Identität $\mathrm{id} \in S_n$ in lauter 1-Zykeln zerlegt, hat sie die Zykelzahl n. Die Zykelzahl der (einzigen) Permutation $\emptyset \to \emptyset$ wird als 0 definiert.

1.5.5 Das Signum

Wir bezeichnen hier mit I_A die Menge der 2-elementigen Teilmengen einer Menge A, d.h. $I_A = \{\{i,j\} \subseteq A \,|\, i \neq j\}$. Wir schreiben I_n für $I_{\underline{n}}$. (In der Kombinatorik lernen wir, dass $|I_n| = \frac{n(n-1)}{2}$.)

Definition. Sei $\pi \in S_n$. Das *Signum* von π ist definiert als

$$\operatorname{sgn} \pi := \prod_{\{i,j\} \in I_n} \frac{\pi(i) - \pi(j)}{i - j}.$$

Man beachte, dass $\operatorname{sgn} \pi$ wohldefiniert ist, weil sich jeder einzelne Quotient nicht ändert, wenn man i und j vertauscht.

Beispiel a. Für $\pi = \mathrm{id} \in S_n$ sind alle Faktoren des Produktes gleich 1, also $\operatorname{sgn} \mathrm{id} = 1$. Für $n = 2$ und $\pi = (1\ 2)$ ist $I_n = \{\{1,2\}\}$, also $\operatorname{sgn}(1\ 2) = \frac{2-1}{1-2} = -1$.

Bemerkung a.

(i) Es gilt stets $\operatorname{sgn} \pi = \pm 1$.

(ii) Wir nennen π *gerade*, falls $\operatorname{sgn} \pi = 1$ und *ungerade* falls $\operatorname{sgn} \pi = -1$.

(iii) Es gilt $\operatorname{sgn} \pi = \operatorname{sgn} \pi'$ wobei $\pi' := \pi|_{T_\pi} \in S_{T_\pi}$.

Beweis. (i) Da π bijektiv ist, gilt $\{\{\pi(i), \pi(j)\} \subseteq \underline{n} \,|\, i \neq j\} = I_n$. D.h. wenn $\{i,j\}$ die Menge I_n durchläuft, so durchläuft auch $\{\pi(i), \pi(j)\}$ genau die Menge I_n. Folglich sind $\prod_{\{i,j\} \in I_n} (\pi(i) - \pi(j))$ und $\prod_{\{i,j\} \in I_n} (i - j)|$ (Zähler und Nenner) bis auf Vorzeichen gleich, und somit $|\operatorname{sgn} \pi| = 1$.

(iii) Es sei $T = T_\pi$ (der Träger von π) und $F = \underline{n} \setminus T$ (die Fixpunkte von π). Die Menge I_n partitioniert sich in $I_n = I_T \cup I_F \cup \{\{i,j\} \,|\, i \in T, j \in F\}$.

Folglich zerlegt sich das Produkt aus der Definition von $\operatorname{sgn}\pi$ in die drei Teilprodukte

$$\begin{aligned}
\prod_{\{i,j\}\in I_T} \frac{\pi(i)-\pi(j)}{i-j} &= \operatorname{sgn}\pi',\\
\prod_{\{i,j\}\in I_F} \frac{\pi(i)-\pi(j)}{i-j} &= \prod_{\{i,j\}\in I_F} \frac{i-j}{i-j} = 1,\\
\prod_{i\in T, j\in F} \frac{\pi(i)-\pi(j)}{i-j} &= \prod_{j\in F} \underbrace{\prod_{i\in T} \frac{\pi(i)-j}{i-j}}_{=1} = 1.
\end{aligned}$$

Man beachte in der letzten Gleichung, dass wenn i die Menge T durchläuft, dann auch $\pi(i)$ genau die Menge T durchläuft. Damit ist $\operatorname{sgn}\pi = \operatorname{sgn}\pi'$ gezeigt. □

Beispiel b. Für jede Transposition $\pi = (a\ b) \in S_n$ ist $\operatorname{sgn}\pi = -1$.

Beweis. Es ist $T_\pi = \{a, b\}$ und $\pi' = \pi|_{\{a,b\}} \in S_{\{a,b\}}$. Somit gilt

$$\operatorname{sgn}\pi = \operatorname{sgn}\pi' = \frac{\pi(a)-\pi(b)}{a-b} = \frac{b-a}{a-b} = -1.$$

□

Satz. *Für alle $\pi, \psi \in S_n$ gilt* $\operatorname{sgn}(\pi\circ\psi) = \operatorname{sgn}\pi \cdot \operatorname{sgn}\psi$.

Beweis. Es gilt

$$\begin{aligned}
\operatorname{sgn}(\pi\circ\psi) &= \prod_{\{i,j\}\in I_n} \frac{(\pi\circ\psi)(i)-(\pi\circ\psi)(j)}{i-j}\\
&= \prod_{\{i,j\}\in I_n} \left(\frac{\pi(\psi(i))-\pi(\psi(j))}{\psi(i)-\psi(j)} \cdot \frac{\psi(i)-\psi(j)}{i-j}\right)
\end{aligned}$$

Da mit $\{i, j\}$ auch $\{\psi(i), \psi(j)\}$ genau die Menge I_n durchläuft ist dieses Produkt gleich $\operatorname{sgn}\pi \cdot \operatorname{sgn}\psi$. □

Folgerung a. *Für alle $\pi, \psi \in S_n$ gelten:*

(i) $\operatorname{sgn}\pi^{-1} = \operatorname{sgn}\pi$.

(ii) $\operatorname{sgn}(\psi^{-1}\pi\psi) = \operatorname{sgn}\pi$.

Beweis. Der Satz und die Tatsache $\operatorname{sgn}\mathrm{id} = 1$. □

Bemerkung b. Aus Folgerung a(ii) geht hervor, dass eine Umbennung der Elemente von $\underline{n}$ (hier vorgenommen durch ψ) das Signum von π nicht ändert. Die Definition des Signum stellt sich deshalb (nachträglich) als unabhängig von der Nummerierung innerhalb der Menge $\underline{n}$ heraus.

Diese Tatsache kann man ausnutzen, um die Definition des Signum auf Permutationen $\pi \in S_A$ (anstatt nur $\pi \in S_n$) auszudehnen: wähle eine beliebige Bijektion $\varphi : A \to \underline{n}$ und setze $\operatorname{sgn} \pi := \operatorname{sgn}(\varphi \circ \pi \circ \varphi^{-1})$ (beachte $\varphi \circ \pi \circ \varphi^{-1} \in S_n$).

Folgerung b. *Es sei $\pi \in S_A$.*

(i) Ist $\pi = \tau_1 \circ \cdots \circ \tau_r$ mit Transpositionen τ_i, so gilt $\operatorname{sgn} \pi = (-1)^r$.

(ii) Ist π ein k-Zykel so gilt $\operatorname{sgn} \pi = (-1)^{k-1}$.

(iii) π is genau dann gerade, wenn in jeder Darstellung von π als Produkt von Transpositionen die Anzahl der Transpositionen gerade ist.

Beweis. (i) folgt aus dem Satz und Beispiel b. (ii) folgt aus (i) und Bemerkung 1.5.3(v). (iii) folgt aus (i). □

Beispiel c. Um das Signum der Permutation π aus Beispiel (1.5.2) zu berechnen, benutzen wir die Zerlegung $\pi = (1\ 5\ 2\ 8)(4\ 6\ 7)(10\ 11)$ aus Beispiel (1.5.3). Dann ergibt sich aus dem Satz und Folgerung b(ii):

$$\operatorname{sgn} \pi = \operatorname{sgn}(1\ 5\ 2\ 8)\cdot\operatorname{sgn}(4\ 6\ 7)\cdot\operatorname{sgn}(10\ 11) = (-1)^3\cdot(-1)^2\cdot(-1)^1 = (-1)^6 = 1.$$

Übung. (i) Es seien $\pi \in S_A$ und $\varphi : A \to \underline{n}$ eine Bijektion. Man zeige, dass $\operatorname{sgn}(\varphi \circ \pi \circ \varphi^{-1})$ unabhängig von der Wahl der Bijektion φ ist.

(ii) Es seien $\pi \in S_n$ und $n \leq m$. Fasse π als Element von S_m auf. Hängt $\operatorname{sgn} \pi$ von m ab?

(iii) Man zeige: Hat $\pi \in S_n$ die Zykelzahl z, so gilt $\operatorname{sgn} \pi = (-1)^{n-z}$.

1.6 Relationen

1.6.1 Definition und Beispiele

Relationen drücken Beziehungen zwischen Elementen von zwei Mengen aus, z.B. wäre „liegt in“ eine Relation zwischen {Städte} und {Länder}. In der Informatik werden Relationen z.B. in relationalen Datenbanken verwendet.

Definition. Es seien M und N zwei Mengen.

(i) Eine Teilmenge $R \subseteq M \times N$ heißt *Relation zwischen M und N*, oder kürzer *Relation auf M* falls $N = M$. Für $(x, y) \in R$ schreiben wir auch xRy und sagen „x steht in Relation zu y bzgl. R“.

(ii) Eine Relation $R \subseteq M \times M$ auf M heißt

(R) *reflexiv*, falls xRx für alle $x \in M$,

(R') *antireflexiv*, falls nicht xRx für alle $x \in M$,

(S) *symmetrisch*, falls $xRy \Rightarrow yRx$ für alle $x, y \in M$,

(A) *antisymmetrisch*, falls $(xRy \wedge yRx) \Rightarrow x = y$ für alle $x, y \in M$,

(T) *transitiv*, falls $(xRy \wedge yRz) \Rightarrow xRz$ für alle $x, y, z \in M$.

(iii) Eine Relation, die (R),(S) und (T) erfüllt, heißt *Äquivalenzrelation.*

(iv) Eine Relation, die (R),(A) und (T) erfüllt, heißt *(partielle) Ordnung.*

(v) Eine Ordnung heißt *Totalordnung*, wenn $xRy \vee yRx$ für alle $x, y \in M$.

Beispiel.

(i) $M = \mathbb{R}$ und $R =$„$\leq$“, d.h. $(x, y) \in R$ genau dann, wenn $x \leq y$.
„$\leq$“ ist reflexiv, antisymmetrisch und transitiv, also eine Ordnung.
„$\leq$“ ist sogar eine Totalordnung.

(ii) $M = \mathbb{R}$ und $R =$„$<$“, d.h. $(x, y) \in R \Leftrightarrow x < y$.
„$<$“ ist antisymmetrisch(!) und transitiv, aber weder reflexiv noch symmetrisch.

(iii) $M = \mathrm{Pot}(N)$ und $R =$„$\subseteq$“.
„$\subseteq$“ ist eine Ordnung. Falls $|N| \geq 2$, so ist „$\subseteq$“ jedoch keine Totalordnung, da z.B. für $\{1\}, \{2\} \in \mathrm{Pot}\,\{1, 2\}$ weder $\{1\} \subseteq \{2\}$ noch $\{2\} \subseteq \{1\}$ gilt.

(iv) $M = \mathbb{Z}$. Die *Teilbarkeitsrelation* „$|$“ ist erklärt durch $x|y$ genau dann, wenn ein $z \in \mathbb{Z}$ existiert mit $xz = y$. Sie ist nicht antisymmetrisch, denn $1|-1$ und $-1|1$ obwohl $1 \neq -1$. Also ist „$|$“ keine Ordnung auf $\mathbb{Z}$.

(v) Die *Teilbarkeitsrelation* „$|$“ ist eine Ordnung auf $\mathbb{N}$, aber keine Totalordnung.

(vi) Auf jeder Menge M stellt die *Gleichheit* „$=$“ eine Äquivalenzrelation dar mit $R = \{(x, x) \mid x \in M\}$.

(vii) Auf einer Menge M von Personen können zwei Relationen V und G erklärt werden durch:

$xVy :\Leftrightarrow x$ ist verwandt mit y,
$xGy :\Leftrightarrow x$ hat das gleiche Geburtsdatum (Tag und Monat) wie y.

Beide sind Äquivalenzrelationen. Ersetzt man „verwandt“ durch „erstgradig verwandt“, so ist V nicht mehr transitiv.

(viii) Jede Abbildung $f : N \to M$ kann als Relation zwischen N und M aufgefasst werden:

$$f = \{(x, f(x)) \mid x \in N\}.$$

Abbildungen sind also eine spezielle Art von Relationen.

(ix) Für jede Abbildung $f : N \to M$ kann man eine Relation R_f auf N erklären durch

$xR_fy :\Leftrightarrow f(x) = f(y)$ (d.h. x und y liegen in derselben Faser von f).

R_f ist eine Äquivalenzrelation.

(x) $M = \mathbb{Z}$. Die *Paritätsrelation* „$\equiv_2$“, definiert durch

$$x \equiv_2 y :\Leftrightarrow x - y \text{ gerade}$$

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

Übung. Durch welche Datenstruktur würden Sie eine Relation auf einer endlichen Menge in einem Computerprogramm repräsentieren? Wie prüfen Sie anhand dieser Datenstruktur, ob die Relation reflexiv, symmetrisch bzw. antisymmetrisch ist?

Übung. Es seien R eine Relation auf A und $A' \subseteq A$. Dann ist $R' := R \cap (A' \times A')$ eine Relation auf A'. Man mache sich klar, dass jede der Eigenschaften aus Teil (ii) der Definition beim Übergang von R zu R' erhalten bleibt.

Übung. Welche Bedingung muss eine Relation $R \subseteq N \times M$ erfüllen, damit sie im Sinne von Beispiel (ix) als eine Abbildung von N nach M aufgefasst werden kann? Unter welcher Bedingung ist diese Abbildung injektiv, surjektiv bzw. bijektiv? Welche Relation gehört im bijektiven Fall zur Umkehrabbildung?

1.6.2 Partielle Ordnungen

Es sei $\trianglelefteq$ eine partielle Ordnung auf M.

Definition. Ein Element $m \in M$ heißt *minimal* in M, falls kein $m' \in M$ existiert mit $m' \neq m$ und $m' \trianglelefteq m$. Ein Element $m \in M$ wird ein *Minimum* von M genannt, falls für alle $m' \in M$ gilt: $m \trianglelefteq m'$.

Analog definiert man *maximal* und *Maximum* (Übung).

Bemerkung a. Nach Definition bedeutet

$$\begin{aligned} m \text{ Minimum von } M: &\qquad \text{für alle } x \in M \text{ gilt } m \trianglelefteq x. \\ m \text{ minimal in } M: &\qquad \text{für alle } x \in M \text{ gilt } x \trianglelefteq m \Rightarrow x = m. \end{aligned}$$

Minimal zu sein ist also zu verstehen als „kein anderes ist kleiner“. Minimum zu sein ist also zu verstehen als „alle anderen sind grösser“.

Beispiel. Wir betrachten die Teilbarkeitsrelation „|“ auf $\mathbb{N}$. Minimal zu sein bzgl. „|“ bedeutet „kein anderes ist Teiler“. Minimum zu sein bzgl. „|“ bedeutet „alle anderen sind Vielfache“.

(i) Die Menge $\{2, 3, 4, 6\}$ besitzt kein Minimum, hat aber die minimalen Elemente 2 und 3.

(ii) Die Menge $\{2, 3, 5\}$ besitzt kein Minimum, und jedes Element ist minimal.

(iii) Die Menge $\{2, 4, 6\}$ besitzt das Minimum 2, und 2 ist das einzige minimale Element.

Satz. *Es sei $\trianglelefteq$ eine partielle Ordnung auf M.*

(i) Jedes Minimum von M ist minimal in M.

(ii) Existiert ein Minimum von M, so ist es das einzige minimale Element in M. Insbesondere ist das Minimum eindeutig.

(iii) Bei einer Totalordnung ist jedes minimale Element in M auch Minimum von M (die Begriffe minimal und Minimum sind bei Totalordnungen also identisch).

Beweis. (i) Ist m ein Minimum und $x \trianglelefteq m$, so folgt $x = m$ wegen der Antisymmetrie ($m \trianglelefteq x \wedge x \trianglelefteq m \Rightarrow x = m$).

(ii) Sei m ein Minimum und sei m' minimal. Da m Minimum ist, gilt $m \trianglelefteq m'$. Da m' minimal ist, folgt daraus $m = m'$.

(iii) Sei $\trianglelefteq$ eine Totalordnung auf M und sei $m \in M$ minimal. Zu zeigen ist $m \trianglelefteq x$ für alle $x \in M$. Sei also $x \in M$ beliebig. Bei einer Totalordnung ist $m \trianglelefteq x$ oder $x \trianglelefteq m$. Im ersten Fall sind wir fertig. Im zweiten Fall folgt $x = m$, da m minimal ist, also $x \trianglelefteq m$ wegen der Reflexivität. □

Bemerkung b. Jede Teilmenge von $\mathbb{N}$ hat bzgl. der Ordnung $\leq$ ein Minimum. (Ohne Beweis; das ist ein Axiom der Mengenlehre.)

Übung. Jede endliche Menge mit partieller Ordnung hat ein minimales Element.

Übung. Formuliere Definition, Bemerkung a und Satz für *maximal* und *Maximum* aus.

Übung. Wir können die Begriffe minimal und Minimum auch definieren, wenn die Relation keine Ordnung ist. Zeigen Sie am Beispiel der Teilbarkeitsrelation auf $\mathbb{Z}$ (die keine Ordnung ist), dass dann der Satz nicht mehr gilt.

1.6.3 Äquivalenzrelationen

Definition. Es sei $\sim$ eine Äquivalenzrelation auf M. Für $x \in M$ heißt

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

die *Äquivalenzklasse von* $\sim$ *zu* x. Die Menge aller Äquivalenzklassen von $\sim$ wird mit $M/_\sim$ bezeichnet.

Bemerkung. Es sei $\sim$ eine Äquivalenzrelation auf M. Dann gilt für alle $x, y \in M$:

(i) $x \in [x]_\sim$,

(ii) $y \in [x]_\sim \Leftrightarrow x \in [y]_\sim$,

(iii) $y \in [x]_\sim \Rightarrow [y]_\sim = [x]_\sim$.

Wegen (iii) bezeichnet man jedes Element einer Äquivalenzklasse als ein *Repräsentant* derselben.

Beweis. als Übung. □

Beispiel.

(i) Für die Gleichheitsrelation auf einer Menge M ist $[x]_= = \{x\}$ und $M/_= = \{\{x\} \mid x \in M\}$.

(ii) Für die Äquivalenzrelationen V und G aus Beispiel (1.6.1)(vii) gilt für jede Person P der Menge:

$$[P]_V = \{\text{Verwandte von } P\},$$
$$[P]_G = \{\text{Personen, die am gleichen Tag Geburtstag feiern wie } P\}.$$

(iii) Es sei $f : N \to M$ eine Abbildung und R_f die Äquivalenzrelation aus Beispiel (1.6.1)(ix). Dann ist

$$[x]_{R_f} = \{x' \in N \mid f(x) = f(x')\} = f^{-1}(\{x\}),$$

für jedes $x \in N$, und $M/_{R_f}$ ist die Menge der nicht-leeren Fasern von f.

(iv) Für die Paritätsrelation aus Beispiel (1.6.1)(x) ist

$$[0]_{\equiv_2} = \{a \in \mathbb{Z} \mid a \text{ gerade}\},$$
$$[1]_{\equiv_2} = \{a \in \mathbb{Z} \mid a \text{ ungerade}\},$$

und $M/_{\equiv_2} = \{[0]_{\equiv_2}, [1]_{\equiv_2}\}$.

Offensichtlich „partitioniert“ eine Äquivalenzrelation die Menge.

Satz. *Es sei M eine Menge.*

(i) Ist $\sim$ eine Äquivalenzrelation auf M, so ist $M/_\sim$ eine Partition von M.

(ii) Ist $\mathcal{P}$ eine Partition von M, so existiert eine Äquivalenzrelation $\sim$ auf M mit $M/_\sim = \mathcal{P}$.

Die Äquivalenzrelationen auf M entsprechen also den Partitionen von M.

Beweis.

(i) Sei $\sim$ eine Äquivalenzrelation auf M und setze $\mathcal{P} := M/_\sim$. Wegen $x \in [x]_\sim$ sind alle Äquivalenzklassen nicht leer und ihre Vereinigung ist ganz M. Es bleibt zu zeigen, dass die Äquivalenzklassen paarweise disjunkt sind (vgl. Definition (1.2.4)(iv)). Betrachte also zwei beliebige Klassen $[x]_\sim, [y]_\sim$ mit $x, y \in M$. Zu zeigen ist:

$$[x]_\sim \neq [y]_\sim \Rightarrow [x]_\sim \cap [y]_\sim = \emptyset,$$

bzw. die Kontraposition

$$[x]_\sim \cap [y]_\sim \neq \emptyset \Rightarrow [x]_\sim = [y]_\sim.$$

Ist aber $z \in [x]_\sim \cap [y]_\sim$, so folgt daraus nach Teil (iii) der Bemerkung $[x]_\sim = [z]_\sim = [y]_\sim$.

(ii) Durch die Vorschrift

$$x \sim y :\Leftrightarrow \ x \text{ und } y \text{ liegen in demselben Teil der Partition}$$

wird eine Äquivalenzrelation definiert. (Man überprüfe das!)
Die Äquivalenzklassen sind offensichtlich genau die Teile von $\mathcal{P}$.

□

Kapitel 2

Algebraische Strukturen

2.1 Gruppen

2.1.1 Strukturen und Verknüpfungen

Definition. Eine *Verknüpfung* oder *Operation* auf einer Menge M ist eine Abbildung $M \times M \to M$. Eine *algebraische Struktur* ist eine Menge mit ein oder mehreren Verknüpfungen.

Beispiel a.

(i) $- : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, (a, b) \mapsto a - b$ ist eine Verknüpfung auf $\mathbb{Z}$.

(ii) Für jede Menge N ist $\circ$ eine Verknüpfung auf $\mathrm{Abb}(N, N)$.

(iii) $\wedge$ ist eine Verknüpfung auf $B := \{\mathrm{w}, \mathrm{f}\}$.

(iv) Es sei A eine beliebige Menge. Sind $a_1, \ldots, a_n \in A$, so nennen wir $a_1 \cdots a_n$ ein *Wort* der Länge n über dem Alphabet A. Es bezeichne ϵ das Wort der Länge 0, und A^* die Menge aller Wörter über A einschliesslich ϵ. Dann wird durch

$$a_1 \cdots a_n || b_1 \cdots b_m := a_1 \cdots a_n b_1 \cdots b_m$$

eine Verknüpfung auf A^* definiert, die *Verkettungsoperation*.

Schreibweise. Es seien M eine Menge, $\bullet$ eine Verknüpfung auf M, $m \in M$, und $A, B \subseteq M$.

(i) $m \bullet A := \{m \bullet a \mid a \in A\} \subseteq M$

(ii) $A \bullet m := \{a \bullet m \mid a \in A\} \subseteq M$

(iii) $A \bullet B := \{a \bullet b \mid a \in A, b \in B\} \subseteq M$

Beispiel b.

$$7\mathbb{Z} = \{7a \,|\, a \in \mathbb{Z}\} = \{\ldots, -14, -7, 0, 7, 14, \ldots\},$$
$$2 + 7\mathbb{Z} = \{2 + 7a \mid a \in \mathbb{Z}\} = \{..., -12, -5, 2, 9, 16, ...\}.$$

2.1.2 Monoide

Definition a. Es sei M eine Menge mit einer Verknüpfung $\bullet : M \times M \to M, (x, y) \mapsto x \bullet y$. Wir nennen $(M, \bullet)$ ein *Monoid*, wenn folgende Axiome gelten:

(G1) $(x \bullet y) \bullet z = x \bullet (y \bullet z)$ für alle $x, y, z \in M$.

(G2) Es existiert $e \in M$ mit $e \bullet x = x \bullet e = x$ für alle $x \in M$.

Das Monoid heißt *abelsch*, wenn zusätzlich gilt:

(G4) $x \bullet y = y \bullet x$ für alle $x, y \in G$.

Man nennt (G1) das Assoziativgesetz und (G4) das Kommutativgesetz.

Bemerkung. Das Element e in (G2) ist eindeutig und wird das *neutrale Element* von M genannt.

Beweis. Sind $e, e' \in G$ zwei Elemente, die (G2) erfüllen, so gilt einerseits $e \bullet e' = e$ und andererseits $e' \bullet e = e'$, also $e = e'$. □

Schreibweise.

(i) In einem Monoid $(M, \bullet)$ gilt $a_1 \bullet a_2 \bullet \cdots \bullet a_n := (\cdots((a_1 \bullet a_2) \bullet a_3) \bullet \cdots a_n)$ (oder jede andere Klammerung).

(ii) In einem abelschen Monoid benutzt man häufig $+$ als Verknüpfungszeichen, schreibt 0 statt e und na $(n \in \mathbb{N})$ als Abkürzung für $\underbrace{a + a + \cdots + a}_{n\text{-mal}}$.

(iii) Falls $\cdot$ als Verknüpfungszeichen benutzt wird, schreibt man häufig 1 statt e und a^n $(n \in \mathbb{N})$ als Abkürzung für $\underbrace{a \cdot a \cdot \ldots \cdot a}_{n\text{-mal}}$.

Beispiel a. Es sei A eine beliebige Menge, $B := \{\text{w,f}\}$.

(i) $(\mathbb{N}, +)$ ist kein Monoid, da (G2) nicht gilt.

(ii) $(\mathbb{Z}, -)$ ist kein Monoid, da (G1) nicht gilt.

(iii) $(\mathbb{N}_0, +)$ ist ein abelsches Monoid mit neutralem Element 0.

(iv) $(\mathbb{R}, \cdot)$ ist ein abelsches Monoid mit neutralem Element 1.

(v) Für jede nicht-leere Menge A ist $(\mathrm{Abb}(A, A), \circ)$ ein Monoid mit neutralem Element id_A.

(vi) $(B, \wedge)$ ist ein abelsches Monoid mit neutralem Element w.

(vii) $(B, \vee), (B, \mathrm{XOR})$ sind abelsche Monoide mit neutralem Element f.

(viii) $(B, \Rightarrow)$ ist kein Monoid, da (G1) nicht gilt (man prüfe nach, dass z.B. $(\mathrm{f} \Rightarrow \mathrm{f}) \Rightarrow \mathrm{f}$ ungleich $\mathrm{f} \Rightarrow (f \Rightarrow f)$ ist).

(ix) $(A^*, ||)$ ist Monoid mit neutralem Element ϵ.

Übung. Es sei A eine nicht-leere Menge. Man zeige, dass $(\mathrm{Abb}(A, A), \circ)$ genau dann abelsch ist wenn $|A| = 1$.

2.1.3 Inverse und Einheiten

Definition. Es seien $(M, \bullet)$ ein Monoid mit neutralem Element e und $a \in M$.

(i) Gibt es $b \in M$ mit $a \bullet b = e$, so heißt a *rechtsinvertierbar* und b *rechtsinvers* zu a bzw. b ein *Rechtsinverses* von a.

(ii) Gibt es $b \in M$ mit $b \bullet a = e$, so heißt a *linksinvertierbar* und b *linksinvers* zu a bzw. b ein *Linksinverses* von a.

(iii) Ist a sowohl links- als auch rechtsinvertierbar, so heißt a eine *Einheit*.

(iv) Gibt es $b \in M$ mit $b \bullet a = a \bullet b = e$, so heißt a *invertierbar* und b *invers* zu a bzw. b ein *Inverses* von a.

Bemerkung. Es seien $(M, \bullet)$ ein Monoid und $a \in M$. Dann ist a genau dann eine Einheit, wenn a invertierbar ist. In diesem Fall ist jedes Linksinverse von a auch Rechtsinverses, und umgekehrt. Weiter ist das Inverse von a eindeutig durch a bestimmt und wird mit a^{-1} bezeichnet. Wir bezeichnen die Menge der Einheiten von M mit $M^\times$.

Beweis. Per Definition ist jedes invertierbare Element eine Einheit. Sei umgekehrt a eine Einheit, etwa $b, b' \in M$ mit $b \bullet a = e$ und $a \bullet b' = e$. Dann folgt $b = b \bullet e = b \bullet (a \bullet b') = (b \bullet a) \bullet b' = e \bullet b' = b'$. Also ist $b = b'$ und somit a invertierbar. Mit $b = b'$ sind auch alle weiteren Aussagen der Bemerkung gezeigt. □

Beispiel. Es sei A eine nicht-leere Menge. Wir betrachten ein Element $f : A \to A$ des Monoids $(\mathrm{Abb}(A, A), \circ)$.

(i) f ist genau dann rechtsinvertierbar, wenn f surjektiv ist.

(ii) f ist genau dann linksinvertierbar, wenn f injektiv ist.

(iii) f ist genau dann invertierbar, wenn f bijektiv ist.

Übung a. Es seien $(M, \bullet)$ ein Monoid, $a \in M$, und m_a die Abbildung

$$m_a : M \to M, x \mapsto a \bullet x.$$

Man zeige:

(i) m_a ist genau dann surjektiv, wenn a rechtsinvertierbar ist.

(ii) Ist a linksinvertierbar, so ist m_a injektiv.

Man gebe ein Beispiel dafür an, dass die Umkehrung von (ii) nicht gilt.

Übung b. Es seien $(M, \bullet)$ ein Monoid, $a, a' \in M^\times, b, c \in M$, und b invers zu a.

(i) Es gilt $a^{-1} \in M^\times$ und $(a^{-1})^{-1} = a$.

(ii) Es gilt $a \bullet a' \in M^\times$ und $(a \bullet a')^{-1} = a'^{-1} \bullet a^{-1}$.

(iii) Die Gleichung $a \bullet x = c$ hat eine eindeutige Lösung für $x \in M$.

(iv) Die Gleichung $x \bullet a = c$ hat eine eindeutige Lösung für $x \in M$.

(v) Aus $a \bullet c = e$ folgt $c = a^{-1}$.

(vi) Aus $c \bullet a = e$ folgt $c = a^{-1}$.

2.1.4 Gruppen

Definition a. Eine Monoid $(G, \bullet)$, in dem alle Elemente invertierbar sind, heißt *Gruppe*. D.h. in einer Gruppe gilt:

(G3) Für alle $x \in G$ existiert $x' \in G$ mit $x \bullet x' = x' \bullet x = e$.

Beispiel a.

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

(ii) $(\mathbb{N}_0, +)$ ist keine Gruppe, da (G3) nicht gilt.

(iii) $(\mathbb{R}, \cdot)$ ist keine Gruppe, da (G3) nicht gilt.

(iv) $(\mathbb{R}\backslash\{0\}, \cdot)$ und $(\mathbb{R}_{>0}, \cdot)$ sind abelsche Gruppen.

(v) $(\mathbb{Z}\backslash\{0\}, \cdot)$ und $(\mathbb{N}, \cdot)$ sind keine Gruppen.

(vi) Für jede nicht-leere endliche Menge A ist $(S_A, \circ)$ eine Gruppe.

(vii) $(B, \wedge), (B, \vee)$ sind keine Gruppen.

(viii) (B, XOR) ist eine Gruppe.

(ix) $(A^*, ||)$ ist keine Gruppe, da (G3) nicht gilt.

Schreibweise.

(i) In einer abelschen Gruppe benutzt man häufig $+$ als Verknüpfungszeichen, schreibt $-a$ für das Inverse von a, und benutzt die Abkürzungen: $a - b := a + (-b), (-n)a := n(-a)$ für $n \in \mathbb{N}$, $0a := 0$.

(ii) Falls $\cdot$ als Verknüpfungszeichen benutzt wird, schreibt man a^{-1} für das Inverse von a, 1 statt e, lässt $\cdot$ einfach weg, und benutzt die Abkürzungen: $a^{-n} := (a^{-1})^n$ für $n \in \mathbb{N}$, $a^0 := 1$. Falls die Gruppe abelsch ist, kann man auch a/b für ab^{-1} schreiben.

Bemerkung. Ist $(M, \bullet)$ ein Monoid, so ist $(M^\times, \bullet)$ eine Gruppe. Die Gruppe $(M^\times, \bullet)$ wird *Einheitengruppe* von M genannt.

Beweis. Zu zeigen ist, dass $\bullet$ eine Verknüpfung auf $M^\times$ ist, und dass für $a \in M^\times$ auch das Inverse von a in $M^\times$ liegt. Beides wurde in Bemerkung 2.1.3 gezeigt. □

Satz. *Es sei $(G, \cdot)$ eine Gruppe und $a, b \in G$.*

(i) Für alle $c \in G$ gilt: $a = b \Leftrightarrow a \cdot c = b \cdot c \Leftrightarrow c \cdot a = c \cdot b$.
(„Multiplikation“ von links oder rechts in einer Gruppe ist eine Äquivalenzumformung.*)*

(ii) Die Gleichung $a \cdot x = b$ hat eine eindeutige Lösung $x \in G$ (ebenso die Gleichung $x \cdot a = b$).

Beweis.

(i) Die Implikation $a = b \Rightarrow a \cdot c = b \cdot c$ ist trivial. Damit folgt aber auch $a \cdot c = b \cdot c \Rightarrow (a \cdot c) \cdot c^{-1} = (b \cdot c) \cdot c^{-1}$, und die rechte Seite lautet $a = b$. Die Äquivalenz $a = b \Leftrightarrow c \cdot a = c \cdot b$ verläuft entsprechend mit Multiplikation von c^{-1} auf der linken Seite.

(ii) Nach (i) gilt $a \cdot x = b$ genau dann, wenn $x = a^{-1} \cdot (a \cdot x) = a^{-1} \cdot b$ ist. Entsprechend gilt $x \cdot a = b$ genau dann, wenn $x = (x \cdot a) \cdot a^{-1} = b \cdot a^{-1}$ ist.

□

Beispiel b. Gesucht ist ein $\sigma \in S_3$ mit $(123) \circ \sigma = (12)$. Nach Aussage des Satzes gilt:

$$\begin{aligned}(1\ 2\ 3) \circ \sigma = (1\ 2) &\Leftrightarrow (1\ 2\ 3)^{-1} \circ (1\ 2\ 3) \circ \sigma = (1\ 2\ 3)^{-1} \circ (1\ 2) \\ &\Leftrightarrow \sigma = (3\ 2\ 1) \circ (1\ 2) = (2\ 3).\end{aligned}$$

Somit ist $\sigma = (2\ 3)$ die einzige Lösung.

Übung a. Bestimmen Sie zu allen Beispielen von Monoiden und Gruppen die neutralen bzw. inversen Elemente.

Übung b. Es sei A eine nicht-leere endliche Menge. Man zeige, dass S_A genau dann abelsch ist, wenn $|A| \leq 2$.

Übung c. Es seien $(G, \cdot)$ eine Gruppe und $a \in G$. Ist die Abbildung $\lambda_a : G \to G, x \mapsto a \cdot x$ injektiv, surjektiv, bijektiv?

2.1.5 Untergruppen

Definition. Es sei $(G, \cdot)$ eine Gruppe. Eine Teilmenge $H \subseteq G$ heißt *Untergruppe von* G, geschr. $H \leq G$, wenn gilt:

(U1) $e \in H$.

(U2) Für alle $x, y \in H$ ist auch $x \cdot y^{-1} \in H$. (H ist *abgeschlossen* bzgl. $\cdot$)

In diesem Fall ist H selbst eine Gruppe bzgl. der Verknüpfung $\cdot$ aus G.

Beispiel a.

(i) Für jedes $n \in \mathbb{N}_0$ ist $n\mathbb{Z} := \{nz \mid z \in \mathbb{Z}\}$ eine Untergruppe von $(\mathbb{Z}, +)$. (z.B. $3\mathbb{Z} = \{\ldots, -9, -6, -3, 0, 3, 6, 9, \ldots\}$)

(ii) $\mathbb{N}$ ist keine Untergruppe von $(\mathbb{Z}, +)$.

(iii) $H := \{\pi \in S_n \mid \pi(n) = n\}$ ist eine Untergruppe von $(S_n, \circ)$.

(iv) $\mathbb{Q}_{>0}$ ist eine Untergruppe von $(\mathbb{R}_{>0}, \cdot)$.

(v) $\mathbb{N}$ ist keine Untergruppe von $(\mathbb{R}_{>0}, \cdot)$.

Beweis.

(i) (U1): $e = 0 = n \cdot 0 \in n\mathbb{Z}$.
(U2): $nx - ny = n(x - y) \in n\mathbb{Z}$.

(ii) (U1) gilt nicht, denn $e = 0 \notin \mathbb{N}$.

(iii) (U1): $e = \mathrm{id}_{\underline{n}}$ lässt n fest, also $\mathrm{id}_{\underline{n}} \in H$.
(U2): Seien $\sigma, \pi \in H$, d.h. $\sigma(n) = n$ und $\pi(n) = n$. Aus $\pi(n) = n$ folgt $\pi^{-1}(n) = n$. Weiter ergibt sich $\sigma \circ \pi^{-1}(n) = \sigma(\pi^{-1}(n)) = \sigma(n) = n$, d.h. $\sigma \circ \pi^{-1} \in H$.

(iv) (U1): $e = 1 \in \mathbb{Q}_{>0}$.
(U2): Sind $x, y \in \mathbb{Q}_{>0}$, so ist auch $xy^{-1} \in \mathbb{Q}_{>0}$.

(v) (U2) gilt nicht, da z.B. $2^{-1} \notin \mathbb{N}$.

□

Übung.

(i) Ist $\{\pi \in S_n \mid \operatorname{sgn} \pi = 1\}$ eine Untergruppe von $(S_n, \circ)$?

(ii) Ist $\{\pi \in S_n \mid \operatorname{sgn} \pi = -1\}$ eine Untergruppe von $(S_n, \circ)$?

2.1.6 Kartesische Produkte

Satz. *Es seien $(G, \cdot)$ eine Gruppe und M eine Menge. Die Menge $\mathrm{Abb}(M, G) = \{f : M \to G\}$ wird zu einer Gruppe $(\mathrm{Abb}(M, G), \bullet)$, wenn man die Verknüpfung*

$$\bullet : \mathrm{Abb}(M, G) \times \mathrm{Abb}(M, G) \to \mathrm{Abb}(M, G), (f, g) \mapsto f \bullet g$$

durch

$$(f \bullet g)(x) := f(x) \cdot g(x) \text{ für alle } x \in M$$

definiert. Da $\bullet$ durch $\cdot$ definiert ist schreibt man in der Regel $(\mathrm{Abb}(M, G), \cdot)$. Ist $(G, \cdot)$ abelsch, so ist auch $(\mathrm{Abb}(M, G), \cdot)$ abelsch.

Beispiel. Es sei $(G, \cdot)$ eine Gruppe. Die Gruppe $(G^{\underline{n}}, \cdot)$ ist dann die Menge

$$G^n = \{n\text{-Tupel über } G\} = \{(a_1, ..., a_n) \mid a_i \in G\}$$

mit *komponentenweiser* Verknüpfung, d.h.

$$(a_1, ..., a_n) \cdot (b_1, ..., b_n) := (a_1 \cdot b_1, ..., a_n \cdot b_n).$$

G^n wird das *n-fache kartesische Produkt* von G genannt.

Übung. Es seien $(G, \bullet)$ und $(G', \circ)$ zwei Gruppen. Man zeige, dass die Menge $G \times G'$ mit der Verknüpfung

$$(g_1, g_2') \cdot (g_2, g_2') := (g_1 \bullet g_2, g_1' \circ g_2')$$

wieder eine Gruppe ist.

2.2 Ringe

2.2.1 Definition und Beispiele

In $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ gibt es zwei Verknüpfungen $+$ und $\cdot$, die mittels der Distributivgesetze miteinander verbunden sind. Die entsprechende Abstraktion der Rechenregeln führt zu den Begriffen Ring und Körper.

Definition. Eine Menge R mit zwei Verknüpfungen

$$+ : R \times R \to R, \quad \text{und} \quad \cdot : R \times R \to R$$

heißt *Ring*, wenn folgende Bedingungen erfüllt sind:

(R1) $(R, +)$ ist eine abelsche Gruppe.

(R2) $(R, \cdot)$ ist ein Monoid.

(R3) $x \cdot (y + z) = x \cdot y + x \cdot z$ und $(x + y) \cdot z = x \cdot z + y \cdot z$ für alle $x, y, z \in R$.

Der Ring heißt *kommutativ*, wenn zusätzlich gilt:

(R4) $x \cdot y = y \cdot x$ für alle $x, y \in R$.

Beispiel.

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

(ii) $R = \{0\}$ mit $0 + 0 := 0$ und $0 \cdot 0 := 0$ bildet den trivialen Ring.

(iii) Nicht-kommutative Ringe begegnen uns in der Linearen Algebra, z.B. der „Matrizenring“ und der „Endomorphismenring“.

Bemerkung. Die Gleichungen aus (R3) heißen *Distributivgesetze*. Man vereinbart in einem Ring, dass $\cdot$ stärker bindet als $+$, d.h. $a \cdot b + c$ steht für $(a \cdot b) + c$, und $a + b \cdot c$ für $a + (b \cdot c)$. Dies spart Klammern und wurde in obiger Formulierung von (R3) bereits benutzt! Ferner wird vereinbart, dass $\cdot$ weggelassen werden kann, d.h. ab steht für $a \cdot b$. Das neutrale Element der Gruppe $(R, +)$ wird mit 0 bezeichnet und *Nullelement* bzw. kurz *Null* von R genannt. Das neutrale Element des Monoid $(R, \cdot)$ wird mit 1 bezeichnet und *Einselement* bzw. kurz *Eins* von R genannt. Wir nennen $-a$ das *additive Inverse* oder *negative* Element von a.

Übung a. Es sei R ein Ring. Man zeige:

(i) $0 \cdot a = a \cdot 0 = 0$ für alle $a \in R$.

(ii) $-a = (-1) \cdot a$ und $a = (-1) \cdot (-a)$ für alle $a \in R$.

(iii) $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$ für alle $a, b \in R$.

Übung b. Es sei R ein Ring. Für jedes $n \in \mathbb{N}$ und $a \in R$ definieren wir

$$na := \underbrace{a + \ldots + a}_{n\text{-mal}}.$$

Man zeige: $-(na) = n(-a)$. Wie definiert man sinnvoll na für alle $n \in \mathbb{Z}$?

Übung c. Man zeige: Ist R ein Ring mit $1 = 0$, so ist $R = \{0\}$.

Übung d. Es seien R, S zwei Ringe, $n \in \mathbb{N}$ und M eine Menge. Wie sind die Verknüpfungen zu definieren, mit denen auch $R \times S, R^n$ und $\mathrm{Abb}(M, R)$ zu einem Ring werden?

2.2.2 Einheitengruppe

Definition. Es sei R ein Ring. Die Begriffe *invertierbar*, *Einheit*, *Einheitengruppe* und die Notation $R^\times$ beziehen sich auf das Monoid $(R, \cdot)$.

Beispiel.

(i) $\mathbb{Z}^\times = \{1, -1\}$.

(ii) In jedem Ring R ist $1, -1 \in R^\times$. (1 und -1 können aber gleich sein, wie wir an den Beispielen $\mathbb{F}_2$ und $\mathbb{F}_4$ unten sehen werden.)

Übung a. Es seien R kommutativ, $a \in R$, und m_a bezeichne die Abbildung $m_a : R \to R, x \mapsto ax$. Man zeige die Äquivalenz folgender Ausssagen:

(i) a ist Einheit.

(ii) m_a ist bijektiv.

(iii) Die Gleichung $ax = b$ ist für alle $b \in R$ eindeutig lösbar.

Insbesondere gilt für jedes $a \in R^\times$: $ax = 0 \Rightarrow x = 0$.

Übung b. Es seien R kommutativ, $a, b \in R$. Man zeige: $ab \in R^\times \Leftrightarrow a \in R^\times \wedge b \in R^\times$. Hieraus folgt, dass auch $R \setminus R^\times$ unter der Multiplikation abgeschlossen ist.

2.2.3 Körper

Definition. Ein kommutativer Ring R heißt *Körper*, wenn $1 \neq 0$ und $R^\times = R\backslash\{0\}$ gilt.

Ein Körper ist also ein nicht-trivialer Ring, in dem jedes von 0 verschiedene Element invertierbar ist.

Beispiel a. $(\mathbb{Q}, +, \cdot), (\mathbb{R}, +, \cdot)$ und $(\mathbb{C}, +, \cdot)$ sind Körper. Im Unterschied zu $\mathbb{Q}$ erfüllt $\mathbb{R}$ noch die „Vollständigkeitsaxiome“ und die „Anordnungsaxiome“, die man in der Analysis-Vorlesung lernt. $(\mathbb{Z}, +, \cdot)$ ist kein Körper.

Es gibt aber auch endliche Körper.

Beispiel b. Definiert man auf der Menge $\{0, 1\}$ zwei Abbildungen $+, \cdot$ durch die *Verknüpfungstafeln*

$+$	0	1
0	0	1
1	1	0

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

so entsteht ein Körper (man prüfe alle Axiome nach!). Wir bezeichnen diesen Körper mit $\mathbb{F}_2$.

Identifiziert man 0 mit „falsch“ und 1 mit „wahr“, dann stellt man ausserdem fest, dass $+$ gerade der Verknüpfung XOR entspricht, und $\cdot$ der Verknüpfung $\wedge$.

Beispiel c. Die Menge $\mathbb{F}_4 := \{0, 1, a, b\}$ mit den Verknüpfungstafeln

$+$	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

$\cdot$	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

bildet einen Körper.

Beweis. Übung. □

Bemerkung.

(i) Die Tafeln in Beispiel c sind bis auf Benennung der Elemente a, b eindeutig, d.h. es gibt genau einen Körper mit 4 Elementen (siehe Vorlesung oder vgl. [3], §2.2, 37-38, leicht lesbar).

(ii) Es gibt für jede Primzahlpotenz p^n genau einen Körper mit p^n Elementen (ohne Beweis). Dieser wird mit $\mathbb{F}_{p^n}$ bezeichnet (das $\mathbb{F}$ steht hier für „field“, engl. für Körper). Für $n = 1$ werden diese Körper in §2.4 unten konstruiert: $\mathbb{F}_p$ ist identisch mit dem dort eingeführten „Restklassenring“ $\mathbb{Z}_p$.
Achtung: $\mathbb{F}_{p^n}$ für $n > 1$ wird in dieser Vorlesung nicht behandelt und ist insbesondere nicht identisch mit $\mathbb{Z}_{p^n}$, denn $\mathbb{Z}_{p^n}$ ist für $n > 1$ kein Körper.

(iii) Endliche Körper sind für die Informatik von besonderer Bedeutung, etwa in der Kodierungstheorie. Es sei daran erinnert, dass man ein Bit als Element des Körper $\mathbb{F}_2$ auffassen kann, ein Byte als Element des Körpers $\mathbb{F}_{256}$, usw.

Übung. Sind K, L zwei Körper, so ist der Ring $K \times L$ (mit komponentenweisen Operationen) *kein* Körper.

2.2.4 Teilbarkeitsrelation

Es sei R ein kommutativer Ring.

Definition. Es seien $a, b \in R$. Wir sagen a *teilt* b bzw. b *ist Vielfaches von* a, geschr. $a|b$, wenn $x \in R$ existiert mit $ax = b$.

Bemerkung a. Die Relation $|$ auf R ist reflexiv und transitiv. Für alle $a, b, c \in R$ und alle $u, v \in R^\times$ gelten:

(i) $a|b \Rightarrow a|bc$,

(ii) $(a|b \wedge a|c) \Rightarrow a|b + c$,

(iii) $a|0$,

(iv) $0|a \Leftrightarrow a = 0$,

(v) $a|b \Leftrightarrow ua|vb$.

Beweis. Siehe Vorlesung. □

Übung a. Es seien $a, b \in R$. Man zeige: Falls $a|b$, dann gilt für alle $c \in R$: $a|b + c \Leftrightarrow a|c$.

Bemerkung b. Wir nennen $a, b \in R$ *assoziiert*, geschr. $a \sim b$, wenn ein $u \in R^\times$ existiert mit $au = b$. Aus $a \sim b$ folgt offensichtlich $a|b$ und $b|a$.

Beispiel. In $\mathbb{Z}$ gilt: $a \sim b \Leftrightarrow |a| = |b|$. Die Relation $|$ auf $\mathbb{Z}$ ist also nicht antisymmetrisch.

Übung b. (i) Man zeige, dass $\sim$ eine Äquivalenzrelation auf R ist.

(ii) Die Äquivalenzklassen $[a]_\sim$ bzgl. $\sim$ heißen die *Assoziiertenklassen* von R. Wie sehen die Assoziiertenklassen von $\mathbb{Z}$ aus?

(iii) Die Assoziiertenklasse von 1 ist $R^\times$.

(iv) Nach Teil (v) von Bemerkung a hängt die Relation $a|b$ nur von den Assoziiertenklassen von a und b ab. Die Relation $|$ lässt sich also als eine Relation $|_\sim$ auf der Menge $R/_\sim$ der Assoziiertenklassen von R auffassen. Man zeige, dass $|_\sim$ reflexiv und transitiv ist.

Übung c. Man zeige: Ist b eine Einheit in R und $a|b$, so ist auch a eine Einheit.

2.2.5 Ideale

Es sei R ein kommutativer Ring.

Definition. Eine Teilmenge $I \subseteq R$ von R heißt *Ideal* von R, falls gilt:

(i) I ist Untergruppe der additiven Gruppe $(R, +)$.

(ii) $RI \subseteq I$, das heißt $ra \in I$ für alle $r \in R$ und $a \in I$.

Bemerkung. Für Elemente $a_1, \ldots, a_k \in R$ definieren wir

$$(a_1, \ldots, a_k) = \{r_1 a_1 + \ldots + r_k a_k \mid \ r_1, \ldots, r_k \in R\}.$$

Dann ist $(a_1, \ldots, a_k)$ das kleinste Ideal, das $a_1, \ldots, a_k$ enthält, und wird *das von $a_1, \ldots, a_r$ erzeugte Ideal* genannt. Ideale, die von einem Element erzeugt werden, d.h. Ideale von der Form (a), heißen *Hauptideale*. Für alle $a, b \in R$ gelten:

(i) $a|b \Leftrightarrow (a) \supseteq (b)$

(ii) $a \sim b \Rightarrow (a) = (b)$

Beweis. Als Übung. □

Übung. Man zeige, dass mit zwei Idealen $I, J \subseteq R$ auch $I \cap J$ ein Ideal von R ist.

2.2.6 Nullteiler

Es sei R ein kommutativer Ring.

Definition. Ein Element $a \in R$ heißt *Nullteiler* von R, wenn $b \in R\backslash\{0\}$ existiert mit $ab = 0$. Der Ring R heißt *nullteilerfrei*, wenn er keine Nullteiler ausser 0 enthält.

Bemerkung. Sei $a \in R$ und bezeichne m_a die Abbildung $m_a : R \to R, x \mapsto ax$. Dann sind äquivalent:

(i) a ist kein Nullteiler von R.

(ii) Für alle $b \in R$ gilt: $ab = 0 \Rightarrow b = 0$.

(iii) Für alle $b, b' \in R$ gilt: $ab = ab' \Rightarrow b = b'$. (Kürzungsregel)

(iv) m_a ist injektiv.

Insbesondere sind Einheiten keine Nullteiler (nach Übung 2.2.2 ist m_a für Einheiten bijektiv). Weiter ist R genau dann nullteilerfrei, wenn für alle $a, b \in R$ gilt:

$$ab = 0 \Rightarrow (a = 0 \vee b = 0).$$

Beweis. Übung. □

Beispiel. $\mathbb{Z}$, alle Körper sowie der triviale Ring sind nullteilerfrei.

Beweis. als Übung. □

Übung a. Man zeige: 0 ist genau dann kein Nullteiler, wenn R der triviale Ring ist.

Übung b. Man zeige für alle $a, b \in R$: Ist a ein Nullteiler, so auch ab. Gilt auch die Umkehrung?

Übung c. Es sei R nullteilerfrei. Man zeige, dass für alle $a, b \in R$ gilt: $(a) = (b) \Leftrightarrow a \sim b$. Es gibt also eine Bijektion zwischen $R/_\sim$ und der Menge der Hauptideale $\mathcal{P}$ von R, die die Relation $|_\sim$ in die partielle Ordnung $\supseteq$ auf $\mathcal{P}$ überführt. Insbesondere ist $|_\sim$.

2.3 Der Euklidische Algorithmus

Wir betrachten den Ring $(\mathbb{Z}, +, \cdot)$, der kommutativ und nullteilerfrei ist.

2.3.1 Division mit Rest in $\mathbb{Z}$

Satz. *Für alle $a, b \in \mathbb{Z}$ mit $b \neq 0$ existieren eindeutige $q, r \in \mathbb{Z}$ mit $a = qb + r$ und $0 \leq r < |b|$.*

Beweis. Seien $a, b \in \mathbb{Z}$ beliebig. Wegen $a = qb + r \Leftrightarrow a = (-q)(-b) + r$ können wir oBdA $b \geq 0$ annehmen. Eindeutigkeit: Angenommen, wir haben $q, q', r, r' \in \mathbb{Z}$ mit $qb + r = q'b + r'$ und $0 \leq r, r' < b$. Nach Annahme ist $(q - q')b = r' - r$, also $b | r' - r$. Ebenfalls nach Annahme ist $0 \leq r' - r < b$. Es folgt $r' - r = 0$ bzw. $r' = r$. Da $\mathbb{Z}$ nullteilerfrei ist und $b \neq 0$, folgt aus $(q - q')b = 0$ auch $q = q'$.

Existenz: Wähle q maximal mit $qb \leq a$ und setze $r := a - qb$. (Die Wahl von q bedeutet $q := \lfloor \frac{a}{b} \rfloor$.) Damit ist $r \geq 0$ klar. Wir zeigen $r < b$ mit einem Widerspruchsbeweis. Angenommen $r \geq b$. Dann ist $a = r + qb \geq (q + 1)b$. Das steht im Widerspruch zur Maximalität von q, also ist die Annahme $r \geq b$ falsch und die Behauptung $r < b$ bewiesen. □

Beispiel. $-237 = (-12) \cdot 21 + 15, 0 \leq 15 < 21$.
Man beachte $(-11) \cdot 21 = -231 > -237$ und $(-12) \cdot 21 = -252 \leq -237$.

Übung. Zeigen Sie, dass es für jedes Ideal $I \subseteq \mathbb{Z}$ ein $g \in \mathbb{Z}$ mit $I = (g)$ gibt.
Tip: Wenn $I \neq \{0\}$, sei $g \in \mathbb{N}$ minimal mit $g \in I$. Dann ist $I = (g)$.
Bemerkung: Nullteilerfrei Ringe mit $1 \neq 0$, in denen jedes Ideal ein Hauptideal ist, werden *Hauptidealringe* genannt.

2.3.2 Der ggT

Definition. Der *größte gemeinsame Teiler* von $a, b \in \mathbb{Z}$ ist definiert durch

$$\mathrm{ggT}(a, b) := \max\{d \in \mathbb{N} \mid d|a, d|b\},$$

falls a, b nicht beide gleich 0 sind, und $\mathrm{ggT}(0, 0) := 0$. Wir nennen a und b *teilerfremd*, wenn $\mathrm{ggT}(a, b) = 1$.

Bemerkung. Für alle $a, b \in \mathbb{Z}$ gilt:

(i) $\mathrm{ggT}(a, b) = \mathrm{ggT}(b, a)$,

(ii) $\mathrm{ggT}(a, b) = \mathrm{ggT}(|a|, |b|)$,

(iii) $\mathrm{ggT}(a, 0) = |a|$,

(iv) $a = qb + r \Rightarrow \mathrm{ggT}(a, b) = \mathrm{ggT}(b, r)$.

Beweis. Sei $a = qb + r$ bzw. $r = a - qb$. Nach Bemerkung (2.2.4)(i) und (ii) gilt sowohl $d|a, b \Rightarrow d|r$ und $d|b, r \Rightarrow d|a$. Die gemeinsamen Teiler von a, b sind also identisch mit den gemeinsamen Teilern von b, r. □

Übung. Es seien $a, b \in \mathbb{Z}$. Nach Übung 2.3.1 ist $(a, b) = (g)$ für ein $g \in \mathbb{N}_0$. Man zeige:

(i) Die gemeinsamen Teiler von a, b sind genau die Teiler von g.

(ii) $g = \mathrm{ggT}(a, b)$.

2.3.3 Das kgV

Definition. Der *kleinste gemeinsame Vielfache* von $a, b \in \mathbb{Z}$ ist definiert durch

$$\mathrm{kgV}(a, b) := \min\{m \in \mathbb{N}_0 \mid a|m, b|m\}.$$

Bemerkung. Für alle $a, b \in \mathbb{Z}$ gilt:

(i) $\mathrm{kgV}(a, b) = \mathrm{kgV}(b, a)$,

(ii) $\mathrm{kgV}(a, b) = \mathrm{kgV}(|a|, |b|)$,

(iii) $\mathrm{kgV}(a, 0) = 0$,

Übung. Es seien $a, b \in \mathbb{Z}$. Nach Übung 2.3.1 ist $(a) \cap (b) = (k)$ für ein $k \in \mathbb{N}_0$. Man zeige:

(i) Die gemeinsamen Vielfachen von a, b sind genau die Vielfachen von k.

(ii) $k = \mathrm{kgV}(a, b)$.

Man folgere

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

2.3.4 Der Euklidische Algorithmus

Beispiel. Wie lautet $\mathrm{ggT}(91, 168)$?
Rechnung:

$$\begin{aligned} 168 &= 1 \cdot 91 + 77 \\ 91 &= 1 \cdot 77 + 14 \\ 77 &= 5 \cdot 14 + 7 \\ 14 &= 2 \cdot 7 + 0. \end{aligned}$$

Nach Bemerkung (2.3.2) gilt somit

$$\mathrm{ggT}(168, 91) = \mathrm{ggT}(91, 77) = \mathrm{ggT}(77, 14) = \mathrm{ggT}(14, 7) = \mathrm{ggT}(7, 0) = 7.$$

Rückwärts Einsetzen:

$$\begin{aligned} 7 &= 77 - 5 \cdot 14 \\ &= 77 - 5 \cdot (91 - 1 \cdot 77) = -5 \cdot 91 + 6 \cdot 77 \\ &= -5 \cdot 91 + 6 \cdot (168 - 1 \cdot 91) = 6 \cdot 168 - 11 \cdot 91. \end{aligned}$$

Somit gilt $\mathrm{ggT}(91, 168) = (-11) \cdot 91 + 6 \cdot 168$.

Algorithmus. *Es seien $a, b \in \mathbb{Z}$ mit $b \neq 0$. Die folgende Prozedur liefert $d, \lambda, \mu \in \mathbb{Z}$ mit $d = \mathrm{ggT}(a, b) = \lambda a + \mu b$.*

EUKLID(a, b)
1 bestimme q, r mit $a = qb + r$ und $0 \leq r < |b|$
2 **if** $r = 0$
3 **then return** $(b, 0, 1)$
4 **else** $(d, \lambda, \mu) \leftarrow$ EUKLID(b, r)
5 **return** $(d, \mu, \lambda - q\mu)$

Beweis. 1. Es sei $a = qb + r$.
3. Falls $r = 0$, dann $b|a$, also $\mathrm{ggT}(a, b) = b = 0 \cdot a + 1 \cdot b$.
4. Sei $r > 0$ und $d = \mathrm{ggT}(b, r) = \lambda b + \mu r$.
5. Nach Bemerkung (2.3.2) ist $d = \mathrm{ggT}(a, b)$. Ausserdem gilt $d = \lambda b + \mu(a - qb) = \mu a + (\lambda - q\mu)b$. □

Bemerkung. Der größte gemeinsame Teiler wurde ohne Verwendung des Begriffs „Primzahl“ definiert und kann mit dem Euklidischen Algorithmus ohne Kenntnis der Primfaktorzerlegung berechnet werden.

Übung a. Es seien $a, b \in \mathbb{N}$. Die Koeffizienten λ, μ in der Darstellung $\mathrm{ggT}(a, b) = \lambda a + \mu b$ sind nicht eindeutig. Geben Sie ein Beispiel an. Zeigen Sie weiter, dass λ, μ unter der Zusatzbedingung $-\frac{b}{d} < |\lambda| \leq 0$ und $0 \leq |\mu| < \frac{a}{d}$ eindeutig werden.

2.4 Restklassenringe

2.4.1 Kongruenz modulo n

Definition. Für jedes $n \in \mathbb{N}$ definieren wir auf $\mathbb{Z}$ eine Relation $\equiv_n$ durch

$$a \equiv_n b :\Leftrightarrow n|a - b.$$

Statt $a \equiv_n b$ schreibt man auch $a \equiv b \pmod{n}$ und sagt „*a kongruent b modulo n*“.

Bemerkung. Es gilt $a \equiv_n b$ genau dann, wenn a und b bei Division durch n denselben Rest lassen.

Beweis. Seien $a = qn + r$ und $b = q'n + r'$ mit $0 \leq r, r' < n$. Dann ist $a - b = (q - q')n + (r - r')$ und $|r - r'| < n$. Nach Bemerkung (2.2.4) gilt $n|a - b$ genau dann, wenn $n|r - r'$. Wegen $|r - r'| < n$ ist das genau dann der Fall, wenn $r - r' = 0$. □

Beispiel. Ist 14 kongruent 23 modulo 3 ($14 \equiv_3 23$)? Ja, weil $14 - 23 = -9$ Vielfaches von 3 ist. Alternativ kann man die Reste bei Division durch 3 vergleichen: $14 = 4 \cdot 3 + 2$ und $23 = 7 \cdot 4 + 2$. Sie stimmen überein (beide $= 2$).

Satz. *Für jedes $n \in \mathbb{N}$ ist die Relation $\equiv_n$ eine Äquivalenzrelation.*

Beweis. Leichte Übung unter Verwendung von Bemerkung (2.2.4). □

2.4.2 Restklassen modulo n

Es sei $n \in \mathbb{N}$ in diesem Abschnitt fest gewählt.

Definition a. Die Äquivalenzklasse von $a \in \mathbb{Z}$ bzgl. $\equiv_n$ wird mit $\overline{a}$ bezeichnet und wird die *Restklasse von a modulo n* genannt.

Bemerkung. Die Restklasse $\overline{a}$ besteht aus allen ganzen Zahlen, die bei Division durch n denselben Rest lassen wie a. Es gilt

$$\overline{a} = a + n\mathbb{Z} = \{\ldots, a - 2n, a - n, a, a + n, a + 2n, \ldots\}.$$

Dividiert man a durch n mit Rest, etwa $a = qn + r$ mit $0 \leq r < n$, so ist $\overline{a} = \overline{r}$. Der Rest r ist weiterhin der kleinste nicht-negative Repräsentant von $\overline{a}$. Folglich hat jede Restklasse modulo n genau einen Repräsentanten zwischen 0 und $n - 1$ (den Rest r). Es gibt also genau n verschiedene Restklassen modulo n: $\overline{0}, \overline{1}, \ldots, \overline{n-1}$.

Beispiel a. Für $n = 3$ ist $\overline{14} = \{\ldots, 5, 8, 11, 14, 17, 20, 23, \ldots\} = \overline{23}$. Wegen $14 = 4 \cdot 3 + 2$ ist $\overline{14} = \overline{2}$, und 2 ist der kleinste nicht-negative Repräsentant von $\overline{14}$.

Definition b. Die Menge der Restklassen modulo n wird mit $\mathbb{Z}_n$ bezeichnet, also $\mathbb{Z}_n := \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$. Es gilt $|\mathbb{Z}_n| = n$.

Beispiel b. $\mathbb{Z}_3 = \{\overline{0}, \overline{1}, \overline{2}\}$.

2.4.3 Rechnen mit Restklassen

Wir möchten auf der Menge der Restklassen modulo n zwei Verknüpfungen $+$ und $\cdot$ einführen mittels der Definition

$$\overline{a} + \overline{b} := \overline{a+b} \quad \text{und} \quad \overline{a} \cdot \overline{b} := \overline{a \cdot b}. \tag{$*$}$$

Das Problem in dieser Definition ist, dass sie – auf den ersten Blick – von der Wahl der Repräsentanten a und b abzuhängen scheint. Der folgende Satz zeigt, dass dem nicht so ist. Nur aufgrund des Satzes handelt es sich bei ($*$) überhaupt um eine gültige Definition.

Satz. *Sei $n \in \mathbb{N}$ fest. Sind $a, a', b, b' \in \mathbb{Z}$ mit $\overline{a} = \overline{a'}$ und $\overline{b} = \overline{b'}$, so gilt:*

(i) $\overline{a+b} = \overline{a'+b'}$,

(ii) $\overline{a \cdot b} = \overline{a' \cdot b'}$.

Beweis. Nach Voraussetzung ist $n|a-a'$ und $n|b-b'$. Nach Bemerkung (2.2.4) (ii) teilt n auch $(a-a')+(b-b') = (a+b)-(a'+b')$, also gilt (i). Nach Bemerkung (2.2.4) (i) und (ii) teilt n auch $(a-a')b'+(b-b')a = (a \cdot b - a' \cdot b')$, also gilt (ii). □

Folgerung. *Die Menge $\mathbb{Z}_n$ bildet bzgl. der Verknüpfungen aus ($*$) einen kommutativen Ring.*

Beweis. Addition und Multiplikation in $\mathbb{Z}_n$ sind über die entsprechenden Operationen aus $\mathbb{Z}$ definiert. Daher werde Assoziativ-, Kommutativ- und Distributivgesetze von $\mathbb{Z}$ „geerbt“. Weiter ist die 0 in $\mathbb{Z}_n$ die Restklasse $\overline{0}$, das negative Element zu $\overline{a}$ ist $\overline{-a}$, und die 1 in $\mathbb{Z}_n$ ist die Restklasse $\overline{1}$. Damit prüft man alle Axiome leicht nach. □

Definition. Der Ring $(\mathbb{Z}_n, +, \cdot)$ mit den Verknüpfungen aus ($*$) wird *Restklassenring modulo n* genannt.

Beispiel.

(i) $\mathbb{Z}_2 = \{\overline{0}, \overline{1}\}$, wobei

$$\overline{0} = 2\mathbb{Z} = \{\text{gerade ganze Zahlen}\},$$
$$\overline{1} = 2\mathbb{Z} = \{\text{ungerade ganze Zahlen}\}.$$

Die Verknüpfungstafeln von $\mathbb{Z}_2$ lauten (beachte $\overline{1}+\overline{1} = \overline{1+1} = \overline{2} = \overline{0}$):

$+$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

und

$\cdot$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$

Aus der Tabelle für + liest man z.B. ab, dass „gerade+ungerade immer ungerade ergibt“ und dass „ungerade+ungerade immer gerade ergibt“. Diese Aussagen sind hiermit auch bewiesen (genauer durch obigen Satz)!

Identifiziert man $\overline{0}$ mit falsch und $\overline{1}$ mit wahr, so entspricht + gerade XOR und $\cdot$ entspricht $\wedge$. Damit ist gezeigt, dass auch $(B, \mathrm{XOR}, \wedge)$ einen kommutativen Ring bildet, und dass dieser als identisch mit dem Ring $(\mathbb{Z}_2, +, \cdot)$ angesehen werden kann.

(ii) Die Verknüpfungstafeln von $\mathbb{Z}_4$ lauten

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

und

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

(iii) In $\mathbb{Z}_7$ gilt:

$$\begin{aligned}
\overline{3} + \overline{5} &= \overline{8} = \overline{1},\\
\overline{3} - \overline{5} &= \overline{3} + (\overline{-5}) = \overline{3} + \overline{-5} = \overline{3-5} = \overline{-2} = \overline{5},\\
\overline{6} \cdot \overline{5} &= \overline{30} = \overline{2},\\
\overline{6} \cdot \overline{5} &= \overline{-1} \cdot \overline{5} = \overline{-5} = \overline{2},\\
\overline{6}^{100000} &= \overline{-1}^{100000} = \overline{(-1)^{100000}} = \overline{1}.
\end{aligned}$$

(iv) In $\mathbb{Z}_6$ gilt $\overline{3} \cdot \overline{2} = \overline{6} = \overline{0}$, aber $\overline{3} \neq \overline{0}$ und $\overline{2} \neq \overline{0}$. Die Restklasse $\overline{0}$ ist aber die 0 in $\mathbb{Z}_6$, d.h. $\mathbb{Z}_6$ ist nicht nullteilerfrei! $\mathbb{Z}_6$ ist auch ein Gegenbeispiel zur Kürzungsregel (vgl. Bemerkung 2.2.6): $\overline{2} \cdot \overline{3} = \overline{4} \cdot \overline{3}$, aber $\overline{2} \neq \overline{4}$.

(v) In $\mathbb{Z}_6$ ist $\overline{5}$ eine Einheit, denn $\overline{5} \cdot \overline{5} = \overline{1}$ und $\overline{1}$ ist die 1. Neben $\overline{1}$ ist $\overline{5}$ sogar die einzige Einheit (man prüfe das nach!), also $\mathbb{Z}_6^\times = \{\overline{1}, \overline{5}\}$. Man beachte $\overline{5} = -\overline{1}$.

Übung. Was für ein Ring ist $\mathbb{Z}_1$?

2.4.4 Gleichungen in $\mathbb{Z}_n$

Beispiel a. Für welche $b \in \mathbb{Z}$ ist $\overline{9} \cdot x = \overline{b}$ in $\mathbb{Z}_{15}$ lösbar?

x	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{7}$	$\overline{8}$	$\overline{9}$	$\overline{10}$	$\overline{11}$	$\overline{12}$	$\overline{13}$	$\overline{14}$
$\overline{9} \cdot x$	$\overline{0}$	$\overline{9}$	$\overline{3}$	$\overline{12}$	$\overline{6}$	$\overline{0}$	$\overline{9}$	$\overline{3}$	$\overline{12}$	$\overline{6}$	$\overline{0}$	$\overline{9}$	$\overline{3}$	$\overline{12}$	$\overline{6}$

Antwort: Es gibt genau dann eine Lösung, wenn $\overline{b} = \overline{0}, \overline{3}, \overline{3}, \overline{9}, \overline{12}$. Für $b = 3$ gibt es z.B. die Lösungen $x = \overline{2}, \overline{7}, \overline{12}$.

Satz. *Es seien $n \in \mathbb{N}$ und $a, b \in \mathbb{Z}$ gegeben. Die Gleichung $\overline{a} \cdot x = \overline{b}$ in $\mathbb{Z}_n$ ist genau dann lösbar, wenn $\mathrm{ggT}(a, n) | b$.*

Beweis. Sei $\overline{a} \cdot x = \overline{b}$ lösbar, etwa $\lambda \in \mathbb{Z}$ mit $\overline{a} \cdot \overline{\lambda} = \overline{b}$. D.h. $n | \lambda a - b$. Aus $\mathrm{ggT}(a, n) | \lambda a$ und $\mathrm{ggT}(a, n) | \lambda a - b$ folgt $\mathrm{ggT}(a, n) | b$ (vgl. Übung 2.2.4a).

Sei umgekehrt $\mathrm{ggT}(a, n) | b$, etwa $c \in \mathbb{Z}$ mit $\mathrm{ggT}(a, n) \cdot c = b$. Nach Algorithmus 2.3.4 gibt es $\lambda, \mu \in \mathbb{Z}$ mit $\mathrm{ggT}(a, n) = \lambda a + \mu n$. Multiplikation mit c liefert $b = (c\lambda)a + (c\mu)n$. In $\mathbb{Z}_n$ bedeutet das $\overline{b} = \overline{c\lambda} \cdot \overline{a}$, d.h. $x = c\lambda$ ist eine Lösung. □

Beispiel b. Löse $\overline{6} \cdot x = \overline{9}$ in $\mathbb{Z}_{15}$. Rechung: Mit dem euklidischen Algorithmus berechnet man $\mathrm{ggT}(6, 15) = 3 = \overline{1 \cdot 15 - 2 \cdot 6}$. Multiplikation mit 3 liefert $9 = 3 \cdot 15 - 6 \cdot 6$. Modulo 15 ergibt sich $\overline{9} = \overline{0} - \overline{6} \cdot \overline{6}$. Folglich ist $x = -\overline{6} = \overline{-6} = \overline{9}$ eine Lösung. Die Lösung ist nicht eindeutig, z.B. ist auch $\overline{6} \cdot \overline{4} = \overline{24} = \overline{9}$ oder $\overline{6} \cdot \overline{14} = \overline{6} \cdot \overline{-1} = \overline{-6} = \overline{9}$.

Folgerung. *Es seien $n \in \mathbb{N}$ und $a \in \mathbb{Z}$.*

(i) $\overline{a} \in \mathbb{Z}_n^\times \Leftrightarrow \mathrm{ggT}(a, n) = 1$.

(ii) $\mathbb{Z}_n$ ist genau dann ein Körper, wenn n eine Primzahl ist.

Beweis. Übung unter Verwendung des Satzes. □

Beispiel c. $\mathbb{Z}_9^\times = \{\overline{1}, \overline{2}, \overline{4}, \overline{5}, \overline{7}, \overline{8}\}$. Es gilt:

$$\overline{1}^{-1} = \overline{1}, \qquad \overline{8}^{-1} = \overline{8},$$
$$\overline{2}^{-1} = \overline{5}, \qquad \overline{7}^{-1} = \overline{4},$$
$$\overline{4}^{-1} = \overline{7}, \qquad \overline{5}^{-1} = \overline{2}.$$

Übung a. Wie lauten alle Einheiten von $\mathbb{Z}_{11}$ und ihre Inversen? Ist $\mathbb{Z}_{11}$ ein Körper?

Übung b. Man zeige, dass für teilerfremde $a, b \in \mathbb{Z}$ stets gilt: $a | bc \Rightarrow a | c$.
Hinweis: Man rechne in $\mathbb{Z}_a$.

Übung c. Wieviele Einheiten hat $\mathbb{Z}_{p^n}$, wenn p eine Primzahl ist?

Übung d. Es sei $n > 1$. Man zeige, dass $\overline{a} \in \mathbb{Z}_n$ genau dann Nullteiler ist, wenn $\mathrm{ggT}(a, n) \neq 1$.

Übung e. Man zeige, dass in $\mathbb{Z}_n$ jedes Element entweder Einheit oder Nullteiler ist.

Übung f. Man prüfe folgende Aussage aus Übung 2.2.6c in verschiedenen $\mathbb{Z}_n$ nach: $(a) = (b) \Leftrightarrow a \sim b$?

2.4.5 Die Euler'sche Funktion

Definition. Für $n \in \mathbb{N}$ definiere

$$\varphi(n) := |\mathbb{Z}_n^\times| = |\{a \in Z \,|\, 0 \leq a < n, \mathrm{ggT}(a,n) = 1\}.$$

Die Abbildung $\varphi : \mathbb{N} \to \mathbb{N}$ heißt die *Euler'sche φ-Funktion.*

Bemerkung.

(i) Für alle $m, n \in \mathbb{N}$ mit $\mathrm{ggT}(m,n) = 1$ gilt $\varphi(mn) = \varphi(m)\varphi(n)$.

(ii) Für alle Primzahlen p gilt $\varphi(p^k) = p^{k-1}(p-1)$.

Beweis. (i) ohne Beweis. (ii) als Übung. (Kombinatorik!) □

Beispiel. $\varphi(9) = \varphi(3^2) = 3^1(3-1) = 3 \cdot 2 = 6$.
$\varphi(20) = \varphi(4) \cdot \varphi(5) = 2^1(2-1) \cdot 5^0(5-1) = 2 \cdot 4 = 8$.

2.5 Das RSA-Kryptosystem

2.5.1 Kryptosysteme

Ein *Kryptosystem* besteht aus einer endlichen Menge M von Nachrichten, einer endlichen Menge $K \subseteq P \times S$ von Schlüsselpaaren, einer Familie von Abbildungen $c_p : M \to M$ (crypt) für alle $p \in P$, und einer Familie von Abbildungen $d_s : M \to M$ (decrypt) für alle $s \in S$, so dass für jedes Schlüsselpaar $(p,s) \in K$ gilt: $d_s \circ c_p = \mathrm{id}_M$. Da die Menge M endlich ist folgt, dass alle c_p und d_s notwendigerweise bijektiv sind und $d_s = c_p^{-1}$ gilt. Der Sender einer Nachricht $m \in M$ und Besitzer des Schlüsselteils p berechnet $m' := c_p(m)$ und verschickt m'. Der Empfänger einer verschlüsselten Nachricht m' und Besitzer des Schlüsselpaares (p,s) berechnet $d_s(m') = d_s(c_p(m)) = m$.

Für die praktische Berechenbarkeit von $c_p(m)$ und $d_s(m')$ ist es sinnvoll, dass M eine algebraische Struktur hat und die Abbildungsvorschriften von c_p und d_s durch Rechenoperationen gegeben sind.

Beispiel. Es sei $(G, \cdot)$ eine Gruppe. Wähle $M = P = S = G$. Nach Übung 2.1.4c ist $\lambda_a : G \to G, x \mapsto ax$ für jedes $a \in G$ bijektiv und hat die Umkehrabbildung $(\lambda_a)^{-1} = \lambda_{a^{-1}}$. Daher können wir wählen: $K = \{(a, a^{-1}) \,|\, a \in G\}$ und $c_a = d_a = \lambda_a$ für alle $a \in G$. Konkrete Beispiele sind:

(i) $G = (\mathbb{Z}_{26}, +)$. Die Elemente von G können z.B. mit den Buchstaben `A`–`Z` identifiziert werden. Es ist $K = \{(a, -a) \,|\, a \in \mathbb{Z}_{26}\}$. Für $(p,s) = (a,-a) \in K$ bedeutet die Abbildung $c_p = \lambda_a : \mathbb{Z}_{26} \to \mathbb{Z}_{26}, m \to m + a$,

dass die Buchstaben gemäß ihrer Reihenfolge im Alphabet zyklisch um a verschoben werden, bei $a = 2$ etwa: `A` $\mapsto$ `C`, `B` $\mapsto$ `D`,...,`Y` $\mapsto$ `A`,`Z` $\mapsto$ `B`. Die Abbildung d_s lautet $d_s = \lambda_{-a} : \mathbb{Z}_{26} \to \mathbb{Z}_{26}, m \to m - a$.

(ii) $G = (\{0,1\}^n, \oplus)$, die Menge der n-bit-Wörter mit bitweisem xor. Für alle $a \in G$ gilt $-a = a$, also $K = \{(a,a) \mid a \in G\}$. Für $(p,s) = (a,a) \in K$ sind also c_p und d_s identisch gleich $\{0,1\}^n \to \{0,1\}^n, w \mapsto w \oplus a$.

(iii) $G = (K^\times, \cdot)$ für einen endlichen Körper K. Als Körper kann z.B. $K = \mathbb{F}_{256}$ gewählt werden, dessen Elemente mit Bytes identifiziert werden können. Dieser spielt als Baustein auch im Kryptografie-Standard AES (advanced encryption standard, 2001) eine Rolle.

Das Kryptosystem heißt *asymmetrisch* oder *public key*, wenn es praktisch nicht möglich ist, aus der Kenntnis von p auf die Abbildungsvorschrift von d_s zu schliessen. Dadurch braucht p nicht geheim gehalten werden, sondern kann öffentlich gemacht werden. Anderenfalls heißt das Kryptosystem *symmetrisch* oder *secret key*. Die Kryptosysteme des Beispiels sind alle als symmetrisch einzustufen.

RSA ist ein *asymmetrisches* Kryptosystem, das 1978 von den Mathematikern Rivest, Shamir und Adleman erfunden wurde. Es benutzt $M = \mathbb{Z}_n$ und die Potenzierungsabbildung

$$c_k : M \to M, m \mapsto m^k.$$

Übung a. Können die Abbildungen $c_k : \mathbb{Z}_{26} \to \mathbb{Z}_{26}, m \to km$ für ein Kryptosystem verwendet werden? Können die Abbildungen $c_k : \mathbb{Z}_{26} \to \mathbb{Z}_{26}, m \to 3m - k$ für ein Kryptosystem verwendet werden?

Übung b. Man erläutere, wie man in den Kryptosystemen des Beispiels für jedes Schlüsselpaar (p,s) von p auf s schliessen kann.

2.5.2 Potenzen in $\mathbb{Z}_n$

Beispiel. Wie sehen die Abbildungen $c : \mathbb{Z}_{11} \to \mathbb{Z}_{11}, m \mapsto m^k$ für verschiedene k aus? Für welche k ist c bijektiv? Wie sieht die Umkehrabbildung aus?

x	0	1	2	3	4	5	6	7	8	9	10	bijektiv
x^2	0	1	4	9	5	3	3	5	9	4	1	—
x^3	0	1	8	5	9	4	7	2	6	3	10	$\surd$
x^4	0	1	5	4	3	9	9	3	4	5	1	—
x^5	0	1	10	1	1	1	10	10	10	1	10	—
x^6	0	1	9	3	4	5	5	4	3	9	1	—
x^7	0	1	7	9	5	3	8	6	2	4	10	$\surd$
x^8	0	1	3	5	9	4	4	9	5	3	1	—
x^9	0	1	6	4	3	9	2	8	7	5	10	$\surd$
x^{10}	0	1	1	1	1	1	1	1	1	1	1	—
x^{11}	0	1	2	3	4	5	6	7	8	9	10	$\surd$

Für $c : x \mapsto x^3$ ist $c^{-1} : x \mapsto x^7$. Für $c : x \mapsto x^9$ ist $c^{-1} : x \mapsto x^9$.

Beispiel. Wie sehen die Abbildungen $c : \mathbb{Z}_{10} \to \mathbb{Z}_{10}, m \mapsto m^k$ für verschiedene k aus? Für welche k ist c bijektiv? Wie sieht die Umkehrabbildung aus?

x^1	0	1	2	3	4	5	6	7	8	9	bijektiv
x^2	0	1	4	9	6	5	6	9	4	1	—
x^3	0	1	8	7	4	5	6	3	2	9	$\surd$
x^4	0	1	6	1	6	5	6	1	6	1	—
x^5	0	1	2	3	4	5	6	7	8	9	$\surd$

Für $c : x \mapsto x^3$ ist $c^{-1} = c$.

2.5.3 Der kleine Satz von Fermat

Satz (Fermat)**.** *Ist p eine Primzahl so gilt in $\mathbb{Z}_p$:*

$$x^{p-1} = 1 \text{ für alle } x \neq 0.$$

Daraus folgt $x^p = x$ für alle $x \in \mathbb{Z}_p$.

Beweis. Es sei $x \in \mathbb{Z}_p \setminus \{0\}$. Da p eine Primzahl ist, ist x invertierbar ($\mathbb{Z}_p$ ist ein Körper). Also ist die Abbildung $\mathbb{Z}_p \to \mathbb{Z}_p, y \mapsto xy$ bijektiv. Die Liste von Elementen $\overline{1} \cdot x, \overline{2} \cdot x, \ldots, \overline{p-1} \cdot x$ stimmt also bis auf Reihenfolge mit $\overline{1}, \overline{2}, \ldots, \overline{p-1}$ überein. Somit gilt in $\mathbb{Z}_p$:

$$\overline{(p-1)!} = \overline{1} \cdot \overline{2} \cdot \ldots \cdot \overline{p-1} = (\overline{1} \cdot x) \cdot (\overline{2} \cdot x) \cdot \ldots \cdot (\overline{p-1} \cdot x) = \overline{(p-1)!} \cdot x^{p-1}.$$

Es gilt $p \nmid (p-1)!$, da alle Faktoren von $(p-1)!$ kleiner als p sind. D.h. $\overline{(p-1)!} \neq 0$. Da $\mathbb{Z}_p$ nullteilerfrei ist, folgt $x^{p-1} = 1$ (Kürzungsregel). □

Beispiel. $p = 7, a = 3 : a^6 = 729 = 104 \cdot 7 + 1$.
$p = 11, a = 2 : a^{10} = 1024 = 93 \cdot 11 + 1$.
Die Kongruenz kann auch gelten, wenn p keine Primzahl ist:
$p = 341 = 2 \cdot 11 \cdot 17, a = 2 : a^{340} \equiv 1 \pmod{341}$.
$p = 91 = 7 \cdot 13, a = 3 : a^{90} \equiv 1 \pmod{91}$.

Bemerkung. Der kleine Satz von Fermat kann als Primzahltest verwendet werden. Für gegebenes $p \in \mathbb{N}$ testet man die Kongruenz $a^p \equiv a \pmod{p}$ für verschiedene Werte von a. Das ist rechnerisch leicht, da nur in $\mathbb{Z}_p$ gearbeitet wird. Gilt sie für ein a nicht, so ist p keine Primzahl. Gilt sie für mehrere a, so ist p mit sehr hoher Wahrscheinlichkeit eine Primzahl.

2.5.4 RSA

Es sind n, k und l so zu finden, dass

$$x^{kl} = x \tag{2.1}$$

für alle $x \in \mathbb{Z}_n$. Dann sind die Abbildungen

$$c : \mathbb{Z}_n \to \mathbb{Z}_n, \quad x \mapsto x^k$$
$$d : \mathbb{Z}_n \to \mathbb{Z}_n, \quad x \mapsto x^l$$

bijektiv mit $d \circ c = \mathrm{id}$ und $c \circ d = \mathrm{id}$, denn $d(c(x)) = (x^k)^l = x^{kl} = x$ und $c(d(x)) = (x^l)^k = x^{kl} = x$.

Satz. *Es seien p und q zwei verschiedene Primzahlen. Setze*

$$n := pq, \quad m := (p-1)(q-1).$$

Es sei weiter k eine beliebige Zahl teilerfremd zu m und l eine Zahl mit

$$kl \equiv 1 \pmod{m}.$$

Dann erfüllen n, k, l die Bedingung (2.1).

Beweis. Vorbetrachtung: Für alle $a, b \in \mathbb{Z}$ gilt $\overline{a} = \overline{b}$ in $\mathbb{Z}_n$ genau dann, wenn $\overline{a} = \overline{b}$ in $\mathbb{Z}_p$ und in $\mathbb{Z}_q$ gilt. In der Tat, da $n = pq$ und p und q verschiedene Primzahlen sind, gilt $n|a-b$ genau dann, wenn $p|a-b$ und $q|a-b$ (das folgt aus der Primfaktorzerlegung).

Zu zeigen ist (2.1) für alle $x \in \mathbb{Z}_n$. Nach der Vorbetrachtung reicht es, (2.1) in $\mathbb{Z}_p$ und in $\mathbb{Z}_q$ nachzuweisen. Wir rechnen im folgenden nur noch in $\mathbb{Z}_p$ (in $\mathbb{Z}_q$ geht es genauso). Falls $x = 0$, so gilt $x^{kl} = x$ trivialerweise. Falls $x \neq 0$, so gilt nach dem Satz von Fermat $x^{p-1} = 1$. Da m ein Vielfaches von $p-1$ ist, gilt somit auch $x^m = 1$. Nach Wahl von l ist $kl = 1 + \lambda m$ für ein $\lambda \in \mathbb{Z}$. Es folgt $x^{kl} = x^{1+\lambda m} = x(x^m)^\lambda = x$. Damit ist (2.1) gezeigt. □

Beispiel. Wähle $p = 11$ und $q = 13$. Dann ist $n = 11 \cdot 13 = 143$ und $m = 10 \cdot 12 = 120$. Wähle ein (möglichst kleines) k, dass teilerfremd zu 120 ist, etwa $k = 7$. Berechne das Inverse von 7 modulo 120: wegen $7 \cdot 17 = 119 \equiv -1 \pmod{120}$ ist $l = 120 - 17 = 103$ eine Lösung. Es ergeben sich

$$c : \mathbb{Z}_{143} \to \mathbb{Z}_{143}, \quad x \mapsto x^7$$
$$d : \mathbb{Z}_{143} \to \mathbb{Z}_{143}, \quad x \mapsto x^{103}$$

2.6 Polynome

In diesem Abschnitt sei K ein Körper.

2.6.1 Definition und Beispiele

Definition.

(i) Ein *Polynom* über K in der *Unbestimmten* X ist ein Ausdruck der Form

$$f = \sum_{i=0}^{n} a_i X^i$$

mit $a_i \in K$ für alle $i = 0, \ldots, n$. Die a_i heißen die *Koeffizienten* des Polynoms, insbesondere heißt a_0 der *absolute Koeffizient*.
(Koeffizienten, die gleich 0 sind können beliebig hinzugefügt oder weggelassen werden, ohne den Ausdruck zu verändern.)

(ii) Zwei Polynome $f = \sum_{i=0}^{n} a_i X^i$ und $g = \sum_{i=0}^{n} b_i X^i$ sind genau dann *gleich*, wenn $a_i = b_i$ für alle $i = 0, \ldots, n$.

(iii) Die Menge aller Polynome über K wird mit $K[X]$ bezeichnet.

(iv) Sind alle Koeffizienten von $f \in K[X]$ gleich 0, so heißt f das *Nullpolynom*, geschr. $f = 0$.

(v) Ist $f \in K[X]$ nicht das Nullpolynom, dann wird das größte $i \in \mathbb{N}_0$, für das $a_i \neq 0$ ist, der *Grad* von f genannt und mit $\deg f$ bezeichnet. Für das Nullpolynom setzen wir $\deg 0 := -\infty$.

(vi) Ist $\deg f = n \geq 0$, so heißt a_n der *Hauptkoeffizient* von f.

(vii) Ein Polynom heißt *normiert*, wenn der Hauptkoeffizient gleich 1 ist.

(viii) Ein Polynom f heißt *linear*, wenn $\deg f = 1$.

(ix) Ein Polynom f heißt *konstant*, wenn $\deg f \leq 0$.

Schreibweise. Der Kürze halber schreibt man X^i statt $1X^i$, X statt X^1, a_0 statt a_0X^0, und $0X^i$ lässt man weg.

Beispiel.

(i) $f = 1X^4 + 0X^3 - \frac{1}{3}X^2 + 1X^1 - 2X^0 = X^4 - \frac{1}{3}X^2 + X - 2 \in \mathbb{R}[X]$.

(ii) $g = \overline{1}X^2 + \overline{1}X^1 + \overline{0}X^0 = X^2 + X \in \mathbb{Z}_2[X]$.

Bemerkung a. Jedes Polynom $f \in K[X]$ definiert eine Abbildung $K \to K$ dadurch, dass man das „Einsetzen“ in die Unbestimmte als Zuordnungsvorschrift wählt. Diese Abbildung bezeichnen wir ebenfalls mit f, sprechen aber zur Unterscheidung von der *Polynomfunktion* zu f. Für jedes $a \in K$ nennen wir $f(a)$ den *Wert von* f an der Stelle a.

Die Polynome f und g aus dem Beispiel haben die Polynomfunktionen

$$f : \mathbb{R} \to \mathbb{R}, \quad a \mapsto f(a) = a^4 - \frac{1}{3}a^2 + a^1 - 2a^0 = a^4 - \frac{1}{3}a^2 + a - 2.$$

$$g : \mathbb{Z}_2 \to \mathbb{Z}_2, \quad a \mapsto g(a) = a^2 - a = 0.$$

(Man beachte, dass $a^2 - a = 0$ für alle $a \in \mathbb{Z}_2$.)

Verschiedene Polynome können dieselbe Polynomfunktion haben (z.B. g und das Nullpolynom). Aus diesem Grund sind Polynomfunktionen und Polynome zu unterscheiden.

Bemerkung b. Jedes $a \in K$ kann als konstantes Polynom aX^0 aufgefasst werden. Auf diese Weise wird K zu einer Teilmenge von $K[X]$.

2.6.2 Der Polynomring

Für Polynome gibt es eine natürliche Addition und Multiplikation, die aus der Menge $K[X]$ einen Ring macht.

Definition. Für beliebige Polynome $f = \sum_{i=0}^{n} a_i X^i$ und $g = \sum_{i=0}^{n} b_i X^i$ aus $K[X]$ wird deren Summe und Produkt definiert als:

$$f + g := \sum_{i=0}^{n} (a_i + b_i) X^i,$$

$$f \cdot g := \sum_{i=0}^{n} c_i X^i \text{ mit } c_i := \sum_{k=0}^{i} a_k b_{i-k}.$$

Bemerkung.

(i) Mit dieser Addition und Multiplikation wird $K[X]$ ein kommutativer Ring. (Man prüfe die Ringaxiome nach!) Das neutrale Element der Addition ist das Nullpolynom, und das neutrale Element der Multiplikation ist das konstante Polynom $1 = 1X^0$.

(ii) Man kann Polynome auch als endliche Folgen auffassen, etwa das Polynom $-3X^2 + X + 2$ als die Folge $(2, 1, -3, 0, 0, \ldots)$. Somit haben wir die Inklusion $K[X] \subseteq \mathrm{Abb}(\mathbb{N}, K)$. Dabei stimmt die Addition in $K[X]$ mit der punktweisen Addition in $\mathrm{Abb}(\mathbb{N}, K)$ überein, nicht aber die Multiplikation.

(iii) K ist ein „Teilring“ von $K[X]$.

(iv) Für jedes $a \in K$ ist die Abbildung

$$\tau_a : K[X] \to K, \quad f \mapsto f(a)$$

ein Ringhomomorphismus (der *Einsetzungshomomorphismus*).

(v) Es gelten die Gradformeln

$$\begin{aligned} \deg(f+g) &\leq \max\{\deg f, \deg g\}, \\ \deg(f \cdot g) &= \deg f + \deg g. \end{aligned}$$

(vi) $K[X]$ ist nullteilerfrei.

(vii) In $K[X]$ gilt die Kürzungsregel, d.h. für alle $f, g, h \in K[X]$ mit $f \neq 0$ gilt:

$$fg = fh \Rightarrow g = h.$$

(viii) Die Einheitengruppe des Ringes $K[X]$ lautet

$$\begin{aligned} K[X]^\times &= \{f \in K[X] \mid \deg f = 0\} = \{\text{konstante Polynome} \neq 0\} \\ &= K^\times = K \backslash \{0\}. \end{aligned}$$

(ix) Betrachte die Teilbarkeitsrelation $|$ auf $K[X]$:

$$x - 1 | x^2 - 2x + 1, \quad x^2 + x + 1 | x^3 - 1, \quad 2x + 4 | x + 2, \quad \ldots$$

Auf der Menge der normierten Polynome aus $K[X]$ bildet $|$ eine partielle Ordnung. Aus $f|g$ folgt offensichtlich $\deg f \leq \deg g$.

Beweis. Übung. □

2.6.3 Polynomdivision

Satz. *Es seien $f, g \in K[X]$ mit $g \neq 0$. Dann existieren eindeutige $q, r \in K[X]$ mit $f = qg + r$ und $\deg r < \deg g$.*

Beweis. Eindeutigkeit: Angenommen, wir haben $q, q', r, r' \in K[X]$ mit $qg + r = q'g + r'$ und $\deg r, r' < \deg g$. Dann ist $(q - q')g = r' - r$, also

$$\deg(q - q') + \deg g = \deg(r' - r) \leq \max\{\deg r', \deg r\} < \deg g.$$

Es folgt $\deg(q - q') < 0$, d.h. $q - q' = 0$. Somit ist $q = q'$ und $r = r'$.

Existenz: Wir können $\deg f \geq \deg g$ annehmen, denn sonst ist $f = 0 \cdot g + f$ und $\deg f < \deg g$. Zusammen mit der Voraussetzung $g \neq 0$ haben wir $\deg f \geq \deg g \geq 0$ und können somit eine vollständige Induktion nach $\deg f$ führen.
Induktionsanfang ($\deg f = 0$): Dann ist auch $\deg g = 0$, d.h. f und g sind beide konstant und ungleich 0. Für $f = a_0$ und $g = b_0$ mit $a_0, b_0 \in K\backslash\{0\}$ gilt aber $f = \frac{a_0}{b_0} \cdot g + 0$ und $\deg 0 = -\infty < 0 = \deg g$. Damit ist der Induktionsanfang erledigt.
Induktionsschritt: Sei jetzt $\deg f = n > 0$ und sei die Existenz von q und r für alle f mit $\deg f < n$ bereits bewiesen (Ind.Vor.). Es seien $f = \sum_{i=0}^{n} a_i X^i$ und $g = \sum_{i=0}^{m} b_i X^i$ mit $a_n, b_m \neq 0$ und $m \leq n$. Setzt man $f' := f - \frac{a_n}{b_m} X^{n-m} g$, so ist $\deg f' < n$. Nach Induktionsvoraussetzung gibt es $q', r \in K[X]$ mit $f' = q'g + r$ und $\deg r < \deg g$. Es folgt $f = (\frac{a_n}{b_m} X^{n-m} + q')g + r$, d.h. $q := \frac{a_n}{b_m} X^{n-m} + q'$ und r sind wie gewünscht. □

Beispiel. $f = 2X^3 - 9X^2 + 4X, g = X^2 - 3X - 4 \in \mathbb{Q}[X]$. Wir dividieren f durch g mit Rest:

$$\begin{array}{rrrl}
(2X^3 & -9X^2 + & 4X) & : (X^2 - 3X - 4) = 2X - 3 \\
-(2X^3 & -6X^2 - & 8X) & \\
\hline
 & -3X^2 + & 12X & \\
 & -(-3X^2 + & 9X + 12) & \\
\hline
 & & 3X - 12 &
\end{array}$$

Also

$$f = \underbrace{(2X - 3)}_{q} \cdot g + \underbrace{3X - 12}_{r}, \quad \deg r = 1 < 2 = \deg g.$$

Folgerung. *Für alle $f, g \in K[X]$ hat die Menge der normierten gemeinsamen Teiler von f und g*

$$\{h \in K[X] \mid h \text{ normiert}, h|f, h|g\}$$

bzgl. der Ordnung $|$ ein Maximum.

Beweis. Es sei

$$M = \{h \in K[X] \mid h \text{ normiert}, h|f, h|g\}.$$

Behauptung: Hat M ein Element der Form $d = \lambda f + \mu g$ mit $\lambda, \mu \in K[X]$, so ist d ein Maximum von M bzgl. $|$. Aus dem Euklidischen Algorithmus folgt die Existenz eines solchen Elementes, und damit die Aussage des Korollars. Es bleibt, die Behauptung zu zeigen. Sei dazu d' ein beliebiges Element aus M. Zu zeigen ist $d'|d$. Aus $d'|f$ und $d'|g$ folgt aber $d'|\lambda f + \mu g = d$ nach Bemerkung (2.2.4). □

Definition. Das Maximum wird der größte gemeinsame Teiler von f und g genannt, geschr. $\mathrm{ggT}(f, g)$.

Beispiel (Fortsetzung)**.** Wir dividieren g durch r mit Rest:

$$\begin{array}{rrl} (X^2 & -3X & -4) : (3X - 12) = \frac{1}{3}X + \frac{1}{3} = \frac{1}{3}(X+1) \\ -(X^2 & -4X) & \\ \hline & X & -4 \\ & -(X & -4) \\ \hline & & 0 \end{array}$$

D.h. $g = \frac{1}{3}(X+1) \cdot r + 0$. Damit ist r bis auf Normierung der ggT von f und g, also $\mathrm{ggT}(f, g) = X - 4$. Rückwärtseinsetzen liefert weiterhin die Darstellung:

$$\begin{aligned} \mathrm{ggT}(f, g) = X - 4 &= \frac{1}{3}(3X - 12) = \frac{1}{3}(f - (2X - 3)g) \\ &= \frac{1}{3} \cdot f - \frac{1}{3}(2X - 3) \cdot g. \end{aligned}$$

2.6.4 Nullstellen

Definition. Wir sagen $a \in K$ ist *Nullstelle* eines Polynoms $f \in K[X]$, wenn $f(a) = 0$ gilt, d.h. wenn die durch f definierte Polynomfunktion an der Stelle a den Wert 0 hat.

Satz. *Es seien $f \in K[X]$ und $a \in K$. Dann gilt:*

$$f(a) = 0 \Leftrightarrow X - a \text{ teilt } f.$$

Beweis. $\Leftarrow$: Es sei $f = (X - a) \cdot g$ mit $g \in K[X]$. Da τ_a (das Einsetzen von a) ein Homomorphismus ist, folgt $f(a) = (a - a) \cdot g(a) = 0$.

$\Rightarrow$: Es sei $f(a) = 0$. Nach Polynomdivision gibt es eindeutig bestimmte $q, r \in K[X]$ mit $f = q \cdot (X - a) + r$ und $\deg r < \deg(X - a) = 1$. Das

bedeutet, dass r konstant ist, also $r = r_0 \in K$. Da τ_a ein Homomorphismus ist, folgt $0 = f(a) = q(a)(a-a) + r(a) = q(a) \cdot 0 + r(a) = r_0$. Somit ist r das Nullpolynom und $f = (X-a) \cdot q$. □

Definition. Es seien $0 \neq f \in K[X]$ und $a \in K$. Die Teiler von f der Form $X - a$ werden *Linearfaktoren* von f genannt. Weiter heißt

$$\max\{n \in \mathbb{N}_0 \mid (X-a)^n \text{ teilt } f\}$$

die *Vielfachheit von* a als Nullstelle von f.

Bemerkung. Wegen der Gradformel aus Bemerkung (2.6.2) ist die Vielfachheit stets $\leq \deg f$, also insbesondere endlich. Der Satz besagt, dass a genau dann Nullstelle von $f \neq 0$ ist, wenn a Vielfachheit ≥ 1 hat.

2.6.5 Zerlegung in Linearfaktoren

Satz. *Es sei $0 \neq f \in K[X]$. Sind $a_1, \ldots, a_l$ paarweise verschiedene Nullstellen von f mit den Vielfachheiten $n_1, \ldots, n_l$, so gilt*

$$f = (X-a_1)^{n_1} \cdots (X-a_l)^{n_l} \cdot g \tag{2.2}$$

für ein $0 \neq g \in K[X]$ mit $g(a_1), \ldots, g(a_l) \neq 0$.

Beweis. Wenn f die Zerlegung (2.2) hat, dann folgt $g(a_i) \neq 0$ aus der Maximalität der n_i. In der Tat, falls $g(a_i) = 0$ dann würde g nach Satz (2.6.4) von $X - a_i$ geteilt werden, woraus $(X-a_i)^{n_i+1} | f$ folgt.

Wir zeigen nun per Induktion nach l, dass die Zerlegung (2.2) existiert. Für $l = 1$ folgt das aus der Definition der Vielfachheit. Sei also $l > 1$ und die Behauptung für $l-1$ bereits bewiesen. Dann gibt es $0 \neq g \in K[X]$ mit $f = (X-a_1)^{n_1} \cdots (X-a_{l-1})^{n_{l-1}} \cdot g$. Setzen wir $h := (X-a_1)^{n_1} \cdots (X-a_{l-1})^{n_{l-1}}$, so erhalten wir $f = hg$. Da die $a_1, \ldots, a_l$ paarweise verschieden sind, ist $h(a_l) = (a_l - a_1)^{n_1} \cdots (a_l - a_{l-1})^{n_{l-1}} \neq 0$. Nach Voraussetzung gilt $(X-a_l)^{n_l} | f = hg$. Das folgende Lemma zeigt, dass dann g von $(X-a_l)^{n_l}$ geteilt wird. Damit ist die Behauptung bewiesen. □

Lemma. *Es seien $g, h \in K[X], a \in K$ und $n \in \mathbb{N}$. Aus $(X-a)^n | hg$ und $h(a) \neq 0$ folgt $(X-a)^n | g$.*

Beweis. Induktion nach n. $n = 1$: Wegen $X - a | hg$ gilt $h(a)g(a) = (hg)(a) = 0$, also $g(a) = 0$ denn $h(a) \neq 0$. Nach Satz (2.6.4) bedeutet das $X - a | g$.

Sei nun $n > 1$ und die Behauptung für n bereits bewiesen. Sei $(X-a)^n | hg$. Da insbesondere $X - a | hg$, so folgt nach der Überlegung für $n = 1$, dass $X - a | g$. Sei $g = (X-a) \cdot g'$, also $(X-a)^n | hg = (X-a)hg'$. Mit der Kürzungsregel in $K[X]$ folgt $(X-a)^{n-1} | hg'$, und daraus nach Induktionsvoraussetzung $(X-a)^{n-1} | g'$. Insgesamt also $(X-a)^n | g$. □

Folgerung. *Es sei $0 \neq f \in K[X]$. Sind $a_1, \ldots, a_l$ paarweise verschiedene Nullstellen von f mit den Vielfachheiten $n_1, \ldots, n_l$, so gilt $\sum_{i=1}^{l} n_i \leq \deg f$.*

Das heißt, jedes Polynom f hat höchstens $\deg f$ viele Nullstellen, wenn jede Nullstelle mit ihrer Vielfachheit gezählt wird.

Beweis. Folgt sofort aus dem Satz. □

Definition. Es sei $0 \neq f \in K[X]$. Wir sagen *f zerfällt vollständig in Linearfaktoren* (über K), wenn es paarweise verschiedenen Nullstellen $a_1, \ldots, a_l$ gibt, deren Vielfachheiten $\sum_{i=1}^{l} n_i = \deg f$ erfüllen. Das ist genau dann der Fall, wenn es eine Zerlegung

$$f = c(X - a_1)^{n_1} \cdots (X - a_l)^{n_l}$$

gibt mit $c \in K$ konstant.

2.6.6 Fundamentalsatz der Algebra

Satz. *Jedes Polynom $f \in \mathbb{C}[X]$ zerfällt vollständig in Linearfaktoren.*

Beispiel.

$$\begin{aligned} f(X) &= X^4 - 1 = (X^2 - 1)(X^2 + 1) \\ &= (X + 1)(X - 1)(X^2 + 1) \\ &= (X + 1)(X - 1)(X - i)(X + i) \end{aligned}$$

Folgerung. *Jedes Polynom $f \in \mathbb{R}[X]$ besitzt eine Zerlegung $f = f_1 \ldots f_l$ mit allen $f_i \in \mathbb{R}[X]$ und $\deg f_i \leq 2$.*

Beweis. Für $z = a + bi \in \mathbb{C}$ mit $a, b \in \mathbb{R}$ heißt $\overline{z} = a - bi$ das *konjugierte Element* zu z. Offensichtlich gilt $\overline{z} = z$ genau dann, wenn $z \in \mathbb{R}$. Da die Abbildung $\mathbb{C} \to \mathbb{C}, z \mapsto \overline{z}$ ein Ringisomorphismus ist (man prüfe das nach!), gilt $f(\overline{z}) = \overline{f(z)}$. Folglich ist $f(z) = 0 \Leftrightarrow f(\overline{z}) = 0$. Die komplexen (nichtreellen) Nullstellen treten also in Paaren, bestehend aus z und $\overline{z}$, auf. Somit hat f nach dem Fundamentalsatz eine Zerlegung der Form

$$f = c(X - a_1) \cdots (X - a_r)(X - z_1)(X - \overline{z_1}) \cdots (X - z_s)(X - \overline{z_s})$$

mit $a_1, \ldots, a_r \in \mathbb{R}, z_1, \ldots, z_s \in \mathbb{C}\backslash\mathbb{R}, c \in \mathbb{C}$ und $r + 2s = \deg f$. Man sagt, das Polynom f hat r reelle Nullstellen und s Paare komplex-konjugierter Nullstellen. Die Behauptung folgt nun, weil

$$(X - z)(X - \overline{z}) = X^2 - (z + \overline{z})X + z\overline{z} \in \mathbb{R}[X].$$

(Man prüfe nach, dass $z + \overline{z}$ und $z\overline{z}$ tatsächlich reell sind!) □

2.7 Boole'sche Algebren

2.7.1 Boole'sche Funktionen

Es sei $B = \{f, w\} = \{0, 1\}$ $(0 = f, 1 = w)$.

Definition. Eine Abbildung $B^n \to B$ wird n-stellige *logische Verknüpfung* oder auch *Bool'sche Funktion* genannt.

Beispiel. (i) $\neg : B \to B, 0 \mapsto 1, 1 \mapsto 0$ ist eine 1-stellige logische Verknüpfung.

(ii) $\wedge : B \times B \to B, (x, y) \mapsto x \wedge y$ ist eine 2-stellige logische Verknüpfung.

(iii) Neue logische Verknüpfungen können durch Aufstellen eines *Terms* definiert werden, der Variablen $(X_1, X_2, \ldots)$, Konstanten (0 und 1), Symbole für schon definierte logische Verknüpfungen $(\circ_1, \circ_2, \ldots)$ und Klammern $((,))$ enthält. Z.B. beschreibt $X_1 \vee (X_1 \wedge X_2)$ die gleiche logische Verknüpfung wie X_1 (Tabelle siehe Vorlesung).

(iv) Logische Verknüfpungen können durch Aufstellen einer Wertetabelle definiert werden. Z.B. beschreibt folgende Tabelle einen *Volladdierer* (s steht für Summe, c für Übertrag):

x	y	z	s	c
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

Der Volladdierer vermag drei einstellige Binärzahlen zu addieren.

2.7.2 Universale Systeme

Definition. Eine endliche Menge von logischen Verknüpfungen $\{\circ_1, \ldots, \circ_r\}$ heißt *universal*, wenn sich jede n-stellige Bool'sche Funktion (für beliebiges n) durch einen Term in $X_1, \ldots, X_n$ und $\circ_1, \ldots, \circ_r$ definieren lässt.

Beispiel a. (i) $\{\wedge, \vee, \neg\}$ ist universal, da jede logische Verknüpfung eine *disjunktive Normalform besitzt.* Z.B. gilt für die Ausgaben s und c des Volladdierers:

$$s(x,y,z) = (\neg x \wedge \neg y \wedge z) \vee (\neg x \wedge y \wedge \neg z) \vee (x \wedge \neg y \wedge \neg z) \vee (x \wedge y \wedge z)$$
$$c(x,y,z) = (\neg x \wedge y \wedge z) \vee (x \wedge y \wedge \neg z) \vee (x \wedge \neg y \wedge z) \vee (x \wedge y \wedge z)$$

(ii) $\{\wedge, \neg\}$ ist universal, denn $x \vee y = \neg(\neg x \wedge \neg y)$.

(iii) $\{\mathrm{XOR}, \wedge\}$ ist universal, denn $\neg x = 1 \,\mathrm{XOR}\, x$.

Frage. Welches universelle System ist zu bevorzugen? Dann kann abhängen von

(i) technischen Gegebenheiten (z.B. welche Verknüpfungen lassen sich leicht durch Schaltungen realisieren?),

(ii) der algebraischen Struktur des Systems (z.B. mit welchem System lassen sich Terme leicht vereinfachen, etwa $(x_1 \vee x_2) \wedge \neg x_1$ zu $x_2 \wedge \neg x_1$?).

Bemerkung. $(B, \mathrm{XOR}, \wedge)$ ist ein Ring, der isomorph ist zu $(\mathbb{Z}_2, +, \cdot)$.

Beweis. $x \,\mathrm{XOR}\, y = x + y$ in $\mathbb{Z}_2$, $x \wedge y = x \cdot y$ in $\mathbb{Z}_2$. □

Beispiel b. Wir wollen $\vee$ möglichst einfach durch $\mathrm{XOR}, \wedge$ ausdrücken. Zuerst drücken wir $\vee$ gemäss Beispiel a aus:

$$x \vee y = \neg(\neg x \wedge \neg y) = 1 \,\mathrm{XOR}((1 \,\mathrm{XOR}\, x) \wedge (1 \,\mathrm{XOR}\, y)).$$

Nun vereinfachen mit Hilfe bekannter Rechenregeln für $\mathbb{Z}_2$

$$x \vee y = 1 + ((1+x)(1+y)) = 1 + (1 + x + y + xy) = x + y + xy$$
$$= x \,\mathrm{XOR}\, y \,\mathrm{XOR}(x \wedge y).$$

Frage. Welche algebraische Struktur hat $(B, \wedge, \vee, \neg)$?

Übung. Wir definieren $x \,\mathrm{NAND}\, y := \neg(a \wedge b)$. Man zeige, dass $\{\mathrm{NAND}\}$ universal ist.

2.7.3 Bool'sche Algebren

Definition. Es sei A eine Menge mit zwei 2-stelligen Verknüpfungen $\oplus, \odot$ und einer 1-stelligen Verknüpfung $\bar{}$. Wir nennen $(A, \oplus, \odot, \bar{})$ eine *Bool'sche Algebra*, wenn folgende Axiome gelten:

(B1) $(A, \oplus)$ ist abelsches Monoid mit neutralem Element 0,

(B2) $(A, \oplus)$ ist abelsches Monoid mit neutralem Element 1,

(B3) Für alle $a \in A$ gilt: $a \oplus \overline{a} = 1, a \odot \overline{a} = 0$.

(B3) Für alle $a, b, c \in A$ gilt:

$$a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$$
$$a \oplus (b \odot c) = (a \oplus b) \odot (a \oplus c)$$

In diesem Fall wird $\overline{a}$ das *Komplement* von a genannt.

Bemerkung. Die Axiome sind symmetrisch in $\oplus$ und $\odot$, sowie 0 und 1.

Beispiel a. (i) $(B, \vee, \wedge, \neg)$ ist Boolsche Algebra mit $0 = f$ und $1 = w$, die sog. *Schaltalgebra*.

(ii) Für jede Menge M ist $(\mathrm{Pot}(M), \cup, \cap, \overline{\ })$, wobei $\overline{X} := M \setminus X$, eine Bool'sche Algebra mit $0 = \emptyset$ und $1 = M$. Die Axiome sind dabei offensichtlich erfüllt.

Satz. *In einer Bool'schen Algebra $(A, \oplus, \odot, \overline{\ })$ gelten für alle $a, b, c \in A$ stets:*

(i) $a \oplus a = a, \quad a \odot a = a$ *(Idempotenz)*

(ii) $a \oplus 1 = 1, \quad a \odot 0 = 0$ *(neutrale Elemente)*

(iii) $a \oplus (a \odot b) = a, \quad a \oplus (\bar{a} \odot b) = a \oplus b$ *(Absorption)*
$a \odot (a \oplus b) = a, \quad a \odot (\bar{a} \oplus b) = a \odot b$

(iv) $(a \oplus b = a \oplus c$ *und* $\bar{a} \oplus b = \bar{a} \oplus c) \Leftrightarrow b = c$ *(Kürzen)*
$(a \odot b = a \odot c$ *und* $\bar{a} \odot b = \bar{a} \odot c) \Leftrightarrow b = c$

(v) $(a \oplus b = 1$ *und* $a \odot b = 0) \Leftrightarrow b = \bar{a}$ *(Eindeutiges Komplement)*

(vi) $\bar{\bar{a}} = a$ *(Involution)*

(vii) $\bar{0} = 1, \quad \bar{1} = 0$ *(Konstante)*

(viii) $\overline{a \oplus b} = \bar{a} \odot \bar{b}, \quad \overline{a \odot b} = \bar{a} \oplus \bar{b}$ *(deMorgan'sche Regeln)*

Beweis. (exemplarisch)

(i) $a \oplus a \overset{B2}{=} (a \oplus a) \odot 1 \overset{B3}{=} (a \oplus a) \odot (a \oplus \overline{a}) \overset{B4}{=} a \oplus (a \odot \overline{a}) \overset{B3}{=} a \oplus 0 \overset{B1}{=} a$.
$a \odot a = a$ geht genauso.

(ii) $a \oplus 1 \overset{B3}{=} a \oplus (a \oplus \overline{a}) \overset{B1}{=} (a \oplus a) \oplus \overline{a} \overset{(i)}{=} a \oplus \overline{a} \overset{B3}{=} 1.$
$a \odot 0 = 0$ geht genauso.

(iii) $a \oplus (a \odot b) \overset{B2}{=} (a \odot 1) \oplus (a \odot b) \overset{B4}{=} a \odot (1 \oplus b) \overset{(ii)}{=} a \odot 1 \overset{B2}{=} a.$
$a \odot (a \oplus b) = a$ geht genauso.

Rest als Übung. □

Beispiel b. Mit Hilfe der Regeln des Satzes können Terme in der Bool'schen Algebra $(B, \vee, \wedge, \overline{\ })$ vereinfacht werden. Wir führen dies am Übertrag des Volladdierers vor und zählen dabei nach jeder Umformung die Anzahl der Verknüpfungen. Die disjunktive Normalform von $c(x, y, z)$ lautet nach Beispiel 2.7.2a:

$$c(x, y, z) = (\neg x \wedge y \wedge z) \vee (x \wedge y \wedge \neg z) \vee (x \wedge \neg y \wedge z) \vee (x \wedge y \wedge z)$$

Also

$$\begin{aligned}
c(x, y, z) &= (\overline{x} \odot y \odot z) \oplus (x \odot \overline{y} \odot z) \oplus (x \odot y \odot \overline{z}) \oplus (x \odot y \odot z) && 11 \\
&\overset{B4}{=} (x \odot ((\overline{y} \odot z) \oplus (y \odot \overline{z}) \oplus (y \odot z))) \oplus (\overline{x} \odot y \odot z) && 9 \\
&\overset{B4}{=} (x \odot ((\overline{y} \odot z) \oplus (y \odot (z \oplus \overline{z})))) \oplus (\overline{x} \odot y \odot z) && 8 \\
&\overset{B3}{=} (x \odot ((\overline{y} \odot z) \oplus (y \odot 1))) \oplus (\overline{x} \odot y \odot z) && 7 \\
&\overset{B2}{=} (x \odot ((\overline{y} \odot z) \oplus y)) \oplus (\overline{x} \odot y \odot z) && 6 \\
&\overset{(iii)}{=} (x \odot (y \oplus z)) \oplus (\overline{x} \odot y \odot z) && 5 \\
&\overset{B4}{=} (x \odot y) \oplus (x \odot z) \oplus (\overline{x} \odot y \odot z) && 6 \\
&\overset{B4}{=} (y \odot (x \oplus (\overline{x} \odot z))) \oplus (x \odot z) && 5 \\
&\overset{(iii)}{=} (y \odot (x \oplus z)) \oplus (x \odot z) && 4 \\
&= (y \wedge (x \vee z)) \vee (x \wedge z).
\end{aligned}$$

Übung. Es sei $(A, \oplus, \odot, \overline{\ })$ eine Bool'sche Algebra. Man zeige, dass durch

$$a \leq b :\Leftrightarrow a \odot b = a$$

eine partielle Ordnung auf A definiert wird.

Diskrete Mathematik

Einleitung

In der Mathematik ist der Begriff „**diskret**“ als gegensätzlich zu „kontinuierlich“ zu verstehen. *Diskret* werden solche Strukturen genannt, die endlich sind oder – falls unendlich – zumindest schrittweise abzählbar; als *kontinuierlich* dagegen solche, die nicht schrittweise abzählbar sind. In diesem Sinne ist z.B. die Zahlenmenge der natürlichen Zahlen $\{1, 2, 3, \ldots\}$ diskret, während die Zahlenmenge der reellen Zahlen (Dezimalbrüche) kontinuierlich ist. Letzteres wird veranschaulicht, indem man sich die reellen Zahlen als eine kontinuierliche Zahlengerade (von $-\infty$ bis $+\infty$ mit 0 in der „Mitte“) vorstellt. Auf dieser reellen Zahlengerade sind dann die natürlichen Zahlen als eine abzählbare Folge von Punkten zu finden.

In dieses Schema passen insbesondere die in der Elektro- bzw. Informationstechnik verwendeten Begriffe „digital“ und „analog“. Ein „digitaler Wert“ ist auf einer diskreten Menge definiert (mit den Elementen 0 und 1, also sogar auf einer endlichen Menge), während ein analoger Wert auf einem Kontinuum (z.B. auf einem bestimmten Abschnitt der reellen Zahlengerade) definiert ist.

Unter den mathematischen Disziplinen beschäftigt sich die **Analysis** mit kontinuierlichen Strukturen (insbesondere mit den reellen Zahlen) und die **Diskrete Mathematik** mit diskreten Strukturen. Die diskrete Mathematik, obwohl in der Form des Studiums der natürlichen Zahlen schon im Altertum präsent, wird aber erst seit dem 20. Jahrhundert als eigenständiges Gebiet betrachtet. So wie eine besondere Motivation für die Entwicklung der Analysis auf Anwendungen in der Physik zurückgeht, gilt das gleiche für die diskrete Mathematik und die Informatik. Offensichtlich sind die in der Informatik beschriebenen und untersuchten Objekte wie Digitalcomputer, Programme (Algorithmen), formale Sprachen, etc. diskreter Natur, während die in der klassichen Physik untersuchten Prozesse kontinuierlicher Natur sind (bzw. sich als kontinuierlich vorgestellt werden).

Wichtige diskrete Strukturen bzw. Objekte, die in dieser Vorlesung behandelt werden, sind endliche Mengen und Summen (Kap. Kombinatorik), endliche Graphen (Kap. Graphentheorie), das Zahlsystem der ganzen Zahlen und Polynome (beides Kap. Algebraische Strukturen).

Kapitel 3

Kombinatorik

3.1 Permutationen und Kombinationen

Es sei A in diesem Abschnitt eine endliche Menge mit $|A| = n$.

3.1.1 Permutationen

Definition a. Es sei $k \in \mathbb{N}, k \leq n$. Eine *$k$-Permutation* aus A ist eine geordnete Auswahl von k verschiedenen Elementen aus A. Eine n-Permutation aus A wird auch kurz *Permutation von A* genannt.

Mit „geordneter Auswahl“ ist gemeint, dass es auf die Reihenfolge der Auswahl ankommt. Mathematisch ist eine k-Permutation aus A ein k-Tupel über A, dessen Einträge paarweise verschieden sind. Dementsprechend werden k-Permutationen in derselben Schreibweise wie Tupel notiert.

Eine Permutation von A kann auch als eine „Anordnung“ von A aufgefasst werden.

Beispiel a.

(i) $(4, 3, 2), (4, 2, 3)$ und $(3, 5, 1)$ sind verschiedene 3-Permutationen aus $\underline{5}$.

(ii) $(1, 2, 1)$ ist keine Permutation.

(iii) $(1, 3, 5, 2, 4)$ und $(5, 4, 3, 2, 1)$ sind Permutationen von $\underline{5}$.

(iv) Die Medaillenverteilung nach einem 100m-Lauf mit 8 Läufern ist eine 3-Permutation aus $\underline{8}$.

(v) Die aktuelle Bundesligatabelle ist eine Permutation von $\underline{18}$.

Definition b. Für $n \in \mathbb{N}$ heißt

$$n! := 1 \cdot 2 \cdot \ldots \cdot n$$

die *Fakultät von n*. Wir setzen $0! := 1$.

Satz. *Es sei $k \in \mathbb{N}, k \leq n$. Die Anzahl der k-Permutationen aus A beträgt $\frac{n!}{(n-k)!}$. Die Anzahl der Permutationen von A beträgt n!.*

Beweis. siehe Vorlesung □

Beispiel b.

(i) Es gibt $\frac{3!}{(3-2)!} = 6$ 2-Permutationen aus $\underline{3}$.

(ii) Es gibt $\frac{8!}{(8-3)!} = 6 \cdot 7 \cdot 8 = 336$ mögliche Medaillenverteilungen (Gold, Silber, Bronze) auf 8 Läufer.

(iii) Es gibt $18! \approx 6,4 \cdot 10^{15}$ mögliche Bundesligatabellen aus 18 Mannschaften.

3.1.2 Kombinationen

Definition. Es sei $k \in \mathbb{N}, k \leq n$. Eine *k-Kombination* aus A ist eine ungeordnete Auswahl von k verschiedenen Elementen aus A.

Mit „ungeordneter Auswahl“ ist gemeint, dass es auf die Reihenfolge der Auswahl nicht ankommt. Mathematisch ist eine k-Kombination aus A eine k-elementige Teilmenge von A. Dementsprechend werden k-Kombinationen in derselben Schreibweise wie Mengen notiert.

Beispiel a.

(i) Es sei $A = \underline{5} = \{1, 2, 3, 4, 5\}$. Dann sind $\{4, 3, 2\} = \{4, 2, 3\}$ und $\{3, 5, 1\}$ verschiedene 3-Kombinationen aus A.

(ii) Ein ausgefüllter Lottoschein ist eine 6-Kombination aus $\underline{49}$.

(iii) Die Bundesliga-Absteiger bilden einen 3-Kombination aus $\underline{18}$.

(iv) Eine Skathand ist eine 10-Kombination aus $\underline{32}$.

Satz. *Es sei $k \in \mathbb{N}$ mit $k \leq n$. Die Anzahl der k-Kombinationen aus A beträgt $\frac{n!}{k!(n-k)!}$.*

Beweis. siehe Vorlesung □

Beispiel b.

(i) Es gibt $\frac{4!}{2!(4-2)!} = 6$ 2-Kombinationen aus $\underline{4}$.

(ii) Es gibt $\frac{49!}{6!43!} = 13983816$ Möglichkeiten, einen Lottoschein auszufüllen.

(iii) Es gibt $\frac{18!}{3!15!} = 816$ Möglichkeiten, drei von 18 Mannschaften absteigen zu lassen.

(iv) Es gibt $\frac{32!}{10!22!} \approx 64512240$ mögliche Skathände.

3.1.3 Tupel

Bemerkung. Es sei $k \in \mathbb{N}$. Ein k-Tupel über A ist eine geordnete Auswahl von k beliebigen (nicht notwendigerweise verschiedenen) Elementen aus A.

Beispiel.

(i) Eine natürliche Zahl mit maximal k Dezimalstellen ist ein k-Tupel über $\{0, 1, \ldots 9\}$.

(ii) Eine Teilmenge von $\underline{n}$ ist ein n-Tupel über $\{0, 1\}$. (Der i-te Eintrag ist genau dann 1, wenn i Element der Teilmenge ist.)

Satz. *Es sei $k \in \mathbb{N}$. Die Anzahl der k-Tupel über A beträgt n^k.*

Beweis. klar. □

Folgerung. $|\operatorname{Pot}(A)| = 2^n$.

Beweis. Der Satz und Beispiel (ii). □

3.1.4 Multimengen

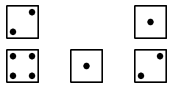
Definition. Es sei $k \in \mathbb{N}$. Eine *k-Multimenge über A* ist eine ungeordnete Auswahl von k beliebigen (nicht notwendigerweise verschiedenen) Elementen aus A.

Schreibweise. Eine Multimenge ist eine „Menge mit Wiederholungen“ und wird mit den modifizierten Mengenklammern $\{^*$ und $^*\}$ notiert.

Bemerkung. Eine k-Multimenge über A kann auch aufgefasst werden als ein n-Tupel über $\mathbb{N}_0$, deren Einträge sich zu k aufsummieren. Dazu nummeriert man die Elemente von A, etwa $A = \{a_1, \ldots, a_n\}$, und gibt im i-ten Eintrag des Tupels an, wie oft a_i in der Multimenge vorkommt. Wir nennen dieses Tupel das *Häufigkeitstupel* der Multimenge A.

Beispiel.

(i) Ein Lostopf ist eine Multimenge, aber in der Regel keine Menge, da gewisse Lose mehrfach vorkommen können (z.B. Nieten).

(ii) Das Resultat eines Kniffel-Wurfs (Wurf mit 5 Würfeln gleichzeitig) ist eine 5-Multimenge über $\underline{6}$. Der Wurf  bedeutet z.B. die Multimenge $\{^*2, 1, 4, 1, 2^*\} = \{^*1, 1, 2, 2, 4^*\}$. Als 6-Tupel über $\mathbb{N}_0$ geschrieben bedeutet dieser Wurf $(2, 2, 0, 1, 0, 0)$.

(iii) Das Resultat einer Klausur mit k Teilnehmern und 11 möglichen Noten (von 1.0 bis 5.0) ist ein k-Tupel über $\underline{11}$. Nummeriert man die Teilnehmer von 1 bis k und ist a_i die Note von Teilnehmer i, dann ist das Resultat das Tupel $(a_1, \ldots, a_k)$.

(iv) Der Notenspiegel einer Klausur mit k Teilnehmern und 11 möglichen Noten (von 1.0 bis 5.0) ist eine k-Multimenge über $\underline{11}$. Nummeriert man die Teilnehmer von 1 bis k und ist a_i die Note von Teilnehmer i, dann ist der Notenspiegel die k-Multimenge $\{^*a_1, \ldots, a_k{}^*\}$. Üblicherweise wird ein Notenspiegel als Tabelle der Häufigkeiten der einzelnen Noten angegeben. Diese Tabelle ist gerade das oben erwähnte Häufigkeitstupel von A, ein 11-Tupel über $\mathbb{N}_0$.

Satz. *Es sei $k \in \mathbb{N}$. Die Anzahl der k-Multimengen über A beträgt $\frac{(n+k-1)!}{k!(n-1)!}$.*

Beweis. Es sei $k \in \mathbb{N}$. Wir zählen die n-Tupel $(l_1, \ldots, l_n)$ über $\mathbb{N}_0$ mit $\sum_{i=1}^{n} l_i = k$. Dazu kodieren wir Tupel dieser Art als Wörter über $\{0, 1\}$, indem wir für jedes l_i genau l_i viele Einsen schreiben und für jedes Komma eine Null. Aus dem Tupel $(2, 2, 0, 1, 0, 0)$ wird z.B. das Wort 1101100100. (Man beachte, dass jeder Null-Eintrag im Tupel zu einem „leeren Teilwort“ wird, gezeichnet durch $\lrcorner$ in der Schreibweise $110110\lrcorner 010\lrcorner 0\lrcorner = 1101100100$.) Offensichtlich gehört zu jedem Tupel, deren Einträge sich zu k aufsummieren, ein Wort mit k Einsen und $n-1$ Nullen. Umgekehrt entsteht jedes Wort mit k Einsen und $n-1$ Nullen aus einem solchen Tupel. Es bleibt also, die Anzahl dieser Wörter zu zählen.

Ein Wort aus k Einsen und $n-1$ Nullen hat die Länge $n+k-1$ und ist eindeutig durch die Positionen der k vielen Einsen gegeben, entspricht also einer k-Kombination aus $\underline{n+k-1}$. Gemäss Satz (3.1.2) lautet die gesuchte Anzahl somit $\frac{(n+k-1)!}{k!(n-1)!}$. □

3.2 Binomialkoeffizienten

Es seien in diesem Abschnitt $n, k \in \mathbb{N}_0$.

3.2.1 Definition und Binomischer Lehrsatz

Definition. Für $k \leq n$ heißt

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}$$

der *Binomialkoeffizient* „n über k“.

Nach Satz (3.1.2) ist $\binom{n}{k}$ gleich der Anzahl der k-Kombinationen aus einer n-elementigen Menge. Insbesondere ist $\binom{n}{k}$ stets eine ganze Zahl. Es gilt $\binom{n}{0} = \binom{n}{n} = 1$ für alle $n \in \mathbb{N}_0$ und $\binom{n}{1} = \binom{n}{n-1} = n$ für alle $n \in \mathbb{N}$.

Satz (Binomischer Lehrsatz)**.** *Für $a, b \in \mathbb{R}$ und $n \in \mathbb{N}_0$ gilt*

$$\begin{aligned}(a+b)^n &= \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \\ &= \binom{n}{0} b^n + \binom{n}{1} a^1 b^{n-1} + \ldots + \binom{n}{n-1} a^{n-1} b^1 + \binom{n}{n} a^n.\end{aligned}$$

Beweis. Siehe Vorlesung. □

Übung. Man zeige mit Hilfe des Binomischen Lehrsatzes die Identität

$$\sum_{k=1}^{n} (-1)^k \binom{n}{k} = -1 \quad (n \geq 1).$$

3.2.2 Das Pascal'sche Dreieck

Satz. *Für alle $n, m \in \mathbb{N}_0$ gelten:*

(i) $\binom{n}{k} = \binom{n}{n-k}$ *für alle* $0 \leq k \leq n$,

(ii) $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ *für alle* $1 \leq k \leq n-1$,

(iii) $\sum_{i=0}^{k} \binom{m}{i}\binom{n}{k-i} = \binom{m+n}{k}$ *für alle* $0 \leq k \leq n, m$. *(Vandermonde-Identität)*

Beweis. Siehe Vorlesung. □

Die Binomialkoeffizienten lassen sich im sog. *Pascal'schen Dreieck* anordnen:

$n=0$:						1						
$n=1$:					1		1					
$n=2$:				1		2		1				
$n=3$:			1		3		3		1			
$n=4$:		1		4		6		4		1		
$n=5$:	1		5		10		10		5		1	
$n=6$:	1	6		15		20		15		6		1

Definiert man $\binom{n}{k} := 0$ für $k < 0$ und $k > n$, also „ausserhalb des Dreiecks“, so gelten die Aussagen des Satzes uneingeschränkt für alle $k, n, m \in \mathbb{N}_0$.

Übung.

(i) Man zeige Teil (ii) des Satzes durch direktes Einsetzen der Definition und Umformung.

(ii) Man zeige mittels vollständiger Induktion, dass $\binom{n}{k}$ eine ganze Zahl ist. *Hinweis:* Verwende den Satz.

(iii) Man zeige den Binomischen Lehrsatz mittels vollständiger Induktion. *Hinweis:* Verwende den Satz.

(iv) Man zeige die Identität

$$\sum_{k=1}^{n} (-1)^k \binom{n}{k} = -1 \quad (n \geq 1).$$

mittels vollständiger Induktion. *Hinweis:* Verwende den Satz.

(v) Man zeige die Vandermonde-Identität mit einem kombinatorischen Beweis. *Hinweis:* Verallgemeinere den Beweis von Teil (ii) des Satzes.

(vi) Man zeige die Vandermonde-Identität mittels vollständiger Induktion.

3.3 Kombinatorische Beweisprinzipien

Wir formulieren nun systematisch einige kombinatorische Beweisprinzipien. Zum Teil wurden diese Prinzipien in den Beweisen der §§1–2 und in den Übungen schon angewendet.

3.3.1 Summenregel

Prinzip. Für disjunkte, endliche Mengen A und B gilt stets

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

Das Prinzip lässt sich sofort auf endlich viele Mengen verallgemeinern: Für paarweise disjunkte, endliche Mengen $A_1, \ldots, A_r$ gilt stets

$$|\bigcup_{i=1}^{r} A_i| = \sum_{i=1}^{r} |A_i|.$$

Beispiel.

(i) Der Beweis von Satz (3.2.2)(ii).

(ii) Ist $A \subseteq M$, so hat die *Komplementärmenge* $M \setminus A$ die Mächtigkeit $|M| - |A|$.

Übung.

(i) Wieviele Teilmengen von $\underline{6}$ gibt es, die höchstens 4 Elemente enthalten?

(ii) Man zeige mit Hilfe der Summenregel, dass $\sum_{i=0}^{n} \binom{n}{i} = 2^n$.

3.3.2 Produktregel

Prinzip. Für zwei beliebige endliche Mengen A und B gilt stets

$$|A \times B| = |A| \cdot |B|.$$

Das Prinzip lässt sich sofort auf endlich viele Mengen verallgemeinern: Für endliche Mengen $A_1, \ldots, A_r$ gilt stets

$$|A_1 \times \ldots \times A_r| = \prod_{i=1}^{r} |A_i|.$$

Insbesondere gilt für jede endliche Menge und jedes $n \in \mathbb{N}$:

$$|A^n| = |A|^n.$$

Beispiel.

(i) Der Beweis von Satz (3.1.2).

(ii) Der Beweis von Satz (3.1.4).

Übung a. Wieviele Tippreihen mit genau 4 Richtigen gibt es für eine feste Lotto-Ziehung?

Satz. *Es sei $\mathcal{A}$ eine Multimenge, die r verschiedene Elemente $a_1, \ldots, a_r$ enthält, wobei a_i mit Häufigkeit k_i auftritt. Sei $k = k_1 + \ldots + k_r$, die „Mächtigkeit" von $\mathcal{A}$. Die Anzahl der Anordnungen von $\mathcal{A}$ beträgt dann:*

$$\frac{k!}{k_1! \cdots k_r!},$$

1. Beweis. Wir betrachten statt $\mathcal{A}$ zunächst die Menge

$$A = \{a_{11}, \ldots, a_{1k_1}, a_{21}, \ldots, a_{2k_2}, \ldots, a_{r1}, \ldots, a_{rk_r}\},$$

in der die a_{ij} als verschieden angenommen werden. Offensichtlich ist $|A| = k$. Nach Satz 3.1.1 gibt es $k!$ verschiedene Anordnungen von A. Jede Anordnung von $\mathcal{A}$ entsteht aus einer Anordnung von A, indem man, für jedes i, alle a_{ij} durch a_i ersetzt. Diese Ersetzung, durchgeführt für ein festes i, macht genau verschiedene $k_i!$ Anordnungen von A gleich. Nach der Produktregel macht diese Ersetzung, durchgeführt für alle i, also genau $k_1! \cdots k_r!$ verschiedene Anordnungen von A gleich. Daraus ergibt sich die Formel $\frac{k!}{k_1! \cdots k_r!}$ für die Zahl der Anordnungen von $\mathcal{A}$. □

2. Beweis. Jede Anordnung von $\mathcal{A}$ entsteht auf eindeutige Weise aus folgendem Prozess: Wir wählen eine k_1-Kombination von $\underline{k}$; diese gibt die Positionen in der Anordnung an, an denen wir a_1 eintragen. (Es muss genau k_1 Positionen in der Anordnung geben, an denen a_1 steht.) Wir wählen dann eine k_2-Kombination aus den verbleibenden $k - k_1$ Positionen, um dort a_2 einzutragen, usw. Nach Produktregel gibt es für diesen Prozess genau $\binom{k}{k_1}\binom{k-k_1}{k_2}\binom{k-k_1-k_2}{k_3} \ldots \binom{k-k_1-\ldots-k_{r-1}}{k_r}$. (Der letzte Faktor ist identisch $\binom{k_r}{k_r} = 1$.) Durch Einsetzen und Kürzen ergibt sich die Formel. □

Übung b. (i) Wieviele verschiedene Wörter kann man durch Anordnung der Buchstaben `P,I,Z,Z,A` gewinnen?

(ii) Wieviele Möglichkeiten gibt es, aus 25 Fussballspielern zwei Mannschaftsaufstellungen (erste und zweite Mannschaft) mit je 11 Spielern zu machen?

(iii) Auf einem Kongress gibt es einen Hauptredner, der 3x vortragen soll, und drei Nebenredner, die je 2x vortragen sollen. Wieviele Vortragsprogramme sind möglich?

3.3.3 Inklusions-Exklusions-Prinzip

Prinzip. Für zwei beliebige endliche Mengen A und B gilt stets

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Das Prinzip lässt sich auf endlich viele Mengen verallgemeinern:

Satz. *Für endliche Mengen $A_1, \ldots, A_r$ gilt die Formel*

$$\begin{aligned} |\bigcup_{i=1}^{r} A_i| &= \sum_{k=1}^{r} (-1)^{k-1} \sum_{1 \le i_1 < \ldots < i_k \le r} |A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}| \\ &= \sum_{k=1}^{r} (-1)^{k-1} \sum_{I \subseteq \underline{r}, |I|=k} |\bigcap_{i \in I} A_i|. \end{aligned}$$

Beweis. Setze $A := \bigcup_{i=1}^{r} A_i$. Wir rechnen nach, dass jedes Element $a \in A$ auf der rechten Seite der Formel tatsächlich genau einmal gezählt wird. Sei also a ein beliebiges fest gewähltes Element aus A. Definiere I_a als die Menge der Indizes i aller Mengen A_i, die a enthalten, d.h.

$$I_a := \{i \in \underline{r} \,|\, a \in A_i\}.$$

In der Formel werden Ausdrücke der Form $|\bigcap_{i \in I} A_i|$ für bestimmte Indexmengen $I \subseteq \underline{r}$ aufsummiert. Sei $I \subseteq \underline{r}$ eine beliebige solche Indexmenge. Dann wird das Element a in $|\bigcap_{i \in I} A_i|$ genau 1-mal gezählt, wenn $a \in \bigcap_{i \in I} A_i$, sonst 0-mal. Weiter gilt $a \in \bigcap_{i \in I} A_i$ genau dann wenn $i \in I_a$ für alle $i \in I$, also genau dann wenn $I \subseteq I_a$. Der Anteil von a an dem Ausdruck

$$\sum_{I \subseteq \underline{r}, |I|=k} |\bigcap_{i \in I} A_i|$$

für festes k beträgt somit

$$\sum_{I \subseteq I_a, |I|=k} 1 + \sum_{I \not\subseteq I_a, |I|=k} 0 = \sum_{I \subseteq I_a, |I|=k} 1,$$

also genau die Anzahl der k-elementigen Teilmengen von I_a. Diese Zahl hängt nur von $|I_a|$ ab, und beträgt $\binom{|I_a|}{k}$ falls $k \le |I_a|$ und 0 falls $k > |I_a|$. Der Anteil von a an der gesamten rechten Seite beträgt somit

$$\sum_{k=1}^{|I_a|} (-1)^{k-1} \binom{|I_a|}{k}.$$

Nach Übung 3.2 gilt für alle $m \in \mathbb{N}$:

$$\sum_{k=1}^{m} (-1)^{k-1} \binom{m}{k} = -1.$$

Damit ist gezeigt, dass a auf der gesamten rechten Seite genau einmal gezählt wurde. □

Für $r = 3$ ergibt sich z.B.

$$\begin{aligned} |A_1 \cup A_2 \cup A_3| = &+ |A_1| + |A_2| + |A_3| \\ &- |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| \\ &+ |A_1 \cap A_2 \cap A_3|. \end{aligned}$$

Beispiel. Wieviele Zahlen zwischen 1 und 100 sind durch 2, 3 oder 5 teilbar? Wir haben die Menge

$$A = \{i \in \mathbb{N} \mid i \leq 100, 2|i \vee 3|i \vee 5|i\}$$

zu zählen. Leicht zählbar sind die Mengen

$$A_n := \{i \in \mathbb{N} \mid i \leq 100, n|i\},$$

für alle $n \in \mathbb{N}$ ist nämlich $|A_n| = \lfloor \frac{100}{n} \rfloor$. Da A die Vereinigung $A = A_2 \cup A_3 \cup A_5$ ist, ergibt sich nach dem Inklusions-Exklusions-Prinzip

$$|A| = |A_2| + |A_3| + |A_5| - |A_2 \cap A_3| - |A_2 \cap A_5| - |A_3 \cap A_5| + |A_2 \cap A_3 \cap A_5|.$$

Es bleibt, die verschiedenen Durchschnitte zu zählen. Nun ist jede natürliche Zahl i genau dann durch 2 und 3 teilbar, wenn sie durch 6 teilbar ist. D.h. $A_2 \cap A_3 = A_6$. Analog ergibt sich $A_2 \cap A_5 = A_{10}, A_3 \cap A_5 = A_{15}, A_2 \cap A_3 \cap A_5 = A_{30}$. (Man beachte, dass 2, 3 und 5 Primzahlen sind; allgemein gilt $A_n \cap A_m = A_{\mathrm{kgV}(n,m)}$ für beliebige $n, m \in \mathbb{N}$.) Also

$$\begin{aligned} |A| &= |A_2| + |A_3| + |A_5| - |A_6| - |A_{10}| - |A_{15}| + |A_{30}| \\ &= 50 + 33 + 20 - 16 - 10 - 6 + 3 = 74. \end{aligned}$$

Übung. Die Bevölkerung von Aachen, die arbeitet oder studiert, betrage 150000. Wenn davon 20% studieren und 90% arbeiten, wieviele Aachener Studenten arbeiten dann neben ihrem Studium?

Übung. Man zeige für alle $a, b \in \mathbb{N}$: $\binom{a}{2} + \binom{b}{2} \leq \binom{a+b-1}{2}$.
Hinweis für einen kombinatorischen Beweis: Seien A, B zwei Mengen mit $|A| = a, |B| = b$ und $|A \cap B| = 1$.

3.3.4 Schubfachprinzip

Prinzip. Verteilt man n Elemente auf m Schubladen und ist $n > m$, so enthält eine Schublade mindestens zwei Elemente.

Beispiel.

(i) In jeder Menge von 13 Personen gibt es zwei, die im gleichen Monat Geburtstag haben.

(ii) Auf jeder Party gibt es zwei Personen, die die gleiche Anzahl von Leuten kennen.

3.4 Stirling'sche Zahlen

Die Binomialkoeffizienten wurden eingeführt, da sie beim Zählen von Teilmengen bzw. Multimengen fester Mächtigkeit auftreten. Die Stirling'schen Zahlen stellen zwei weitere Arten von Zählkoeffizienten dar. Sie treten auf beim Zählen von Partitionen mit fester Anzahl von Teilen bzw. beim Zählen von Permutationen mit fester Zykelzahl.

3.4.1 Stirling-Zahlen zweiter Art

Definition. Es seien $n, k \in \mathbb{N}_0$. Wir definieren

$$S_{n,k} := \text{Anzahl der Partitionen von } \underline{n} \text{ mit genau } k \text{ Teilen.}$$

Die Zahlen $S_{n,k}$ heißen *Stirling-Zahlen zweiter Art*. Partitionen mit k Teilen nennen wir auch kurz *k-Partitionen*.

Beispiel. Wieviele Möglichkeiten gibt es, n Studenten auf k Tutoriengruppen aufzuteilen, wobei keine Gruppe leer bleiben soll? Eine solche Aufteilung ist eine k-Partition von $\underline{n}$, somit gibt es $S_{n,k}$ Möglichkeiten.

Bemerkung. Für alle $n, k \in \mathbb{N}_0$ gelten:

(i) $S_{n,n} = 1$,

(ii) $S_{n,0} = 0$ falls $n > 0$,

(iii) $S_{n,k} = 0$ falls $k > n$.

Beweis.

(i) Es gibt genau eine n-Partition von $\underline{n}$. Das gilt auch für $n = 0$, da es genau eine Partition der leeren Menge gibt, und die hat 0 Teile.

(ii) Eine Partition einer nicht-leeren Menge muss mindestens 1 Teil haben.

(iii) Eine Partition von $\underline{n}$ kann höchstens n Teile haben.

□

Satz. *Für alle $n \in \mathbb{N}_0$ und $k \in \mathbb{N}$ gilt $S_{n+1,k} = S_{n,k-1} + kS_{n,k}$.*

Beweis. Es sei $T_1 \cup \ldots \cup T_k = \underline{n+1}$ eine k-Partition von $\underline{n+1}$. Wir nehmen o.B.d.A. an, dass $n+1$ in T_k liegt (die Nummerierung der Teile spielt keine Rolle). Entfernt man $n+1$ aus allen Mengen, so bekommt man $T_1 \cup \ldots \cup T_{k-1} \cup T_k \setminus \{n+1\} = \underline{n}$. Je nachdem, ob $T_k \setminus \{n+1\}$ leer ist oder nicht, ist dies eine $(k-1)$-Partition oder eine k-Partition von $\underline{n}$. Umgekehrt entsteht also jede k-Partition von $\underline{n+1}$ auf eine der folgenden Arten:

a) Hinzufügen des Teiles $\{n+1\}$ zu einer $(k-1)$-Partition von $\underline{n}$,

b) Hinzufügen des Elementes $n+1$ zu einem der Teile einer k-Partition von $\underline{n}$.

Keine Partition kann auf beide Arten entstehen, denn bei a) liegt $n+1$ stets in einem Teil der Mächtigkeit 1, bei b) stets in einem Teil der Mächtigkeit > 1. Folglich ist die Anwendung der Summenregel erlaubt. Bei a) gibt es $S_{n,k-1}$ viele $(k-1)$-Partitionen von $\underline{n}$, die jeweils auf eindeutige Art um den Teil $\{n+1\}$ ergänzt werden. Bei b) gibt es $S_{n,k}$ viele k-Partitionen von $\underline{n}$, bei denen auf jeweils k verschiedene Arten das Element $n+1$ zu einem der Teile hinzugefügt wird. Nach Produktregel entstehen durch b) also $kS_{n,k}$ viele k-Partitionen von $\underline{n+1}$. Mit der Summenregel ergibt sich schliesslich die Formel $S_{n+1,k} = S_{n,k-1} + kS_{n,k}$. □

Die Zahlen $S_{n,k}$ lassen sich im sog. *Stirling-Dreieck zweiter Art* anordnen:

$n = 0$:						1						
$n = 1$:					0		1					
$n = 2$:				0		1		1				
$n = 3$:			0		1		3		1			
$n = 4$:		0		1		7		6		1		
$n = 5$:	0		1		15		25		10		1	
$n = 6$:	0	1		31		90		65		15		1

Übung. Man zeige:

(i) Die Anzahl der surjektiven Abbildungen $\underline{n} \to \underline{k}$ beträgt $k! \cdot S_{n,k}$.

(ii) Es gilt $\sum_{k=0}^{m} S_{n,k} \cdot \frac{m!}{(m-k)!} = n^m$.
Tip: n^m ist die Anzahl aller Abbildungen $\underline{m} \to \underline{n}$.

3.4.2 Stirling-Zahlen erster Art

Definition. Es seien $n, k \in \mathbb{N}_0$. Wir definieren

$$s_{n,k} := \text{Anzahl der Permutationen aus } S_n \text{ mit Zykelzahl } k.$$

Die Zahlen $s_{n,k}$ heißen *Stirling-Zahlen erster Art.*

Beispiel. Bei einem Treffen von n Philosophen teilen sich diese in k Diskussionsgruppen auf (Gruppen mit nur einer Person sind erlaubt). Die Teilnehmer jeder Gruppe setzen sich im Kreis hin und philosophieren über ein Thema. Wieviele mögliche Sitzordnungen gibt es? Antwort: $s_{n,k}$.

Bemerkung. Für alle $n, k \in \mathbb{N}_0$ gelten:

(i) $s_{n,n} = 1$,

(ii) $s_{n,0} = 0$ falls $n > 0$,

(iii) $s_{n,k} = 0$ falls $k > n$.

Beweis.

(i) Hat $\pi \in S_n$ die Zykelzahl n, so müssen alle Zykeln die Länge 1 haben, also ist $\pi = \mathrm{id}$. Das gilt auch für $n = 0$, denn das einzige Element aus S_0 hat Zykelzahl 0 (vgl. Bemerkung 1(1.5.4)).

(ii) Die Zykelzahl eines Elementes von S_n mit $n > 0$ ist stets > 0.

(iii) Die Zykelzahl eines Elementes von S_n kann höchstens n betragen.

□

Satz. *Für alle $n \in \mathbb{N}_0$ und $k \in \mathbb{N}$ gilt $s_{n+1,k} = s_{n,k-1} + n s_{n,k}$.*

Beweis. Eine Modifikation des Beweises von Satz (3.4.1) (Übung). □

Die Zahlen $s_{n,k}$ lassen sich im sog. *Stirling-Dreieck erster Art* anordnen:

$n=0$:						1						
$n=1$:					0		1					
$n=2$:				0		1		1				
$n=3$:			0		2		3		1			
$n=4$:		0		6		11		6		1		
$n=5$:	0		24		50		35		10		1	
$n=6$:	0	120		274		225		85		15		1

Übung.

(i) Man führe den Beweis des Satzes aus.

(ii) Man zeige $\sum_{k=0}^{n} s_{n,k} = n!$.

Kapitel 4

Graphentheorie

4.1 Grundbegriffe

4.1.1 Ungerichtete Graphen

Definition a. Ein (ungerichteter) *Graph* ist ein Paar $G = (V, E)$, bestehend aus einer endlichen Menge V und einer Menge E von zweielementigen Teilmengen von V. Die Elemente von V werden *Knoten* (engl. *vertex*) genannt, die Elemente von E *Kanten* (engl. *edge*). Es heißt $n_G := |V|$ die *Knotenzahl* und $m_G := |E|$ die Kantenzahl von G.

Bemerkung a. Das mathematische Modell für eine Kante zwischen den Knoten $u, v \in V$ ist hier die *zwei*elementige Teilmenge $\{u, v\} = \{v, u\} \subseteq V$. Das bedeutet, dass unsere Definition keine sog. „Schlingen“ zulässt, d.h. Kanten von einem Knoten zu sich selbst. Für die Kante $\{u, v\}$ verwenden wir alternativ auch die Schreibweise uv bzw. vu.

Ein weiteres mögliches mathematisches Modell für die Kanten ist, die Kantenmenge als eine symmetrische, antireflexive Relation auf der Knotenmenge aufzufassen.

Erlaubt ist der Graph $G = (\emptyset, \emptyset)$.

Bemerkung b. Andere verbreitete Definitionen von Graphen erlauben gerichtete Kanten, Schlingen, Mehrfachkanten, gewichtete Kanten, gefärbte Kanten, usw. Entsprechend muss das mathematische Modell für die Kantenmenge variiert werden.

Übung a. Jede Relation auf einer Menge V kann als ein gerichteter Graph (mit erlaubten Schlingen) veranschaulicht werden. Man mache sich klar, was jede einzelne der folgenden Eigenschaften der Relation für das Aussehen dieses Graphen bedeuten: symmetrisch, antisymmetrisch, reflexiv, antireflexiv, transitiv, Äquivalenzrelation, Totalordnung.

Übung b. Was wäre ein mathematisches Modell für einen ungerichteten Graphen mit Mehrfachkanten bzw. mit gewichteten Kanten?

In diesem und den folgenden Abschnitten sei $G = (V, E)$ stets ein Graph.

Definition b.

(i) Ist $uv \in E$ eine Kante, so werden u und v die *Endknoten* von uv genannt. In diesem Fall heißen u und v *adjazent* oder *benachbart*, sowie u *Nachbar* von v und umgekehrt.

(ii) Die Menge aller Nachbarn von $v \in V$ wird mit $\Gamma(v) := \Gamma_G(v)$ bezeichnet.

(iii) G heißt *vollständig*er Graph, wenn je zwei beliebige Knoten adjazent sind, also genau dann, wenn $m_G = \binom{n_G}{2}$.

(iv) Eine Kante $e \in E$ heißt *inzident* zu einem Knoten $v \in V$, wenn v ein Endknoten von e ist.

(v) Zwei verschiedene Kanten heißen *inzident*, wenn sie einen gemeinsamen Endknoten haben.

Übung c. In jedem Graph G gilt $m_G \leq \binom{n_G}{2}$.

4.1.2 Datenstruktur für Graphen

Wir nehmen $V = \{1, \ldots, n\}$ an.

Definition. Die *Adjazenzmatrix* von G ist das quadratische Schema

$$A := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \text{ mit } a_{ij} := \begin{cases} 1 & \text{falls } ij \in E, \\ 0 & \text{falls } ij \notin E. \end{cases}$$

Die *Adjazenzliste* von G ist die Liste $\Gamma := (\Gamma(1), \Gamma(2), \ldots, \Gamma(n))$.

Bemerkung. Die Adjazenzmatrix enthält 0 entlang der Diagonalen von a_{11} bis a_{nn} und ist spiegelsymmetrisch zu dieser Diagonalen.

Beispiel. Siehe Vorlesung.

4.1.3 Teilgraphen

Definition. Ein Graph $G' = (V', E')$ wird *Teilgraph* von G genannt, geschr. $G' \leq G$, wenn $V' \subseteq V$ und $E' \subseteq E$.

Beispiel. Ist $V' \subseteq V$, so wird durch $E' := E \cap \{uv \mid u, v \in V'\}$ ein Teilgraph (V', E') von G definiert. Dieser wird der *auf V' induzierte Teilgraph von G* genannt, geschr. $G|_{V'}$.

4.1.4 Der Grad

Definition. Wir definieren den *Grad* von $v \in V$ als $\deg(v) := |\Gamma(v)|$, also die Anzahl der Nachbarn von v bzw. die Anzahl der zu v inzidenten Kanten. Knoten mit Grad 0 heißen *isoliert*.

Bemerkung. $\sum_{v \in V} \deg(v) = 2m_G$.

Korollar. *In jedem Graphen ist die Anzahl der Knoten mit ungeradem Grad gerade.*

Beispiel. Die Anzahl der Personen auf einer Party, die einer ungeraden Zahl von Gästen die Hand geben, ist gerade. (Aufgrund dieses Beispiel wird das Korollar auch „Handschlagslemma“ genannt.)

4.1.5 Kantenzüge, Pfade, Kreise, Touren

Definition. Es sei $l \in \mathbb{N}_0$.

(i) Ein *Kantenzug der Länge l in G* ist ein Tupel $(v_0, v_1, \ldots, v_l)$ von Knoten mit $v_i v_{i+1} \in E$ für alle $i = 0, \ldots, l-1$. Zu einem Kantenzug $(v_0, \ldots, v_l)$ sagen wir auch genauer *Kantenzug von v_0 nach v_l* oder *v_0-v_l-Kantenzug*, und die Knoten v_0, v_l werden sein *Anfangs-* bzw. *Endknoten* genannt. Der Kantenzug heißt *geschlossen* falls $v_0 = v_l$.

(ii) Ein Kantenzug $(v_0, \ldots, v_l)$ heißt *Pfad der Länge l in G*, falls die Knoten $v_0, \ldots, v_l$ paarweise verschieden sind. Zu einem Pfad $(v_0, \ldots, v_l)$ sagen wir auch genauer *Pfad von v_0 nach v_l* oder *v_0-v_l-Pfad*, und die Knoten v_0, v_l werden sein *Anfangs-* bzw. *Endknoten* genannt,

(iii) Ein *Kreis der Länge l in G* ist ein geschlossener Kantenzug $(v_0, \ldots, v_l)$, für den $l \geq 3$ und $(v_0, \ldots, v_{l-1})$ ein Pfad ist.

(iv) Eine *Tour der Länge l in G* ist ein geschlossener Kantenzug $(v_0, \ldots, v_l)$, für den die Kanten $v_0v_1, v_1v_2, \ldots, v_{l-1}v_l, v_lv_0$ paarweise verschieden sind.

Bemerkung.

(i) Für jeden Knoten $v \in V$ ist (v) ein v-v-Pfad der Länge 0.

(ii) Jeder Kreis ist eine Tour, aber nicht umgekehrt.

(iii) Ist $(v_0, \ldots, v_l)$ ein Kreis, so ist auch $(v_1, \ldots, v_{l-1}, v_0, v_1)$ ein Kreis. Diese beiden Kreise sind formal verschieden! Liest man das Tupel $(v_0, \ldots, v_l)$ aber als Zykel $(v_0\ v_1\ \ldots\ v_{l-1})$, also als eine Permutation von V, so liefern beide Kreise denselben Zykel.

(iv) Ist $(v_0, \ldots, v_l)$ ein Kreis, so ist auch $(v_l, \ldots, v_0)$ ein Kreis. Diese beiden Kreise sind formal verschieden!

Beispiel. Siehe Vorlesung.

Übung. Ist eine Kante $e \in E$ *Teil* von zwei *verschiedenen* Kreisen von $G = (V, E)$, so besitzt auch $(V, E \setminus \{e\})$ einen Kreis. Hier ist zunächst geeignet zu definieren, was es heißt, dass e Teil eines Kreises ist, und wann zwei Kreise als gleich anzusehen sind.

4.1.6 Zusammenhang

Definition. Die *Zusammenhangsrelation* $\sim$ auf V wird definiert durch

$$u \sim v :\Leftrightarrow \text{ es gibt einen } u\text{-}v\text{-Kantenzug in } G.$$

G heißt *zusammenhängend*, falls $u \sim v$ für alle $u, v \in V$, anderenfalls *unzusammenhängend*.

Bemerkung. Offensichtlich ist $\sim$ eine Äquivalenzrelation (Übung). Wir lesen $u \sim v$ auch als „u ist verbunden mit v“ oder „u und v hängen zusammen“. Für alle $u, v \in V$ gilt:

$$u \sim v \Leftrightarrow \text{ es gibt einen } u\text{-}v\text{-Pfad in } G.$$

Beweis. $\Rightarrow$: Sei $u \sim v$ und sei $(v_0, v_1, \ldots, v_l)$ mit $v_0 = u$ und $v_l = v$ ein u-v-Kantenzug in G von minimaler Länge l. Angenommen $(v_0, v_1, \ldots, v_l)$ ist kein Pfad, d.h. $v_i = v_j$ für geeignete $1 \leq i < j \leq l$. Dann ist $(v_0, \ldots, v_i, v_{j+1}, \ldots, v_l)$ ein u-v-Kantenzug der Länge $l - (j - i) < l$ im Widerspruch zur Minimaliät von l. Also ist die Annahme falsch und $(v_0, v_1, \ldots, v_l)$ ein Pfad.
$\Leftarrow$: trivial. □

Lemma. *Besitzt G einen Knoten vom Grad $n_G - 1$, so ist G zusammenhängend.*

Satz. *Ist G unzusammenhängend, so gilt $m_G \leq \binom{n_G-1}{2}$.*

Beweis. Sei G unzusammenhängend. Dann existieren $u, v \in V$ mit $u \not\sim v$. Für jedes $w \in \Gamma(v)$ ist $w \notin \Gamma(u)$ (sont wäre $u \sim v$). Aus diesem Grund ist die Menge

$$E' := E \cup \{uw \mid w \in \Gamma(v)\} \setminus \{vw \mid w \in \Gamma(v)\}$$

gleichmächtig mit E. Ausserdem ist v in (V, E') isoliert. Für den Graph $G' = (V \setminus v, E')$ gilt also $m_G = m_{G'} \leq \binom{n_{G'}}{2} = \binom{n_G-1}{2}$. □

Beweis durch Induktion. Sei $n = n_G$. Für $n = 1$ ist die Aussage trivial, für $n = 2$ lautet sie $0 \leq 0$ (Induktionsanfang). Sei nun $n \geq 3$. Sei oBdA $V = \underline{n}$. Falls n isoliert ist, so folgt $m_G \leq \binom{n-1}{2}$ aus der Betrachtung von $G|_{\underline{n-1}}$. Sei also n nicht isoliert und G unzusammenhängend. Dann ist a) $\deg n \leq n-2$ (Lemma), und b) $G|_{\underline{n-1}}$ unzusammenhängend (sonst G zusammenhängend). Mit Induktionsvoraussetzung, angewendet auf $G' := G|_{\underline{n-1}}$, folgt $m_G = m_{G'} \leq \binom{n-2}{2} + (n-2) = \frac{(n-2)(n-3)}{2} + (n-2) = \binom{n-1}{2}$. □

4.1.7 Zusammenhangskomponenten

Definition. Die *Zusammenhangskomponenten* oder kurz *Komponenten* von G sind die induzierten Teilgraphen $G|_U$, wobei U die Äquivalenzklassen von V bzgl. $\sim$ durchläuft. Die Anzahl der Äquivalenzklassen von $\sim$ bezeichnen wir als *Komponentenzahl* r_G von G. Es heißt $G|_{[v]_\sim}$ die *Zusammenhangskomponente* von $v \in V$. Komponenten, die aus einem einzelnen Knoten bestehen, nennen wir *trivial*.

Beispiel. Siehe Vorlesung.

Bemerkung. G ist genau dann zusammenhängend, wenn $r_G \leq 1$. Eine Komponente ist genau dann trivial, wenn sie keine Kanten enthält. Ein Knoten ist genau dann isoliert, wenn seine Zusammenhangskomponente trivial ist.

Lemma. *Für alle $u, v \in V$ gilt:*

(i) $r_{(V,E)} - 1 \leq r_{(V,E\cup\{uv\})} \leq r_{(V,E)}$.

(ii) $r_{(V,E\setminus uv)} - 1 \leq r_{(V,E)} \leq r_{(V,E\setminus uv)}$.

Beweis. a) Die neue Kante uv kann höchstens zwei Komponenten verbinden. b) folgt aus a). □

Satz. *Für jeden Graph G gilt $n_G \leq m_G + r_G$.*

Beweis. als Übung (Induktion nach m_G). □

Korollar. *In jedem zusammenhängenden Graphen gilt $n_G \leq m_G + 1$.*

4.1.8 Brücken

Bemerkung a. Es seien $e = uv \in E, G' = (V, E \setminus e)$. Folgende Aussagen äquivalent:

(i) $u \not\sim v$ in G',

(ii) $r_{G'} > r_G$.

Definition. Eine Kante $e = uv \in E$ heißt *Brücke* von G, wenn die Bedingungen aus Bemerkung a erfüllt sind, sonst *Nicht-Brücke* von G.

Beispiel. Ist $\deg u = 1$, so ist die einzige zu u inzidente Kante eine Brücke. Weitere Beispiele inkl. Bilder siehe Vorlesung.

Bemerkung b. Es seien $e = uv \in E, G' = (V, E \setminus e)$. Folgende Aussagen äquivalent:

(i) e ist keine Brücke von G,

(ii) $u \sim v$ in $(V, E \setminus \{e\})$,

(iii) $r_{G'} = r_G$,

(iv) es gibt einen u-v-Kantenzug in G, der nicht über e führt,

(v) es gibt einen u-v-Pfad in G, der nicht über e führt,

(vi) e ist Teil eines Kreises in G.

Beweis. Die Äquivalenz (iv)⇔(v) benutzt Bemerkung (4.1.6). Der Rest ist trivial. □

Satz. *Ist $u \in V$ zu l Brücken inzident ist ($l \in \mathbb{N}$), so besitzt G mindestens l von u verschiedene Knoten von ungeradem Grad.*

Folgerung. *Haben in einem Graphen alle Knoten geraden Grad, so besitzt er keine Brücken.*

Beweis des Satzes. Seien $e_1, \dots e_l \in E$ zu u inzidente Brücken in G, $e_i = uv_i$. Setze $G' = (V, E \setminus \{e_1, \dots, e_r\})$. Behauptung: jede der Komponenten G'_{v_i} enthält einen Knoten von ungeradem Grad in G. In der Tat, falls $\deg_G v_i$ gerade ist, so ist $\deg_{G'}(v_i)$ ungerade. Nach dem Handschlagslemma, angewendet auf G'_{v_i}, enthält dann G'_{v_i} einen weiteren Knoten $v'_i \neq v_i$ mit $\deg_{G'}$ ungerade. Wegen $v'_i \neq v_i$ ist aber $\deg_G v'_i = \deg_{G'} v'_i$, also ungerade. □

4.2 Distanz und gewichtete Graphen

4.2.1 Distanz

Es sei $G = (V, E)$ ein Graph.

Definition. Für alle $v, w \in V$ mit $v \sim w$ definieren wir die *Distanz* zwischen v und w als

$$d(v, w) := \min\{l \in \mathbb{N}_0 \,|\, \text{in } G \text{ existiert ein } v\text{-}w\text{-Pfad der Länge } l\} \in \mathbb{N}_0.$$

Für alle $v, w \in V$ mit $v \not\sim w$ wird $d(v, w) := \infty$ gesetzt.

Bemerkung. Für alle $v, w \in V$ gelten:

(i) $d(v, w) = 0 \Leftrightarrow v = w$,

(ii) $d(v, w) < \infty \Leftrightarrow v \sim w$.

G ist genau dann zusammenhängend, wenn $d(v, w) < \infty$ für alle $v, w \in V$.

4.2.2 Breitensuche

Die *Breitensuche* ist ein Algorithmus, der, beginnend bei einer Wurzel $w \in V$, alle Knoten der Zusammenhangskomponente von w mit aufsteigender Distanz durchläuft. Er eignet sich also zur Berechnung der Zusammenhangskomponenten von G, insbesondere zur Bestimmung der Brücken und zur Prüfung des Graphen auf Zusammenhang. Ausserdem können mit der Breitensuche die Distanzen $d(v, w)$ sowie kürzeste Pfade von v nach w für jeden Knoten v bestimmt werden. Die kürzesten Pfade von jedem v zu w können dadurch angegeben werden, dass man jedem $v \in V$ einen *Vorgänger* mit kleinerer Distanz zu w zuordnet. Aus den Vorgängern erhalt man dann umgekehrt einen kürzesten Kantenzug von v nach w, indem man, ausgehend von v, sukzessive zum jeweiligen Vorgänger übergeht.

Algorithmus. *Es sei G ein Graph mit Knotenmenge $V = \{1, \ldots, n\}$, gegeben als Adjazenzliste Γ, und es sei $w \in V$. Die in Abbildung 4.2.2 dargestellte Prozedur* Breitensuche *berechnet zu jedem $v \in V$ die Distanz $d(v) := d(v, w)$ sowie einen Vorgänger $p(v)$ in einem kürzesten v-w-Pfad.*

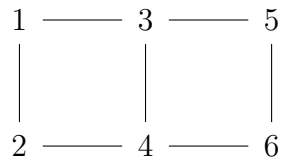
Die verwendete Datenstruktur `queue` ist eine Warteschlange im „First-in-first-out“-Modus. Der Aufruf Insert(Q, x) hängt das Element x am Ende der Warteschlange ein, der Aufruf Extract(Q) entnimmt das Element, das am Anfang der Warteschlange steht.

```
BREITENSUCHE(Γ, w)
 1  initialisiere array d[1, ..., n] mit allen Einträgen gleich ∞
 2  initialisiere array p[1, ..., n] mit allen Einträgen gleich NIL
 3  initialisiere leere queue Q (FIFO)
 4  d[w] ← 0
 5  INSERT(Q, w)
 6  while Q ist nicht leer
 7  do v ← EXTRACT(Q)
 8      for u ∈ Γ[v]
 9      do if d[u] = ∞
10            then INSERT(Q, u)
11                 d[u] ← d[v] + 1
12                 p[u] ← v
13  return d, p
```

Abbildung 4.1: Prozedur Breitensuche

Bemerkung a. Da der Verlauf der Breitensuche davon abhängt, in welcher Reihenfolge Knoten in die Warteschlange aufgenommen werden, spielt die Anordnung der Adjazenzlisten eine Rolle, die bestimmt in welcher Reihenfolge die Knoten in der `for`-Schleife bearbeitet werden. An folgendem Beispiel wird deutlich, wie die Anordnung der Adjazenzlisten den Verlauf und das Ergebnis für p, nicht aber das Ergebnis für d beeinflusst.

Beispiel. Betrachte folgenden Graph mit $V = \underline{6}$ und Wurzel $w = 1$:



Die erste Tabelle zeigt den Ablauf der Breitensuche, wenn die Adjazenzlisten mit aufsteigender Nummerierung angeordnet sind. Jede Zeile entspricht dabei einem Durchlauf der `while`-Schleife und gibt folgendes an: die Zustände der Datenstrukturen d, p, Q zu Beginn der `while`-Schleife, das von EXTRACT gelieferte v, dessen Adjazenzliste $\Gamma(v)$, und die Teilliste der $u \in \Gamma(v)$ mit

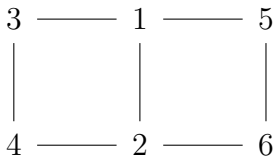
$d[u] = \infty$.

d	p	Q	v	$\Gamma(v)$	$d[u] = \infty$
$(0, \infty, \infty, \infty, \infty, \infty)$	$(-, -, -, -, -, -)$	(1)	1	$(2, 3)$	$(2, 3)$
$(0, 1, 1, \infty, \infty, \infty)$	$(-, 1, 1, -, -, -)$	$(2, 3)$	2	$(1, 4)$	(4)
$(0, 1, 1, 2, \infty, \infty)$	$(-, 1, 1, 2, -, -)$	$(3, 4)$	3	$(1, 4, 5)$	(5)
$(0, 1, 1, 2, 2, \infty)$	$(-, 1, 1, 2, 3, -)$	$(4, 5)$	4	$(2, 3, 6)$	(6)
$(0, 1, 1, 2, 2, 3)$	$(-, 1, 1, 2, 3, 4)$	$(5, 6)$	5	$(3, 6)$	$()$
$(0, 1, 1, 2, 2, 3)$	$(-, 1, 1, 2, 3, 4)$	(6)	6	$(4, 5)$	$()$
$(0, 1, 1, 2, 2, 3)$	$(-, 1, 1, 2, 3, 4)$	$()$			

Die nächste Tabelle zeigt den Ablauf, wenn die Adjazenzlisten mit absteigender Nummerierung angeordnet sind.

d	p	Q	v	$\Gamma(v)$	$d[u] = \infty$
$(0, \infty, \infty, \infty, \infty, \infty)$	$(-, -, -, -, -, -)$	(1)	1	$(3, 2)$	$(3, 2)$
$(0, 1, 1, \infty, \infty, \infty)$	$(-, 1, 1, -, -, -)$	$(3, 2)$	3	$(5, 4, 1)$	$(5, 4)$
$(0, 1, 1, 2, 2, \infty)$	$(-, 1, 1, 3, 3, -)$	$(2, 5, 4)$	2	$(4, 1)$	$()$
$(0, 1, 1, 2, 2, \infty)$	$(-, 1, 1, 3, 3, -)$	$(5, 4)$	5	$(6, 3)$	(6)
$(0, 1, 1, 2, 2, 3)$	$(-, 1, 1, 3, 3, 5)$	$(4, 6)$	4	$(6, 3, 2)$	$()$
$(0, 1, 1, 2, 2, 3)$	$(-, 1, 1, 3, 3, 5)$	(6)	6	$(5, 4)$	$()$
$(0, 1, 1, 2, 2, 3)$	$(-, 1, 1, 3, 3, 5)$	$()$			

Übung a. Wir betrachten den folgenden Graph mit $V = \underline{6}$ und Wurzel $w = 1$:



Man beschreibe den Verlauf der Breitensuche mit einer Tabelle wie im Beispiel, wobei die Adjazenzlisten mit aufsteigender Nummerierung angeordnet sind.

Übung b. Geben Sie eine Schleifeninvariante für die while-Schleife in der Breitensuche an.

Bemerkung b. Die *Tiefensuche* wird realisiert, wenn man die queue (FIFO) durch einen stack (LIFO=„Last-in-first-out“) ersetzt. Geht es nur um die Bestimmung der Zusammenhangskomponente von w bzw. um die Prüfung des gesamten Graphen auf Zusammenhang, dann spielt es keine Rolle, ob Breiten- oder Tiefensuche verwendet wird.

4.2.3 Dijkstra's Algorithmus

Definition. Ein (ungerichteter) *gewichteter Graph* ist ein Tripel $G = (V, E, f)$, wobei (V, E) ein Graph ist und w eine *Gewichtsfunktion* $f : E \to \mathbb{R}_{\geq 0}$. Für jede Teilmenge $T \subseteq E$ und jeden Kantenzug $z = (v_0, \ldots, v_l)$ in G definieren wir deren *Gewichte* als $f(T) := \sum_{e \in T} f(e)$ bzw. $f(z) := \sum_{i=1}^{l} f(v_{i-1}v_i)$.

Für alle $v, w \in V$ mit $v \sim w$ definieren wir die *Distanz* zwischen v und w als

$$d(v, w) := \min\{f(z) \mid z \text{ ist } v\text{-}w\text{-Pfad in } G\} \in \mathbb{R}_{\geq 0}.$$

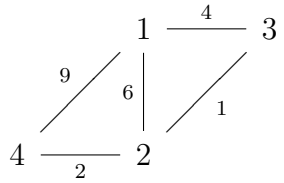
Für alle $v, w \in V$ mit $v \not\sim w$ wird $d(v, w) := \infty$ gesetzt.

Der Algorithmus von *Dijkstra* (1959) ist eine modifizierte Form der Breitensuche, die, beginnend bei einer Wurzel $w \in V$, für jeden Knoten der Zusammenhangskomponente von w die Distanz $d(v, w)$ sowie einen v-w-Pfad z mit minimalem Gewicht, d.h. mit $f(z) = d(v, w)$, berechnet.

Algorithmus. *Es sei $G = (V, E, f)$ ein gewichteter Graph mit Knotenmenge $V = \{1, \ldots, n\}$, gegeben als Adjazenzliste Γ, und es sei $w \in V$. Die in der Abbildung unten dargestellte Prozedur* Dijkstra *berechnet zu jedem $v \in V$ die Distanz $d(v) := d(v, w)$ sowie einen Vorgänger $p(v)$ in einem v-w-Pfad von minimalem Gewicht.*

Die verwendete Datenstruktur `priority queue` ist eine sog. *Vorrangwarteschlange*, bei der jedem ihrer Element ein *Prioritätswert* zugeordnet ist. Der Aufruf Insert(Q, x, n) fügt das Element x in die Warteschlange ein und ordnet x die Priorität $n \geq 0$ zu. Falls x bereits in der Warteschlange enthalten ist, wird nur die Priorität neu auf n gesetzt. Der Aufruf ExtractMin(Q) entnimmt das Element mit der niedrigsten Priorität.

Beispiel. Betrachte folgenden gewichteten Graph mit $V = \underline{4}$ und Wurzel $w = 1$:



Die erste Tabelle zeigt den Ablauf des Dijkstra-Algorithmus, wenn die Adjazenzlisten mit aufsteigender Nummerierung angeordnet sind. Jede Zeile entspricht dabei einem Durchlauf der `while`-Schleife und gibt folgendes an: die Zustände der Datenstrukturen d, p, Q zu Beginn der `while`-Schleife, das von ExtractMin gelieferte v, dessen Adjazenzliste $\Gamma(v)$, und die Teilliste der $u \in \Gamma(v)$ mit $d[v] + f(uv) < d[u]$. Die Vorrangwarteschlange Q wird jetzt

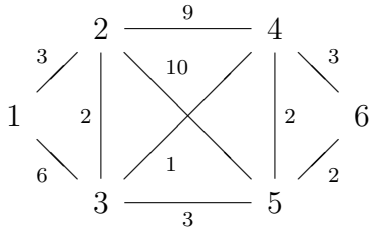
```
DIJKSTRA(Γ, w)
 1  initialisiere array d[1, ..., n] mit allen Einträgen gleich ∞
 2  initialisiere array p[1, ..., n] mit allen Einträgen gleich NIL
 3  initialisiere priority queue Q mit Elementen 1, ..., n und allen Prioritäten = ∞
 4  d[w] ← 0
 5  INSERT(Q, w, d[w])
 6  while Q nicht leer
 7  do v ← EXTRACTMIN(Q)
 8      for u ∈ Γ[v]
 9      do if d[v] + f(uv) < d[u]
10            then d[u] ← d[v] + f(uv)
11                 p[u] ← v
12                 INSERT(Q, w, d[w])
13  return d, p
```

Abbildung 4.2: Prozedur Dijkstra

als Menge geschrieben. Die Prioritäten in Q brauchen nicht extra aufgelistet werden, da sie mit den Werten $d[v]$ übereinstimmen.

d	p	Q	v	$\Gamma(v)$	$d[v] + f(uv) < d[u]$
$(0, \infty, \infty, \infty)$	$(-,-,-,-)$	$\{1,2,3,4\}$	1	$(2,3,4)$	$(2,3,4)$
$(0,6,4,9)$	$(-,1,1,-)$	$\{2,3,4\}$	3	$(1,2)$	(2)
$(0,5,4,9)$	$(-,3,1,-)$	$\{2,4\}$	2	$(1,2,4)$	(4)
$(0,5,4,8)$	$(-,3,1,2)$	$\{4\}$	4	$(1,2)$	$()$
$(0,5,4,8)$	$(-,3,1,2)$	$\{\}$			

Übung a. Betrachte folgenden gewichteten Graph mit $V = \underline{6}$ und Wurzel $w = 1$:



Man beschreibe den Verlauf des Dijkstra-Algorithmus mit einer Tabelle wie im Beispiel, wobei die Adjazenzlisten mit aufsteigender Nummerierung angeordnet sind.

Übung b. Geben Sie eine Schleifeninvariante für die while-Schleife im Dijkstra-Algorithmus an und versuchen Sie damit, die Korrektheit zu beweisen.

4.3 Hamiltonkreise und Eulertouren

Es sei $G = (V, E)$ ein Graph.

4.3.1 Definition und Beispiele

Definition.

(i) Ein Kreis der Länge n in G heißt *Hamiltonkreis*.

(ii) Eine Tour der Länge m in G heißt *Eulertour*.

Bemerkung.

(i) Ein geschlossener Kantenzug $(v_0, \ldots, v_l)$ ist genau dann ein Hamiltonkreis, wenn in der Auflistung $v_0, \ldots, v_{l-1}$ jeder Knoten aus V genau einmal vorkommt. Existiert ein Hamiltonkreis, so ist G zusammenhängend und jeder Knoten hat Grad ≥ 2.

(ii) Ein geschlossener Kantenzug $(v_0, \ldots, v_l)$ ist genau dann eine Eulertour, wenn in der Auflistung $v_0v_1, v_1v_2, \ldots, v_{l-1}v_l$ jede Kante aus E genau einmal vorkommt. Existiert eine Eulertour, so hat G höchstens eine nicht-triviale Komponente.

(iii) Ein (nicht notwendig geschlossener) Kantenzug $(v_0, \ldots, v_l)$ heißt *Eulerzug*, wenn in der Auflistung $v_0v_1, v_1v_2, \ldots, v_{l-1}v_l$ jede Kante aus E genau einmal vorkommt.

Beispiel a. Jeder Graph ohne Kanten hat eine Eulertour der Länge 0. Der Graph „Haus vom Nikolaus“ besitzt einen Eulerzug, aber keine Eulertour. Weitere Beispiele inkl. Bilder siehe Vorlesung.

Beispiel b. Das Straßennetz einer Stadt sei durch einen Graphen modelliert, in dem die Knoten den Kreuzungen entsprechen und die Kanten den Straßenabschnitten. Der Fahrer eines Schneeräumfahrzeuges sucht dann eine Eulertour.

Übung. Ist jeder Hamiltonkreis eine Eulertour?

4.3.2 Eulertouren

Bemerkung. Existiert in G eine Eulertour, so gelten:

(i) alle Knoten haben geraden Grad,

(ii) höchstens eine Komponente ist nicht-trivial.

Beweis. Hat man v über eine Kante erreicht, dann muss man v über eine andere Kante wieder verlassen. Hat man v nicht erreicht, so ist v isoliert, also $\deg v = 0$ gerade. □

Unser Ziel ist es nun, die Umkehrung dieser Bemerkung zu zeigen unter der notwendigen Voraussetzung, dass G höchstens eine nicht-triviale Komponente hat.

Satz. *Jeder Graph, der höchstens eine nicht-triviale Komponente besitzt (z.B. ein zusammenhängender Graph) und genau zwei Knoten u, v mit ungeradem Grad, besitzt auch einen u-v-Eulerzug.*

Beweis. Es sei G ein solcher Graph. Da $\deg u$ und $\deg v$ ungerade (also > 0) sind, liegen u und v in der einzigen nicht-trivialen Komponente von G. Nach Satz 4.1.8 ist u zu höchstens einer Brücke inzident. Wir führen Induktion nach m_G. Für $m_G = 1$ ist $E = \{uv\}$ und die Aussage trivial (Induktionsanfang). Sei nun $m_G > 1$ und die Behauptung für kleineres m_G bereits bewiesen. Wähle eine Kante $e = uw$, die entweder keine Brücke oder die einzige zu u inzidente Kante ist. Betrachte $G' := (V, E \setminus \{e\})$. Auch G' hat höchstens eine nicht-triviale Komponente, da entweder e keine Brücke oder u in G' isoliert ist. In G' sind ausserdem w und v die einzigen Knoten mit ungeradem Grad. Nach Induktionsvoraussetzung besitzt G' also einen Eulerzug von w nach v. Beginned mit $e = uw$ liefert dies einen Eulerzug von u nach v in G. □

Korollar. *Ein Graph besitzt genau dann eine Eulertour, wenn er höchstens eine nicht-triviale Komponente besitzt (z.B. wenn er zusammenhängend ist) und alle Knoten geraden Grad haben.*

4.3.3 Hamiltonkreise

Bemerkung. Existiert in G ein Hamiltonkreis, so ist G zusammenhängend und $n_G \geq 3$.

Satz. *Es sei $n \geq 3$ und G zusammenhängend. Falls für alle $u, v \in V$ mit $u \neq v$ und $uv \notin E$ gilt*

$$\deg u + \deg v \geq n,$$

so besitzt G einen Hamiltonkreis.

Beweis. Es sei $(v_1, \ldots, v_n)$ eine beliebige Permutation der Knotenmenge V, $n \geq 3$. Wir fassen $(v_1, \ldots, v_n, v_1)$ als ein Kreis der Länge n in dem vollständigen Graphen mit Knotenmenge V auf. Von den n Kanten dieses Kreises seien r

in E. Wenn $r = n$, dann ist der Kreis ein Hamiltonkreis in G. Sei also $r < n$, o.B.d.A. etwa $v_1v_2 \notin E$.

Behauptung: Es existiert ein $i \in \{3, \ldots, n\}$ so, dass $v_1v_{i-1}, v_2v_i \in E$. Dann ist $(v_1, v_{i-1}, v_{i-2}, \ldots, v_2, v_i, v_{i+1}, \ldots, v_n, v_1)$ ein Kreis im vollständigen Graphen, von dessen Kanten mindestens $r+1$ in E liegen. In der Tat wurden $v_1v_2, v_{i-1}v_i$ ersetzt durch v_1v_{i-1}, v_2v_i. Nach endlich vielen Schritten kommen wir zu einem Kreis der Länge n im vollständigen Graphen, dessen sämtliche Kanten in E liegen, d.h. zu einem Hamiltonkreis in G.

Wir zeigen nun die Behauptung. Die Bedingung an $i \in \{3, \ldots, n\}$ lautet $v_i \in \Gamma(v_2)$ und $v_{i-1} \in \Gamma(v_1)$. Setze $S := \Gamma(v_2)$ und $T := \{v_j \mid v_{j-1} \in \Gamma(v_1)\}$. Zu zeigen ist: $S \cap T \neq \emptyset$. Es gilt $|S| = \deg v_2$ und $|T| = \deg v_1$. Wegen $v_1v_2 \notin E$ gilt nach Voraussetzung $|S| + |T| \geq n$. Aus $n = |S \cup T| = |S| + |T| - |S \cap T|$ (Inklusions-Exklusions-Prinzip) folgt demnach $|S \cap T| \geq 1$. □

Bemerkung. Erfüllt ein Graph die Voraussetzungen des Satzes und ist n gerade, so gilt $m_G \geq \frac{n^2}{4}$. Wegen $\frac{n^2}{4} > \frac{n(n-1)}{4} = \frac{1}{2}\binom{n}{2}$ bedeutet das, dass der Graph mindestens halb soviele Kanten enthält, wie der vollständige Graph mit gleicher Knotenzahl.

Beweis. Es bezeichne S die Gradsumme des Graphen G, also $S = 2m_G$. Die Behauptung ist dann $S \geq \frac{n^2}{2}$. Es sei k der minimale Grad der unter allen Knoten auftritt. Ist $k \geq \frac{n}{2}$, dann folgt $S \geq n \cdot \frac{n}{2} = \frac{n^2}{2}$ wie gefordert. Sei also $k = \frac{n}{2} - l$ mit $l \in \mathbb{N}$. Wähle einen Knoten v mit Grad k. Partitioniere die Knotenmenge in $V = V_0 \cup V_1$, wobei $V_0 := \Gamma(v) \cup \{v\}$ und $V_1 := V \setminus V_0$. Für alle $w \in V_1$ gilt nach Voraussetzung $n \leq \deg v + \deg w = k + \deg w$, also $\deg w \geq n - k = \frac{n}{2} + l$. Nach Wahl von k gilt $\deg w \geq k = \frac{n}{2} - l$ für alle $w \in V_0$. Ausserdem ist $|V_0| = k + 1 = \frac{n}{2} - l + 1$ und $|V_1| = n - (k+1) = \frac{n}{2} + l - 1$. Es folgt

$$\begin{aligned} S &\geq |V_0|(\frac{n}{2} - l) + |V_1|(\frac{n}{2} + l) = \\ &= ((\frac{n}{2} - l) + 1)(\frac{n}{2} - l) + ((\frac{n}{2} + l) - 1)(\frac{n}{2} + l) \\ &= (\frac{n}{2} - l)^2 + (\frac{n}{2} + l)^2 + (\frac{n}{2} - l) - (\frac{n}{2} + l) \\ &= 2\frac{n^2}{4} + 2l^2 - 2l = \frac{n^2}{2} + 2(l^2 - l) \geq \frac{n^2}{2}. \end{aligned}$$

□

Algorithmus (Fleury, „Schneeräumen, 1883“)**.** *Es sei $G = (V, E)$ ein zusammenhängender Graph, dessen Knoten geraden Grad haben. Die in der Abbildung unten dargestellte Prozedur* `FLEURY` *berechnet einen Eulerzug.*

```
FLEURY(V, E)
 1  initialisiere leere Liste T
 2  v ← beliebiger Knoten aus V
 3  APPEND(T, v)
 4  while E ist nicht leer
 5  do if deg v = 1
 6        then w ← einziger Nachbar von v
 7        else  w ← ein Nachbar von v mit vw keine Brücke
 8     APPEND(T, v)
 9     E ← E \ {vw}
10     v ← w
11  return T
```

Abbildung 4.3: Prozedur Fleury

Der Aufruf Append(T, x) hängt das Element x am Ende der Liste T an.

Übung. Man zeige: Erfüllt ein Graph die Voraussetzungen des Satzes, so gilt $d_w(v) \leq 2$ für alle $v, w \in V$.

Übung. Man zeige: Erfüllt ein Graph die Voraussetzungen des Satzes und ist n ungerade, so gilt $m_G \geq \frac{n^2-1}{4}$. *Tip: Modifiziere den Beweis der Bemerkung.*

Übung. Geben Sie einen Graphen mit 6 Knoten und 9 Kanten an, der die Voraussetzungen des Satzes erfüllt.

Übung. Versuchen Sie, für kleines n, einen Graphen mit ungerader Knotenzahl n und Kantenzahl $\frac{n^2-1}{4}$ anzugeben, der die Voraussetzungen des Satzes erfüllt.

4.4 Bäume

Es sei $G = (V, E)$ ein Graph und $n_G > 0$.

4.4.1 Definition und Beispiele

Definition. G heißt *kreisfrei* bzw. *Wald*, falls G keine Kreise enthält. Ein zusammenhängender Wald heißt *Baum*. Die Knoten eines Waldes mit Grad ≤ 1 heißen Blätter.

Beispiel. Siehe Vorlesung.

Bemerkung.

(i) Ein Graph ist genau dann kreisfrei, wenn jede Kante eine Brücke ist.

(ii) Jeder Baum mit mehr als einem Knoten hat mindestens zwei Blätter.

(iii) Jeder Baum mit mehr als zwei Knoten hat höchstens $n-1$ Blätter.

Beweis.

(i) Bemerkung 4.1.8b.

(ii) Es seien G ein Baum und $(v_0, v_1, \ldots, v_l)$ ein beliebiger maximaler Pfad in G. (Ein maximaler Pfad ist einer, der sich nicht „verlängern“ lässt.) Wenn $\deg v_0 > 1$, dann hat v_0 einen Nachbarn $w \neq v_1$. Wäre $w = v_i$ für ein $2 \leq i \leq l$, so gäbe es einen Kreis in G, im Widerspruch dazu, dass G ein Baum ist. Somit ist auch $(w, v_0, \ldots, v_l)$ ein Pfad, im Widerspruch zur Maximalität von $(v_0, \ldots, v_l)$. Folglich ist $\deg v_0 \leq 1$, d.h. v_0 ist ein Blatt, und dasselbe gilt aus Symmetriegründen für v_l.

(iii) Es sei G ein Baum mit n Blättern, d.h. jeder Knoten ist Blatt. Dann ist $n \geq \sum_{v \in V} \deg v = 2m_G \geq 2(n-1) = 2n-2$, wobei die zweite Ungleichung nach Korollar (4.1.7) gilt. Daraus folgt $n \leq 2$.

□

4.4.2 Kantenzahl

Ist r die Komponentenzahl von G, so gilt nach Satz (4.1.7) $r \geq n_G - m_G$. Als Zusatz erhalten wir:

Satz. *Es gilt $r = n_G - m_G$ genau dann, wenn G kreisfrei ist.*

Lemma. *Es sei $e \in E$. Für Komponentenzahl l von $(V, E \setminus \{e\})$ gilt dann $l \leq r+1$. Weiter ist $l = r+1$ genau dann, wenn e eine Brücke ist.*

Beweis. Nach Lemma (4.4.2), angewendet auf $(V, E \setminus \{e\})$, ist $r \geq l-1$, d.h. $l \leq r+1$. Per Definition ist e genau dann eine Brücke, wenn $l > r$, also genau dann, wenn $l = r+1$. □

Beweis. Zunächst sei G kreisfrei. Wir zeigen $r = n_G - m_G$ per Induktion nach m_G (genauso wie im Beweis von Satz (4.1.7)). Ist $m_G = 0$, so besteht jede Komponente aus einem einzelnen Knoten, also gilt $r = n_G$. Sei nun $m_G > 0$ und die Behauptung für kleineres m_G bereits bewiesen. Wähle ein $e \in E$ und setze $G' := (V, E \setminus \{e\})$. Sei l die Komponentenzahl von G'. Da G kreisfrei ist, ist e eine Brücke und nach dem Lemma gilt $l = r+1$. Da G kreisfrei ist,

ist auch G' kreisfrei und nach Induktionsvoraussetzung (angewendet auf G') gilt $l = n_G - (m_G - 1) = n_G - m_G + 1$. Zusammen also $r = n_G - m_G$.

Nun sei G nicht kreisfrei. Dann existiert eine nicht-Brücke e. Der Graph $G' := (V, E \setminus \{e\})$ hat dann dieselbe Komponentenzahl wie G. Nach Satz (4.1.7) (angewendet auf G') gilt also $r \geq n_G - (m_G - 1) > n_G - m_G$. □

Korollar. *Ein Graph ist genau dann ein Baum, wenn mindestens zwei der folgenden Bedingungen erfüllt sind.*

(i) G *ist kreisfrei,*

(ii) G *ist zusammenhängend,*

(iii) $m_G = n_G - 1$.

Beweis. Zu zeigen ist, dass aus je zwei der Bedingungen die dritte folgt. Das wird offensichtlich, wenn man Bedingung (i) dem Satz gemäß durch die Bedingung $r = n_G - m_G$ ersetzt, sowie (ii) durch die Bedingung $r = 1$. □

Bemerkung. Jeder zusammenhängende Graph erfüllt nach Satz (4.1.7) $m_G \geq n_G - 1$. Jeder kreisfreie Graph erfüllt nach Satz (4.4.2) $m_G = n_G - r \leq n_G - 1$. Ein Baum ist also ein zusammenhängender Graph mit minimal möglicher Kantenzahl und ein kreisfreier Graph mit maximal möglicher Kantenzahl.

4.4.3 Spannbäume

Definition. Ein Teilgraph $G' = (V', E')$ von G heißt *Spannbaum* von G (engl. *spanning tree*), wenn G' ein Baum ist und $V' = V$.

Beispiel. Siehe Vorlesung.

Satz. *Jeder zusammenhängende Graph hat einen Spannbaum.*

Beweis. Die Breitensuche mit beliebiger Wurzel w liefert für jedes $v \in V \setminus \{w\}$ einen „Vorgänger“ $p(v)$, der kleinere Distanz zu w hat. Die Kantenmenge $E' := \{vp(v) \mid v \in V, v \neq w\}$ liefert dann einen Spannbaum (V, E') von G: Einerseits ist (V, E') zusammenhängend weil jeder Knoten über seine Vorgänger mit w verbunden ist. Andererseits sind die Kanten $vp(v)$ mit $v \in V \setminus \{w\}$ paarweise verschieden, also $|E'| = n_G - 1$. Nach Korollar (4.4.2) ist (V, E') ein Baum. □

Bemerkung. Die Blätterzahl der Spannbäume eines Graphen ist durch den Graphen nicht eindeutig festgelegt.

Zwei weitere Algorithmen zur Generierung eines Spannbaumes bieten sich an.

Algorithmus a (Sukzessives Entfernen von Kanten)**.** *Es sei (V, E) ein zusammenhängender Graph. Beginnend mit der Kantenmenge $B := E$ werden sukzessive solche Kanten aus B entfernt, die keine Brücken in (V, B) sind. Wenn das nicht mehr möglich ist, dann ist (V, B) ein Spannbaum.*

Beweis. Zu Beginn des Algorithmus ist $(V, B) = (V, E)$, also (V, B) zusammenhängend. Für jedes $e \in B$ sind dann äquivalent:

(i) e ist keine Brücke,

(ii) $(V, B \setminus \{e\})$ ist zusammenhängend.

(Das ist die Definition von Brücke.) Entfernt man aus B stets Kanten e dieser Art, so bleibt (V, B) während des gesamten Verlaufs des Algorithmus zusammenhängend. Wenn das Abbruchkriterium erfüllt ist, d.h. wenn es keine Kanten dieser Art mehr gibt, dann ist (V, B) nach Bemerkung (4.4.1) kreisfrei. Somit ist (V, B) dann ein Baum mit Knotenmenge V, also Spannbaum von (V, E). □

Algorithmus b (Sukzessives Hinzufügen von Kanten)**.** *Beginnend mit der leeren Kantenmenge $B := \emptyset$ werden sukzessive solche Kanten zu B hinzugefügt, deren Endknoten in verschiedenen Komponenten von (V, B) liegen. Wenn das nicht mehr möglich ist, dann ist (V, B) ein Spannbaum.*

Beweis. Zu Beginn des Algorithmus ist (V, B) kreisfrei ($B = \emptyset$). Für jedes $e \in E$ sind dann äquivalent:

(i) e hat Endknoten in verschiedenen Komponenten von (V, B),

(ii) $(V, B \cup \{e\})$ ist kreisfrei.

(Dies ist Bemerkung 4.1.8b.) Fügt man zu B stets Kanten e dieser Art hinzu, so bleibt (V, B) während des gesamten Verlaufs des Algorithmus kreisfrei. Wenn das Abbruchkriterium erfüllt ist, d.h. wenn es keine Kanten dieser Art mehr gibt, so ist (V, B) zusammenhängend. Somit ist (V, B) ein Baum mit Knotenmenge V, also Spannbaum von (V, E). □

Beispiel. Siehe Vorlesung.

4.4.4 Minimale Spannbäume

Es sei nun $G = (V, E)$ ein zusammenhängender Graph mit einer *Gewichtsfunktion*

$$w : E \to \mathbb{R}_{\geq 0}.$$

Für jede Teilmenge $T \subseteq E$ definieren wir

$$w(T) := \sum_{e \in T} w(e).$$

Definition. Ein *minimaler Spannbaum* von G ist ein Spannbaum (V, B) von G mit minimalem Gewicht $w(B)$ unter allen Spannbäumen von G.

Beispiel. Siehe Vorlesung.

Satz (Kruskal)**.** *Die „Greedy-Version" von Algorithmus 4.4.3b, die in jedem einzelnen Schritt unter allen möglichen hinzufügbaren Kanten eine mit geringstem Gewicht hinzufügt, liefert einen minimalen Spannbaum von G.*

Die „Greedy-Version" von Algorithmus 4.4.3b wird auch *Algorithmus von Kruskal* genannt.

Lemma (Austauschlemma)**.** *Es seien (V, A) und (V, B) zwei Bäume mit derselben Knotenmenge V. Für jedes $a \in A \setminus B$ gibt es ein $b \in B \setminus A$ so, dass $(V, B \cup \{a\} \setminus \{b\})$ auch ein Baum ist.*

Beweis. (Es reicht sogar, dass (V, A) kreisfrei ist.) Nach Korollar (4.4.2) ist $|B| = n_G - 1$. Sei $a \in A \setminus B$. Dann ist $(V, B \cup \{a\})$ zusammenhängend, aber wegen $|B \cup \{a\}| = |B| + 1 > n_G - 1$ kein Baum, enthält also einen Kreis. Wähle einen Kreis in $(V, B \cup \{a\})$ und darin eine Kante b, die nicht in A liegt. (Da (V, A) kreisfrei ist, können nicht alle Kanten des Kreises in A liegen.) Da b Teil eines Kreises in $(V, B \cup \{a\})$ ist, ist auch $(V, B \cup \{a\} \setminus \{b\})$ zusammenhängend (vgl. Bemerkung 4.1.8b). Wegen $|B \cup \{a\} \setminus \{b\}| = |B| = n_G - 1$ ist $(V, B \cup \{a\} \setminus \{b\})$ nach Korollar (4.4.2) ein Baum. □

Beweis des Satzes. Es sei (V, A) der vom Greedy-Algorithmus produzierte Spannbaum, und es sei (V, B) ein minimaler Spannbaum von G. Weiter sei $a_1, \ldots, a_{n-1}$ die Aufzählung der Kanten aus A in der Reihenfolge, wie sie vom Greedy-Algorithmus ausgewählt wurden. Falls $A \neq B$, so existiert $1 \leq i \leq n - 1$ mit $a_1, \ldots, a_{i-1} \in B$ und $a_i \notin B$. Wir nehmen weiter an, dass (V, B) unter allen minimalen Spannbäumen ein solcher ist, für den i maximal ist. Nach dem Austauschlemma gibt es ein $b \in B \setminus A$ so, dass $(V, B \cup \{a_i\} \setminus \{b\})$ auch ein Baum ist. Aus der Maximalität von i folgt, dass $(V, B \cup \{a_i\} \setminus \{b\})$ kein minimaler Spannbaum ist.Also ist $w(a_i) > w(b)$. Da $(V, \{a_1, \ldots, a_{i-1}, b\})$ als Teilgraph von (V, B) kreisfrei ist, hätte der Greedy-Algorithmus also im i-ten Schritt b anstatt a_i anwählen muss. Da dies ein Widerspruch ist, muss $A = B$ sein. □

Beispiel. Siehe Vorlesung.

Lineare Algebra I

Einleitung

In der **Algebra** geht es um Gleichungen mit „Unbekannten“ und darum, wie man sie umformt bzw. sogar löst. Mit „Lösen“ meinen wir hier immer exakte Lösungen, im Gegensatz zu angenäherten Lösungen, wie man sie etwa in der *Numerik* betrachtet. In der Regel werden Gleichungen immer über einem bestimmten Zahlbereich betrachtet, etwa über $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ oder einem endlichen Zahlbereich (z.B. $\mathbb{Z}_p$). Typisch für die Algebra ist das *Abstrahieren* von konkreten Zahlbereichen, dass es ermöglicht, mit den gleichen Methoden unabhängig vom Zahlbereich arbeiten zu können. Soweit möglich werden wir auch algorithmische Aspekte des Lösens von Gleichungen hervorheben.

Man nennt einen „Ausdruck“ oder eine „Gleichung“ mit Unbekannten **linear**, wenn die Unbekannten (in ihrer Gesamtheit) linear (insbesondere mit dem Exponenten 1) auftreten, sonst *nicht-linear*.

Beispiel. Es seien x, y, z Unbekannte und a eine Konstante. Die folgenden Gleichungen bezeichnet man als *linear*:

$$3x = 2y, \quad a^2x = 2y, \quad ax - y = 7z.$$

Dagegen sind *nicht-linear*:

$$3x^2 = 2y^2, \quad 3^x = 2y, \quad \sin x = 2y.$$

Die *Lineare Algebra* beschäftigt sich mit linearen Gleichungen, in einem gewissen Sinne also mit der „einfachsten“ Art von Gleichungen. Im Vergleich zu anderen mathematischen Disziplinen bietet sie dafür auch die mit Abstand erfolgreichsten Lösungsansätze. Probleme aus der Praxis, die nicht-linear sind, werden in der Regel zuerst versucht zu „linearisieren“.

Beispiel.

(i) $3x = 2y$ mit $x, y \in \mathbb{R}$. Es gilt $3x = 2y$ genau dann, wenn $y = \frac{3}{2}x$. Die Lösungsmenge $\mathbb{L} := \{(x, y) \mid 3y = 2y, x, y \in \mathbb{R}\}$ ergibt sich also zu

$$\mathbb{L} = \{(x, \frac{3}{2}x) \mid x \in \mathbb{R}\}.$$

Diese Menge kann als eine Gerade im 2-dimensionalen Raum aufgefasst werden.

(ii) $x - 4 = 0$ mit $x \in \mathbb{R}$. Die Lösungsmenge $\mathbb{L} := \{x \mid x - 4 = 0, x \in \mathbb{R}\}$ ergibt sich zu $\mathbb{L} = \{4\}$. Diese Menge kann als ein Punkt im 1-dimensionalen Raum aufgefasst werden.

Ein Zusammenhang zwischen Algebra und Geometrie besteht also darin, dass Lösungsmengen von Gleichungen mit Unbekannten als geometrische Objekte im „Raum“ aufgefasst werden können. Dabei entspricht die Anzahl der Unbekannten gerade der **Dimension** des Raumes. Dieser Zusammenhang ist für die Einführung des zentralen Begriffes der Linearen Algebra verantwortlich: dem **Vektorraum**.

Beispiel. Es sei $f : N \to M$ eine Abbildung. Weiter sei $x \in N$ eine Unbekannte und $c \in M$ eine Konstante. Für die (nicht notwendigerweise lineare) Gleichung $f(x) = c$ und ihre Lösungsmenge $\mathbb{L} := \{x \in N \mid f(x) = c\}$ gilt:

(i) $f(x) = c$ ist genau dann lösbar, wenn c im Bild von f liegt.

(ii) $\mathbb{L} = f^{-1}(\{c\}) = \{\text{Urbilder von } c \text{ unter } f\}$, d.h. $\mathbb{L}$ ist die Faser von f zu c.

(iii) f ist genau dann injektiv, wenn $f(x) = c$ für jedes $c \in M$ höchstens eine Lösung hat.

(iv) f ist genau dann surjektiv, wenn $f(x) = c$ für jedes $c \in M$ mindestens eine Lösung hat.

(v) f ist genau dann bijektiv, wenn $f(x) = c$ für jedes $c \in M$ genau eine Lösung hat.

Offensichtlich kann man Gleichungen also auch mit Abbildungen in Verbindung bringen. Im Fall von linearen Gleichungen sind das die **linearen Abbildungen**.

Kapitel 5

Lineare Gleichungssysteme

5.1 Lineare Gleichungssysteme und Matrizen

5.1.1 Lineare Gleichungssysteme

Definition. Ein *lineares Gleichungssystem* (über $\mathbb{R}$), kurz *LGS*, hat die Form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & ... & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & ... & + & a_{2n}x_n & = & b_2 \\ & & & & \vdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & ... & + & a_{mn}x_n & = & b_m \end{array}$$

mit $a_{ij}, b_j \in \mathbb{R}$ (die *Koeffizienten* des LGS). Das sind m Gleichungen in den n Unbekannten $x_1, ..., x_n$.

Eine *Lösung* des LGS ist ein n-Tupel $(s_1, ..., s_n)$ mit $s_i \in \mathbb{R}$ derart, dass alle m Gleichungen erfüllt sind, wenn s_i für x_i eingesetzt wird $(i = 1, ..., n)$. Die Menge aller Lösungen wird mit $\mathbb{L}$ bezeichnet. Das LGS heißt *homogen*, wenn $b_1 = b_2 = ... = b_m = 0$, sonst *inhomogen*.

Problem: Gegeben a_{ij} und b_i, bestimme alle Lösungen!

Beispiel a. $n = 2$, statt x_1, x_2 nimm x, y.

$$x^2 + y^2 = 1 \qquad \text{und} \qquad xy = 1 \qquad \text{sind nicht linear.}$$

Beispiel b. $n = 2, m = 2$.

$$\text{(i)} \begin{array}{l} x + y = 2 \\ x - y = 0 \end{array} \qquad \text{(ii)} \begin{array}{l} x + y = 2 \\ x + y = 0 \end{array} \qquad \text{(iii)} \begin{array}{r} x + y = 2 \\ 3x + 3y = 6 \end{array}$$

Lösung:

(i) Aus $x - y = 0$ folgt $x = y$. Einsetzen in $x + y = 2$ liefert $2x = 2$, also $x = 1$. Ergebnis: $\mathbb{L} = \{(1, 1)\}$. (eine Lösung)

(ii) Es folgt der Widerspruch $0 = 2$. Ergebnis: $\mathbb{L} = \emptyset$. (keine Lösung)

(iii) Aus $x + y = 2$ folgt $y = 2 - x$. Einsetzen in $3x + 3y = 6$ liefert $3x + 6 - 3x = 6$, also $6 = 6$. Das ist redundant und x bleibt „frei“. Ergebnis: $\mathbb{L} = \{(x, 2 - x) | \ x \in \mathbb{R}\}$. (unendlich viele Lösungen)

Die gezeigten Lösungswege mittels Auflösen und Einsetzen nennt man *algebraische Lösungswege*. Es gibt auch *geometrische Lösungswege* (Bilder siehe Vorlesung). Man stellt fest, dass die Lösungsmenge stets der Schnitt zweier Geraden ist. Folglich gibt es **immer** entweder gar keine, eine oder unendlich viele Lösungen!

5.1.2 Äquivalenzumformungen

Satz. *Die Lösungsmenge eines LGS ändert sich nicht, wenn*

(i) zwei Gleichungen vertauscht werden, oder

(ii) eine Gleichung mit einem $c \neq 0$ ($c \in \mathbb{R}$), multipliziert wird, oder

(iii) das c-fache ($c \in \mathbb{R}$) einer Gleichung zu einer anderen addiert wird.

Diese Umformungen heißen Äquivalenzumformungen.

Beweis. Allgemein gilt: Sind $(*)$ und $(*')$ zwei Gleichungssysteme mit $(*) \Rightarrow (*')$, so gilt für die Lösungsmengen: $\mathbb{L}((*)) \subseteq \mathbb{L}((*'))$.
zu (iii): Bei zwei Gleichungen $l_1 = r_1$ und $l_2 = r_2$ (hier steht l für „linke Seite“ und r für „rechte Seite“) haben wir die Implikationen:

$$(*) \quad \begin{array}{lcl} l_1 & = & r_1 \\ l_2 & = & r_2 \ | \cdot c \end{array} \quad \Longrightarrow \quad (*') \quad \begin{array}{lcl} l_1 & = & r_1 \\ l_2 + c \cdot l_1 & = & r_2 + c \cdot r_1 \ | \cdot (-c) \end{array}$$

$$\Longrightarrow \quad (*) \quad \begin{array}{lcl} l_1 & = & r_1 \\ l_2 & = & r_2 \end{array}$$

Also gilt $\mathbb{L}((*)) \subseteq \mathbb{L}((*')) \subseteq \mathbb{L}((*))$, d.h. $\mathbb{L}((*)) = \mathbb{L}((*'))$. □

Beispiel. Äquivalenzumformungen am Beispiel 5.1.1b:

$$\begin{array}{lcl} x + y & = & 2 \ | \cdot (-1) \\ x - y & = & 0 \end{array} \quad \Longleftrightarrow \quad \begin{array}{lcl} x + y & = & 2 \\ -2y & = & -2 \ | \cdot (-\frac{1}{2}) \end{array}$$

$$\Longleftrightarrow \quad \begin{array}{lcl} x + y & = & 2 \\ y & = & 1 \ | \cdot (-1) \end{array} \quad \Longleftrightarrow \quad \begin{array}{lcl} x & = & 1 \\ y & = & 1 \end{array}$$

Die Lösungsmenge lautet also $\mathbb{L} = \{(1, 1)\}$.

Bemerkung.

(i) Äquivalenzumformungen sind eine (bessere) Alternative zum „Auflösen und Einsetzen“.

(ii) Wir haben in dem Beispiel nur mit den Koeffizienten des LGS gerechnet. Wir können uns sparen, die Unbekannten mit aufzuschreiben, wenn wir die Koeffizienten am „richtigen Platz“ belassen. (→ Begriff Matrix)

(iii) Wir verwenden nur die arithmetischen Operationen Addition, Subtraktion, Multiplikation, Division sowie die Distributivgesetze. Das ist nicht nur im Zahlbereich $\mathbb{R}$ möglich, sondern z.B. auch in $\mathbb{Q}$ und $\mathbb{C}$. (→ Begriff Körper)

5.1.3 Matrizen

Es sei K ein beliebiger Körper.

Definition.

(i) Eine $(m \times n)$-Matrix A über K ist ein rechteckiges „Schema“ von $m \cdot n$ Elementen $a_{ij} \in K$ der Form

$$A = (a_{ij})_{\substack{1 \le i \le m, \\ 1 \le j \le n}} := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Die $a_{ij} \in K$, $1 \le i \le m$, $1 \le j \le n$, heißen die *Koeffizienten* oder *Einträge* von A.

(ii) Zwei $(m \times n)$-Matrizen $A = (a_{ij})$ und $B = (b_{ij})$ über K heißen *gleich*, geschr. $A = B$, wenn $a_{ij} = b_{ij}$ für alle $1 \le i \le m$ und alle $1 \le j \le n$ Die Menge aller $(m \times n)$-Matrizen über K wird mit $K^{m \times n}$ bezeichnet.

(iii) Sei $A = (a_{ij}) \in K^{m \times n}$.

Die $(1 \times n)$-Matrix $z_i := \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix}$ heißt i-te *Zeile* von A.

Die $(m \times 1)$-Matrix $s_j := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$ heißt j-te *Spalte* von A.

(iv) Eine $(1\times n)$-Matrix wird auch (Zeilen-)n-Tupel und eine $(m\times 1)$-Matrix wird (Spalten-)n-Tupel genannt. Wir setzen (vgl. Definition 1.2.3(v)):

$$K^n := K^{n\times 1} = \text{Menge aller Spalten-}n\text{-Tupel über } K.$$

(v) Eine $(m \times n)$-Matrix $A = (a_{ij})$ mit allen $a_{ij} = 0$ wird *Nullmatrix* genannt, geschr. $A = 0$.

Bemerkung.

(i) Im Index gilt „Zeile vor Spalte“, d.h. a_{ij} steht in der i-ten Zeile und j-ten Spalte.

(ii) Eine $(m \times n)$-Matrix $A = (a_{ij})$ über K kann als Abbildung

$$a : \underline{m} \times \underline{n} \to K, (i, j) \mapsto a(i, j) := a_{ij}$$

aufgefasst werden. Das steht in Analogie zu den n-Tupeln, die man ebenfalls als Abbildung auffassen kann (vgl. Bemerkung 1.4.1(ii)).

Schreibweise. Sind $z_1, \ldots, z_m$ die Zeilen und $s_1, \ldots, s_n$ die Spalten von A, so schreiben wir auch:

$$A = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix} = \begin{pmatrix} s_1 & s_2 & \ldots & s_n \end{pmatrix} = (s_1, s_2, \ldots, s_n).$$

Diese Vereinbarung ist Teil einer flexiblen Schreibweise, nach der eine Matrix aus Blöcken, die selbst Matrizen sind, zusammengebaut werden kann. Man kann z.B.

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

bilden, wenn A und B ebenso wie C und D jeweils gleich viele Zeilen haben, und A und C ebenso wie B und D jeweils gleich viele Spalten.

Beispiel.

(i) $\begin{pmatrix} 2 & -1 \\ 4 & 0 \\ 5 & 3 \end{pmatrix}$ ist eine (3×2)-Matrix.

(ii) $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ ist die (2×3)-Nullmatrix.

(iii) $\underbrace{\begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix}}_{(3\times 1)} \neq \underbrace{\begin{pmatrix} 2 & 4 & 5 \end{pmatrix}}_{(1\times 3)}$.

5.1.4 Die Koeffizientenmatrix

Es sei K ein beliebiger Körper.

Definition. Gegeben sei das LGS über K:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & ... & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & ... & + & a_{2n}x_n & = & b_2 \\ & & & & \vdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & ... & + & a_{mn}x_n & = & b_m \end{array}$$

mit $a_{ij}, b_i \in K$ für alle $1 \leq i \leq m, 1 \leq j \leq n$. Die Matrix $A := (a_{ij}) \in K^{m\times n}$ heißt die *Koeffizientenmatrix*, und das Spalten-m-Tupel $b := (b_i) \in K^m$ heißt die *rechte Seite* des LGS. Als *erweiterte Koeffizientenmatrix* bezeichnen wir die Matrix

$$(A, b) = \begin{pmatrix} a_{11} & ... & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & ... & a_{mn} & b_m \end{pmatrix} \in K^{m\times(n+1)}.$$

Für die Lösungsmenge des LGS schreiben wir $\mathbb{L}(A, b)$.

Bemerkung.

(i) Eine *Lösung* des LGS ist ein Spalten-n-Tupel $\begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \in K^n$ mit

$$\sum_{j=1}^{n} a_{ij}s_j = b_i \text{ für jedes } i = 1, ..., m.$$

(ii) $\mathbb{L}(A, b) \subseteq K^n$.

(iii) Die „Namen“ der Unbekannten spielen jetzt keine Rolle mehr.

Beispiel. $K = \mathbb{Q}$ und $n = m = 4$. Das LGS

$$\begin{array}{rcrcrcrcl} x_1 & + & 2x_2 & & & + & x_4 & = & 1 \\ x_1 & + & 2x_2 & + & 2x_3 & + & 3x_4 & = & 5 \\ 2x_1 & + & 4x_2 & & & + & 3x_4 & = & 5 \\ & & & & 3x_3 & + & 2x_4 & = & 3 \end{array}$$

hat die erweiterte Koeffizientenmatrix

$$\begin{pmatrix} 1 & 2 & 0 & 1 & 1 \\ 1 & 2 & 2 & 3 & 5 \\ 2 & 4 & 0 & 3 & 5 \\ 0 & 0 & 3 & 2 & 3 \end{pmatrix} \in \mathbb{Q}^{4\times 5}.$$

Man zeigt mit Äquivalenzumformungen (Rechnung siehe Vorlesung):

$$\mathbb{L} = \left\{ \begin{pmatrix} -2-2t \\ t \\ -1 \\ 3 \end{pmatrix} \middle| \, t \in \mathbb{Q} \right\} = \left\{ \begin{pmatrix} -2 \\ 0 \\ -1 \\ 3 \end{pmatrix} + t \cdot \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \middle| \, t \in \mathbb{Q} \right\} \subseteq \mathbb{Q}^4.$$

5.2 Der Gauß-Algorithmus

Es sei K ein beliebiger Körper.

5.2.1 Zeilentransformationen

Definition. Es sei $m \in \mathbb{N}$. Eine *(m-reihige) elementare Zeilentransformation* ist eine Familie von Abbildungen $g = (g_n)_{n\in\mathbb{N}}$,

$$g_n : K^{m\times n} \to K^{m\times n}, \quad A \mapsto g(A),$$

von einem der drei Typen τ, α, μ, wobei $1 \le i, j \le m$ und $c \in K$:

(i) τ_{ij}: vertauscht die i-te und j-te Zeile von A.

(ii) $\alpha_{ij}(c), i \neq j$: addiert das c-fache der j-ten Zeile zur i-ten Zeile von A.

(iii) $\mu_i(c)$ mit $c \neq 0$: multipliziert die i-te Zeile von A mit c.

Wir schreiben $A \rightsquigarrow B$, wenn die Matrix B aus A durch eine endliche Folge von elementaren Zeilentransformationen hervorgeht.

Beispiel. $K = \mathbb{Q}, m = 3, n = 4$.

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 1 \\ -1 & -1 & 5 & 6 \end{pmatrix} \overset{\tau_{23}}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & 3 & 4 \\ -1 & -1 & 5 & 6 \\ 0 & 0 & 1 & 1 \end{pmatrix} \overset{\alpha_{12}(2)}{\rightsquigarrow} \begin{pmatrix} -1 & 0 & 13 & 16 \\ -1 & -1 & 5 & 6 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

$$\overset{\mu_2(-1)}{\rightsquigarrow} \begin{pmatrix} -1 & 0 & 13 & 16 \\ 1 & 1 & -5 & -6 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Bemerkung.

(i) Jede elementare Zeilentransformation g ist *umkehrbar*, d.h. es gibt eine elementare Zeilentransformation h (mit gleichem m) so, dass für jedes $n \in \mathbb{N}$ gilt: $g_n \circ h_n = h_n \circ g_n = \mathrm{id}_{K^{m\times n}}$.

(ii) Die Relation $\rightsquigarrow$ ist eine Äquivalenzrelation auf $K^{m\times n}$. Gilt $A \rightsquigarrow B$, so nennen wir A und B *Gauß-äquivalent*.

Beweis. Übung (vgl. Satz (5.1.2)). □

Satz. *Es seien $(A,b),(A',b') \in K^{m\times(n+1)}$ die erweiterten Koeffizientenmatrizen zweier linearer Gleichungssysteme. Es gilt:*

$$(A,b) \rightsquigarrow (A',b') \quad \Longrightarrow \quad \mathbb{L}(A,b) = \mathbb{L}(A',b').$$

Beweis. Elementare Zeilentransformationen der erweiterten Koeffizientenmatrix stellen Äquivalenzumformungen des LGS im Sinne von (5.1.2) dar. Wegen Satz (5.1.2) gilt demnach

$$\mathbb{L}(A,b) = \mathbb{L}(\tau_{ij}(A,b)) = \mathbb{L}(\alpha_{ij}(c)(A,b)) = \mathbb{L}(\mu_i(c)(A,b)).$$

Die Behauptung ergibt sich durch Induktion nach der Anzahl der angewendeten elementaren Zeilentransformationen. □

Übung. Gilt auch die Umkehrung des Satzes, d.h. folgt aus $\mathbb{L}(A,b) = \mathbb{L}(A',b')$, dass $(A,b) \rightsquigarrow (A',b')$?

Frage. Es seien $A, A' \in K^{m\times n}$. Folgt aus $\mathbb{L}(A,0) = \mathbb{L}(A',0)$, dass $A \rightsquigarrow A'$?

5.2.2 Zeilenstufenform

Definition. Es sei $A \in K^{m\times n}$. Für $i = 1,\dots,m$ bezeichne z_i die i-te Zeile von A. Definiere $k_i \in \{1,\dots,n+1\}$ als die Anzahl der führenden Nullen von z_i plus 1. Dann sagen wir A hat *Zeilenstufenform*, wenn

$$k_1 < k_2 < \ldots < k_r < k_{r+1} = \ldots = k_m = n+1$$

für ein $0 \le r \le m$. Wir nennen r die *Stufenzahl* von A und $k_1,\dots,k_r$ die *Stufenindizes*.

Bemerkung. Die Definition von k_i bedeutet, dass z_i die Form

$$z_i = \begin{pmatrix} 0 & \ldots 0 & \blacksquare & \star & \ldots & \star \end{pmatrix}$$

hat, wobei $\blacksquare$ und $\star$ beliebige Einträge aus K sind, aber $\blacksquare \neq 0$ ist, und $\blacksquare$ genau an der k_i-ten Stelle steht. Enthält z_i nur Nullen, so ist $k_i = n+1$.

Eine Matrix hat demnach Zeilenstufenform, wenn sie so aussieht:

$$\left(\begin{array}{ccc|cccccccccc}
0 & \cdots & 0 & \blacksquare & \star & \cdots & \star & \star & \star & \cdots & \star & \star & \cdots & \star \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \blacksquare & \star & \cdots & \star & \star & \cdots & \star \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & & & \vdots & & & \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & & \star & \vdots & & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \blacksquare & \star & \cdots & \star \\
\hline
0 & \cdots & 0 & 0 & \cdots & & 0 & 0 & \cdots & & 0 & & \cdots & 0 \\
\vdots & & \vdots & \vdots & & & \vdots & \vdots & & & \vdots & & \vdots & \\
0 & \cdots & 0 & 0 & \cdots & & 0 & 0 & \cdots & & 0 & & \cdots & 0
\end{array}\right)$$

Die ■ bilden die „Stufen“ und k_i ist der Spaltenindex der i-ten Stufe. Es gilt

$$r = \text{ Anzahl Stufen} = \text{Anzahl nicht-Null-Zeilen.}$$

Null-Zeilen dürfen in der Zeilenstufenform nur am unteren Ende der Matrix vorkommen, und es gibt genau $m - r$ davon.

Frage. Es seien $A, A' \in K^{m\times n}$ in Zeilenstufenform. Folgt aus $A \rightsquigarrow A'$, dass A und A' gleiche Stufenzahl (Stufenindizes) haben?

5.2.3 Gauß-Algorithmus I

Satz. *Jede Matrix $A \in K^{m\times n}$ kann durch eine Folge elementarer Zeilentransformationen (vom Typ τ und α) auf Zeilenstufenform gebracht werden.*

Bemerkung a. Der Satz besagt, dass jede Matrix A Gauß-äquivalent zu einer Matrix in Zeilenstufenform ist. Die Zeilenstufenform ist allerdings nicht eindeutig. Jede Matrix, die Gauß-äquivalent zu A und in Zeilenstufenform ist, nennen wir *eine Zeilenstufenform von A.*

Algorithmus (Gauß)**.** *Es sei $A = (a_{ij}) \in K^{m\times n}$. Für $j = 1, \ldots, n$ bezeichne s_j die j-te Spalte von A. Die folgenden Schritte überführen A in Zeilenstufenform.*

1. *Wenn A die Nullmatrix oder eine 1×1-Matrix ist, dann Ende.*
2. *Setze $k := \min\{j \mid 1 \le j \le n, s_j \neq 0\}$.*
3. *Wähle ein i mit $a_{ik} \neq 0$ und wende τ_{1i} an. (τ_{11} ist erlaubt.)*
4. *Für jedes $i = 2, \ldots, m$ wende $\alpha_{i1}(-\frac{a_{ik}}{a_{1k}})$ an.*
5. *Führe Schritt 1 rekursiv mit dem Block $(a_{ij})_{\substack{2\le i\le m\\ k<j\le n}}$ aus.*

Bemerkung b.

(i) Der Gauß-Algorithmus ist ein Algorithmus, der Matrizen auf Zeilenstufenform bringt. Das Lösen von linearen Gleichungssystemen ist eine wichtige Anwendung, die wir in den Abschnitten (5.2.4) und (5.2.5) herausarbeiten werden, aber bei weitem nicht die einzige Anwendung.

(ii) Der Gauß-Algorithmus verändert nicht die Größe einer Matrix. Insbesondere dürfen Null-Zeilen (streng genommen) nicht einfach weggelassen werden. Beim Lösen von (homogenen und inhomogenen) linearen Gleichungssystemen ist das aber trotzdem sinnvoll, da Null-Zeilen redundante Gleichungen repräsentieren.

(iii) Es folgt eine Erläuterung der einzelnen Schritte:

1. Jede Nullmatrix und jede 1×1-Matrix ist in Zeilenstufenform.
2. Die k-Spalte ist die erste Spalte von links, die nicht komplett aus Nullen besteht.
3. Falls in der k-ten Spalte ganz oben eine Null steht, dann tausche die oberste Zeile gegen eine andere, so dass das nicht mehr der Fall ist.
4. Addiere geeignete Vielfache der obersten Zeile zu allen anderen Zeilen, so dass alle anderen Zeilen Null-Einträge in der k-Spalte bekommen.
5. Mache rekursiv weiter mit der Teilmatrix, die in der zweiten Zeile und der $k+1$-ten Spalte beginnt.

(iv) Die k's aus allen rekursiven Durchläufen sind genau die Stufenindizes $k_1, \ldots, k_r$ der Zeilenstufenform, die am Ende herauskommt. Insbesondere durchläuft der Algorithmus genau r Rekursionsschritte.

Beispiel.

$$\begin{pmatrix} 1 & -2 & 3 & 4 & 2 \\ 2 & -4 & 6 & 9 & 1 \\ -1 & 2 & -1 & -3 & -6 \\ 1 & -2 & 5 & 4 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -2 & 3 & 4 & 2 \\ 0 & 0 & 2 & 1 & -4 \\ 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(Rechnung siehe Vorlesung)

5.2.4 Homogene LGS

Anwendung (Lösungsverfahren für homogene LGS)**.**
Gegeben sei ein homogenes LGS mit Koeffizientenmatrix $A \in K^{m\times n}$.

1. Bringe A mittels elementarer Zeilentransformationen auf Zeilenstufenform (z.B. nach Algorithmus (5.2.3)).
2. Die r Unbekannten, die zu den Spalten $k_1, \ldots, k_r$ gehören, werden *abhängig* genannt, die anderen $n - r$ Unbekannten werden *frei* genannt.
3. Ersetze die freien Unbekannten durch Parameter $t_1, \ldots, t_{n-r}$.
4. Löse von unten nach oben nach den abhängigen Unbekannten auf. (*Rückwärtssubstitution*)

Beispiel. Für die Matrix $A \in \mathbb{Q}^{4\times 5}$ aus Beispiel (5.2.3) ergibt sich:

$$A \rightsquigarrow \begin{pmatrix} 1 & -2 & 3 & 4 & 2 \\ 0 & 0 & 2 & 1 & -4 \\ 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \mathbb{L}(A,0) = \left\{ \left. \begin{pmatrix} 2t_1 - \frac{31}{2}t_2 \\ t_1 \\ \frac{1}{2}t_2 \\ 3t_2 \\ t_2 \end{pmatrix} \right| t_1, t_2 \in \mathbb{Q} \right\}.$$

(Rechnung siehe Vorlesung)

Bemerkung.

(i) Ein homogenes LGS hat immer eine Lösung, nämlich die *triviale Lösung* $0 \in K^n$.

(ii) Hat ein homogenes LGS weniger Gleichungen als Unbekannte ($m < n$), so gibt es nicht-triviale Lösungen.

Erklärung: In Zeilenstufenform ist immer $r \leq m$. Aus $m < n$ folgt also $r < n$ bzw. $n - r > 0$. Da $n - r$ die Anzahl der freien Unbekannten ist, gibt es mehr als eine Lösung.

(iii) Für ein homogenes LGS sind folgende Aussagen äquivalent:

- das LGS ist nicht-trivial lösbar,
- $\mathbb{L} \neq \{0\}$,
- das LGS ist nicht eindeutig lösbar,
- es gibt freie Unbekannte ($n - r > 0$).

Vorsicht bei der Aussage „das LGS hat unendlich viele Lösungen“: der Körper kann endlich sein!

Übung. Es seien $A, A' \in K^{m\times n}$ in Zeilenstufenform mit $A \rightsquigarrow A'$. Man zeige als teilweise Antwort auf Frage 5.2.2: Hat A die Stufenzahl n, so hat auch A' die Stufenzahl n.

5.2.5 Inhomogene LGS

Bemerkung. Nicht jedes inhomogene LGS hat eine Lösung. Über jedem Körper ist z.B. $0\cdot x = 1$ unlösbar. Allgemein ist die lineare Gleichung $a\cdot x = b$ genau dann lösbar, wenn $a \neq 0 \vee b = 0$.

Anwendung (Lösungsverfahren für inhomogene LGS)**.**
Gegeben sei ein homogenes LGS mit erweiterter Koeffizientenmatrix $(A, b) \in K^{m\times(n+1)}$. Man bringe (A, b) mittels elementarer Zeilentransformationen auf Zeilenstufenform (z.B. nach Algorithmus (5.2.3)).
Lösungsentscheidung. Es seien $k_1, \ldots, k_r$ die Stufenindizes der Zeilenstufenform. Die Lösbarkeit kann am Index k_r abgelesen werden: Ist $r > 0$ und $k_r = n + 1$, so ist das LGS unlösbar. In der Tat hat dann die r-te Zeile, welche die unterste Nicht-Null-Zeile ist, die Form $\begin{pmatrix} 0 & \ldots & 0 & \blacksquare \end{pmatrix}$. Sie entspricht einer nach der Bemerkung unlösbaren Gleichung $0x_1 + \ldots + 0x_n = b \neq 0$. Ist dagegen $r = 0$ oder $k_r \leq n$, so ist das LGS lösbar.
Lösungsmenge. Man betrachtet zunächst nur das homogene System (d.h. man ignoriert die Spalte b bzw. setzt sie gleich 0). Gemäß Anwendung (5.2.4) definiert man freie und abhängige Unbekannte und bestimmt die Lösungsmenge $\mathbb{L}(A, 0)$. Weiter bestimmt man eine beliebige Lösung $s \in \mathbb{L}(A, b)$, z.B. indem alle freien Unbekannten gleich 0 gesetzt werden. Die Lösungsmenge ergibt sich dann als

$$\mathbb{L}(A, b) = \{s + u \,|\, u \in \mathbb{L}(A, 0)\} = s + \mathbb{L}(A, 0). \tag{5.1}$$

Beweis. Die Lösungsentscheidung ist klar. Der formale Beweis für die Gleichung (5.1) wird erst in Bemerkung 5.3.5 mit Hilfe der Matrizenrechnung erbracht. Man beachte auch Folgerung 5.3.5. □

Beispiel. $n = m = 4$.

$$A = \begin{pmatrix} 1 & -2 & 3 & 4 \\ 2 & -4 & 6 & 9 \\ -1 & 2 & -1 & -3 \\ 1 & -2 & 5 & 4 \end{pmatrix} \in \mathbb{Q}^{4\times 4}, \qquad b = \begin{pmatrix} 2 \\ 1 \\ 1 \\ -6 \end{pmatrix} \in \mathbb{Q}^4.$$

Wie in Beispiel (5.2.3) haben wir

$$(A,b) \rightsquigarrow \begin{pmatrix} 1 & -2 & 3 & 4 & 2 \\ 0 & 0 & 2 & 1 & -4 \\ 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Damit ergibt sich

$$\mathbb{L}(A,b) = \left\{ \begin{pmatrix} 2t + \frac{31}{2} \\ t \\ -\frac{1}{2} \\ -3 \end{pmatrix} \middle| \, t \in \mathbb{Q} \right\} = \left\{ \begin{pmatrix} \frac{31}{2} \\ 0 \\ -\frac{1}{2} \\ -3 \end{pmatrix} + t \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \middle| \, t \in \mathbb{Q} \right\}.$$

(Rechnung siehe Vorlesung) Wie in (5.1) schreiben wir auch:

$$\mathbb{L}(A,b) = \underbrace{\begin{pmatrix} \frac{31}{2} \\ 0 \\ -\frac{1}{2} \\ -3 \end{pmatrix}}_{\text{spezielle Lsg.}} + \underbrace{\mathbb{Q} \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}}_{\mathbb{L}(A,0)}$$

5.2.6 Reduzierte Zeilenstufenform

Beim Lösen von (homogenen oder inhomogenen) LGS mit den vorgestellten Verfahren kann man auch die Rückwärtssubstitution durch elementare Zeilentransformationen darstellen.

Beispiel. Wir formen die Zeilenstufenform aus Beispiel (5.2.5) weiter um:

$$(A,b) \rightsquigarrow \begin{pmatrix} 1 & -2 & 3 & 4 & 2 \\ 0 & 0 & 2 & 1 & -4 \\ 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -2 & 0 & 0 & \frac{31}{2} \\ 0 & 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(Rechnung siehe Vorlesung)
Hieraus kann man die Lösungsmenge ohne weitere Rechnung direkt ablesen:

$$\mathbb{L}(A,b) = \begin{pmatrix} \frac{31}{2} \\ 0 \\ \frac{1}{2} \\ -3 \end{pmatrix} + \mathbb{Q} \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Um das zu systematisieren machen wir die folgende

Definition. Es sei $A \in K^{m\times n}$.

(i) A hat *reduzierte Zeilenstufenform*, wenn A Zeilenstufenform hat (vgl. (5.2.2)) und zusätzlich gilt:

$$\text{Für alle } 1 \le j \le r \text{ ist } a_{1k_j} = a_{2k_j} = \ldots = a_{j-1,k_j} = 0, a_{jk_j} = 1$$

(ii) A hat *Normalform*, wenn A reduzierte Zeilenstufenform hat und zusätzlich gilt:

$$\text{Für alle } 1 \le i \le r \text{ ist } k_i = i.$$

Bemerkung.

(i) Eine Matrix hat reduzierte Zeilenstufenform, wenn sie so aussieht:

$$\left(\begin{array}{ccc|cccccccccc}
0 & \cdots & 0 & 1 & \star & \cdots & \star & 0 & \star & \cdots & 0 & \star & \cdots & \star \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \star & \cdots & 0 & \star & \cdots & \star \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & & & \vdots & & & \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & & 0 & \vdots & & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \star & \cdots & \star \\
\hline
0 & \cdots & 0 & 0 & \cdots & & 0 & 0 & \cdots & & 0 & & \cdots & 0 \\
\vdots & & \vdots & \vdots & & & \vdots & \vdots & & & \vdots & & \vdots & \\
0 & \cdots & 0 & 0 & \cdots & & 0 & 0 & \cdots & & 0 & & \cdots & 0
\end{array}\right)$$

wobei $\star$ beliebige Einträge aus K sind.

(ii) Eine Matrix hat Normalform, wenn sie so aussieht:

$$\left(\begin{array}{ccccc|c}
1 & 0 & 0 & \cdots & 0 & \\
0 & 1 & 0 & \cdots & 0 & \\
0 & 0 & \ddots & & \vdots & \star \\
\vdots & \vdots & & 1 & 0 & \\
0 & 0 & \cdots & 0 & 1 & \\
\hline
 & & 0 & & & 0 \\
 & & & & &
\end{array}\right)$$

wobei $\star$ ein beliebiger „Block“ ist.

5.2.7 Gauß-Algorithmus II

Satz. *Jede Matrix $A \in K^{m\times n}$ kann durch eine Folge elementarer Zeilentransformationen (vom Typ τ, α und μ) auf reduzierte Zeilenstufenform gebracht werden. Mit Spaltenvertauschungen kann A weiter auf Normalform gebracht werden.*

Übung. Man schreibe die einzelnen Schritte eines Algorithmus auf, der eine gegebene Matrix in Zeilenstufenform auf reduzierte Zeilenstufenform bringt (mittels elementarer Zeilentransformationen).

Bemerkung. Beim Lösen von (homogenen und inhomogenen) linearen Gleichungssystemen darf man auch Spalten vertauschen, wenn man über die Zuordnung zwischen Spalten und Unbekannten in geeigneter Weise Buch führt und die „b-Spalte“ an ihrer Stelle belässt. Spaltenvertauschungen gehören üblicherweise nicht zum Gauß-Algorithmus.

Beispiel a. Spaltenvertauschungen können die Rechnung abkürzen. Z.B. kann man

$$(A,b) := \begin{pmatrix} x_1 & x_2 & x_3 & b \\ 2 & 1 & -1 & 2 \\ -2 & 0 & 1 & -6 \\ 1 & 0 & 0 & 3 \end{pmatrix}$$

allein durch Spaltenvertauschungen auf die Zeilenstufenform

$$\begin{pmatrix} x_2 & x_3 & x_1 & b \\ 1 & -1 & 2 & 2 \\ 0 & 1 & -2 & -6 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

bringen. Weiter kommt man in zwei Schritten zur reduzierten Zeilenstufenform:

$$\begin{pmatrix} x_2 & x_3 & x_1 & b \\ 1 & -1 & 2 & 2 \\ 0 & 1 & -2 & -6 \\ 0 & 0 & 1 & 3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} x_2 & x_3 & x_1 & b \\ 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Diese ist gleichzeitig Normalform, und man liest als Lösungsmenge ab:

$$\mathbb{L}(A,b) = \left\{ \begin{pmatrix} 3 \\ -4 \\ 0 \end{pmatrix} \right\}.$$

(Man achte auf die Reihenfolge der Einträge!)

Beispiel b. Über $K = \mathbb{Q}$ sei die folgende erweiterte Koeffizientenmatrix in Normalform gegeben:

$$(A, b) = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 2 & 1 & 4 \\ 0 & 0 & 1 & 0 & -1 & 6 \end{pmatrix}.$$

Die Lösungsmenge kann man direkt ohne jede Rechnung ablesen:

$$\mathbb{L}(A, b) = \begin{pmatrix} 2 \\ 4 \\ 6 \\ 0 \\ 0 \end{pmatrix} + \mathbb{Q} \cdot \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \mathbb{Q} \cdot \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ -1 \end{pmatrix}.$$

(Erläuterung in der Vorlesung.)

5.3 Matrix-Arithmetik

Ab jetzt betrachten wir auch Matrizen über kommutativen Ringen anstatt nur über Körpern. In dem ganzen Abschnitt ist R ein kommutativer Ring mit $1 \neq 0$ und $R^{m \times n}$ die Menge der $m \times n$-Matrizen über R.

5.3.1 Die Grundrechenarten

Schreibweise. Es sei $A \in R^{m \times n}$. Für $1 \leq i \leq m$ und $1 \leq j \leq n$ bezeichnen wir mit A_{ij} den *i-j-Eintrag* von A, d.h. den Eintrag in der i-ten Zeile und j-ten Spalte (Merke: „Zeile vor Spalte“).

Sei umgekehrt a eine Abbildung $a : \underline{m} \times \underline{n} \to R, (i, j) \mapsto a(i, j)$. Dann bezeichnen wir mit

$$(a(i, j)) := (a(i, j))_{ij} := (a(i, j))_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

diejenige Matrix $A \in R^{m \times n}$ mit $A_{ij} = a(i, j)$ für all $1 \leq i \leq m, 1 \leq j \leq n$.

Definition. Es seien $A \in R^{m \times n}$ und $r \in R$.

(i) $A^t := (A_{ji})_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}} \in R^{n \times m}$ heißt *Transponierte von A*.

(ii) $r \cdot A := (r \cdot A_{ij})_{ij} \in R^{m \times n}$ heißt *(skalares) Vielfaches* von A.

(iii) Für jedes $B = (b_{ij}) \in R^{m \times n}$ definieren wir die *Summe* $A + B := (A_{ij} + B_{ij})_{ij} \in R^{m \times n}$.

(iv) Für jedes $B = (b_{ij}) \in R^{n\times l}, l \in \mathbb{N}$, definieren wir das *Produkt* $A \cdot B := (c_{ij})_{ij} \in R^{m\times n}$ durch

$$c_{ij} := \sum_{k=1}^{n} a_{ik}b_{kj} \text{ für alle } i, j.$$

Beispiel a.

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

$$\begin{pmatrix} 2 & 3 & 0 \\ -1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & -1 & 0 & 3 \\ 0 & 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 8 & -3 & 2 & 9 \\ -1 & 4 & -1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 3 & 0 \\ -1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix} \text{ nicht definiert}$$

Bemerkung a.

(i) Die Zeilen von A^t erhält man aus den Spalten von A (in gleicher Reihenfolge), und umgekehrt.

(ii) Die Summe von Matrizen haben wir in Spezialfällen bereits benutzt, z.B. wenn wir Zeilen addiert haben im Gauß-Algorithmus und wenn wir Spalten addiert haben in der Schreibweise $\mathbb{L}(A, b) = s + \mathbb{L}(A, 0)$.

(iii) Das Produkt ist nur definiert, wenn Spaltenzahl von A gleich Zeilenzahl von B ist.

$$\cdot : R^{m\times n} \times R^{n\times l} \to R^{m\times l}$$

Spezialfälle:

$\cdot : R^{m\times n} \times R^{n} \to R^{m}$	$l = 1$(Matrix·Spalte=Spalte)
$\cdot : R^{1\times n} \times R^{n\times l} \to R^{1\times l}$	$m = 1$(Zeile·Matrix=Zeile)
$\cdot : R^{1\times n} \times R^{n} \to R = R^{1\times 1}$	$l = m = 1$(Skalarprodukt)
$\cdot : R^{m} \times R^{1\times l} \to R^{m\times l}$	$n = 1$(Spalte·Zeile=Matrix)

Der Fall $l = m = 1$ ist das Skalarprodukt aus der Schule, nur dass hier einer Vektoren als Zeile geschrieben wird.

(iv) Wir identifizieren $R^{1\times 1}$ mit R, also die 1×1-Matrix (a) über R mit dem Ringelement $a \in R$.

(v) Es seien $A \in R^{m\times n}$ und $B \in R^{n\times l}$. Bezeichnet z_i die i-te Zeile von A und s_j die j-te Spalte von B, so gilt

$$A \cdot B = (z_i \cdot s_j)_{\substack{1\le i\le m,\\ 1\le j\le l}} \in R^{m\times l}.$$

Hier bezeichnet $\cdot$ in $z_i \cdot s_j$ die Matrixmultiplikation (also das Skalarprodukt), und die 1×1-Matrix $z_i \cdot s_j$ wird ihrem einzelnen Eintrag identifiziert.

Beispiel b.

(i) $l = 1$: $\begin{pmatrix} 1 & 0 & -2 \\ 3 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 11 \end{pmatrix}$

(ii) $m = 1$: $\begin{pmatrix} 1 & 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -2 & -1 \end{pmatrix}$

(iii) $l = m = 1$ (Skalarprodukt): $\begin{pmatrix} 1 & 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} = 3 + 0 + 0 = 3$

(iv) $n = 1$: $\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 3 & 0 & -6 \\ 1 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}$

Bemerkung b. Es seien K ein Körper, $A \in K^{m\times n}$ und $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$.

Nach Definition der Matrixmultiplikation (Spezialfall $l = 1$) ist

$$A \cdot x = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in K^m \quad \text{mit } b_i = \sum_{j=1}^{n} A_{ij}x_j \text{ für } i = 1, \ldots, m.$$

Aus diesem Grund schreiben wir das LGS über K mit erweiterter Koeffizientenmatrix $(A, b) \in K^{m\times(n+1)}$ formal als Matrixgleichung

$$A \cdot x = b,$$

wobei $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ein Spalten-n-Tupel ist, das aus Unbekannten besteht. Eine Lösung von $A \cdot x = b$ ist ein Element $s \in K^n$ mit $A \cdot s = b$

Beispiel c. Das LGS

$$\begin{array}{rcrcrcl} 2x_1 & + & x_2 & - & x_3 & = & 5 \\ x_1 & - & x_2 & & & = & 1 \end{array}$$

wird als Matrixgleichung geschrieben:

$$\underbrace{\begin{pmatrix} 2 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}}_{A} \cdot \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 5 \\ -1 \end{pmatrix}}_{b}.$$

Übung. Es sei $A \in R^{m\times n}$. Man mache sich klar, dass sowohl $(A_{ji})_{ij}$ als auch $(A_{ij})_{ji}$ die Transponierte A^t bezeichnet. Wo liegt der Unterschied?

5.3.2 Quadratische Matrizen

Definition. Es sei $n \in \mathbb{N}$.

(i) Eine $n \times n$-Matrix heißt *quadratisch*.

(ii) Die n-reihige *Einheitsmatrix* ist definiert als $E_n := (\delta_{ij})_{1\le i,j\le n}$ mit

$$\delta_{ij} := \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases}$$

Es gilt $E_n = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix} \in R^{n\times n}$, z.B. $E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

(iii) Quadratische Matrizen der Formen

$$\begin{pmatrix} \star & & 0 \\ & \ddots & \\ 0 & & \star \end{pmatrix}, \quad \begin{pmatrix} \star & \dots & \star \\ & \ddots & \\ 0 & \dots & \star \end{pmatrix}, \quad \text{bzw.} \quad \begin{pmatrix} \star & & 0 \\ \vdots & \ddots & \\ \star & \dots & \star \end{pmatrix}$$

mit beliebigen Einträgen $\star \in R$ heißen *Diagonalmatrix*, *obere Dreiecksmatrix*, bzw. *untere Dreiecksmatrix*.

5.3.3 Der Matrizenring

Satz. *Es seien $n, m, l, p \in \mathbb{N}$. Es bezeichne 0 die $m \times n$-Nullmatrix. Für alle $A, A' \in R^{m\times n}, B, B' \in R^{n\times l}, C \in R^{l\times p}$ und $r \in R$ gilt:*

(i) $(R^{m\times n}, +)$ *ist abelsche Gruppe mit neutralem Element* 0.

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

(iii) $E_m \cdot A = A = A \cdot E_n$

(iv) $(A + A') \cdot B = A \cdot B + A' \cdot B$

(v) $A \cdot (B + B') = A \cdot B + A \cdot B'$

(vi) $r \cdot (A \cdot B) = (r \cdot A) \cdot B = A \cdot (r \cdot B)$

(vii) $(A^t)^t = A$

(viii) $(A + A')^t = A^t + (A')^t$

(ix) $(A \cdot B)^t = B^t \cdot A^t$

Beweis. (i) ist klar, weil $+$ einträgeweise definiert ist (vgl. §2.1.6).
(ii) Auf beiden Seiten ergibt sich der i-j-Eintrag $\sum_{\alpha=1}^{n}\sum_{\beta=1}^{l} A_{i\alpha}B_{\alpha\beta}C_{\beta j}$ (Rechnung als Übung).
(iii) Nach Bemerkung 5.3.1v ist $E_m \cdot A = (z_i \cdot s_j)_{ij}$, wobei z_i die i-te Zeile von E_m ist und s_j die j-te Spalte von A. Es gilt

$$z_i = (0, ..., 0, \underbrace{1}_{\text{Pos. } i}, 0, ..., 0) \quad \text{und} \quad s_j = \begin{pmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{pmatrix},$$

also

$$z_i \cdot s_j = 0 \cdot A_{1j} + ... + 1 \cdot A_{ij} + 0 + ... + 0 = A_{ij}.$$

Damit ist $E_m \cdot A = (A_{ij}) = A$ gezeigt. Genauso verfährt man mit $A \cdot E_n = A$.
(iv)

$$\begin{aligned}(A + B) \cdot C &= \left(\sum_{k=1}^{n}(A_{ik} + B_{ik})C_{kj}\right)_{ij} \\ &= \left(\sum_{k=1}^{n} A_{ik}C_{kj} + \sum_{k=1}^{n} B_{ik}C_{kj}\right)_{ij} \\ &= \left(\sum_{k=1}^{n} A_{ik}C_{kj}\right)_{ij} + \left(\sum_{k=1}^{n} B_{ik}C_{kj}\right)_{ij} = AC + BC.\end{aligned}$$

(v) genauso wie (iv).
(vi) Übung (Ansatz wie in (iv)).
(vii) und (viii) sind klar.
(ix)

$$\begin{aligned} (A \cdot B)^t &= \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right)^t_{ij} = \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right)_{ji} \\ & \qquad\qquad\qquad\qquad\qquad\qquad \| \\ B^t \cdot A^t &= (b_{ji})_{ij} \cdot (a_{ji})_{ij} = \left(\sum_{k=1}^{n} b_{ki} a_{jk} \right)_{ij} \end{aligned}$$

□

Übung. Für welche Teile des Satzes braucht man, dass A kommutativ ist?

Folgerung. *Es sei $n \in \mathbb{N}$. Dann wird $R^{n\times n}$ mit der Matrix-Addition und Matrix-Multiplikation aus Definition (5.3.1) zu einem Ring, dem* Matrizenring. *Die neutralen Elemente sind $0 \in R^{n\times n}$ bzgl. der Addition und $E_n \in R^{n\times n}$ bzgl. der Multiplikation.*

Beweis. Die Eigenschaften (i)–(v) aus Satz 5.3.3. □

Bemerkung.

(i) $R^{1\times 1}$ kann mit R identifiziert werden.

(ii) $R^{n\times n}$ ist für $n \geq 2$ nicht kommutativ. Für $n = 2$ sieht man das an

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

und ein solches Beispiel läßt sich für jedes $n \geq 2$ finden.

(iii) $R^{n\times n}$ ist für $n \geq 2$ nicht nullteilerfrei (sogar wenn R ein Körper ist). Es gibt sogar $A \in R^{n\times n}, A \neq 0$, mit $A^2 = 0$, wie man an dem Beispiel $A = \begin{pmatrix} \dots & 0 & 1 \\ & 0 & 0 \\ & & \vdots \end{pmatrix}$ sieht. Insbesondere ist $R^{n\times n}$ für $n \geq 2$ kein Körper.

(iv) Für mehrfache Produkte in $R^{n\times n}$ können wir das Falk-Schema ausdehnen:

	B	C	...
A	$A \cdot B$	$A \cdot B \cdot C$	...

(v) $R^{n\times n}$ ist auch mit komponentenweiser Multiplikation ein Ring (sogar ein kommutativer Ring). Dieser Ring ist aber nicht besonders interessant. Mit komponentenweiser Multiplikation ist man nicht auf quadratische Matrizen beschränkt, auch $R^{m\times n}$ wird damit zu einem Ring.

5.3.4 Die lineare Gruppe

Definition. Die Einheitengruppe des Matrizenringes $R^{n\times n}$ (vgl. §2.2.2) wird die *lineare Gruppe* über R vom Grad n genannt, geschr.

$$\mathrm{GL}_n(R) := (R^{n\times n})^\times = \{A \in R^{n\times n} \mid A \text{ invertierbar}\}.$$

Die invertierbaren Matrizen heißen auch *regulär*. Das inverse Element zu $A \in \mathrm{GL}_n(K)$ wird die *inverse Matrix* zu A genannt, oder die *Inverse* von A.

Beispiel. $A = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \in \mathbb{Q}^{2\times 2}$ ist regulär:

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

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

Also ist $A^{-1} = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}$.

Bemerkung a. Mit $A \in \mathrm{GL}_n(K)$ ist auch $A^t \in \mathrm{GL}_n(K)$ und $(A^t)^{-1} = (A^{-1})^t$.

Beweis. Nach Satz 5.3.3ix gilt

$$A^t \cdot (A^{-1})^t = (A^{-1} \cdot A)^t = E_n^t = E_n,$$
$$(A^{-1})^t \cdot A^t = (A \cdot A^{-1})^t = E_n^t = E_n.$$

□

Übung. Es seien $A, B \in R^{n\times n}$.

(i) Kann man aus $A \cdot B = E_n$ schliessen, dass A regulär und B die Inverse von A ist?

(ii) Wenn A als regulär vorausgesetzt wird, ist dann B notwendigerweise die Inverse von A? Was hat das mit Übung 2.1.3 zu tun?

5.3.5 Matrixmultiplikation und LGS

Es seien K ein Körper, $A \in K^{m\times n}$ und $b \in K^m$. In Matrixschreibweise gilt (wie in Beispiel 5.3.1c):

$$\mathbb{L}(A,b) = \{s \in K^n \mid As = b\}.$$

Schreibweise. Wir schreiben

(i) φ_A für die Abbildung $\varphi_A : K^n \to K^m, x \mapsto A \cdot x$.

(ii) $Ax = b$ für das lineare Gleichungssystem mit erweiterter Koeffizientenmatrix (A, b).

Bemerkung.

(i) Für jedes $s \in \mathbb{L}(A,b)$ gilt

$$\mathbb{L}(A,b) = s + \mathbb{L}(A,0) := \{s + u \mid u \in \mathbb{L}(A,0)\}.$$

(ii) Das Bild von φ_A lautet $\varphi_A(K^n) = \{b \in K^m \mid Ax = b \text{ lösbar}\}$.

(iii) Die Faser von φ_A zu $b \in K^m$ lautet

$$\varphi_A^{-1}(\{b\}) = \{s \in K^n \mid As = b\} = \mathbb{L}(A,b).$$

Beweis. (i) Es sei $s \in \mathbb{L}(A,b)$, d.h. $s \in K^n$ mit $As = b$. Für ein beliebiges $t \in K^n$ folgt unter Benutzung von Satz 5.3.3(v):

$$\begin{aligned} t \in \mathbb{L}(A,b) &\Leftrightarrow At = b \Leftrightarrow At = As \\ &\Leftrightarrow A(t-s) = 0 \Leftrightarrow t - s \in \mathbb{L}(A,0) \Leftrightarrow t \in s + \mathbb{L}(A,0). \end{aligned}$$

□

Folgerung. *Es sei A' eine Zeilenstufenform von A. Folgende Aussagen sind äquivalent:*

1.) $Ax = b$ hat für jedes $b \in K^m$ höchstens eine Lösung.

2.) $Ax = 0$ ist eindeutig lösbar (nur trivial).

3.) A' hat Stufenzahl n.

4.) φ_A is injektiv.

Insbesondere ist dann $m \geq n$.

Beweis. 1.)⇒2.) Setze $b := 0$. 2.)⇒3.) Da es keine freien Unbekannten geben kann, muss A Stufenzahl n haben. 3.)⇒1.) Da A Stufenzahl n hat, gibt es keine freien Unbekannten, also höchstens eine Lösung.
1.)⇔4.) ist klar aus der Definition von φ_A. □

Übung. Wie sieht die reduzierte Zeilenstufenform von A aus, wenn die Aussagen der Folgerung gelten?

5.4 Reguläre Matrizen über Körpern

Es sei K in diesem Abschnitt ein Körper.

5.4.1 Reguläre Koeffizientenmatrizen

Es sei A die Koeffizientenmatrix eines linearen Gleichungssystems. Wir nehmen an, A ist quadratisch, d.h. das System hat genauso viele Unbekannte wie Gleichungen. Sei $A \in K^{n\times n}$.

Bemerkung. Wenn A regulär ist, dann ist $A \cdot x = b$ für jedes $b \in K^n$ eindeutig lösbar, und die Lösung lautet $x = A^{-1}b$. Insbesondere gelten alle Aussagen aus Satz (5.3.5).

Beweis. Eindeutigkeit: Aus $Ax = b = Ax'$ folgt $x = E_n x = (A^{-1}A)x = A^{-1}(Ax) = A^{-1}(Ax') = (A^{-1}A)x' = E_n x' = x'$.
Existenz: $A(A^{-1})x = (AA^{-1})x = E_n x = x$. □

Beispiel. Löse $\begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \cdot x = b$ für verschiedene $b \in K^2$. Da A regulär ist (vgl. Beispiel 5.3.4), ist $Ax = b$ für jedes $b \in K^2$ eindeutig lösbar. Die Lösung erhält man einfach durch Multiplikation mit $A^{-1} = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}$:

$$\begin{aligned} Ax = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad &\Rightarrow \quad x = A^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \\ Ax = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad &\Rightarrow \quad x = A^{-1}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}. \\ Ax = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad &\Rightarrow \quad x = A^{-1}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}. \end{aligned}$$

Beweis. Übung. □

Wir werden in Abschnitt 5.4.3 die Umkehrung der Bemerkung beweisen, also zeigen: Wenn $Ax = 0$ nur trivial lösbar ist, dann ist A regulär. Ausserdem wird in 5.4.3 ein Verfahren zur Bestimmung der Inversen hergeleitet.

5.4.2 Elementarmatrizen

Es seien K ein Körper und $m \in \mathbb{N}$. Jede Matrix $G \in K^{m\times m}$ definiert für jedes $n \in \mathbb{N}$ eine Abbildung

$$g_n : K^{m\times n} \to K^{m\times n}, \quad A \mapsto GA \tag{5.2}$$

Definition a. Eine Familie $g = (g_n)$ von Abbildungen vom Typ (5.2) nennen wir eine *m-reihige Zeilentransformation*, und die Matrix G aus (5.2) eine *definierende Matrix* von g. Weiter heißt g *umkehrbar*, wenn es eine m-reihige Zeilentransformation h gibt mit $g_n \circ h_n = h_n \circ g_n = \mathrm{id}_{K^{n\times n}}$ für alle $n \in \mathbb{N}$.

Bemerkung a. Es seien g, h zwei m-reihige Zeilentransformationen, $A \in K^{m\times n}$. Im Folgenden schreiben wir kurz $g(A)$ für $g_n(A)$, $g \circ h$ für die Zeilentransformation mit $(g \circ h)_n = g_n \circ h_n$, und $g = \mathrm{id}$ statt $g \circ h = \mathrm{id}_{K^{m\times n}}$ für alle $n \in \mathbb{N}$.

Beispiel a. Jede elementare Zeilentransformation (Definition 5.2.1) lässt sich als Multiplikation von links mit einer geeigneten Matrix auffasen, ist also eine elementare Zeilentransformationen im Sinne dieser Definition (sogar umkehrbar).

Übung a. Wie sehen definierende Matrizen der elementaren Zeilentransformationen τ, α, μ aus? Geben Sie eine Zeilentransformation an, die nicht umkehrbar ist.

Folgerung. *Es seien $A, A' \in K^{m\times n}$. Falls $A \rightsquigarrow A'$ so gibt es $G \in \mathrm{GL}_m(K)$ mit $A' = GA$.*

Beispiel b.

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -4 & 2 & -1 \end{pmatrix} \overset{\alpha_{21}(2)}{\mapsto} \begin{pmatrix} 2 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \overset{\alpha_{12}(-1)}{\mapsto} \begin{pmatrix} 2 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = A'$$

Für $G := A_{12}(-1) \cdot A_{21}(2) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 2 & 1 \end{pmatrix}$ muss nach der Folgerung gelten: $G \cdot A = A'$. (Man prüfe das nach!)

Lemma. *Es seien g, h zwei m-reihige Zeilentransformationen mit definierenden Matrizen $G, H \in K^{m\times m}$, und $A \in K^{m\times n}, B \in K^{n\times l}$. Dann gelten:*

(i) $g(A)B = g(AB)$

(ii) $G = g(E_m)$. *Die definierende Matrix von* g *ist insbesondere eindeutig.*

(iii) *Die definierende Matrix von* $g \circ h$ *lautet* GH.

(iv) $g = id \Leftrightarrow G = E_m$

(v) g *ist genau dann umkehrbar, wenn* G *regulär ist. In diesem Fall ist auch* g^{-1} *eine Zeilentransformation und hat definierende Matrix* G^{-1}.

Beweis. (i) $g(A)B = (GA)B = G(AB) = g(AB)$. (ii) $g(E_m) = GE_m = G$. (iii) $g \circ h(E_m) = g(h(E_m)) = g(H) = GH$. (iv) folgt aus (ii). (v) Falls g umkehrbar ist, etwa $g \circ h = h \circ g = \mathrm{id}$, so folgt aus (ii), (iii) und (iv) für die definierende Matrix H von h, dass $GH - HG = E_m$, also G regulär und $H = G^{-1}$. Falls G regulär ist, etwa $GH = HG = E_m$, so folgt aus (ii), (iii) und (iv) für die Zeilentransformation h mit definierender Matrix H, dass $g \circ h = h \circ g = \mathrm{id}$, also g umkehrbar. □

Definition b. Die definierenden Matrizen von elementaren Zeilentransformationen (siehe Beispiel a) werden *Elementarmatrizen* genannt.

Bemerkung b. Nach Teil (ii) des Lemmas entstehen die Elementarmatrizen durch Anwendung einer einzigen elementaren Zeilentransformation auf die $m \times m$-Einheitsmatrix, haben also die Form $T_{ij} := \tau_{ij}(E_m), A_{ij}(c) := \alpha_{ij}(c)(E_m)$ oder $M_i(c) := \mu_i(c)(E_m)$. Es gilt dann

$$\begin{aligned} \tau_{ij}: &\quad K^{m\times n} \to K^{m\times n}, &\quad A \mapsto T_{ij} \cdot A \\ \alpha_{ij}(c): &\quad K^{m\times n} \to K^{m\times n}, &\quad A \mapsto A_{ij}(c) \cdot A \\ \mu_i(c): &\quad K^{m\times n} \to K^{m\times n}, &\quad A \mapsto M_i(c) \cdot A \end{aligned}$$

Nach Teil (v) des Lemmas sind die Elementarmatrizen regulär und ihre Inversen sind selbst wieder Elementarmatrizen.

Übung b. Was passiert, wenn man Elementarmatrizen von rechts anstatt von links an eine Matrix dran multipliziert?

5.4.3 Charakterisierungen

Wir charakterisieren die Begriffe regulär und Gauß-äquivalent.

Satz. *Es seien* $A \in K^{n\times n}$ *und* A' *eine beliebige Zeilenstufenform von* A. *Folgende Aussagen sind äquivalent:*

(i) A *ist regulär.*

(ii) $A \cdot x = 0$ *ist eindeutig lösbar (nur trivial lösbar).*

(iii) $A' = E_n$.

(iv) $A \rightsquigarrow E_n$.

(v) *A ist das Produkt von Elementarmatrizen.*

Beweis. Wir machen einen Ringschluß.
(i)⇒(ii) Da A regulär ist gilt: $Ax = 0 \Rightarrow x = E_n x = A^{-1}Ax = A^{-1}0 = 0$.
(ii)⇔(iii) $Ax = 0$ is genau dann eindeutig lösbar, wenn es keine freien Unbekannten gibt, also genau dann wenn A' genau n Stufen hat, also genau dann wenn $A' = E_n$. (iii)⇒(iv) $A \rightsquigarrow A' = E_n$.
(iv)⇒(v) Wegen $A \rightsquigarrow E_n$ gibt es nach Definition 5.4.2b Elementarmatrizen $G_1, \ldots, G_r$ so, dass $E_n = G_r \cdots G_1 A$. Es folgt $A = G_1^{-1} \cdots G_r^{-1}$ und die G_i^{-1} sind wieder Elementarmatrizen.
(v)⇒(i) Da $\mathrm{GL}_n(K)$ eine Gruppe ist, sind mit Elementarmatrizen auch deren Produkte wieder regulär. □

Hiermit lässt sich nun Folgerung (5.4.2) verbessern:

Folgerung. *Zwei Matrizen $A, A' \in K^{m\times n}$ sind genau dann Gauß-äquivalent, wenn ein $G \in \mathrm{GL}_m(K)$ existiert mit $GA = A'$.*

Beweis. Übung. □

Anwendung (Matrix-Inversion)**.** Mittels Gauß-Algorithmus können wir prüfen, ob eine gegebene Matrix $A \in K^{n\times n}$ regulär ist, denn dies ist nach Aussage des Satzes an jeder reduzierten Zeilenstufenform von A zu erkennen. Wie berechnet man nun A^{-1}? Falls A regulär ist, so können wir folgendes Schema aufstellen:

A	E_n
$\vdots$	$\vdots$
E_n	B

Hier werden die elementaren Zeilentransformationen des Gauß-Algorithmus, die A in E_n überführen (linke Seite des Schemas), parallel dazu auf E_n angewendet (rechte Seite des Schemas). Es sei g die Komposition dieser elementaren Zeilentransformationen, sodass also g eine Zeilentransformation ist mit $g(A) = E_n$. Da rechts dieselbe Zeilentransformation stattfindet wir links, gilt auch $g(E_n) = B$. Nach Lemma 5.4.2(ii) ist B die definierende Matrix von g, also $E_n = g(A) = BA$. Nach Übung 5.3.4 folgt $B = A^{-1}$.

Beispiel. $A = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}$, wie in Beispiel (5.3.4).

A		E_2	
1	2	1	0
-1	-1	0	1
1	2	1	0
1	0	1	1
1	0	-1	-2
0	1	1	1
E_2		A^{-1}	

Übung a. Man mache sich folgendes klar: Das Schema in der Anwendung Matrix-Inversion ist identisch mit dem Schema für das Lösen von inhomogenen linearen Gleichungssystemen, nur dass es hier mehrere rechte Seiten gibt (eine für jede Spalte auf der rechten Seite). Jede Spalte von B lässt sich daher als Lösung eines inhomogenen LGS auffassen. Aus dieser Interpretation lässt sich die Gleichung $AB = E_n$ ablesen.

Übung b. Man prüfe A auf Regularität und berechne die Inverse:

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

5.4.4 Matrixmultiplikation und LGS II

Satz. *Es seien $A \in K^{m\times n}$ und A' eine Zeilenstufenform von A. Folgende Aussagen sind äquivalent:*

1.) $Ax = b$ hat für jedes $b \in K^m$ mindestens eine Lösung.

2.) A' hat Stufenzahl m.

3.) φ_A ist surjektiv.

Beweis. Nach Folgerung 5.4.2 gibt es $G \in \mathrm{GL}_m(K)$ mit $A' = GA$. 1.)⇒2.) Setze $b := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$. Sei x eine Lösung von $Ax = G^{-1}b$. Dann ist $A'x = GAx = GG^{-1}b = b$. Das geht nur, wenn A' genau m Stufen hat.
2.)⇒1.) Sei $b \in K^m$ beliebig. Da A' genau m Stufen hat, gibt es $x \in K^m$ mit $A'x = Gb$. Dann ist $Ax = G^{-1}A'x = G^{-1}Gb = b$.
1.)⇔3.) ist klar nach Bemerkung 5.3.5. □

Übung a. Wie sieht die Normalform von A aus, wenn die Aussagen des Satzes gelten?

Übung b. Man zeige, dass die Aussagen des Satzes äquivalent dazu sind, dass $e_1, \ldots, e_m$ im Bild von φ_A liegen.

Übung c. Es sei $n = m$. Man zeige, dass die Aussagen des Satzes sowie die Aussagen aus Folgerung 5.3.5 äquivalent sind denen aus Satz 5.4.3, und mache sich die Bedeutung klar.

5.5 Matrixgleichungen

Es seien in diesem ganzen Abschnitt K ein Körper und $m \in \mathbb{N}$. Für $i \leq m$ bezeichne e_i das Element $\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \in K^n$ mit dem 1-Eintrag an der i-ten Stelle.

5.5.1 Matrixgleichungen

Definition. Es sei $A \in K^{m\times n}$. Gleichungen der Form

$$A \cdot X = B \quad \text{und} \quad X \cdot A = B$$

mit gegebener Matrix B und unbekannter Matrix X bezeichnen wir als *Matrixgleichungen*. Dabei ist im ersten Fall $B \in K^{m\times l}$ und $X \in K^{n\times l}$ für ein $l \in \mathbb{N}$, und im zweiten Fall $B \in K^{l\times n}$ und $X \in K^{l\times m}$.

Bemerkung. Es seien $b_1, ..., b_l \in K^m$ die Spalten von $B \in K^{m\times l}$.

(i) Ein wichtiger Spezialfall sind die Gleichungen $A \cdot X = E_m$ und $X \cdot A = E_n$.

(ii) Die Lösungen von $A \cdot X = B$ sind genau diejenigen Matrizen $X \in K^{n\times l}$, deren Spalten $x_1, ..., x_l \in K^n$

$$A \cdot x_i = b_i$$

lösen für jedes $1 \leq i \leq l$.

Die Gleichung $A \cdot X = B$ kann daher als Zusammenfassung mehrerer linearer Gleichungssysteme interpretiert werden, und zwar eines für jede Spalte von B. Dabei haben alle linearen Gleichungssysteme dieselbe

Koeffizientenmatrix A, aber verschiedene rechte Seiten. Entsprechend wird $A \cdot X = B$ genauso gelöst wie einzelne lineare Gleichungssysteme: mit dem Gauß-Algorithmus.

(iii) Ist $A \in K^{n\times n}$, so kommt das Schema zur Matrix-Inversion dem Lösen der Matrixgleichung $A \cdot X = E_n$ gleich. Nach (ii) entspricht diese Gleichung wiederum den einzelnen Gleichungssystemen $Ax = e_i$ für $i = 1, \ldots, n$.

(iv) Die Gleichung $A \cdot X = B$ ist genau dann eindeutig lösbar, wenn $Ax_i = b_i$ für jedes $1 \leq i \leq l$ eindeutig lösbar ist. Gemäß Satz 5.3.5 ist das genau dann der Fall, wenn $A \cdot X = B$ lösbar ist und $A \cdot x = 0$ nur trivial lösbar.

(v) Die Gleichung $X \cdot A = B$ kann durch Transposition in eine Gleichung der Form $A' \cdot X = B'$ überführt werden, denn sie ist äquivalent zu $A^t \cdot X^t = B^t$. Hat man X^t gefunden (wie in (ii) beschrieben), so erhält man X durch transponieren, denn $X = (X^t)^t$.

Beispiel a. Es sei $K = \mathbb{R}$. Löse

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

Eine Lösung ist $X = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & -2 \end{pmatrix}$. Alle Lösungen erhält man, indem zu jeder Spalte von X beliebige Vielfache von $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ addiert werden.

(Rechnung siehe Vorlesung.)

Beispiel b. Es sei $K = \mathbb{R}$. Löse

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

Wir lösen stattdessen $A^t \cdot Y = B^t$. Gemäß Beispiel a ist $Y = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & -2 \end{pmatrix}$ eine Lösung. Somit ist $X = Y^t = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \end{pmatrix}$ eine Lösung der Ausgangsgleichung. Alle Lösungen erhält man, indem zu jeder Zeile von X beliebige Vielfache von $\begin{pmatrix} 1 & 2 & 0 \end{pmatrix}$ addiert werden.

Übung. Man zeige, dass $AX = B$ genau dann eindeutig lösbar ist, wenn $m = n$ und A regulär ist. In diesem Fall ist $X = A^{-1}B$ die eindeutige Lösung.

5.5.2 Links- und Rechtsinverse

Definition. Es sei $A \in K^{m\times n}$.

(i) Gibt es eine Matrix $B \in K^{n\times m}$ mit $A \cdot B = E_m$, so heißt A *rechtsinvertierbar* und B eine *Rechtsinverse* zu A.

(ii) Gibt es eine Matrix $B \in K^{n\times m}$ mit $B \cdot A = E_n$, so heißt A *linksinvertierbar* und B eine *Linksinverse* zu A.

Bemerkung. Links- und Rechtsinverse müssen nicht existieren, und wenn sie existieren müssen sie nicht eindeutig sein. Die Bedeutung ihrer Existenz wird im folgenden Satz klar.

Satz. *Es sei $A \in K^{m\times n}$ und $b \in K^m$.*

(i) A besitzt genau dann eine Rechtsinverse R, wenn die Aussagen aus Folgerung 5.3.5 gelten. In diesem Fall ist $n \geq m$, φ_R eine rechtsseitige Umkehrabbildung von φ_A und $R \cdot b$ eine Lösung von $A \cdot x = b$.

(ii) A besitzt genau dann eine Linksinverse R, wenn die Aussagen aus Satz 5.4.4 gelten. In diesem Fall ist $m \geq n$, φ_L eine linksseitige Umkehrabbildung von φ_A und $L \cdot b$ die einzig mögliche Lösung von $A \cdot x = b$.

Beweis. (ii) Ist R eine Rechtsinverse von A, so ist $A \cdot (R \cdot b) = (A \cdot R) \cdot b = E_m \cdot b = b$, d.h. $R \cdot b$ ist eine Lösung von $A \cdot x = b$. Damit gelten auch die Aussagen aus Satz 5.4.4. Umgekehrt findet man die i-te Spalte von R als Lösung von $Ax = e_i$ (vgl. Bemerkung 5.5.1).

(i) Ist L eine Linksinverse von A und $A \cdot x = b$ lösbar, so folgt $x = E_n \cdot x = (L \cdot A) \cdot x = L \cdot (A \cdot x) = L \cdot b$. Damit gelten auch die Aussagen aus Folgerung 5.3.5. Wir zeigen nun die Umkehrung, also sei $Ax = 0$ eindeutig lösbar. Nach Übung 5.3.5 hat A die reduzierte Zeilenstufenform $\left(\frac{E_n}{0}\right)$. Nach Folgerung 5.4.2 gibt es $G \in \mathrm{GL}_m(K)$ mit $GA = \left(\frac{E_n}{0}\right)$. Die oberen n Zeilen von G bilden somit eine Linkinverse von A. □

Beispiel a (Rechtsinverse)**.** Es sei $A = \begin{pmatrix} 2 & -1 & 1 \\ -4 & 2 & -1 \end{pmatrix} \in \mathbb{Q}^{2\times 3}$. Wir bestimmen zunächst eine Rechtsinverse $R \in \mathbb{Q}^{3\times 2}$ von A. Die i-te Spalte von R

ist eine Lösung von $Ax = e_i$. Wir berechnen:

$$\mathbb{L}(A, \begin{pmatrix} 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} -1/2 \\ 0 \\ 2 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$
$$\mathbb{L}(A, \begin{pmatrix} 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} -1/2 \\ 0 \\ 1 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Somit existiert R, ist aber nicht eindeutig. Als eine ganzzahlige Lösung lesen wir z.B.

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

ab. Wir berechnen nun Lösungen von $A \cdot x = b$ für $b = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$ und $b = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$ mit Hilfe der Rechtsinversen:

$$x = R \begin{pmatrix} 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \quad y = R \begin{pmatrix} 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ -2 \end{pmatrix}.$$

Als 2×3-Matrix kann A keine Linksinverse besitzen.

Beispiel b (Linksinverse)**.** Es sei $A = \begin{pmatrix} 2 & -4 \\ -1 & 2 \\ 1 & -1 \end{pmatrix} \in \mathbb{Q}^{2\times 3}$. Wir bestimmen eine Linksinverse $L \in \mathbb{Q}^{2\times 3}$ von A, indem wir $A^t X = E_2$ lösen und $L := X^t$ setzen. Ein solches X wurde bereits in Beispiel a berechnet. Daraus bekommen wir $L = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}$, dieses L ist aber nicht eindeutig. Wir berechnen nun Lösungen von $A \cdot x = b$ für $b = e_1, e_2, e_3, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ mit Hilfe der Linksinversen. Lb ist der einzige Kandidat für eine Lösung und wird zur Probe

eingesetzt:

$$Le_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \text{Probe: } A\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$Le_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \text{Probe: } A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ \star \\ \star \end{pmatrix} \neq \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$Le_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad \text{Probe: } A\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$L\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 8 \\ 5 \end{pmatrix}, \qquad \text{Probe: } A\begin{pmatrix} 8 \\ 5 \end{pmatrix} = \begin{pmatrix} -4 \\ \star \\ \star \end{pmatrix} \neq \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Also gilt

$$\mathbb{L}(A, e_1) = \mathbb{L}(A, e_2), \mathbb{L}(A, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}) = \emptyset, \mathbb{L}(A, e_3) = \{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\}.$$

Als 3×2-Matrix kann A keine Rechtsinverse besitzen.

Folgerung. *Für $A \in K^{n\times n}$ sind äquivalent:*

(i) *A ist regulär.*

(ii) *A besitzt eine Linksinverse.*

(iii) *A besitzt eine Rechtsinverse.*

Beweis. Der Satz und Übung 5.4.4c. □

Übung. Man zeige, dass jede Matrix A, die sowohl eine Links- als auch eine Rechtsinverse besitzt, quadratisch und regulär ist, und dass dann die Links- und Rechtsinversen eindeutig sind und mit A^{-1} übereinstimmen.

Kapitel 6

Vektorräume und lineare Abbildungen

6.1 Vektorräume

6.1.1 Definition und Beispiele

Es sei K ein Körper.

Definition. Es sei $(V, +)$ eine abelsche Gruppe. V heißt K-*Vektorraum* oder *Vektorraum über* K, wenn eine *skalare Multiplikation*

$$\cdot : K \times V \to V, (\lambda, v) \mapsto \lambda \cdot v = \lambda v$$

definiert ist, sodaß für alle $\lambda, \mu \in K$ und $v, w \in V$ gelten:

(V1) $(\lambda + \mu)v = \lambda v + \mu v$

(V2) $\lambda(v + w) = \lambda v + \lambda w$

(V3) $\lambda(\mu v) = (\lambda\mu)v$

(V4) $1v = v$

Die Elemente von V heißen *Vektoren*, die Elemente von K heißen *Skalare*.

Schreibweise. Um das Nullelement von V, den *Nullvektor*, und das Nullelement aus K zu unterscheiden, schreiben wir ersteres $\mathfrak{o}$ und letzteres 0.

Folgerung. *Es sei V ein K-Vektorraum. Für alle $\lambda \in K, v \in V$ gelten:*

(W1) $0v = \mathfrak{o}$

(W2) $\lambda\mathfrak{o} = \mathfrak{o}$

(W3) $-v = (-1)v$

(W4) $(-\lambda)v = -(\lambda v)$

(W5) $\lambda v = \mathfrak{o} \Leftrightarrow \lambda = 0$ *oder* $v = \mathfrak{o}$

Beweis. (W1) $0v = (0+0)v \overset{(V1)}{=} 0v + 0v \overset{Satz\ 2.1.4}{\Longrightarrow} 0v = \mathfrak{o}$

(W2) $\lambda\mathfrak{o} \overset{(W1)}{=} \lambda(0\mathfrak{o}) \overset{(V3)}{=} (\lambda 0)\mathfrak{o} = 0\mathfrak{o} \overset{(W1)}{=} \mathfrak{o}$

(W3) $v+(-1)v \overset{(V4)}{=} 1v+(-1)v \overset{(V1)}{=} (1+(-1))v = 0v \overset{(W1)}{=} \mathfrak{o}$, also $-v = (-1)v$ nach Satz 2.1.4.

(W4) $\lambda v + (-\lambda)v \overset{(V1)}{=} (\lambda + (-\lambda))v = 0v \overset{(W1)}{=} \mathfrak{o}$

(W5) „$\Leftarrow$“: (W1) und (W2).
„$\Rightarrow$“: Sei $\lambda v = \mathfrak{o}$ und $\lambda \neq 0$. Zu zeigen: $v = \mathfrak{o}$.

$$v \overset{(V4)}{=} 1v \overset{\lambda\neq 0}{=} (\lambda^{-1}\lambda)v \overset{(V3)}{=} \lambda^{-1}(\lambda v) \overset{Vor.}{=} \lambda^{-1}\mathfrak{o} \overset{(W2)}{=} \mathfrak{o}.$$

□

Übung. In welcher der Folgerungen wird benutzt, dass K ein Körper ist statt nur ein Ring?

Beispiel.

(i) $V = \{\mathfrak{o}\}$ ist ein K-Vektorraum (für jeden Körper K) mit der skalaren Multiplikation $\lambda \cdot \mathfrak{o} = \mathfrak{o}$ für alle $\lambda \in K$. Er wird der *triviale* K-Vektorraum genannt.

(ii) Sind $K \subseteq L$ zwei beliebige Körper (insbesondere auch für $K = L$), dann ist L ein K-Vektorraum mit

$$\cdot : K \times L \to L, (\lambda, a) \mapsto \lambda a$$

(Die skalare Multiplikation ist hier gerade die Multiplikation in L.)
z.B.: K ist K-Vektorraum, $\mathbb{C}$ ist sowohl $\mathbb{R}$-Vektorraum als auch $\mathbb{Q}$-Vektorraum, $\mathbb{R}$ ist $\mathbb{Q}$-Vektorraum, ...

(iii) $(K^{m\times n}, +)$ ist K-Vektorraum mit

$$\cdot : K \times K^{m\times n} \to K^{m\times n}, (\lambda, A) \mapsto \lambda A$$

(Die skalare Multiplikation ist das skalare Vielfache von Matrizen aus Definition 5.3.1.)
speziell: die Elemente von $K^n = K^{n\times 1}$ und $K^{1\times n}$ heißen *Spaltenvektoren* bzw. *Zeilenvektoren*.

(iv) $\mathbb{R}^3$ der 3-dimensionale „euklidische Raum“, in dem man sich Vektoren als Pfeile ausgehend von einem „Ursprung“ vorstellt. Die Addition von Vektoren entspricht dem „Hintereinanderhängen“ von Pfeilen, die skalare Multiplikation dem „Verlängern“. Nicht jeder Vektorraum erlaubt jedoch eine geometrische Vorstellung.

(v) Sei M beliebige Menge. Nach Satz 2.1.6 ist $(\mathrm{Abb}(M,K),+)$ eine abelsche Gruppe. $\mathrm{Abb}(M,K)$ wird zu einem K-Vektorraum mit der skalaren Multiplikation:

$$\ldots$$

Wichtige Beispiele hiervon sind die $\mathbb{R}$-Vektorräume:

$\mathrm{Abb}(\mathbb{R},\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R}\}$ = reelle Funktionen
$\mathrm{Abb}(\mathbb{N},\mathbb{R}) := \{f : \mathbb{N} \to \mathbb{R}\} = \{(a_0, a_1, a_2, \ldots) \mid a_i \in \mathbb{R}, i \in \mathbb{N}\}$ = reelle Folgen.

Übung. Wie ist die skalare Multiplikation in Beispiel (v) zu definieren? Man vergleiche Beispiel (iii) mit dem Spezialfall $M = \underline{m} \times \underline{n}$ von Beispiel (v).

6.1.2 Untervektorräume

Definition. Es sei V ein K-Vektorraum, $W \subseteq V$. W heißt *Untervektorraum* bzw. *Unterraum* von V, geschrieben $W \leq V$, wenn gelten:

(UV1) $W \neq \emptyset$

(UV2) $w + w' \in W$ für alle $w, w' \in W$

(UV3) $\lambda w \in W$ für alle $\lambda \in K, w \in W$

Bemerkung. Ein Unterraum ist also abgeschlossen unter Addition und unter skalarer Multiplikation. Jeder Unterraum von V enthält $\mathfrak{o}$ und ist selbst ein K-Vektorraum bzgl. der Addition und der skalaren Multiplikation von V. (Man zeige dies zur Übung!)

Beispiel. Es sei V ein K-Vektorraum,

(i) $\{\mathfrak{o}\} \leq V$ und $V \leq V$.

(ii) Für jedes $v \in V$ ist $K \cdot v := \{\lambda v \mid \lambda \in K\} \leq V$.

(iii) Für $W := \{(a_1, ..., a_n) \in K^{1\times n} \mid \sum_{i=1}^{n} a_i = 0\}$ ist $W \leq K^{1\times n}$

(iv) Definiere

$$\begin{aligned} C(\mathbb{R}) &:= \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ stetig } \} \\ C^\infty(\mathbb{R}) &:= \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ beliebig oft stetig diffbar } \} \\ \mathrm{Pol}(\mathbb{R}) &:= \{f : \mathbb{R} \to \mathbb{R}, x \mapsto a_n x^n + ... + a_1 x^1 + a_0 \mid a_i \in \mathbb{R}, n \in \mathbb{N}_0\} \end{aligned}$$

Dann ist $\mathrm{Pol}(\mathbb{R}) \leq C^\infty(\mathbb{R}) \leq C(\mathbb{R}) \leq \mathrm{Abb}(\mathbb{R}, \mathbb{R})$.

(v) $V = \mathbb{R}^2$ (euklidische Ebene)
Geraden durch $\mathfrak{o}$ sind Untervektorräume von V. Geraden, die nicht durch $\mathfrak{o}$ gehen, sind keine Untervektorräume von V.

(vi) Sei V ein K-Vektorraum und $W_1, W_2 \leq V$. Dann ist $W_1 + W_2 \leq V$ und $W_1 \cap W_2 \leq V$.

(Nach Schreibweise 2.1.5 ist $W_1 + W_2 = \{w_1 + w_2 \mid w_i \in W_i\} \subseteq V$.)

(vii) Für jede Matrix $A \in K^{m \times n}$ ist $\mathbb{L}(A, 0)$ ein Unterraum von K^n.

Bemerkung. Es sei $A \in K^{m \times n}$. Der Unterraum $\mathbb{L}(A, 0)$ von K^n heißt *Nullraum* von A.

6.2 Basis und Dimension

6.2.1 Linearkombinationen und Erzeugnis

Es sei K ein Körper, V ein K-Vektorraum.

Definition.

(i) Seien $v_1, \ldots, v_n \in V, n \in \mathbb{N}_0$. Eine *Linearkombination von* $(v_1, \ldots, v_n)$ ist ein formaler Ausdruck der Form $\lambda_1 v_1 + \ldots + \lambda_n v_n$. Die Elemente $\lambda_1, \ldots, \lambda_n \in K$ heißen die *Koeffizienten* der Linearkombination. Ist $v \in V$ mit $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$, so sagen wir v wird durch die Linearkombination *dargestellt*. Die Linearkombination heißt *trivial*, wenn $\lambda_1 = \ldots = \lambda_n = 0$, sonst *nicht-trivial*.

(ii) Sei $M \subseteq V$, $n \in \mathbb{N}_0$. Eine *Linearkombination aus* M ist eine Linearkombination von $(v_1, \ldots, v_n)$ mit paarweise verschiedenen $v_1, \ldots, v_n \in M$. Die Menge $\langle M \rangle$ aller Linearkombinationen aus M heißt die *lineare Hülle* von M oder das *Erzeugnis* von M oder der von M *erzeugte* oder *aufgespannte* Unterraum.

(iii) Gibt es zu gegebenem $v \in V$ eine Linearkombination von $(v_1, \ldots, v_n)$ (bzw. aus M), die v darstellt, so sagen wir auch v *lässt sich aus* $(v_1, \ldots, v_n)$ (bzw. *aus* M) *linear kombinieren.*

Bemerkung. Da wir Linearkombinationen als formale Ausdrücke auffassen, können verschiedene Linearkombinationen denselben Vektor ausdrücken. Beispielsweise sind $1v+(-1)v$ und $0v+0v$ zwei Linearkombinationen des Tupels (v, v), die als formale Ausdrücke verschieden sind, aber beide den Nullvektor darstellen.

Gleiches gilt für Linearkombinationen aus Mengen: $6v$ und $3(2v)$ sind zwei Linearkombinationen aus $\{v, 2v\}$, die als formale Ausdrücke verschieden sind, aber den gleichen Vektor darstellen. Ein weiteres Beispiel bilden die Linearkombinationen $1v + 1(-v)$ und $0v$ aus $\{v, -v\}$.

Schreibweise. $< v_1, \ldots, v_n >:=< \{v_1, \ldots, v_n\} >$.

Beispiel.

(i) $\langle \emptyset \rangle = \{\mathfrak{o}\}$.

(ii) Es sei V der „euklidische Raum“, also der $\mathbb{R}$-Vektorraum $\mathbb{R}^3$. Für jedes $v \in V \backslash \{\mathfrak{o}\}$ ist das Erzeugnis $\langle v \rangle$ ein Gerade durch den Ursprung. Ist weiter $w \in V$ und $w \notin \langle v \rangle$, so ist $\langle v, w \rangle$ eine Ebene.

(iii) $V = \mathbb{R}^3, v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$.

$v_1+v_2 = \begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix}, v_1-v_2 = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}$ sind Linearkombination von (v_1, v_2).

$$\langle v_1, v_2 \rangle = \{\lambda \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R}\} = \{\begin{pmatrix} \lambda - \mu \\ -\lambda - \mu \\ 2\mu \end{pmatrix} \mid \lambda, \mu \in \mathbb{R}\}.$$

(iv) $K = \mathbb{R}, V = C^\infty(\mathbb{R}), v_1 = \mathrm{id}_\mathbb{R}, v_2 = \sin$.

$$\begin{aligned} \langle v_1, v_2 \rangle &= \{a \cdot \mathrm{id}_\mathbb{R} + b \cdot \sin \mid a, b \in \mathbb{R}\} \\ &= \{f : \mathbb{R} \to \mathbb{R}, x \mapsto ax + b\sin(x) \mid a, b \in \mathbb{R}\} \end{aligned}$$

Übung a. Es seien v_1, v_2 wie in Beispiel (iii). Wie prüft man, ob ein gegebenes $v \in V$ in $\langle v_1, v_2 \rangle$ liegt? Man zeige weiter:

$$\langle v_1, v_2 \rangle = \{ \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \mid a_1 + a_2 + a_3 = 0 \}.$$

Satz. *Es sei $M \subseteq V$.*

(i) $M \subseteq \langle M \rangle$.

(ii) $\langle M \rangle \leq V$.

(iii) $M \subseteq W \leq V \Rightarrow \langle M \rangle \subseteq W$.
D.h. $\langle M \rangle$ ist der kleinste Untervektorraum von V, der M enthält.
D.h. $\langle M \rangle$ ist das Minimum (bzgl. der Relation $\subseteq$, vgl. Definition 1.6.2) aller Untervektorräume von V, die M enthalten.

(iv) $M \leq V \Leftrightarrow M = \langle M \rangle$.

(v) $\langle \langle M \rangle \rangle = \langle M \rangle$.

Beweis. (i) v ist Linearkombination von (v).

(ii) Wegen $\mathfrak{o} \in M$ ist $M \neq \emptyset$. Mit $w, w' \in \langle M \rangle$ und $\lambda \in K$ sind offenbar auch $w + w'$ und λw Linearkombination aus M.

(iii) Es sei $M \subseteq W \leq V$. Jede Linearkombination $\lambda_1 v_1 + \ldots \lambda_r v_r$ mit $v_1, \ldots, v_r \in M$ liegt dann in W, d.h. $\langle M \rangle \subseteq W$.

(iv) Ist $M \leq V$ so kann in (iii) $W = M$ gewählt werden, also $M \subseteq \langle M \rangle \subseteq M$, also $M = \langle M \rangle$. Ist $M = \langle M \rangle$, so gilt $M \leq V$ nach (ii).

(v) folgt aus (iii) mit Wahl $W = \langle M \rangle$.

□

Übung b. Man zeige, dass für alle Teilmengen $M \subseteq V$ und alle $v \in V$ gilt: $v \in \langle M \rangle \iff \langle M \cup \{v\} \rangle = \langle M \rangle$.

6.2.2 Zeilenraum und Spaltenraum

Es seien K ein Körper und $A = \in K^{m \times n}$ mit Zeilen $z_1, ..., z_m \in K^{1 \times n}$ und Spalten $s_1, ..., s_n \in K^m$.

Bemerkung.

(i) Sei $V = K^m$, also $s_1, ..., s_n \in V$. Wir haben

$$Ax = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} x_i s_i \text{ für jedes } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n,$$

d.h. Ax ist die Linearkombination von $(s_1, ..., s_n)$ mit Koeffizienten $x_1, \ldots, x_n$.

(ii) Sei $W = K^{1\times n}$, also $z_1, ..., z_m \in W$. Wir haben

$$yA = (y_1, ..., y_m)A = \sum_{i=1}^{m} y_i z_i \text{ für jedes } y = (y_1, ..., y_m) \in K^{1\times m},$$

d.h. yA ist die Linearkombination von $(z_1, ..., z_m)$ mit Koeffizienten $y_1, \ldots, y_m$.

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

$$A \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} + (-1) \cdot \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} - \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

$$\begin{aligned} \begin{pmatrix} 2 & -1 \end{pmatrix} A &= 2 \cdot \begin{pmatrix} 1 & 0 & -2 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 3 & 2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 & -4 \end{pmatrix} - \begin{pmatrix} 3 & 2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -1 & -2 & -4 \end{pmatrix} \end{aligned}$$

Definition.

(i) $\mathrm{ZR}(A) := \langle \{z_1, ..., z_m\} \rangle \leq K^{1\times n}$ heißt *Zeilenraum von A.*

(ii) $\mathrm{SR}(A) := \langle \{s_1, ..., s_n\} \rangle \leq K^m$ heißt *Spaltenraum von A.*

Übung.

(i) $\mathrm{SR}(A) = \{Ax \mid x \in K^n\}$.

(ii) $b \in \mathrm{SR}(A) \iff Ax = b$ lösbar.

6.2.3 Lineare Abhängigkeit

Es sei K ein Körper, V ein K-Vektorraum.

Definition. Es seien $n \in \mathbb{N}_0, \mathcal{T} = (v_1, ..., v_n)$ ein n-Tupel über V, und $M \subseteq V$. Eine *lineare Abhängigkeit von* $\mathcal{T}$ (bzw. M) ist eine nicht-triviale Linearkombination von $\mathcal{T}$ (bzw. *aus* M), die den Nullvektor darstellt. Wir nennen $\mathcal{T}$ (bzw. M) *linear abhängig*, falls eine lineare Abhängigkeit von $\mathcal{T}$ (bzw. M) existiert, sonst *linear unabhängig*.

Bemerkung.

(i) $(v_1, \dots, v_n)$ ist genau dann linear abhängig, wenn $\lambda_1, ..., \lambda_n \in K$ existieren, nicht alle $\lambda_i = 0$, mit $\sum_{i=1}^{n} \lambda_i v_i = \mathfrak{o}$.

(ii) $(v_1, \dots, v_n)$ ist genau dann linear unabhängig, wenn jede Linearkombination von $(v_1, \dots, v_n)$, die $\mathfrak{o}$ darstellt, trivial ist. D.h. wenn gilt:

$$\sum_{i=1}^{n} \lambda_i v_i = \mathfrak{o} \Rightarrow \lambda_1 = ... = \lambda_n = 0.$$

(iii) M ist genau dann linear abhängig, wenn paarweise verschiedene $v_1, \dots, v_n \in M$ existieren ($n \in \mathbb{N}$) sowie $\lambda_1, \dots, \lambda_n \in K \backslash \{0\}$ mit $\sum_{i=1}^{n} \lambda_i v_i = \mathfrak{o}$.

(iv) M linear abhängig $\Rightarrow$ jedes $M' \supseteq M$ linear abhängig

(v) M linear unabhängig $\Rightarrow$ jedes $M' \subseteq M$ linear unabhängig

Beispiel.

(i) $\mathfrak{o} \in M \Rightarrow M$ linear abhängig

(ii) $(..., v, ..., v, ...)$ linear abhängig

(iii) $v \neq \mathfrak{o} \Rightarrow \{v\}$ linear unabhängig

(iv) $\emptyset$ ist linear unabhängig

(v) Es sei $K = \mathbb{Q}$ und $V = \mathbb{Q}^2$. Dann ist $\{\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix}\}$ linear unabhängig:

Seien $\lambda_1, \lambda_2 \in \mathbb{Q}$ mit $\lambda_1 \cdot \begin{pmatrix}1\\2\end{pmatrix} + \lambda_2 \cdot \begin{pmatrix}3\\4\end{pmatrix} = \mathfrak{o}$, d.h. $\begin{pmatrix}1 & 3\\2 & 4\end{pmatrix} \cdot \begin{pmatrix}\lambda_1\\\lambda_2\end{pmatrix} = \mathfrak{o}$.

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \overset{\text{Gauß}}{\rightsquigarrow} \begin{pmatrix} 1 & 3 \\ 0 & -2 \end{pmatrix}$$

$\Rightarrow \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \cdot x = 0$ ist eindeutig lösbar (d.h. nur trivial lösbar).

Also folgt $\lambda_1 = \lambda_2 = 0$.

Dagegen ist $\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 6 \end{pmatrix}\}$ linear abhängig:

$$-\begin{pmatrix} 1 \\ 2 \end{pmatrix} + 2\begin{pmatrix} 3 \\ 4 \end{pmatrix} - \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

(vi) Die Spalten einer Matrix $A \in K^{m \times n}$ sind genau dann linear unabhängig, wenn $Ax = 0$ nur trivial lösbar ist.

Übung a. Es sei $A \in K^{m \times n}$. Man zeige:

(i) Ist A in Zeilenstufenform, und seien $z_1, ..., z_r \in K^{1 \times m}$, die nicht-Null-Zeilen von A. Dann ist $(z_1, ..., z_r)$ linear unabhängig.

(ii) Ist A in Zeilenstufenform, und seien $s_1, ..., s_r \in K^n$ die Spalten von A, die zu den Stufenindizes gehören. Dann ist $(s_1, ..., s_r)$ linear unabhängig.

(iii) Die Zeilen von E_n sind linear unabhängig.

(iv) Die Spalten von E_n sind linear unabhängig.

Übung b. Man definiere eine „Spaltenstufenform“ von A und zeige, dass die nicht-Null-Spalten einer Matrix in Spaltenstufenform linear unabhängig sind.

Übung c. Man zeige, dass es für jede linear abhängige Menge $M \subseteq V$ ein $v \in M$ gib nach mit $\langle M \setminus \{v\} \rangle = \langle M \rangle$.

Übung d. Es seien $K = \mathbb{R}, V = \mathbb{R}^2, u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} -3 \\ -3 \end{pmatrix}$. Man zeige, dass die Menge $M = \{u_1, u_2, u_3\}$ linear abhängig ist. Für welche $w \in M$ gilt $\langle M \rangle = \langle M \setminus \{w\} \rangle$?

Übung e. Es seien $u, v \in V$. Wann ist (u, v) linear abhängig? Wann ist $\{u, v\}$ linear abhängig?

6.2.4 Basen

Es sei K ein Körper, V ein K-Vektorraum.

Definition. Eine Teilmenge $M \subseteq V$ heißt *Erzeugendensystem von V*, wenn $V = \langle M \rangle$. Eine *Basis von V* ist ein linear unabhängiges Erzeugendensystem von V.

Die *Länge* eines Erzeugendensystems M ist definiert als die Zahl $|M|$ falls $|M| < \infty$, und sonst als ∞ (unendlich). Wenn es ein endliches Erzeugendensystem gibt, so nennen wir den Vektorraum *endlich erzeugt*.

Beispiel.

(i) Die leere Menge $\emptyset$ ist Basis des trivialen K-Vektorraums $\{\mathfrak{o}\}$.

(ii) $V = K^n$. Für $1 \leq i \leq m$ sei

$$e_i := \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \text{ der } i\text{-te Einheitsvektor (1 an der } i\text{-ten Stelle).}$$

Dann ist $\{e_1, ..., e_n\}$ Basis von K^n, genannt die *Standardbasis*.

(iii) $V = K^{m \times n}$. Für $1 \leq i \leq m$, $1 \leq j \leq n$ sei $E_{ij} \in K^{m \times n}$ die Matrix mit einer 1 an Position (i, j) und Nullen sonst. Dann ist

$$\{E_{11}, ..., E_{1n}, E_{21}, ..., E_{2n}, E_{31}, ..., E_{mn}\}$$

eine Basis von V, genannt die *Standardbasis*.

(iv) $\{1, i\}$ ist eine Basis des $\mathbb{R}$-Vektorraums $\mathbb{C}$.

(v) $\{1\}$ ist eine Basis des $\mathbb{C}$-Vektorraums $\mathbb{C}$.

(vi) Der $\mathbb{Q}$-Vektorraum $\mathbb{R}$ ist nicht endlich-erzeugt.

(vii) Der $\mathbb{R}$-Vektorraum $\mathrm{Pol}(\mathbb{R})$ hat die unendliche Basis $\{1, x, x^2, x^3, \ldots\}$.

Satz. *Für $B \subseteq V$ sind äquivalent:*

(i) B ist eine Basis von V.

(ii) B ist ein minimales Erzeugendensystem von V.
(D.h. keine echte Teilmenge von B ist Erzeugendensystem von V.)

(iii) *B ist eine maximale linear unabhängige Teilmenge von V.*
(D.h. keine echte Obermenge von B in V ist linear unabhängig.)

Wir benötigen für den Beweis des Satzes das folgende Lemma, welches einen Zusammenhang zwischen den Begriffen Erzeugnis und lineare Abhängigkeit herstellt.

Lemma. *Es sei $M \subseteq V$, $v \notin M$.*

(i) *Wenn $v \in \langle M \rangle$, dann $M \cup \{v\}$ linear abhängig.*

(ii) *Wenn M linear unabhängig und $M \cup \{v\}$ linear abhängig, dann $v \in \langle M \rangle$.*

Beweis. (i) Sei $v = \sum_{i=1}^{n} \lambda_i v_i$ mit $v_i \in M$. Durch Zusammenfassung von Summanden können wir die v_i als paarweise verschieden annehmen. Nach Voraussetzung $v \notin M$ ist auch v von allen v_i verschieden. Es folgt, dass $1v - \sum_{i=1}^{n} \lambda_i v_i = \mathsf{o}$ eine lineare Abhängigkeit von $M \cup \{v\}$ ist.

(ii) Sei $\sum_{i=1}^{n} \lambda_i v_i$ eine lineare Abhängigkeit von $v_i \in M \cup \{v\}$. Da M linear unabhängig ist, kommt v in dieser Linearkombination mit Koeffizient $\neq 0$ vor (sonst wäre schon M linear abhängig). Wir nehmen oBdA $v = v_1$ an. Dann folgt $v = -\sum_{i=2}^{n} \frac{\lambda_i}{\lambda_1} v_i$. Wegen $v_2, \ldots, v_n \neq v$ sind $v_2, \ldots, v_n \in M$, also $v \in \langle M \rangle$. □

Übung a. Es sei $M \subseteq V$ linear unabhängig. Gilt für jedes $v \in V$: $M \cup \{v\}$ linear abhängig $\Leftrightarrow v \in \langle M \rangle$?

Beweis des Satzes. (i)⇒(ii): Sei B eine Basis von V. Dann ist B Erzeugendensystem von V. Um zu zeigen, dass B minimal mit dieser Eigenschaft ist, wählen wir ein beliebiges $v \in B$, setzen $M := B \setminus \{v\}$ und zeigen, dass M kein Erzeugendensystem von V ist. In der Tat, wegen $v \notin M$ and $M \cup \{v\} = B$ linear unabhängig, gilt nach Lemma (i) (Kontraposition): $v \notin \langle M \rangle$.
(ii)⇒(i): Sei B ein minimales Erzeugendensystem. Annahme: B ist linear abhängig. Dann gibt es nach Übung 6.2.3c ein $v \in B$ mit $\langle B \setminus \{v\} \rangle = \langle B \rangle$. Dies steht im Widerspruch zur Minimalität von B, also ist die Annahme falsch, d.h. B ist linear unabhängig.
(i)⇒(iii) Sei B eine Basis von V. Dann ist B linear unabhängige Teilmenge von V. Um zu zeigen, dass B maximal mit dieser Eigenschaft ist, wählen wir ein beliebiges $v \in V \setminus B$ und zeigen, dass $B \cup \{v\}$ linear abhängig ist. Das ist gerade die Aussage von Lemma (i).
(iii)⇒(ii): Sei B maximale linear unabhängige Teilmenge von V. Um zu zeigen, dass B auch Erzeugendensystem von V ist, wählen wir ein beliebiges $v \in V$ und zeigen $v \in \langle B \rangle$. Sei oBdA $v \notin B$ (sonst ist $v \in B \subseteq \langle B \rangle$ klar). Nach Voraussetzung ist dann $B \cup \{v\}$ linear abhängig, also $v \in \langle B \rangle$ nach Lemma (ii). □

Folgerung. *(Basisauswahl) Jedes endliche Erzeugendensystem von V enthält eine Basis von V. Insbesondere hat jeder endlich-erzeugte Vektorraum eine endliche Basis.*

Beweis. Es sei M ein endliches Erzeugendensystem von V, also $M \subseteq V$ mit $\langle M \rangle = V$. Wenn M keine Basis ist, dann ist M nach dem Satz kein minimales Erzeugendensystem. Also gibt es eine echte Teilmenge $M' \subsetneq M$ mit $\langle M' \rangle = V$. Da M endlich ist, kommt man nach endlich vielen Wiederholungen dieses Schlusses zu einem minimalen Erzeugendensystem, also zu einer Basis. □

Bemerkung. Die Folgerung gilt auch ohne die Annahme, dass M endlich ist. Insbesondere hat jeder Vektorraum eine Basis. Für den Beweis benötigt man allerdings das *Lemma von Zorn* (siehe Vorlesung *Mathematische Logik I*).

6.2.5 Dimension

Es sei K ein Körper, V ein **endlich-erzeugter** K-Vektorraum, B eine endliche Basis von V (existiert nach Folgerung 6.2.4), und $n = |B|$.

Lemma. *Für jede Teilmenge $M \subseteq V$ gilt:*

(i) $|M| > n \Rightarrow M$ *linear abhängig*

(ii) M *linear unabhängig* $\Rightarrow |M| \leq n$

(iii) M *erzeugt* $V \Rightarrow |M| \geq n$

Insbesondere ist jede Basis von V endlich.

Beweis. (i) Sei $B = \{v_1, ..., v_n\}$. Wegen $|M| > n$ gibt es paarweise verschiedene $w_1, \ldots, w_{n+1} \in M$. Schreibe jedes w_j als Linearkombination der Basisvektoren $v_1, \ldots, v_n$:

$$w_j = \sum_{i=1}^{n} a_{ij} v_i, \quad a_{ij} \in K, j = 1, ..., n+1.$$

Betrachte das homogene LGS über K

$$\sum_{j=1}^{n+1} a_{ij} x_j = 0 \quad \text{für } i = 1, ..., n,$$

bestehend aus n Gleichungen in $n+1$ Unbekannten $x_1, \ldots, x_{n+1}$. Da es mehr Unbekannte als Gleichungen gibt, besitzt es eine nicht-triviale Lösung (Bemerkung 5.2.4(ii)), d.h. es gibt $c_1, \ldots, c_{n+1} \in K$, nicht alle $c_j = 0$, mit

$$\sum_{j=1}^{n+1} a_{ij} c_j = 0 \quad \text{für } i = 1, ..., n.$$

Es folgt

$$\sum_{j=1}^{n+1} c_j w_j = \sum_{j=1}^{n+1} c_j \sum_{i=1}^{n} a_{ij} v_i \overset{(V2)}{=} \sum_{j=1}^{n+1} \sum_{i=1}^{n} c_j (a_{ij} v_i)$$
$$\overset{(V3)}{=} \sum_{i=1}^{n} \sum_{j=1}^{n+1} (c_j a_{ij}) v_i \overset{(V1)}{=} \sum_{i=1}^{n} (\sum_{j=1}^{n+1} c_j a_{ij}) v_i = \sum_{i=1}^{n} 0 v_i = \mathfrak{o}.$$

Da nicht alle $c_j = 0$ und die w_j paarweise verschieden sind, ist M linear abhängig.

(ii) ist die Kontraposition von (i). Dies impliziert, dass jede Basis von V endlich ist.

(iii) Wir nehmen oBdA an, dass M endlich ist (sonst ist $|M| \geq n$ klar). Als endliches Erzeugendensystem von V enthält M eine (endliche) Basis B' von V (Folgerung 6.2.4). Teil (ii) mit B als M und B' als B besagt: B linear unabhängig $\Rightarrow |B| \leq |B'|$. Also $n = |B| \leq |B'| \leq |M|$. □

Satz. *Es gilt*

$$\begin{aligned} n &= \max\{|M| \,|\, M \subseteq V \text{ linear unabhängig}\} \\ &= \min\{|M| \,|\, M \subseteq V, \langle M \rangle = V\} \end{aligned}$$

Insbesondere haben alle Basen die gleiche Länge.

Beweis. Bezeichne das angegebene Maximum mit l und das Minimum mit k. Da B linear unabhängiges Erzeugendensystem ist, gilt $k \leq n \leq l$. Nach dem Lemma gilt $l \leq n \leq k$. Zusammen folgt die Gleichheit. □

Definition. Die Zahl n wird *Dimension* von V genannt, geschr. $\dim V$ bzw. genauer $\dim_K V$. Für nicht endlich-erzeugte Vektorräume setzen wir $\dim_K V := \infty$.

Folgerung. *Für jede Teilmenge $M \subseteq V$ sind äquivalent:*

(i) M ist Basis von V.

(ii) M ist linear unabhängig und $|M| = n$.

(iii) M erzeugt V und $|M| = n$.

Beweis. Das folgt aus dem Satz und der Charakterisierung in Satz 6.2.4. (Details als Übung) □

Beispiel.

(i) $\dim_K\{0\} = 0$. (Basis: $\emptyset$.)

(ii) $\dim_K K^n = n$. (Standardbasis $e_1, \ldots, e_n$.)

(iii) $\dim_K K^{n\times m} = nm$. (Basis: $\{E_{ij}\}1 \leq i \leq m, 1 \leq i \leq n$.)

(iv) $\dim_{\mathbb{Q}} \mathbb{R} = \dim_{\mathbb{Q}} \mathbb{C} = \infty$. (Basis: unbekannt.)

(v) Die $\mathbb{R}$-Vektorräume $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$, $C(\mathbb{R})$, $C^\infty(\mathbb{R})$, $\mathrm{Pol}(\mathbb{R})$, $\mathrm{Abb}(\mathbb{R}, \mathbb{N})$ haben $\dim_{\mathbb{R}} = \infty$.

(vi) $\dim_{\mathbb{R}} \mathbb{C} = 2$, $\dim_{\mathbb{C}} \mathbb{C} = 1$.

$\{1, i\}$ ist Basis von $\mathbb{C}$ als $\mathbb{R}$-Vektorraum, aber $\{1, i\}$ ist linear abhängig über $\mathbb{C}$:
$\lambda \cdot 1 + \mu \cdot i = 0 \overset{\lambda,\mu\in\mathbb{R}}{\Rightarrow} \lambda = \mu = 0$.
$\lambda \cdot 1 + \mu \cdot i = 0$ für $\lambda = i, \mu = -1 \in \mathbb{C}$.

(vii) Es seien $v_1, ..., v_n \in V$ paarweise verschieden.

$\{v_1, ..., v_n\}$ linear unabhängig $\Rightarrow \dim_K\langle v_1, ..., v_n\rangle = n$ (Basis: $\{v_1, \ldots, v_n\}$.)

Übung a. Man überlege sich ein Beispiel, in dem V zugleich Vektorraum über zwei verschiedenen Körpern ist und dabei unterschiedliche Dimensionen hat.

Übung b. Es sei M ein endliches Erzeugendensystem von V. Man zeige: $n = \max\{|M'| \,|\, M' \subseteq M, M' \text{ linear unabhängig}\}$. Wir sagen dazu: „$n$ ist die Maximalzahl linear unabhängiger Elemente von M“.

6.2.6 Basisergänzung

Es seien K ein Körper und V ein **endlich-dimensionaler** K-Vektorraum.

Satz. *Jede linear unabhängige Teilmenge von V lässt sich zu einer Basis ergänzen.*

Beweis. Es sei $M \subseteq V$ linear unabhängig. Wenn M keine Basis ist, dann ist M nach Satz 6.2.4 nicht maximal linear unabhängig, besitzt also eine echte Obermenge $M' \supseteq M$ in V, die linear unabhängig ist. Da $|M'| \leq \dim V < \infty$ (Teil (ii) von Satz 6.2.5), gelangt man nach endlich vielen Wiederholungen dieses Schlusses zu einer maximal linear unabhängigen Teilmenge von V, also zu einer Basis. □

Algorithmus. *Die in der Abbildung unten dargestellte Prozedur* `Basisergänzung` *liefert zu jeder linear unabhängigen Teilmenge $M \subseteq V$ (insbesondere auch zu $M = \emptyset$) eine Obermenge $B \supseteq M$, die Basis von V ist.*

Beweis. Nach Voraussetzung ist B in Schritt 1 linear unabhängig. Wir zeigen, dass 'B linear unabhängig' eine Schleifeninvariante ist: Aus B linear unabhängig und $v \notin \langle B \rangle$ folgt nach Teil (ii) von Lemma 6.2.4, dass auch $B \cup \{v\}$ linear unabhängig ist. Nach Teil (ii) von Satz 6.2.5 ist also stets $|B| \leq \dim V < \infty$. Da B mit jedem Durchlauf größer wird, bricht die Schleife ab. Bei Abbruch ist $\langle B \rangle = V$, also B eine Basis. (Statt $\langle B \rangle = V$ ist als Abbruchkriterium auch $|B| = \dim V$ erlaubt, sofern $\dim V$ bekannt ist.) □

BASISERGÄNZUNG(M)
1 $B \leftarrow M$
2 **while** $\langle B \rangle \neq V$
3 **do** $v \leftarrow$ beliebiges Element aus $V \setminus \langle B \rangle$
4 $B \leftarrow B \cup \{v\}$
5 **return** B

Abbildung 6.1: Prozedur Basisergänzung

Beispiel a. $V := \mathbb{R}^3, M := \{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \}$.

Wähle $v \notin \langle M \rangle$, z.B. $v = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

$M' := M \cup \{v\} = \{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \}$

Wähle $w \notin \langle M' \rangle$, z.B. $w = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

$M'' := M' \cup \{v\} = \{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \}$.

M'' ist Basis weil $|M''| = 3 = \dim V$ (Folgerung 6.2.5).

Folgerung. *Für jeden Unterraum $U \leq V$ gelten:*

(i) $\dim_K U \leq \dim_K V$,

(ii) $\dim_K U = \dim_K V \Rightarrow U = V$.

Beweis. Sei B eine Basis von U. Dann ist $|B| = \dim_K U$, und B ist auch linear unabhängig als Teilmenge von V. Nach Satz 6.2.5 folgt $\dim_K U = |B| \leq \dim_K V$. Falls $|B| = \dim_K U = \dim_K V$, so ist B nach Folgerung 6.2.5 auch Basis von V. Also $U = \langle B \rangle = V$. □

Beispiel b. Mit dem Dimensionsbegriff und der Folgerung kann man sehr leicht alle U Unterräume von $\mathbb{R}^3$ bestimmen:

$\dim U = 0:$	$U = \langle \emptyset \rangle = \{\mathfrak{o}\}$	(Ursprung)
$\dim U = 1:$	$U = \langle v \rangle, v \neq \mathfrak{o}$	(alle Geraden durch $\mathfrak{o}$)
$\dim U = 2:$	$U = \langle v, w \rangle, v, w \neq \mathfrak{o}, w \notin \langle v \rangle$	(alle Ebenen durch $\mathfrak{o}$)
$\dim U = 3:$	$U = V$	(ganz $\mathbb{R}^3$)

6.2.7 Koordinaten

Es sei V ein **endlich-dimensionaler** K-Vektorraum.

Satz. *Eine Teilmenge $M \subseteq V$ ist genau dann Basis von V, wenn jedes $v \in V$ eine eindeutige Darstellung als Linearkombination von Elementen aus M besitzt.*

Beweis. 1. M ist genau dann linear unabhängig, wenn jedes $v \in V$ höchstens eine Darstellung als Linearkombination von Elementen aus M besitzt. (Details siehe Vorlesung).

2. M erzeugt genau dann V, wenn jedes $v \in V$ mindestens eine Darstellung als Linearkombination von Elementen aus M besitzt. (klar) □

Folgerung. *Ist K ein endlicher Körper mit q Elementen und $\dim_K V = n$, so gilt $|V| = q^n$. Im Allgemeinen misst die Dimension die „Größe“ eines Vektorraums (nicht nur für endliche Körper).*

Definition a. Es sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V, $|B| = n$. Ist $v \in V$ und $v = \sum_{i=1}^{n} \lambda_i v_i$ die eindeutige Darstellung von v als Linearkombination der Basisvektoren, dann werden die (eindeutigen) Koeffizienten $\lambda_1, \ldots, \lambda_n$ die *Koordinaten* von v bzgl. B genannt.

Definition b. Es seien $v_1, ..., v_n \in V$. Das n-Tupel $\mathcal{B} = (v_1, ..., v_n)$ heißt *geordnete Basis von V*, wenn $\{v_1, ..., v_n\}$ eine Basis von V ist und $v_1, ..., v_n$ paarweise verschieden sind. In diesem Fall definieren wir die *Koordinatenabbildung* bzgl. $\mathcal{B}$ als

$$\kappa_{\mathcal{B}} : V \to K^n, v \mapsto (\lambda_1, \ldots, \lambda_n),$$

wobei $v = \sum_{i=1}^{n} \lambda_i v_i$. Das Bild $\kappa_{\mathcal{B}}(v)$ heißt der *Koordinatenvektor* von v bzgl. $\mathcal{B}$.

Bemerkung.

(i) Jeder endlich-dimensionale Vektorraum besitzt eine geordnete Basis, also auch eine Koordinatenabbildung.

(ii) Koordinatenabbildungen sind stets bijektiv.

Beispiel.

(i) Betrachte $\mathbb{C}$ als $\mathbb{R}$-Vektorraum mit der geordneten Basis $\mathcal{B} = (1, i)$. Die Koordinatenabbildung zu $\mathcal{B}$ lautet:

$$\kappa_B : \mathbb{C} \to \mathbb{R}^2, a + bi \mapsto \begin{pmatrix} a \\ b \end{pmatrix}.$$

(Daher kommt die Interpretation von $\mathbb{C}$ als „komplexe Ebene“.)

(ii) Betrachte den $\mathbb{R}$-Vektorraum $\mathbb{R}^{2\times 2}$ mit der geordneten Basis

$$B = (E_{11}, E_{12}, E_{21}, E_{22}) = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Die Koordinatenabbildung zu $\mathcal{B}$ lautet:

$$\kappa_B : \mathbb{R}^{2\times 2} \to \mathbb{R}^4, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

(iii) Betrachte den $\mathbb{R}$-Vektorraum $\mathbb{R}^2$ mit der geordneten Basis

$$B = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right).$$

Die Koordinatenabbildung zu $\mathcal{B}$ lautet:

$$\kappa_B : \mathbb{R}^2 \to \mathbb{R}^2, \begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} \frac{a+b}{2} \\ \frac{a-b}{2} \end{pmatrix}.$$

Übung: Man prüfe das nach!

6.2.8 Der Lösungsraum eines homogenen LGS

Es seien K ein Körper und $A \in K^{m\times n}$. Bekanntlich ist $\mathbb{L}(A,0)$ ein Unterraum von K^n. Wir wollen Basis und Dimension dieses *Lösungsraumes* bestimmen.

Satz. *Ist die Matrix A in Normalform, also*

$$A = \left(\begin{array}{cccc|ccc} 1 & 0 & \dots & 0 & c_{1(r+1)} & \dots & c_{1n} \\ 0 & 1 & & 0 & c_{2(r+1)} & \dots & c_{2n} \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & c_{r(r+1)} & \dots & c_{rn} \\ 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & 0 & \dots & 0 \end{array}\right),$$

so bilden die Spalten der Matrix

$$L := \begin{pmatrix} c_{1(r+1)} & \dots & c_{1n} \\ c_{2(r+1)} & \dots & c_{2n} \\ \vdots & \ddots & \vdots \\ c_{r(r+1)} & \dots & c_{rn} \\ -1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & -1 \end{pmatrix}$$

eine Basis von $\mathbb{L}(A,0)$. Insbesondere gilt $\mathbb{L}(A,0) = \mathrm{SR}(L)$ und $\dim \mathbb{L}(A,0) = n-r$.

Beweis. Es seien $v_{r+1},\dots,v_n$ die Spalten von L. Wegen $A\cdot L = 0$ (nachprüfen!) gilt $v_{r+1},\dots,v_n \in \mathbb{L}(A,0)$. An der Form von L erkennt man, dass $B := \{v_{r+1},\dots,v_n\}$ linear unabhängig ist (Teil (ii) von Übung 6.2.3a). und $|B| = n-r$. Wir zeigen nun, dass für ein beliebiges $w \in \mathbb{L}(A,0)\setminus B$ die Menge $B\cup\{w\}$ linear abhängig ist; dann ist gezeigt, dass B eine maximal linear unabhängige Teilmenge von $\mathbb{L}(A,0)$, also Basis von $\mathbb{L}(A,0)$ ist (Satz 6.2.4). Sei also

$$w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{L}(A,0)\setminus B, \quad w' := w + \mathrm{w}_{r+1}v_{r+1} + \ldots + w_n v_n \in \mathbb{L}(A,0).$$

Wir zeigen $w' = 0$; dann folgt, dass $B \cup \{w\}$ linear abhängig ist. Man rechnet nach, dass für geeignete $w'_1, \ldots, w'_r \in K$:

$$w' = \begin{pmatrix} w'_1 \\ \vdots \\ w'_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in K^n, \quad Aw' = \begin{pmatrix} w'_1 \\ \vdots \\ w'_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in K^m.$$

Aus $Aw' = 0$ folgt also wie gewünscht $w' = 0$. □

Folgerung. *Alle Zeilenstufenformen von A haben dieselbe Stufenzahl r und es gilt*

$$\dim \mathbb{L}(A, 0) = n - r.$$

Wir sprechen im folgenden auch einfach von der Stufenzahl von A *(anstatt von der Stufenzahl der Zeilenstufenform von A). Sie ist genau dann gleich n, wenn $Ax = 0$ nur trivial lösbar ist.*

Beispiel. Es sei $A = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$. Dann bilden die Spalten von $L = \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}$ einen Basis von $\mathbb{L}(A, 0)$. Insbesondere gilt $\mathbb{L}(A, 0) = \mathrm{SR}(L)$ und $\dim \mathbb{L}(A, 0) = 3 - 1 = 2$.

Achtung: Hat man, um A auf Normalform zu bringen, Spaltenvertauschungen gemacht, so muss man diese in Form von Zeilenvertauschungen in der Basis von $\mathbb{L}(A, 0)$ wieder rückgängig machen.

Beispiel a. Es sei

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

Dann bilden die Spalten von

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

eine Basis von $\mathbb{L}_0(A)$.

6.3 Unterräume des K^n und $K^{1\times m}$

6.3.1 Fundamentalräume und Rang

Es seien $A, A' \in K^{m\times n}$.

Definition. Mit *Rang von* A, geschr. $\operatorname{Rg} A$, bezeichnen wir Stufenzahl einer (jeder) Zeilenstufenform von A. Es seien $s_1, \ldots, s_n$ die Spalten und $z_1, \ldots, z_m$ die Zeilen von A. Die folgenden vier Unterräume heißen die *Fundamentalräume* von A:

$\mathrm{SR}(A) := \langle s_1, \ldots, s_n\rangle$	Spaltenraum
$\mathrm{ZR}(A) := \langle z_1, \ldots, z_m\rangle$	Zeilenraum
$\mathbb{L}_0(A) := \{x \in K^n \mid Ax = 0\}$	Nullraum
$\mathbb{L}^0(A) := \{y \in K^{1\times m} \mid yA = 0\}$	

Bemerkung. Es gelten

$$\mathrm{SR}(A) \leq K^m, \qquad \mathbb{L}_0(A) \leq K^n,$$
$$\mathrm{ZR}(A) \leq K^{1\times n}, \qquad \mathbb{L}^0(A) \leq K^{1\times m}.$$

Weiter lässt sich jeder Unterraum $U \leq K^m$ als Spaltenraum einer $m \times l$-Matrix über K schreiben, wobei $l = \dim_K U$. Ebenso lässt sich jeder Unterraum $U \leq K^{1\times n}$ als Zeilenraum einer $l \times n$-Matrix über K schreiben, wobei $l = \dim_K U$.

Beweis. Man wählt eine Basis von U aus und trägt die Basisvektoren in die Spalten (bzw. Zeilen) von A ein. □

Übung a. Eine quadratische Matrix $A \in K^{n\times n}$ ist genau dann regulär, wenn $\operatorname{Rg} A = n$.

Frage. Zu den vier Fundamentalräumen stellen sich folgende Fragen: Welche Zusammenhänge bestehen zwischen ihnen? Wie bestimmt Basis und Dimension? Wie hängen die Dimensionen mit dem Rang zusammen? Wie testet man für ein gegebenes Element aus $K^m, K^n, K^{1\times n}$ bzw. $K^{1\times m}$ in dem entsprechenden Fundamentalraum liegt? Läßt sich jeder Unterraum von K^n als Nullraum einer Matrix schreiben?

Einige Antworten kennen wir schon: Basis und Dimension von $\mathbb{L}_0(A)$ wurden z.B. in (6.2.8) berechnet. Die Tests lauten: $b \in \mathrm{SR}(A) \Leftrightarrow Ax = b$ lösbar, $x \in \mathbb{L}_0(A) \Leftrightarrow Ax = 0$, $c \in \mathrm{ZR}(A) \Leftrightarrow xA = c$ lösbar, $y \in \mathbb{L}^0(A) \Leftrightarrow yA = 0$.

Beispiel. Es seien A in Zeilenstufenform, $r = \operatorname{Rg} A$, und $1 \leq k_1 < \ldots < k_r \leq n$ die Stufenindizes. Da die Zeilen linear unabhängig sind (Teil (i) von Übung 6.2.3a), und $\mathrm{ZR}(A)$ erzeugen, gilt $\dim \mathrm{ZR}(A) = r$. Wegen $\mathrm{SR}(A) \leq \langle e_1, \ldots, e_r\rangle$ gilt $\dim \mathrm{SR}(A) \leq \dim\langle e_1, \ldots, e_r\rangle = r$ (Folgerung 6.2.6). Da das Tupel $(s_{k_1}, \ldots, s_{k_r})$ linear unabhängig ist (Teil (ii) von Übung 6.2.3a), gilt $\dim \mathrm{SR}(A) \geq r$. Zusammen also $\dim \mathrm{ZR}(A) = \dim \mathrm{SR}(A) = \operatorname{Rg} A$.

In untenstehendem Lemma wird gezeigt, dass elementare Zeilentransformationen an einer Matrix folgende Dinge nicht ändern: den Nullraum, den Zeilenraum, die linearen (Un)abhängigkeit von Spalten, den Zeilenrang, den Spaltenrang.

Lemma. *Falls $A \rightsquigarrow A'$ (elementare Zeilentransformationen) gelten:*

(i) $\mathbb{L}_0(A) = \mathbb{L}_0(A')$,

(ii) *Sind $s_1, \ldots, s_n$ die Spalten von A, $s'_1, \ldots, s'_n$ die Spalten von A', und $1 \leq i_1, \ldots, i_l \leq n$, dann:*

$$\{s_{i_1}, \ldots, s_{i_l}\} \text{ linear unabhängig} \Leftrightarrow \{s'_{i_1}, \ldots, s'_{i_l}\} \text{ linear unabhängig,}$$

(iii) $\dim \mathrm{SR}(A) = \dim \mathrm{SR}(A')$,

(iv) $\mathrm{ZR}(A) = \mathrm{ZR}(A')$,

(v) $\dim \mathrm{ZR}(A) = \dim \mathrm{ZR}(A')$.

Beweis. (i) ist bekannt aus Satz 5.2.1.
(ii) Durch Herausstreichen von Spalten aus A und A' (und aus dem ganzen Prozess $A \rightsquigarrow A'$) können wir oBdA annehmen, dass $l = n$ und $i_j = j$ für $j = 1, \ldots n$. Laut Teil (vi) von Beispiel 6.2.3 ist die Behauptung äquivalent zu: $Ax = 0$ nur trivial lösbar $\Leftrightarrow A'x = 0$ nur trivial lösbar. Das folgt wiederum aus (i).
(iii) folgt aus (ii) und der Bemerkung.
(iv) und (v) als Übung. □

Übung b. Man mache sich klar, dass elementare Zeilentransformationen an einer Matrix folgende Dinge ändern können: den Spaltenraum, den Raum $\mathbb{L}^0$, die lineare (Un)abhängigkeit von Zeilen.

Satz. *Es gilt stets:*

(i) $\mathrm{Rg}(A) = \dim \mathrm{SR}(A) = \dim \mathrm{ZR}(A)$.

(ii) $\operatorname{Rg} A + \dim \mathbb{L}_0(A) = n$.

Beweis. (i) ist das Beispiel und das Lemma. (ii) wurde in Folgerung 6.2.8 gezeigt. □

Übung c. Man zeige, dass $\operatorname{Rg} A$ genau die Maximalzahl linear unabhängiger Spalten (Zeilen) von A ist. Weiter folgere man: $\operatorname{Rg} A = \operatorname{Rg} A^t$.

Übung d. Man zeige die Korrektheit folgender Methode zur Basisauswahl: Sei $M = \{v_1, \ldots, v_l\} \subseteq K^m$, $U = \langle M \rangle$. Trage $v_1, \ldots, v_l$ in die Spalten einer $m \times l$-Matrix ein und bringe diese auf Zeilenstufenform. Seien $k_1, \ldots, k_r$ die Stufenindizes. Dann ist $\{v_{k_1}, \ldots, v_{k_r}\} \subseteq M$ eine Basis von U.

Übung e. Es seien A, A' beide in Zeilenstufenform und $A \rightsquigarrow A'$. Haben A und A' dann stets identische Stufenindizes?

6.3.2 Nullraum vs. Spaltenraum

Es sei $U \leq K^n$. In 6.3.1 wurde gefragt, ob sich U als Spaltenraum einer Matrix B und als Nullraum einer Matrix A schreiben lässt. Ersteres kommt der Angabe eines Erzeugendensystems gleich (wenn B dabei minimale Spaltenzahl hat, sogar der Angabe einer Basis) und ist offensichtlich möglich. Letzteres stellt eine Beschreibung von U durch *definierende Gleichungen* dar. Da jede Zeile von A eine Gleichung darstellt, ist eine minimale Zeilenzahl gewünscht. Beide Schreibweisen haben ihre Vor- und Nachteile, etwa beim Test $c \in U$ für gegebenes $c \in K^n$.

Beispiel a. Es sei E eine Ebene im euklidischen Raum $\mathbb{R}^3$. Die Schreibweise $E = \operatorname{SR}(B)$ mit $B = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \\ u_3 & v_3 \end{pmatrix}$ ist die *Parameterform* von E, wobei u, v die Richtungsvektoren sind. Die Schreibweise $E = \mathbb{L}_0(A)$ mit $A = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}$ ist die *Hesse'sche Normalenform* von E, wobei $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ der Normalenvektor ist.

Dieser Abschnitt zeigt, wie sich ein gegebener Spaltenraum als ein Nullraum schreiben lässt, womit dann beide Ausgangsfragen positiv beantwortet sind. Es seien $A \in K^{m \times n}$ und $B \in K^{n \times l}$.

Lemma.

(i) $\operatorname{SR}(B) \subseteq \mathbb{L}_0(A) \Leftrightarrow AB = 0$.

(ii) $\operatorname{SR}(B) = \mathbb{L}_0(A) \Leftrightarrow AB = 0 \wedge \operatorname{Rg} A + \operatorname{Rg} B = n$.

Beweis. (i) Offensichtlich gilt $AB = 0$ genau dann, wenn jede Spalte von B in $\mathbb{L}_0(A)$ liegt. Da $\mathbb{L}_0(A)$ abgeschlossen unter Linearkombinationen ist (es ist ein Unterraum), bedeutet letzteres gerade $\mathrm{SR}(B) \subseteq \mathbb{L}_0(A)$.

(ii) Es gilt $\dim \mathrm{SR}(B) = \mathrm{Rg}\, B$, und nach Satz 6.3.1 gilt $\dim \mathbb{L}_0(A) = n - \mathrm{Rg}\, A$. Also haben $\mathrm{SR}(B)$ und $\mathbb{L}_0(A)$ genau dann gleiche Dimension, wenn $\mathrm{Rg}\, A + \mathrm{Rg}\, B = n$ ist. □

Satz. $\mathrm{SR}(B) = \mathbb{L}_0(A) \Leftrightarrow \mathbb{L}_0(B^t) = \mathrm{SR}(A^t)$.

Beweis. Durch zweimalige Anwendung des Lemmas folgt:

$$\begin{aligned} \mathrm{SR}(B) = \mathbb{L}_0(A) &\Leftrightarrow AB = 0 \wedge \mathrm{Rg}\, A + \mathrm{Rg}\, B = n \\ &\Leftrightarrow B^t A^t = 0 \wedge \mathrm{Rg}\, A^t + \mathrm{Rg}\, B^t = n \\ &\Leftrightarrow \mathrm{SR}(A^t) = \mathbb{L}_0(B^t). \end{aligned}$$

□

Beispiel b. Für $B = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ ist A gesucht mit $\mathbb{L}_0(A) = \mathrm{SR}(B)$. Nach dem Satz ist das äquivalent zu $\mathbb{L}_o(B^t) = \mathrm{SR}(A^t)$. Nun kann A^t mit dem Verfahren aus 6.2.8 bestimmt werden. Als Ergebnis erhält man:

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

Übung. Es sei $U \leq K^n$ und $\dim U = d$. Man zeige, dass U Nullraum einer Matrix $A \in K^{(n-d)\times n}$ ist.

6.3.3 Anwendung: Lineare Codes

Ein Sender schickt Bitfolgen über einen Kanal, der evtl. fehlerbehaftet ist, zu einem Empfänger. Über den Kanal werden folgende Annahmen gemacht:

(i) Es gehen keine Bits verloren.

(ii) Die Fehlerwahrscheinlichkeit für ein einzelnes Bit, bei der Übertragung zu kippen, ist $< 1/2$.

Um Fehler erkennen bzw. sogar korrigieren zu können, wird nach folgendem Schema vorgegangen ($n > k$):

Nachricht		Codewort	Kanal	Empfangswort		Nachricht
$\mathbb{F}_2^k$	$\to$	$\mathbb{F}_2^n$	$\to$	$\mathbb{F}_2^n$	$\to$	$\mathbb{F}_2^k$

Die Bitfolge wird also in Nachrichten fester Länge (k bit) zerlegt, und jede Nachricht als ein Codewort (n bit) codiert. Die Menge C aller möglichen Codewörter ist eine echte Teilmenge von $\mathbb{F}_2^n$. Ist das Empfangswort kein Codewort, so kann der Empfänger mit Sicherheit davon ausgehen, dass ein Fehler bei der Übertragung stattgefunden hat (und evtl. eine erneute Sendung anfordern). Übertragungsfehler, bei denen ein Codewort in ein anderes übergeht, können allerdings nicht erkannt werden. Die Idee ist nun, C so zu wählen, dass sich verschiedene Codewörter an hinreichend vielen Stellen unterscheiden. Dadurch wird es unwahrscheinlich, dass ein Codewort durch einen Übertragungsfehler in ein anderes Codewort übergeht.

Definition a. Ein Unterraum $C \leq \mathbb{F}_2^n$ heißt *(binärer) linearer Code* der Länge n. Die Elemente von C heißen *Codewörter*.

Bemerkung a. In einem Code C der Dimension k gibt es 2^k Codewörter.

Codierung. Der Sender schreibt den Code C als Spaltenraum $\mathrm{SR}(G)$ mit $G \in \mathbb{F}_2^{n\times k}$ und minimaler Spaltenzahl. Dann ist $\dim C = k$ und die Spalten von G sind linear unabhängig. Insbesondere ist $Gx = 0$ nur trivial lösbar (vgl. Bemerkung (6.2.2)). Nach Satz (5.5.2) folgt, dass die Abbildung

$$\varphi_G : \mathbb{F}_2^k \to \mathbb{F}_2^n, \quad v \mapsto Gv$$

injektiv ist. Das Bild von φ_G ist per Definition $\mathrm{SR}(G) = C$. Daher kann φ_G als Codierungsabbildung verwendet werden. Die Matrix G wird *Generatormatrix* von C genannt.

Dekodierung. Der Empfänger schreibt den Code C als Nullraum $\mathbb{L}_0(H)$ mit $H \in \mathbb{F}_2^{l\times n}$ und minimaler Zeilenzahl. Dann ist $l = n - \dim C = n - k$. Zur Prüfung des Empfangswortes $w \in \mathbb{F}_2^n$ berechnet der Empfänger $Hw \in \mathbb{F}_2^k$, das Prüfergebnis. Ist $Hw \neq 0$, so liegt mit Sicherheit ein Fehler vor, da w kein Codewort ist. Ist $Hw = 0$, so ist w ein Codewort und der Empfänger geht davon aus, dass kein Fehler vorliegt (was mit einer gewissen Wahrscheinlichkeit richtig ist). Die Matrix H wird *Kontrollmatrix* von C genannt.

Bemerkung b. Weder Generator- noch Kontrollmatrix sind eindeutig durch den Code C definiert. Die Größen beider Matrizen sind aber durch Länge und Dimension von C bestimmt.

Einen Übertragungsfehler auf einem Wort $c \in \mathbb{F}_2^n$ stellen wir uns als Addition im Vektorraum $\mathbb{F}_2^n$ eines *Fehlervektors* $\epsilon \in \mathbb{F}_2^n$ vor. Da die Addition über $\mathbb{F}_2$ der XOR-Verknüpfung entspricht, verändert die Addition von ϵ zu c genau die Einträge von c, die an einer Position stehen, an der ϵ eine 1 enthält.

Bemerkung c. Es bleiben genau die Übertragungsfehler $\epsilon \in \mathbb{F}_2^n$ unerkannt, für die $H\epsilon = 0$ ist, d.h. die selbst Codewörter sind.

Beweis. Sei $c \in C$ das gesendete Codewort und $w = c+\epsilon$ das Empfangswort. Der Fehler ϵ bleibt unentdeckt, wenn das Prüfergebnis $Hw = H(c+\epsilon) = Hc + H\epsilon = 0$ ist. Wegen $c \in C$ ist $Hc = 0$, also $Hw = 0$ genau dann, wenn $H\epsilon = 0$. □

Definition b. Wir nennnen einen Fehlervektor $\epsilon \in \mathbb{F}_2^n$ *einfach*, wenn ϵ genau einen 1-Eintrag enthält und sonst nur Nullen.

Satz.

(i) Sind alle Spalten von H ungleich 0, so werden alle 1-fachen Übertragungsfehler erkannt.

(ii) Sind alle Spalten von H ungleich 0 und paarweise verschieden, so können alle 1-fachen Übertragungsfehler vom Empfänger korrigiert werden.

Beweis. Es sei ϵ ein 1-facher Fehlervektor, d.h. $\epsilon = e_i$ für einen Einheitsvektor e_i. Das Prüfergebnis lautet dann $H(c+\epsilon) = Hc + H\epsilon = He_i = i$-te Spalte von H, und zwar unabhängig vom Codewort $c \in C$. Sind alle Spalten von H ungleich 0, so ist also $He_i \neq 0$ für alle i, d.h. alle 1-fachen Fehler werden erkannt.

Sind die Spalten von H ausserdem paarweise verschieden, so erlaubt das Prüfergebnis He_i einen eindeutigen Rückschluß auf die Stelle i, durch Vergleich des Prüfergebnisses mit allen Spalten von H. □

Beispiel a. (3-facher Wiederholungscode) $C = \{\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\} \leq \mathbb{F}_2^3$ ist ein Code der Länge 3 mit Dimension 1. Eine Generator- und Kontrollmatrix sind z.B.

$$G = \begin{pmatrix}1\\1\\1\end{pmatrix}, \quad H = \begin{pmatrix}1 & 1 & 0\\1 & 0 & 1\end{pmatrix}.$$

An H erkennt man, dass alle 1-fachen Fehler korrigiert werden können. Z.B. für $\epsilon = e_3$:

$$v = (1) \mapsto c = \begin{pmatrix}1\\1\\1\end{pmatrix} \mapsto w = c+\epsilon = \begin{pmatrix}1\\1\\0\end{pmatrix} \mapsto Hw = \begin{pmatrix}0\\1\end{pmatrix} = \text{3-te Spalte von } H.$$

Beispiel b. (Konstruktion von Codes) Wichtige Eigenschaften des Codes sind nach dem Satz an der Kontrollmatrix zu erkennen. Aus diesem Grund konstruieren wir nun einen Code indirekt über seine Kontrollmatrix. Angenommen, die Zahl $n-k$ (also die Zeilenzahl von H) sei fest. Für einen effizienten Code ist die Zahl der Codewörter, 2^k, zu maximieren. Wegen $k = n-(n-k)$ ist also n, die Spaltenzahl von H, zu maximieren.

Wollen wir etwa einen möglichst effizienten Code mit $n-k=3$ konstruieren, der alle 1-fachen Fehler korrigieren kann, so müssen wir in die Spalten von H genau die verschiedenen Spaltenvektoren aus $\mathbb{F}_2^3$ ungleich 0 eintragen. Das sind die Binärzahlen 1 bis 7, also

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

Der Code $C = \mathbb{L}_0(H)$ heißt *Hamming-Code*. Eine Generatormatrix ist z.B.

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

Die Rechnung zur Bestimmung von G wurde bereits in Beispiel (6.2.8) gemacht.

6.4 Lineare Abbildungen

6.4.1 Homomorphismen

Homomorphismen sind „strukturerhaltende Abbildungen“.

Definition a. Es seien $(G, \circ)$ und $(H, \bullet)$ zwei Gruppen. Eine Abbildung $\varphi : G \to H$ heißt *Gruppen-Homomorphismus*, wenn für alle $x, y \in G$ gilt:

$$\varphi(x \circ y) = \varphi(x) \bullet \varphi(y).$$

Definition b. Seien R und S zwei Ringe. Eine Abbildung $\varphi : R \to S$ heißt *Ring-Homomorphismus*, wenn gelten:

(i) φ ist Gruppenhomomorphismus $(R, +) \to (S, +)$

(ii) $\varphi(xy) = \varphi(x)\varphi(y)$ für alle $x, y \in R$

(iii) $\varphi(1) = 1$

Ein „Körper-Homomorphismus“ ist schlicht ein Ringhomomorphismus zwischen zwei Körpern (jeder Körper ist ein kommutativer Ring).

Definition c. Ein Homomorphismus φ zwischen zwei Strukturen heißt *Monomorphismus* wenn er injektiv ist, *Epimorphismus* wenn er surjektiv ist, und *Isomorphismus* wenn er bijektiv ist. Existiert ein Isomorphismus $\varphi : A \to B$ dann heißen A und B *isomorph*, geschr. $A \cong B$.

Bemerkung. Oft untersucht man Strukturen (Gruppen, Ringe, Vektorräume, etc.) nur “bis auf Isomorphie”, d.h. man unterscheidet nicht zwischen isomorphen Strukturen. Dies ist dadurch begründet, dass isomorphe Strukturen im Prinzip durch “Umbenennung der Elemente” (vermittelt durch einen Isomorphismus) auseinander hervorgehen.

Beispiel a.

(i) Für jede Untergruppe U von $(G, \circ)$ ist die Abbildung

$$\varphi : (U, \circ) \to (G, \circ), \quad x \mapsto x$$

ein Gruppen-Monomorphismus, z.B. $(\mathbb{Z}, +) \to (\mathbb{R}, +), x \mapsto x$.

(ii) $\exp : (\mathbb{R}, +) \to (\mathbb{R}_{>0}, \cdot), x \mapsto \exp(x) = e^x$ ist Gruppen-Isomorphismus ($e^{x+y} = e^x \cdot e^y$ für alle $x, y \in \mathbb{R}$). Daher gilt $(\mathbb{R}, +) \cong (\mathbb{R}_{>0}, \cdot)$.

(iii) Es sei $(G, \cdot)$ Gruppe, $n \in \mathbb{N}$. Für jedes $1 \leq i \leq n$ ist

$$\epsilon_i : G \to G^n, \quad x \mapsto (e, ..., e, \underbrace{x}_{i\text{-te Position}}, e, ..., e)$$

ein Gruppen-Monomorphismus und

$$\rho_i : G^n \to G, \quad (x_1, ..., x_n) \mapsto x_i$$

ein Gruppen-Epimorphismus. Es gilt stets $\rho_i \circ \epsilon_i = \mathrm{id}_G$.

(iv) Gegeben seien ein Körper K, eine abelsche Gruppe $(V, +)$ und eine Abbildung

$$\cdot : K \times V \to V, (\lambda, v) \mapsto \lambda \cdot v = \lambda v.$$

Das Vektorraumaxiom $(V2)$ besagt, dass für jedes $\lambda \in K$ die Abbildung

$$f_\lambda : (V, +) \to (V, +), v \mapsto \lambda v$$

ein Gruppen-Homomorphismus ist.
Das Vektorraumaxiom $(V1)$ besagt, dass für jedes $v \in K$ die Abbildung

$$g_v : (K, +) \to (V, +), v \mapsto \lambda v$$

ein Gruppen-Homomorphismus ist.

(v) Es gibt n^n Abbildungen $\mathbb{Z}_n \to \mathbb{Z}_n$, aber nur n Gruppen-Homomorphismen $(\mathbb{Z}_n, +) \to (\mathbb{Z}_n, +)$.

Übung a. Für jeden Gruppenhomomorphismus $\varphi : G \to H$ gelten:

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

(ii) $\varphi(g^{-1}) = \varphi(g)^{-1}$ für alle $g \in G$,

(iii) φ ist genau dann injektiv, wenn für alle $g \in G$ gilt: $\varphi(g) = e_H \Rightarrow g = e_G$.

Beispiel b.

(i) Für jedes $n \in \mathbb{N}$ ist die Abbildung $\varphi : \mathbb{Z} \to \mathbb{Z}_n, a \mapsto \overline{a}$ ein Ring-Epimorphismus.

(ii) Für jedes $a \in \mathbb{Z}$ mit $a \neq 1$ ist die Abbildung

$$m_a : \mathbb{Z} \to \mathbb{Z}, x \mapsto a \cdot x$$

kein Ring-Homomorphismus, da $m_a(1) = a \neq 1$.

(iii) $\varphi : \mathbb{Z} \to \mathbb{Z}, a \mapsto a^3$ ist kein Ring-Homomorphismus, weil z.B. $\varphi(1+1) = 8 \neq 2 = \varphi(1) + \varphi(1)$.

(iv) $\varphi : \mathbb{Z}_3 \to \mathbb{Z}_3, a \mapsto a^3$ ist ein Ring-Isomorphismus.

(v) $\mathbb{F}_p \to \mathbb{F}_p, x \mapsto x^p$ ist ein Ring-Isomorphismus.

(vi) $\mathbb{Q} \to \mathbb{R}, x \mapsto x$ ist ein Ring-Monomorphismus.

(vii) $\mathbb{C} \to \mathbb{C}, z \mapsto \overline{z}$ ist ein Ring-Isomorphismus (komplexe Konjugation: $\overline{a + bi} = a - bi$).

Übung b. Jeder Körperhomomorphismus ist injektiv.

6.4.2 Lineare Abbildungen

Definition. Es seien V, W zwei K-Vektorräume.

(i) Eine Abbildung $\varphi : V \to W$ heißt *lineare Abbildung* oder *Vektorraum-Homomorphismus*, falls für alle $v, v' \in V$ und alle $\lambda \in K$ gelten:

(1) $\varphi(v + v') = \varphi(v) + \varphi(v')$,

(1) $\varphi(\lambda v) = \lambda\varphi(v)$.

Die Menge aller Homomorphismen $V \to W$ wird mit $\mathrm{Hom}(V, W)$ bezeichnet.

(ii) Ein Vektorraum-Homomorphismus $\varphi : V \to V$ heißt *Endomorphismus* von V. Die Menge $\mathrm{Hom}(V, V)$ aller Endomorphismen von V wird mit $\mathrm{End}(V)$ bezeichnet.

Bemerkung.

(i) Für jeden Vektorraum-Homomorphismus $\varphi : V \to W$ gilt $\varphi(\mathfrak{o}) = \mathfrak{o}$.

(ii) Im folgenden meinen wir mit *Homomorphismus* stets einen Vektorraum-Homomorphismus.

Beispiel.

(i) Lineare Abbildungen $\mathbb{R}^2 \to \mathbb{R}^2$ sind z.B.: Drehungen um $\mathfrak{o}$, Spiegelungen an Geraden durch $\mathfrak{o}$, Projektionen auf Koordinatenachsen. Nicht linear sind dagegen Translationen (Verschiebungen).

(ii) Betrachte die Abbildungen $\varphi_1, \ldots, \varphi_4 : \mathbb{R}^3 \to \mathbb{R}^2$ mit

$$\varphi_1 : \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \begin{pmatrix} a \\ b \end{pmatrix}, \qquad \varphi_2 : \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \begin{pmatrix} 1+a \\ b \end{pmatrix},$$

$$\varphi_3 : \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \begin{pmatrix} a+c \\ b \end{pmatrix}, \qquad \varphi_4 : \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \begin{pmatrix} a \\ b^2 \end{pmatrix}.$$

Davon sind φ_1, φ_3 linear, φ_2, φ_4 dagegen nicht. Die Abbildung φ_1 ist gerade die Projektion des $\mathbb{R}^3$ auf die e_1-e_2-Ebene.

(iii) Die Transpositionsabbildung

$$(\cdot)^t : K^{m\times n} \to K^{n\times m}, A \mapsto A^t$$

ist linear. Spezialfälle sind:

$$\begin{aligned}(\cdot)^t &: K^{1\times n} \to K^n, z \mapsto z^t\\ (\cdot)^t &: K^m \to K^{1\times m}, s \mapsto s^t\end{aligned}$$

(iv) Für jede Matrix $A \in K^{m\times n}$ ist die Abbildung $\varphi_A : K^n \to K^m, x \mapsto Ax$ linear.

(v) Koordinatenabbildungen endlich-dimensionaler Vektorräume ($\kappa_B : V \to K^n$) sind stets linear.

(vi) Betrachte die $\mathbb{R}$-Vektorräume $\mathrm{Abb}(\mathbb{R},\mathbb{R})$ (reelle Funktionen) und $\mathbb{R}$. Für jedes fest gewählte $a \in \mathbb{R}$ ist die Abbildung

$$\epsilon_a : \mathrm{Abb}(\mathbb{R},\mathbb{R}) \to \mathbb{R}, \quad f \mapsto f(a)$$

linear und wird *Einsetzungshomomorphismus* genannt.

(vii) Betrachte die $\mathbb{R}$-Vektorräume $C^\infty(\mathbb{R})$ (beliebig oft stetig differenzierbare reelle Funktionen) und $\mathbb{R}$. Die *Ableitungsabbildung*

$$\mathrm{diff} : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}), \quad f \mapsto f'$$

ist linear. (Das ist eine bekannte Ableitungsregel aus der Analysis.)

Übung. Es sei $f : V \to W$.

(i) f ist genau dann linear, wenn für alle $n \in \mathbb{N}, \lambda_i \in K, v_i \in V$ gilt: $f(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i f(v_i)$.

Sei nun f linear.

(ii) Ist $g : W \to U$ linear, so auch $g \circ f : V \to U$.

(iii) Ist f bijektiv (Isomorphismus), so ist auch f^{-1} Isomorphismus.

6.4.3 Kern und Bild

Definition. Es sei $\varphi \in \mathrm{Hom}(V,W)$.

(i) $\mathrm{Kern}\,\varphi := \{v \in V \mid \varphi(v) = \mathfrak{o}\}$ heißt *Kern von* φ.

(ii) $\mathrm{Bild}\,\varphi := \varphi(V) = \{\varphi(v) \mid v \in V\}$ heißt *Bild von* φ.

(Die Bezeichnung kommt von engl. *kernel* und *image*.)

Bemerkung. Für jedes $\varphi \in \mathrm{Hom}(V, W)$ gilt:

(i) $\mathrm{Kern}\,\varphi \leq V$.

(ii) $\mathrm{Bild}\,\varphi \leq W$.

(iii) φ injektiv $\Leftrightarrow \mathrm{Kern}\,\varphi = \{\mathfrak{o}\}$.

(iv) φ surjektiv $\Leftrightarrow \mathrm{Bild}\,\varphi = W$.

(v) Für jedes $v \in V$ und $w = \varphi(v)$ gilt:

$$\varphi^{-1}(\{w\}) = v + \mathrm{Kern}\,\varphi.$$

Beweis. (siehe Vorlesung) □

Beispiel.

(i) Es sei $A \in K^{m\times n}$. Dann ist $\mathrm{Kern}\,\varphi_A = \mathbb{L}_0(A)$ und $\mathrm{Bild}\,\varphi_A = \mathrm{SR}(A)$. Die Bemerkung liefert bereits bekannte Aussagen: $\mathbb{L}_0(A) \leq K^n, \mathrm{SR}(A) \leq K^m$, φ_A injektiv $\Leftrightarrow Ax = 0$ nur trivial lösbar, φ_A surjektiv $\Leftrightarrow \mathrm{SR}(A) = K^m$, $\mathbb{L}(A, b) = s + \mathbb{L}_0(A)$ für jedes $s \in \mathbb{L}(A, b)$.

(ii) Es sei $A \in K^{m\times n}$. Für das Bild unter der Transpositonsabbildung gilt:

$$\mathrm{SR}(A)^t = \mathrm{ZR}(A^t), \quad \mathrm{ZR}(A)^t = \mathrm{SR}(A^t).$$

$$\mathbb{L}_0(A)^t = \mathbb{L}^0(A^t), \quad \mathbb{L}^0(A)^t = \mathbb{L}_0(A^t).$$

(iii) Der Kern der Projektion φ_1 aus Beispiel (6.4.2)(ii) ist die von $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ erzeugte Gerade. Das Bild ist ganz $\mathbb{R}^2$.

(iv) Der Kern des Einsetzungshomomorphismus ϵ_a aus Beispiel (6.4.2)(vi) besteht genau aus denjenigen reellen Funktionen, die bei a eine Nullstelle haben. Das Bild ist ganz $\mathbb{R}$, weil es zu jedem $x \in \mathbb{R}$ eine reelle Funktion gibt, die an der Stelle a den Wert x annimmt (z.B. die konstante Funktion mit dem Wert x).

(v) Der Kern der Ableitungsabbildung aus Beispiel (6.4.2)(vii) besteht genau aus den konstanten reellen Funktionen. (Der Kern ist 1-dimensional!) Das Bild ist ganz $C^\infty(\mathbb{R})$, weil jede stetige reelle Funktion eine Stammfunktion hat.

Zur Berechnung von Kern und Bild mit Hilfe von Koordinaten siehe Abschnitt (6.5.2).

Übung. Für jedes $\varphi \in \mathrm{Hom}(V, W)$ gilt:

(i) $U \leq V \Rightarrow \varphi(U) \leq W$.

(ii) $U \leq W \Rightarrow \varphi^{-1}(U) \leq V$.

(iii) $M \subseteq V \Rightarrow \varphi(\langle M \rangle) = \langle \varphi(M) \rangle$.

(iv) $M \subseteq V$ linear unabhängig, φ injektiv $\Rightarrow \varphi(M)$ linear unabhängig.

(v) $U \leq V \Rightarrow \dim \varphi(U) \leq \dim U$.

(vi) $U \leq V$ und φ injektiv $\Rightarrow \dim \varphi(U) = \dim U$.

(vii) $U \leq \mathrm{Bild}\, \varphi \Rightarrow \dim \varphi^{-1}(U) \geq \dim U$.

Beweis. Es gilt $\varphi(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i \varphi(v_i)$. Daraus lassen sich i)–iv) folgern.
Z.B. iv) Es seien M linear unabhängig und φ injektiv. Sei $\sum_{i=1}^n \lambda_i \varphi(v_i) = \mathfrak{o}_W$ eine lineare Abhängigkeit in $\varphi(M)$, d.h. $\lambda_i \in K$, $v_i \in M$ und $\varphi(v_i)$ paarweise verschieden. Dann sind auch die v_i paarweise verschieden (das ist klar für jede Abbildung φ) und $\varphi(\sum_{i=1}^n \lambda_i v_i) = \mathfrak{o}_W$. Da φ injektiv ist, folgt $\sum_{i=1}^n \lambda_i v_i = \mathfrak{o}_V$. Da M linear unabhängig ist, sind alle $\lambda_i = 0$. Damit ist gezeigt, dass $\varphi(M)$ linear unabhängig ist.
Z.B. v) Wähle Basis B von U. Nach iii) ist $\varphi(U) = \langle \varphi(B) \rangle$, also $\dim \varphi(U) \leq |\varphi(B)| \leq |B| = \dim U$. □

6.4.4 Existenz linearer Abbildungen

Es seien V, W zwei K-Vektorräume.

Frage. Es seien Vektoren $v_1, \ldots, v_n \in V$ und $w_1, \ldots, w_n \in W$ gegeben. Gibt es eine lineare Abbildung $V \to W$ mit $v_i \mapsto w_i$?

Satz a. *Es sei B eine Basis von V. Zu jedem $v \in B$ sei ein $w_v \in W$ gegeben. Dann existiert eine eindeutige lineare Abbildung $\varphi : V \to W$ mit $\varphi(v) = w_v$ für alle $v \in B$.*

Beweis. Eindeutigkeit: Sei $\varphi : V \to W$ linear mit $\varphi(v) = w_v$ für alle $v \in B$. Wir zeigen, dass damit φ schon auf ganz V festgelegt ist. Sei dazu $v \in V$ beliebig, etwa $v = \sum_{i=1}^n \lambda_i v_i$ mit $v_1, \ldots, v_n \in B$ paarweise verschieden und

$\lambda_i \in K \setminus \{0\}$. Nach Satz (6.2.7) sind die λ_i eindeutig bestimmt. Aus der Linearität von φ folgt:

$$\varphi(v) = \sum_{i=1}^{n} \lambda_i \varphi(v_i) = \sum_{i=1}^{n} \lambda_i w_{v_i}. \tag{6.1}$$

Wegen der Eindeutigkeit der λ_i ist damit auch $\varphi(v)$ eindeutig festgelegt.
Existenz: Benutzen wir die Gleichung (6.1) als Definition für ein $\varphi : V \to W$, so bleibt nur noch zu prüfen, dass das so definierte φ auch linear ist (Übung). □

Bemerkung.

(i) Man liest Satz a auch so: Jede Abildung $f : B \to W$ lässt sich eindeutig zu einer linearen Abbildung $\varphi : V \to W$ fortsetzen.

(ii) Die lineare Unabhängigkeit von B wird in Satz a nur bei der Eindeutigkeit gebraucht; die Tatsache, dass B Erzeugendensystem dagegen nur bei der Existenz. Somit gilt: Ist B linear unabhängig (statt Basis), so gibt es in Satz a mindestens ein solches φ (statt genau ein); ist B Erzeugendensystem, so gibt es in Satz a höchstens ein solches φ.

Beispiel. $V = \mathbb{R}^3, B = (e_1, e_2, e_3)$. Wähle $w_1 = e_2, w_2 = -e_1, w_3 = 0$. Nach Satz a gibt es genau einen Endomorphismus von $\mathbb{R}^3$ mit $e_i \mapsto w_i$ für $i = 1, 2, 3$. Frage: Was für eine Abbildung ist das?
Antwort: Sei φ die Projektion auf die e_1-e_2-Ebene gefolgt von einer 90°-Drehung um die e_3-Achse. Wir wissen aus vorherigen Beispielen, dass φ linear ist. Da φ auch $\varphi(e_i) = w_i$ für $i = 1, 2, 3$ erfüllt, muss es der gesuchte Endomorphismus sein (weil er eindeutig ist). Die Abbildungsvorschrift lautet:

$$\varphi : \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \begin{pmatrix} -b \\ a \\ 0 \end{pmatrix}.$$

Satz b. *Sei W ein beliebiger nicht-trivialer K-Vektorraum. Eine Teilmenge $B \subseteq V$ ist genau dann eine Basis von V, wenn sich jede Abbildung $f : B \to W$ eindeutig zu einer linearen Abbildung $\varphi : V \to W$ fortsetzen lässt.*

Beweis. Eine Richtung wurde bereits in Satz a gezeigt. Wir setzen nun voraus, jede Abbildung $f : B \to W$ lasse sich eindeutig zu einer linearen Abbildung $\varphi : V \to W$ fortsetzen, und folgern, dass B Basis ist.
Annahme: B ist linear abhängig, etwa $\sum_{i=1}^{n} \lambda_i v_i = \mathfrak{o}_V$ mit $n \in \mathbb{N}, v_1, \ldots, v_n \in B$ paarweise verschieden und $\lambda_i \in K \setminus \{0\}$. Nach Voraussetzung gibt es

$\varphi \in \mathrm{Hom}(V,W)$ mit $\varphi(v_1) \neq \mathfrak{o}_W$ und $\varphi(v_i) = \mathfrak{o}_W$ für $i = 2, \ldots, n$. Dann folgt $\mathfrak{o}_W = \varphi(\mathfrak{o}_V) = \varphi(\sum \lambda_i v_i) = \sum \lambda_i \varphi(v_i) = \lambda_1 \varphi(v_1) \neq \mathfrak{o}_W$. Da dies ein Widerspruch ist, ist die Annahme falsch, also B linear unabhängig.
Annahme: B ist keine Basis. Ergänze die linear unabhängige Menge B zu einer Basis B' von V und wähle ein $v \in B' \setminus B$. Dann existieren nach Satz a mindestens zwei verschiedene Fortsetzungen φ von f, nämlich ein φ mit $\varphi(v) = \mathfrak{o}_W$ und eins mit $\varphi(v) \neq \mathfrak{o}_W$. Da dies ein Widerspruch zur Voraussetzung ist, ist die Annahme falsch, also B eine Basis von V. □

Übung a. Es sei $\varphi : V \to W$ linear und surjektiv (Epimorphismus). Man zeige, dass eine lineare Abbildung $\psi : W \to V$ existiert mit $\varphi \circ \psi = \mathrm{id}_W$.
Hinweis: Übung 6.4.2 und Satz a.

Übung b. Es sei $\varphi : V \to W$ linear und injektiv (Monomorphismus). Man zeige, dass eine lineare Abbildung $\psi : W \to V$ existiert mit $\psi \circ \varphi = \mathrm{id}_V$.
Hinweis: Übung 6.4.2 und Satz a.

6.4.5 Monomorphismen und Epimorphismen

Es seien V, W zwei beliebige K-Vektorräume und $\varphi \in \mathrm{Hom}(V,W)$.

Satz. *Es sei B eine beliebige Basis von V.*

(i) Folgende Aussagen sind äquivalent:

1. *φ ist Monomorphismus.*
2. *Für jede Teilmenge $M \subseteq V$ gilt:*
   M linear unabhängig $\Rightarrow \varphi(M)$ linear unabhängig.
3. *$\varphi(B)$ ist linear unabhängig und $\varphi|_B$ ist injektiv.*

In diesem Fall gilt $\dim V \leq \dim W$.

(ii) Folgende Aussagen sind äquivalent:

1. *φ ist Epimorphismus.*
2. *Für jede Teilmenge $M \subseteq V$ gilt:*
   M erzeugt $V \Rightarrow \varphi(M)$ erzeugt W.
3. *$\varphi(B)$ ist Erzeugendensystem von W.*

In diesem Fall gilt $\dim V \geq \dim W$.

(iii) Folgende Aussagen sind äquivalent:

1. *φ ist Isomorphismus.*

2. *Für jede Teilmenge* $M \subseteq V$ *gilt:*
   M *Basis von* $V \Rightarrow \varphi(M)$ *Basis von* W.

3. $\varphi(B)$ *ist Basis von* W *und* $\varphi|_B$ *ist injektiv.*

In diesem Fall gilt $\dim V = \dim W$.

Beweis.

(i) 1. $\Rightarrow$ 2. wurde bereits in Übung 6.4.3iv gezeigt.
2. $\Rightarrow$ 1. Sei $\varphi(v) = \mathfrak{o}_W, v \in V$. Wir zeigen $v = \mathfrak{o}_V$. Dann ist nachgewiesen: φ ist injektiv, d.h. Monomorphismus. Wende 2. mit $M = \{v\}$ an. Da $\varphi(M) = \{\mathfrak{o}_W\}$ linear abhängig ist, ist auch $M = \{v\}$ linear abhängig, d.h. $v = \mathfrak{o}_V$.
3.$\Rightarrow$ 1. Sei $\varphi(v) = \mathfrak{o}_W, v \in V$. Wir zeigen $v = \mathfrak{o}_V$; dann ist φ Monomorphismus. Schreibe $v = \sum_{i=1}^{n} \lambda_i v_i$ mit $n \in \mathbb{N}$, $v_1, \ldots, v_n \in B$ paarweise verschieden und $\lambda_i \in K$. Da $\varphi|_B$ injektiv ist, sind auch $\varphi(v_1), \ldots, \varphi(v_n)$ paarweise verschieden. Da $\varphi(B)$ linear unabhängig ist und $\mathfrak{o}_W = \varphi(v) = \sum_{i=1}^{n} \lambda_i \varphi(v_i)$, sind alle $\lambda_i = 0$, also $v = \mathfrak{o}_V$.
1.$\wedge$2.$\Rightarrow$ 3. ist trivial.
Nach Übung 6.4.3vi ist für injektives φ: $\dim V = \dim \varphi(V) \leq \dim W$.

(ii) als Übung.

(iii) 1.$\Leftrightarrow$ 3. und 1.$\Rightarrow$ 2. folgen aus i) und ii).
2.$\Rightarrow$ 1. folgt aus i) und ii) zusammen mit Basisergänzung und Basisauswahl (Details als Übung).
Die Dimensionsgleichung folgt auch aus i) und ii).

□

Übung. Sei V endlich-dimensional.

(i) $\varphi|_B$ injektiv $\Leftrightarrow |\varphi(B)| = |B|$.

(ii) φ injektiv $\Leftrightarrow$ für alle $U \leq V$ gilt $\dim \varphi(U) = \dim U \Leftrightarrow \dim \varphi(V) = \dim V$.

Beispiel.

(i) Im $\mathbb{R}^2$ sind Drehungen um $\mathfrak{o}$ und Spiegelungen an Ursprungsgeraden stets Isomorphismen.

(ii) Die Projektion φ_1 aus Beispiel (6.4.2)(ii) ist Epimorphismus, aber kein Monomorphismus.

(iii) Die Codierungsabbildung $\varphi_G : \mathbb{F}_2^k \to \mathbb{F}_2^n$ (vgl. 6.3.3) ist ein Monomorphismus, aber kein Epimorphismus.

(iv) Koordinatenabbildungen ($\kappa_B : V \to K^n$) sind stets Isomorphismen.

(v) Mit φ Isomorphismus, ist auch φ^{-1} linear (Übung) und somit Isomorphismus.

(vi) Die Transpositionsabbildung

$$^t : K^{m\times n} \to K^{n\times m}, A \mapsto A^t$$

ist ein Isomorphismus.

(vii) Für $\varphi \in \mathrm{Hom}(V, W)$ gilt: φ Monomorphismus $\Leftrightarrow \varphi : V \to \varphi(V)$ Isomorphismus.

6.4.6 Endlich-dimensionale Vektorräume

Es seien V, W zwei endlich-dimensionale K-Vektorräume und $\varphi \in \mathrm{Hom}(V, W)$.

Definition.

$$\begin{aligned} \mathrm{Rg}\,\varphi &:= \dim(\mathrm{Bild}\,\varphi) && \textit{Rang von } \varphi. \\ \mathrm{Def}\,\varphi &:= \dim(\mathrm{Kern}\,\varphi) && \textit{Defekt von } \varphi. \end{aligned}$$

Beispiel. Die Projektion φ_1 aus Beispiel (6.4.2)(ii) hat den Rang 2, da das Bild gleich $\mathbb{R}^2$ ist, und den Defekt 1, da der Kern die Gerade $\langle e_3 \rangle$ ist.

Bemerkung a.

(i) $\mathrm{Rg}\,\varphi \leq \dim W$, und $\mathrm{Rg}\,\varphi = \dim W \Leftrightarrow \varphi$ Epimorphismus.

(ii) $\mathrm{Def}\,\varphi \leq \dim V$, und $\mathrm{Def}\,\varphi = 0 \Leftrightarrow \varphi$ Monomorphismus.

(iii) $\mathrm{Rg}\,\varphi \leq \dim V$ und $\mathrm{Rg}\,\varphi = \dim V \Leftrightarrow \varphi$ Monomorphismus.

Im Spezialfall $\dim V = \dim W = n$ ergibt sich:

$$\mathrm{Rg}\,\varphi = n \Leftrightarrow \mathrm{Def}\,\varphi = 0 \Leftrightarrow \varphi \text{ Isomorphismus},$$

d.h. die Begriffe Monomorphismus, Epimorphismus und Isomorphismus sind in diesem Fall äquivalent.

Beweis. (i) und (ii) sind unmittelbar klar aus den Definitionen der Begriffe. (iii): Es sei B eine beliebige Basis von V. Da $\varphi(V)$ von $\varphi(B)$ erzeugt wird, gilt $\dim \varphi(V) \leq |\varphi(B)|$ mit Gleichheit genau dann, wenn $\varphi(B)$ Basis von $\varphi(V)$ ist. Weiter ist $|\varphi(B)| \leq |B| = \dim V$ mit Gleichheit genau dann, wenn $\varphi|_B$ injektiv ist. Also $\dim \varphi(V) \leq |\varphi(B)| \leq |B| = \dim V$ mit

$$\begin{aligned}\dim \varphi(V) = \dim V &\Leftrightarrow \dim \varphi(V) = |\varphi(B)| \wedge |\varphi(B)| = |B| \\ &\Leftrightarrow \varphi(B) \text{ Basis von } \varphi(V) \wedge \varphi|_B \text{ injektiv.}\end{aligned}$$

Nach Satz 6.4.5 ist letzteres äquivalent zu $\varphi : V \to \varphi(V)$ Isomorphismus, bzw. zu $\varphi : V \to W$ Monomorphismus.

Im Fall $\dim V = \dim W = n$ ergibt sich:

$$\text{Def } \varphi = 0 \overset{(ii)}{\Leftrightarrow} \varphi \text{ Monomorphismus} \overset{(iii)}{\Leftrightarrow} \text{Rg } \varphi = n \overset{(i)}{\Leftrightarrow} \varphi \text{ Epimorphismus.}$$

□

Satz. $V \cong W \Leftrightarrow \dim V = \dim W$.

Beweis. $\Rightarrow$ wurde schon in Satz (6.4.5) (iii) gezeigt. Sei nun $\dim V = \dim W = n$. Wähle beliebige Basen $\{v_1, \ldots, v_n\}$ von V und $\{w_1, \ldots, w_n\}$ von W. Nach Satz (6.4.4) gibt es eine lineare Abbildung $\varphi : V \to W$ mit $v_i \mapsto w_i$ für $i = 1, \ldots, n$. Nach Satz (6.4.5) (iii) ist dieses φ ein Isomorphismus. □

Folgerung. *$V \cong K^n$, wobei $n = \dim V$. Jeder Isomorphismus $V \to K^n$ ist die Koordinatenabbildung κ_B bzgl. einer geeigneten geordneten Basis $\mathcal{B}$ von V.*

Beweis. Wegen $\dim V = \dim K^n$ folgt $V \cong K^n$ aus dem Satz. Es sei $\varphi : V \to K^n$ ein beliebiger Isomorphismus. Dann ist auch $\varphi^{-1} : K^n \to V$ ein Isomorphismus (Beispiel 6.4.5). Da $\{e_1, \ldots, e_n\}$ eine Basis von K^n ist und φ^{-1} ein Isomorphismus, ist nach Satz (6.4.5) $\{\varphi^{-1}(e_1), \ldots, \varphi^{-1}(e_n)\}$ eine Basis von V. Für die geordnete Basis $\mathcal{B} := (\varphi^{-1}(e_1), \ldots, \varphi^{-1}(e_n))$ und die Koordinatenabbildung κ_B gilt dann $\kappa_B : \varphi^{-1}(e_i) \mapsto e_i$ für $i = 1, \ldots, n$. Die linearen Abbildungen κ_B und φ stimmen also auf den Basiselementen von $\mathcal{B}$ überein. Laut Satz (6.4.4) handelt sich daher um dieselben Abbildungen, d.h. $\kappa_B = \varphi$. □

Bemerkung b. Die Folgerung ist von großer Bedeutung. Auf Grund dieser Tatsache lassen sich nämlich sämtliche Rechnungen in endlich-dimensionalen Vektorräumen auf Rechnungen in K^n zurückführen (via Koordinatenabbildungen).

Beispiel.

(i) $\mathbb{C} \cong \mathbb{R}^2$ (als $\mathbb{R}$-Vektorraum).

(ii) $\mathbb{R}^{2\times 2} \cong \mathbb{R}^4$.

(iii) $\mathbb{R}^{n\times m} \cong \mathbb{R}^{nm} \cong \mathbb{R}^{m\times n}$.

(iv) $K^{1\times n} \cong K^n$.

6.4.7 $\mathrm{Hom}(V, W)$ als Vektorraum

Es seien V und W zwei K-Vektorräume.

Satz. *Die Menge* $\mathrm{Hom}(V, W)$ *wird selbst zu einem* K*-Vektorraum, wenn man Addition und skalare Multiplikation* punktweise *definiert, d.h.*

$$(\varphi + \psi)(x) := \varphi(x) + \psi(x),$$
$$(\lambda\varphi)(x) := \lambda\varphi(x).$$

Beweis. Zunächst ist $\mathrm{Abb}(V, W)$ mit punktweiser Addition und punktweiser skalarer Multiplikation ein K-Vektorraum. Es bleibt zu zeigen, dass die Teilmenge $\mathrm{Hom}(V, W) \subseteq \mathrm{Abb}(V, W)$ die Unterraumbedingungen erfüllt. Man rechnet leicht nach, dass mit φ, ψ linear auch $\varphi + \psi$ und $\lambda\varphi$ wieder linear sind (Übung). □

Frage. Welche Dimension hat $\mathrm{Hom}(V, W)$? Wie sieht eine Basis aus?

6.4.8 Der Endomorphismenring

Bemerkung. Die Komposition von linearen Abbildungen ist wieder linear. Genauer, sind U, V, W drei K-Vektorräume, $\varphi \in \mathrm{Hom}(V, W)$ und $\psi \in \mathrm{Hom}(U, V)$, so ist $\varphi \circ \psi \in \mathrm{Hom}(U, W)$.

Beweis. Übung. □

Wir betrachten nun $\mathrm{End}(V) = \mathrm{Hom}(V, V)$ für einen K-Vektorraum V. Gemäß Satz (6.4.7) ist $\mathrm{End}(V)$ selbst ein K-Vektorraum mit der punktweisen Addition und skalaren Multiplikation. Im Falle von $\mathrm{End}(V)$ haben wir zusätzlich noch die Verknüpfung $\circ$ (Komposition).

Satz. $(\mathrm{End}(V), +, \circ)$ *ist ein Ring. Die neutralen Elemente sind die Nullabbildung* $(0 : V \to V, v \mapsto 0)$ *und die Identität* $(1 = id_V)$.

Definition. $(\mathrm{End}(V), +, \circ)$ wird der *Endomorphismenring von* V genannt. Die Einheitengruppe $\mathrm{Aut}(V) := (\mathrm{End}(V)^\times, \circ)$ heißt *Automorphismengruppe*, deren Elemente *Automorphismen*.

6.5 Lineare Abbildungen und Matrizen

In diesem Abschnitt werden alle Vektorräume als endlich-dimensional und nicht-trivial und vorausgesetzt, also $0 < \dim < \infty$. Mit Basen sind immer geordnete Basen gemeint und wir bezeichnen diese mit $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$.

Ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine geordnete Basis von V und $\varphi \in \mathrm{Aut}(V)$, so bezeichne $\varphi(\mathcal{B})$ das Tupel $(\varphi(v_1), \ldots, \varphi(v_n))$. Man beachte, dass die Zuordnung

$$\mathrm{Aut}(V) \to \{\text{geordnete Basen von } V\}, \quad \varphi \mapsto \varphi(\mathcal{B})$$

eine Bijektion ist (das folgt aus den Sätzen 6.4.4 und 6.4.5).

6.5.1 Die Abbildungsmatrix

Es seien V und W zwei K-Vektorräume mit $\dim V = n$ und $\dim W = m$ und mit den geordneten Basen $\mathcal{A} = (v_1, \ldots, v_n)$ von V und $\mathcal{B}$ von W.

Definition. Die *Abbildungsmatrix* von φ bzgl. $\mathcal{A}$ und $\mathcal{B}$ ist definiert als

$$M_\varphi := {}^{\mathcal{B}}M_\varphi^{\mathcal{A}} := (s_1, \ldots, s_n), \quad s_i := \kappa_{\mathcal{B}}(\varphi(v_i)).$$

Ist $V = W$ und $\mathcal{A} = \mathcal{B}$, so sagen wir kurz Abbildungsmatrix bzgl. $\mathcal{A}$ und schreiben $M_\varphi^{\mathcal{A}}$ statt ${}^{\mathcal{A}}M_\varphi^{\mathcal{A}}$. Wir verwenden die Schreibweise M_φ, wenn $\mathcal{A}$ und $\mathcal{B}$ fest gewählt sind.

Bemerkung a. Die Abbildung

$${}^{\mathcal{B}}M^{\mathcal{A}} : \mathrm{Hom}(V, W) \to K^{m \times n}, \varphi \mapsto {}^{\mathcal{B}}M_\varphi^{\mathcal{A}}$$

ist ein K-Vektorraum-Isomorphismus. Insbesondere hat $\mathrm{Hom}(V, W)$ die Dimension nm.

Beweis. Nach Satz 6.4.4a wird φ ein-eindeutig durch das Tupel $\varphi(\mathcal{A}) := (\varphi(v_1), \ldots, \varphi(v_n))$ beschrieben. Da $\kappa_{\mathcal{B}}$ bijektiv ist, wird $\varphi(\mathcal{A})$ wiederum ein-eindeutig durch ${}^{\mathcal{B}}M_\varphi^{\mathcal{A}}$ beschrieben. Die Abbildung ${}^{\mathcal{B}}M^{\mathcal{A}}$ ist somit eine Bijektion und, wie man leicht nachrechnet, auch linear: ${}^{\mathcal{B}}M_{\varphi+\psi}^{\mathcal{A}} = {}^{\mathcal{B}}M_\varphi^{\mathcal{A}} + {}^{\mathcal{B}}M_\psi^{\mathcal{A}}$ und ${}^{\mathcal{B}}M_\varphi^{\mathcal{A}} = \lambda {}^{\mathcal{B}}M_\varphi^{\mathcal{A}}$. □

Satz. *Die Abbildungsmatrix ${}^{\mathcal{B}}M_\varphi^{\mathcal{A}}$ ist die eindeutige Matrix $M \in K^{m \times n}$ mit der Eigenschaft*

$$\kappa_{\mathcal{B}}(\varphi(v)) = M \cdot \kappa_{\mathcal{A}}(v) \text{ für alle } v \in V, \tag{6.2}$$

Bemerkung b. Eine äquivalente Formulierung der Bedingung (6.2) ist:

$$\kappa_{\mathcal{B}} \circ \varphi = \varphi_M \circ \kappa_{\mathcal{A}} \quad \text{bzw.} \quad \varphi = \kappa_{\mathcal{B}}^{-1} \circ \varphi_M \circ \kappa_{\mathcal{A}}.$$

(Diagram siehe Vorlesung.)

Beweis des Satzes. Laut Satz 6.4.4a ist die Bedingung (6.2) äquivalent dazu, dass $\kappa_{\mathcal{B}}(\varphi(v_i)) = M \cdot \kappa_{\mathcal{A}}(v_i)$ für alle Basisvektoren v_i aus $\mathcal{A}$ gilt. Für $1 \leq i \leq n$ ist aber $M\kappa_{\mathcal{A}}(v_i) = Me_i = i$-te Spalte von M, und, per Definition, $\kappa_{\mathcal{B}}(\varphi(v_i)) = s_i$. Somit ist (6.2) äquivalent zu $M = {}^{\mathcal{B}}M_{\varphi}^{\mathcal{A}}$. □

Im folgenden Beispiel wird dargestellt, wie die Abbildungsmatrix das effektive Rechnen mit der linearen Abbildung mittels Koordinaten erlaubt.

Beispiel. Es bezeichne $\mathcal{E}$ die Standardbasis von K^n, d.h. $\mathcal{E} = (e_1, ..., e_n)$. Dann ist die zugehörige Koordinatenabbildung $\kappa_{\mathcal{E}} : K^n \to K^n$ die Identität.

(i) Wie lautet die Abbildungsmatrix S_0 der Spiegelung von $\mathbb{R}^2$ an der e_1-Achse bzgl. $\mathcal{E}$? Wir haben $\mathcal{E} = (e_1, e_2) = (\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix})$. Die Spiegelung an der e_1-Achse ist eine lineare Abbildung $\sigma_0 : \mathbb{R}^2 \to \mathbb{R}^2$ mit den Bildern der Basisvektoren $\sigma_0(e_1) = e_1$ und $\sigma_0(e_2) = -e_2$ (Zeichnung siehe Vorlesung). Die Koordinatenvektoren dieser Bilder lauten

$$\kappa_{\mathcal{E}}(\sigma_0(e_1)) = \sigma_0(e_1) = \begin{pmatrix}1\\0\end{pmatrix}, \quad \kappa_{\mathcal{E}}(\sigma_0(e_2)) = \sigma_0(e_2) = \begin{pmatrix}0\\-1\end{pmatrix},$$

also ist $S_0 = {}^{\mathcal{E}}M_{\sigma_0}^{\mathcal{E}} = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}$. Probe: $\sigma_0(\begin{pmatrix}a\\b\end{pmatrix}) = \begin{pmatrix}a\\-b\end{pmatrix} = S_0 \cdot \begin{pmatrix}a\\b\end{pmatrix}$.

(ii) Wie lautet die Abbildungsmatrix R_α der Drehung von $\mathbb{R}^2$ um α gegen den Uhrzeigersinn? Diese Drehung ist eine lineare Abbildung $\rho_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ mit den Bildern der Basisvektoren $\rho_\alpha(e_1) = \cos\alpha \cdot e_1 + \sin\alpha \cdot e_2$ und $\rho_\alpha(e_2) = -\sin\alpha \cdot e_1 + \cos\alpha \cdot e_2$ (Zeichnung siehe Vorlesung). Die Koordinatenvektoren dieser Bilder lauten

$$\kappa_{\mathcal{E}}(\rho_\alpha(e_1)) = \rho_\alpha(e_1) = \begin{pmatrix}\cos\alpha\\\sin\alpha\end{pmatrix}, \kappa_{\mathcal{E}}(\rho_\alpha(e_2)) = \rho_\alpha(e_2) = \begin{pmatrix}-\sin\alpha\\\cos\alpha\end{pmatrix},$$

also ist $R_\alpha = {}^{\mathcal{E}}M_{\rho_\alpha}^{\mathcal{E}} = \begin{pmatrix}\cos\alpha & -\sin\alpha\\\sin\alpha & \cos\alpha\end{pmatrix}$.

Probe: $\rho_\alpha(\begin{pmatrix}a\\b\end{pmatrix}) = \begin{pmatrix}a\cos\alpha - b\sin\alpha\\a\sin\alpha + b\cos\alpha\end{pmatrix} = R_\alpha \cdot \begin{pmatrix}a\\b\end{pmatrix}$.

(iii) Es sei $A \in K^{m\times n}$. Für $\varphi_A : K^n \to K^m, x \mapsto A \cdot x$ ist

$$ {}^{\mathcal{E}}M^{\mathcal{E}}_{\varphi_A} = A.$$

(iv) Betrachte $\varphi = \mathrm{id}_V$. Für jede Basis $\mathcal{B}$ von V ist

$$ {}^{\mathcal{B}}M^{\mathcal{B}}_{\mathrm{id}_V} = E_n, \quad \text{wobei } n = \dim V.$$

(v) Es sei

$$V = \mathrm{Pol}_n(\mathbb{R}) = \{a_n\mathbf{x}^n + ... + a_1\mathbf{x}^1 + a_0 \mid a_0, ..., a_n \in \mathbb{R}\}$$

V ist $\mathbb{R}$-Vektorraum der Dimension $n+1$ mit Basis $\mathcal{A} = (1, \mathbf{x}, \mathbf{x}^2, ..., \mathbf{x}^n)$. Die *Ableitungsabbildung*

$$\mathrm{diff} : V \to V, f \mapsto f'$$

ist linear (vgl. Beispiel (6.4.2)(vii)). Wir bestimmen die Abbildungsmatrix von diff bezüglich $\mathcal{A}$. Die Bilder der Basisvektoren sind

$$\mathrm{diff}(\mathbf{x}^i) = \begin{cases} 0 & \text{falls } i = 0, \\ i\mathbf{x}^{i-1} & \text{falls } i > 0. \end{cases}$$

Übersetzt in Koordinaten bzgl. $\mathcal{A}$ bedeutet das

$$M_{\mathrm{diff}} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 2 & & \\ \vdots & \vdots & 0 & & \\ & & \vdots & \ddots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}.$$

Probe: Für $p = a_n\mathbf{x}^n + ... + a_1\mathbf{x} + a_0 \in \mathrm{Pol}_n(\mathbb{R})$ ist $p' = na_n\mathbf{x}^{n-1} + \ldots + 2a_2\mathbf{x} + a_1$. Also wie gewünscht

$$\kappa_{\mathcal{A}}(p') = \begin{pmatrix} a_1 \\ 2a_2 \\ \vdots \\ na_n \\ 0 \end{pmatrix} = M_{\mathrm{diff}} \cdot \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix} = M_{\mathrm{diff}} \cdot \kappa_{\mathcal{A}}(p).$$

6.5.2 Berechnung von Kern und Bild

Bemerkung. Es seien Basen $\mathcal{A}$ von V und $\mathcal{B}$ von W fest gewählt mit zugehörigen Koordinatenabbildungen $\kappa_\mathcal{A}$ und $\kappa_\mathcal{B}$.

(i) $\kappa_\mathcal{A}(\operatorname{Kern}\varphi) = \mathbb{L}_0(M_\varphi)$ bzw. $\operatorname{Kern}\varphi = \kappa_\mathcal{A}^{-1}(\mathbb{L}_0(M_\varphi))$.

(ii) $\kappa_\mathcal{B}(\operatorname{Bild}\varphi) = \operatorname{SR}(M_\varphi)$ bzw. $\operatorname{Bild}\varphi = \kappa_\mathcal{B}^{-1}(\operatorname{SR}(M_\varphi))$.

(iii) $\operatorname{Rg}\varphi = \operatorname{Rg} M_\varphi$ und $\operatorname{Def}\varphi = \operatorname{Def} M_\varphi$.

Beweis. Wir zeigen exemplarisch die Aussagen über $\operatorname{Kern}\varphi$ und $\operatorname{Def}\varphi$, die Aussagen über $\operatorname{Bild}\varphi$ und $\operatorname{Rg}\varphi$ gehen ähnlich. (i): Die Aussage ist $v \in \operatorname{Kern}\varphi \Leftrightarrow \kappa_\mathcal{A}(v) \in \mathbb{L}_0(M_\varphi)$. Das folgt aus der Injektivität von $\kappa_\mathcal{B}$ und der Bedingung (6.2):

$$\varphi(v) = 0 \Leftrightarrow \kappa_\mathcal{B}(\varphi(v)) = 0 \Leftrightarrow M_\varphi \kappa_\mathcal{A}(v) = 0.$$

(iii): Als Monomorphismus erhält $\kappa_\mathcal{A}$ die Dimension von Unterräumen (Übung 6.4.5 oder Bemerkung 6.4.6iii). Daher folgt aus (i):

$$\operatorname{Def}\varphi = \dim(\operatorname{Kern}\varphi) = \dim(\mathbb{L}_0(M_\varphi)) = \operatorname{Def} M_\varphi.$$

□

Folgerung. *Für $\varphi \in \operatorname{Hom}(V, W)$ gilt stets:* $\operatorname{Rg}\varphi + \operatorname{Def}\varphi = \dim V$.

Beweis. Wegen Teil (iii) der Bemerkung folgt dies aus der entsprechenden Gleichung für Matrizen aus Satz (6.3.1). □

Beispiel. Wir berechnen mittels Koordinaten Kern und Bild der Ableitungsabbildung

$$\operatorname{diff} : \operatorname{Pol}_n(\mathbb{R}) \to \operatorname{Pol}_n(\mathbb{R}), f \mapsto f'.$$

Wir wählen als Basis $\mathcal{A} = (1, \mathbf{x}, \mathbf{x}^2, ..., \mathbf{x}^n)$. Für die Koordinatenabbildung $\kappa_\mathcal{A}$ gilt dann $\kappa_\mathcal{A}(\mathbf{x}^i) = e_{i+1}$ für $i = 0, \ldots, n$. Die Abbildungsmatrix von diff bzgl. $\mathcal{A}$ wurde in Beispiel (6.5.1) bestimmt und lautet

$$M_{\text{diff}} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 2 & & \\ \vdots & \vdots & 0 & & \\ & & \vdots & \ddots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}.$$

Man sieht sofort

$$\mathbb{L}_0(M_\varphi) = \langle \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \rangle, \quad \text{und} \quad \mathrm{SR}(M_\varphi) = \langle \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ \vdots \\ 0 \end{pmatrix}, \cdots \begin{pmatrix} 0 \\ \vdots \\ n \\ 0 \end{pmatrix} \rangle,$$

also $\mathbb{L}_0(M_\varphi) = \langle e_1 \rangle$ und $\mathrm{SR}(M_\varphi) = \langle e_1, \ldots, e_n \rangle$. Rückübersetzt in Vektoren aus $\mathrm{Pol}_n(\mathbb{R})$ bedeutet das

$$\begin{aligned} \mathrm{Kern}(\mathrm{diff}) &= \langle \kappa_{\mathcal{A}}^{-1}(e_1) \rangle = \langle 1 \rangle = \{a_0 \mid a_0 \in \mathbb{R}\} \text{ (konstante Polynome)} \\ \mathrm{Bild}(\mathrm{diff}) &= \langle \kappa_{\mathcal{A}}^{-1}(e_1), \ldots, \kappa_{\mathcal{A}}^{-1}(e_n) \rangle = \langle 1, \mathbf{x}, ..., \mathbf{x}^{n-1} \rangle \\ &= \{a_0 + a_1\mathbf{x} + ... + a_{n-1}\mathbf{x}^{n-1} \mid a_i \in \mathbb{R}\}. \end{aligned}$$

Man sieht $\mathrm{Rg}(\mathrm{diff}) = n$ und $\mathrm{Def}(\mathrm{diff}) = 1$, die Gleichung $\mathrm{Rg}(\mathrm{diff}) + \mathrm{Def}(\mathrm{diff}) = n + 1 = \dim \mathrm{Pol}_n(\mathbb{R})$ ist also erfüllt.

6.5.3 Der Produktsatz

Es seien $\varphi : V \to W$ und $\psi : U \to V$ zwei lineare Abbildungen zwischen K-Vektorräumen. Von U, V, W seien die geordneten Basen $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ fest gewählt.

Satz. ${}^{\mathcal{C}}M^{\mathcal{A}}_{\varphi\circ\psi} = {}^{\mathcal{C}}M^{\mathcal{B}}_{\varphi} \cdot {}^{\mathcal{B}}M^{\mathcal{A}}_{\psi}$.

Beweis. Man beachte, dass stets $\varphi_A \circ \varphi_B = \varphi_{AB}$ sofern das Matrixprodukt AB existiert. Damit ergibt sich

$$\begin{aligned} \varphi \circ \psi &= \kappa_{\mathcal{C}}^{-1} \circ \varphi_{M_\varphi} \circ \kappa_{\mathcal{B}} \circ \kappa_{\mathcal{B}}^{-1} \circ \varphi_{M_\psi} \circ \kappa_{\mathcal{A}} \\ &= \kappa_{\mathcal{C}}^{-1} \circ (\varphi_{M_\varphi} \circ \varphi_{M_\psi}) \circ \kappa_{\mathcal{A}} = \kappa_{\mathcal{C}}^{-1} \circ \varphi_{M_\varphi \cdot M_\psi} \circ \kappa_{\mathcal{A}}. \end{aligned}$$

Aus Satz 6.5.1 folgt $M_{\varphi\circ\psi} = M_\varphi \cdot M_\psi$. □

Folgerung a. *Falls* $\dim V = \dim W$ *(d.h.* ${}^{\mathcal{C}}M^{\mathcal{B}}_{\varphi}$ *quadratisch) ist, so ist* φ *genau dann ein Isomorphismus, wenn* M_φ *regulär ist. In diesem Fall gilt* ${}^{\mathcal{B}}M^{\mathcal{C}}_{\varphi^{-1}} = ({}^{\mathcal{C}}M^{\mathcal{B}}_{\varphi})^{-1}$.

Beweis. Es sei $\dim V = \dim W = n$, d.h. M_φ eine $n \times n$-Matrix. Nach den Bemerkungen (6.4.6) und (6.5.2)(iii) ist φ genau dann ein Isomorphismus, wenn $\mathrm{Rg}\, M_\varphi = \mathrm{Rg}\, \varphi = n$. Nach Satz 5.4.3 zusammen mit Satz 5.3.5 hat eine quadratische Matrix genau dann vollen Rang (hier: $\mathrm{Rg}\, M_\varphi = n$), wenn sie regulär ist.

Sei nun φ ein Isomorphismus. Aus dem Produktsatz und Beispiel (6.5.1)(iv) folgt ${}^{\mathcal{C}}M^{\mathcal{B}}_{\varphi} \cdot {}^{\mathcal{B}}M^{\mathcal{C}}_{\varphi^{-1}} = {}^{\mathcal{C}}M^{\mathcal{C}}_{\mathrm{id}} = E_n$, also $({}^{\mathcal{C}}M^{\mathcal{B}}_{\varphi})^{-1} = {}^{\mathcal{B}}M^{\mathcal{C}}_{\varphi^{-1}}$. □

Folgerung b. *Der Vektorraum-Isomorphismus*

$$ {}^{\mathcal{B}}M^{\mathcal{B}} : \mathrm{End}(V) \to K^{n\times n}, \quad \varphi \mapsto {}^{\mathcal{B}}M^{\mathcal{B}}_{\varphi} $$

ist auch Ringisomorphismus. Die Einschränkung

$$ {}^{\mathcal{B}}M^{\mathcal{B}} : \mathrm{Aut}(V) \to \mathrm{GL}_n(K), \quad \varphi \mapsto {}^{\mathcal{B}}M^{\mathcal{B}}_{\varphi} $$

ist Gruppenisomorphismus.

Beweis. Um zu sehen, dass M ein Ringhomomorphismus ist, sind $M_{\mathrm{id}} = E_n$, $M_{\varphi+\psi} = M_\varphi + M_\psi$, und $M_{\varphi\circ\psi} = M_\varphi \cdot M_\psi$ zu prüfen. Alle drei Bedingungen wurden bereits gezeigt (die letzte durch den Produktsatz).

Die Einschränkung eines Ringisomorphismus auf die Einheitengruppen ist stets Gruppenisomorphismus. □

Beispiel a. Es sei $\sigma_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der um den Winkel α gegen den Uhrzeigersinn gedrehten e_1-Achse (Zeichnung siehe Vorlesung). Wie lautet die Abbildungsmatrix S_α von σ_α bzgl. $\mathcal{E}$? Idee: Wir schreiben σ_α als die Komposition $\rho_\alpha \circ \sigma_0 \circ \rho_{-\alpha}$ und verwenden den Produktsatz. Die Abbildungsmatrizen von ρ_α und σ_0 sind bereits aus Beispiel (6.5.1) bekannt. Es ergibt sich:

$$ \begin{aligned} S_\alpha = M_{\sigma_\alpha} &= M_{\rho_\alpha} \cdot M_{\sigma_0} \cdot M_{\rho_{-\alpha}} = R_\alpha \cdot S_0 \cdot R_{-\alpha} \\ &= \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos^2\alpha - \sin^2\alpha & 2\sin\alpha\cos\alpha \\ 2\sin\alpha\cos\alpha & \sin^2\alpha - \cos^2\alpha \end{pmatrix} \end{aligned} $$

Für $\alpha = 30°$ erhalten wir wegen $\sin 30° = \sin\frac{\pi}{6} = \frac{1}{2}$ und $\cos 30° = \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}$ die Matrix $S_{30°} = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$.

Beispiel b. Da ρ_α bijektiv ist, ist $R_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ invertierbar. Die Umkehrabbildung von ρ_α ist $\rho_{-\alpha}$, folglich

$$ R_\alpha^{-1} = (M_{\rho_\alpha})^{-1} = M_{\rho_\alpha{}^{-1}} = M_{\rho_{-\alpha}} = R_{-\alpha} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}. $$

Man beachte, dass die Gleichung $R_\alpha \cdot R_{-\alpha} = E_2$ äquivalent ist zu: $\sin^2\alpha + \cos^2\alpha = 1$.

Übung. Warum gilt die Gleichung $\sigma_\alpha = \sigma_0 \circ \rho_{2\alpha}$? Man berechne daraus S_α mit Hilfe des Produktsatzes. Wie ist das Ergebnis im Vergleich zu Beispiel a zu interpretieren?

6.5.4 Die Basiswechselmatrix

Es seien $\mathcal{A}$, $\mathcal{B}$ zwei geordnete Basen von V, $0 < \dim V = n < \infty$.

Definition. Die Matrix ${}^{\mathcal{B}}M^{\mathcal{A}}_{\mathrm{id}_V}$ wird *Basiswechselmatrix* oder *Basistransformationsmatrix* von $\mathcal{A}$ zu $\mathcal{B}$ genannt, geschrieben ${}^{\mathcal{B}}T^{\mathcal{A}}$.

Bemerkung.

(i) In den Spalten von ${}^{\mathcal{B}}T^{\mathcal{A}}$ stehen die Basisvektoren aus $\mathcal{A}$, geschrieben in Koordinaten bzgl. $\mathcal{B}$. Kurz: In den Spalten stehen „die alten Basisvektoren bzgl. der neuen Basis“.

(ii) ${}^{\mathcal{B}}T^{\mathcal{A}}$ ist die eindeutige Matrix $T \in K^{n\times n}$ mit der Eigenschaft $\kappa_{\mathcal{B}} = \varphi_T \circ \kappa_{\mathcal{A}}$, d.h.
$$\kappa_{\mathcal{B}}(v) = T \cdot \kappa_{\mathcal{A}}(v) \quad \text{für alle } v \in V.$$

(iii) Es gilt ${}^{\mathcal{A}}T^{\mathcal{B}} = ({}^{\mathcal{B}}T^{\mathcal{A}})^{-1}$.

(iv) Für jedes $\varphi \in \mathrm{Aut}(V)$ gilt ${}^{\mathcal{B}}M^{\mathcal{B}}_{\varphi} = {}^{\mathcal{B}}T^{\varphi(\mathcal{B})}$.

Beweis. Die Aussagen ergeben sich aus der Definition und den Bemerkungen (6.5.1), sowie aus Folgerung 6.5.3a:
$${}^{\mathcal{A}}T^{\mathcal{B}} = {}^{\mathcal{A}}M^{\mathcal{B}}_{\mathrm{id}} = ({}^{\mathcal{B}}M^{\mathcal{A}}_{\mathrm{id}^{-1}})^{-1} = ({}^{\mathcal{B}}M^{\mathcal{A}}_{\mathrm{id}})^{-1} = ({}^{\mathcal{B}}T^{\mathcal{A}})^{-1}.$$
□

Folgerung. *Für festes $\mathcal{B}$ ist folgende Abbildung eine Bijektion:*
$$\{\textit{geordnete Basen von } V\} \to \mathrm{GL}_n(K), \quad \mathcal{A} \mapsto {}^{\mathcal{B}}T^{\mathcal{A}}.$$

Beweis. Nach Teil (iv) der Bemerkung sind die Zuordnungen $\mathrm{Aut}(V) \to \mathrm{GL}_n(K), \varphi \mapsto {}^{\mathcal{B}}M^{\mathcal{B}}_{\varphi}$ und $\varphi \mapsto {}^{\mathcal{B}}T^{\varphi(\mathcal{B})}$ identisch. Die Aussage folgt aus der Tatsache, dass sowohl $\mathrm{Aut}(V) \to \mathrm{GL}_n(K), \varphi \mapsto {}^{\mathcal{B}}M^{\mathcal{B}}_{\varphi}$ als auch $\mathrm{Aut}(V) \to \{\text{geordnete Basen von } V\}, \varphi \mapsto \varphi(\mathcal{B})$ bijektiv sind. (Diagram siehe Vorlesung.) □

Beispiel. Es sei $V = \mathbb{R}^2$. Wir betrachten die Basen $\mathcal{A} = \mathcal{E} = (e_1, e_2)$ und $\mathcal{B} = (v_1, v_2) = (\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}-2\\1\end{pmatrix})$. Dann gilt:
$$T := {}^{\mathcal{A}}T^{\mathcal{B}} = \begin{pmatrix}1 & -2\\ 2 & 1\end{pmatrix}, \quad {}^{\mathcal{B}}T^{\mathcal{A}} = T^{-1} = \begin{pmatrix}1 & -2\\ 2 & 1\end{pmatrix}^{-1} = \ldots = \frac{1}{5}\begin{pmatrix}1 & 2\\ -2 & 1\end{pmatrix}.$$

Wir bestimmen die Koordinaten von $v := \begin{pmatrix} -1 \\ 3 \end{pmatrix}$ bzgl. $\mathcal{B}$ mittels Basiswechselmatrix. Es gilt $\kappa_{\mathcal{A}}(v) = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$, also Teil (ii) der Bemerkung

$$\kappa_{\mathcal{B}}(v) = {}^{\mathcal{B}}T^{\mathcal{A}} \cdot \kappa_{\mathcal{A}}(v) = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 3 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 5 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Per Definition ist $\kappa_{\mathcal{B}}(v) = \begin{pmatrix} \lambda \\ \mu \end{pmatrix}$ gerade der Lösungsvektor der Gleichung $\lambda v_1 + \mu v_2 = v$. Das liefert uns die Probe: $1 \cdot v_1 + 1 \cdot v_2 = v$.$\surd$

Anmerkung: Natürlich kann $\kappa_{\mathcal{B}}(v)$ als Lösung des linearen Gleichungssystems $\lambda v_1 + \mu v_2 = v$ auch direkt (ohne Verwendung von Basiswechselmatrizen) bestimmt werden. In Matrixschreibweise lautet dieses Gleichungssystem $T \cdot \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$, d.h. die Basiswechselmatrix taucht auch hier wieder auf.

6.5.5 Der Basiswechselsatz

Es seien $\mathcal{A}$, $\mathcal{A}'$ geordnete Basen von V und $\mathcal{B}$, $\mathcal{B}'$ geordnete Basen von W.

Satz. *Für jede lineare Abbildung $\varphi : V \to W$ gilt:*

$${}^{\mathcal{B}'}M_{\varphi}^{\mathcal{A}'} = {}^{\mathcal{B}'}T^{\mathcal{B}} \,.\, {}^{\mathcal{B}}M_{\varphi}^{\mathcal{A}} \,.\, {}^{\mathcal{A}}T^{\mathcal{A}'}.$$

Beweis. Nach Definition (6.5.4) und Satz (6.5.3) gilt

$${}^{\mathcal{B}'}T^{\mathcal{B}} \,.\, {}^{\mathcal{B}}M_{\varphi}^{\mathcal{A}} \,.\, {}^{\mathcal{A}}T^{\mathcal{A}'} = {}^{\mathcal{B}'}M_{\mathrm{id}_W}^{\mathcal{B}} \,.\, {}^{\mathcal{B}}M_{\varphi}^{\mathcal{A}} \,.\, {}^{\mathcal{A}}M_{\mathrm{id}_V}^{\mathcal{A}'} = {}^{\mathcal{B}'}M_{\mathrm{id}_W \circ \varphi \circ \mathrm{id}_V}^{\mathcal{A}'} = {}^{\mathcal{B}'}M_{\varphi}^{\mathcal{A}'}.$$

□

Beispiel. Wir betrachten die lineare Abbildung $\varphi_A : \mathbb{R}^2 \to \mathbb{R}^2$ mit $A = \begin{pmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{pmatrix}$. Wir haben ${}^{\mathcal{E}}M_{\varphi_A}^{\mathcal{E}} = A$ (vgl. Beispiel (6.5.1)(iii)). *Frage:* Gibt es eine Basis $\mathcal{B}$ so, dass ${}^{\mathcal{B}}M_{\varphi_A}^{\mathcal{B}}$ besonders „einfach“ wird? Dies ist in der Tat der Fall für die Basis $\mathcal{B} = (v_1, v_2) = (\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix})$, wie folgende Rechnung zeigt. Aus Beispiel (6.5.4) wissen wir

$${}^{\mathcal{A}}T^{\mathcal{B}} = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}, \quad {}^{\mathcal{B}}T^{\mathcal{A}} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}.$$

Nach obigem Satz gilt folglich

$$\begin{aligned}
{}^{\mathcal{B}}M^{\mathcal{B}}_{\varphi_A} &= {}^{\mathcal{B}}T^{\mathcal{A}} \cdot {}^{\mathcal{A}}M^{\mathcal{A}}_{\varphi_A} \cdot {}^{\mathcal{A}}T^{\mathcal{B}} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}\frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \\
&= \frac{1}{25}\begin{pmatrix} 5 & 10 \\ 10 & -5 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 5 & 0 \\ 0 & -5 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\end{aligned}$$

Die Abbildung φ_A ist also eine Spiegelung an der v_1-Achse! (Bild siehe Vorlesung.) Wie man die „richtige“ Basis $\mathcal{B}$ systematisch auffindet, werden wir erst später sehen, im Kapitel über Eigenvektoren.

6.5.6 Ähnliche Matrizen

Wir betrachten den Basiswechselsatz im Fall von Endomorphismen. Es seien dazu $\mathcal{A}, \mathcal{B}$ Basen von V und $\varphi \in \mathrm{End}(V)$. Wir schreiben für ${}^{\mathcal{A}}M^{\mathcal{A}}_{\varphi}$ kurz $M^{\mathcal{A}}_{\varphi}$. Dann liefert der Basiswechselsatz

$$M^{\mathcal{B}}_{\varphi} = T^{-1}M^{\mathcal{A}}_{\varphi}T, \quad \text{mit } T = {}^{\mathcal{A}}T^{\mathcal{B}}.$$

Dies legt folgende Definition nahe.

Definition. Es seien R ein kommutativer Ring und $A, B \in R^{n\times n}$. Wir nennen A *ähnlich* zu B, wenn ein $T \in \mathrm{GL}_n(R)$ existiert mit $B = T^1AT$.

Übung a. Man zeige, dass die Ähnlichkeit von Matrizen eine Äquivalenzrelation ist.

Übung b. Man zeige, dass ähnliche Matrizen denselben Rang haben.

Der Basiswechselsatz 6.5.5 besagt: die Abbildungsmatrizen einer festen linearen Abbildung bezüglich verschiedener Basen sind ähnlich zueinander. Folgerung 6.5.4 impliziert ferner: Zwei quadratische Matrizen sind genau dann äquivalent, wenn sie dieselbe lineare Abbildung beschreiben, nur bezüglich verschiedener Basen.

Kapitel 7

Determinanten und Eigenvektoren

7.1 Determinanten

In diesem Paragrafen seien R ein kommutativer Ring und $n \in \mathbb{N}$. Wir betrachten quadratische $n \times n$-Matrizen über R und bezeichnen mit $s_1, \ldots, s_n \in R^n$ jeweils die Spalten der Matrix.

7.1.1 Definition und Eigenschaften

Definition. Eine Abbildung $D : R^{n \times n} \to R$ heißt *Determinante*, wenn für alle $1 \leq i, j \leq n$ und $\lambda \in R$ gelten:

a) $D(\ldots, s_j + s'_j, \ldots) = D(\ldots, s_j, \ldots) + D(\ldots, s'_j, \ldots)$.

b) $D(\ldots, \lambda s_j, \ldots) = \lambda D(\ldots, s_j, \ldots)$.

c) $D(\ldots, s, \ldots, s, \ldots) = 0$.

d) $D(E_n) = 1$.

Man sagt, eine Determinante ist *multilinear* (Eigenschaft a) und b)), *alternierend* (c)) und *normiert* (d)).

Bemerkung. Ist $D : R^{n \times n} \to R$ eine Determinante, so gelten:

(i) $D(\ldots, s_i, \ldots, s_j, \ldots) = -D(\ldots, s_j, \ldots, s_i, \ldots)$.

(ii) $D(s_{\pi(1)}, \ldots, s_{\pi(n)}) = \operatorname{sgn}(\pi) \cdot D(s_1, \ldots, s_n)$ für alle $\pi \in S_n$.

(iii) $D(e_{\pi(1)}, \ldots, e_{\pi(n)}) = \operatorname{sgn}(\pi)$ für alle $\pi \in S_n$.

(iv) $D(\ldots, 0, \ldots) = 0$.

(v) $D(\lambda A) = \lambda^n D(A)$.

(vi) $D(\ldots, s_i + \lambda s_j, \ldots, s_j, \ldots) = D(\ldots, s_i, \ldots, s_j, \ldots)$.

Beweis. siehe Vorlesung. □

7.1.2 Die Leibniz-Formel

Satz. *Für jedes $n \in \mathbb{N}$ gibt es genau eine Determinante, geschr.*

$$\det : R^{n\times n} \to R, \quad A \mapsto \det A \text{ oder } |A|.$$

Für $A = (a_{ij})$ gilt die Leibniz-Formel*:*

$$\det(A) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) a_{\pi(1),1} a_{\pi(2),2} \cdots a_{\pi(n),n} = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{\pi(i),i}.$$

Beweis. Eindeutigkeit: Wir zeigen, dass für eine multilineare, alternierende Abbildung $D : R^{n\times n} \to D$ gilt:

$$D(A) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) a_{\pi(1),1} \cdots a_{\pi(n),n} D(E_n). \tag{7.1}$$

Aus der Normiertheit folgt dann die Leibniz-Formel. Beweis von (7.1) in der Vorlesung.

Existenz: Als Übung zeige man, dass die Leibniz-Formel tatsächlich eine Determinante definiert. □

Beispiel. Für $n = 2$ ergibt sich

$$\det A = a_{11}a_{22} - a_{12}a_{21}.$$

Für $n = 3$ ergibt sich die als *Regel von Sarrus* (franz. Mathematiker, 1795-1861) bekannte Formel

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Man zeichne sich zu diesen Formeln ein Bild (siehe Vorlesung)! Achtung: die Bilder lassen sich nicht auf $n \geq 4$ verallgemeinern, da die Leibniz-Formel dann zu viele Summanden hat.

Ein Beispiel für $n = 3$:

$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ -1 & 0 & 5 \end{vmatrix} = (5 - 4 + 0) - (-3 + 0 + 0) = 1 - (-3) = 4.$$

Folgerung. *Es gilt stets* $\det A = \det A^t$. *Somit gelten alle Determinantenregeln aus* (7.1.1) *auch analog für Zeilen statt für Spalten.*

Beweis. Wenn π die Menge S_n durchläuft, so durchläuft auch π^{-1} die Menge S_n. Da ausserdem $\operatorname{sgn}\pi = \operatorname{sgn}\pi^{-1}$ gilt, haben wir nach der Leibniz-Formel $\det A = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^n a_{\pi^{-1}(i),i}$. Wenn i die Zahlen $1,\dots,n$ durchläuft, so durchläuft für festes $\pi \in S_n$ auch $\pi(i)$ genau $1,\dots,n$, weil π eine Bijektion ist. Also gilt $\prod_{i=1}^n a_{\pi^{-1}(i),i} = \prod_{i=1}^n a_{\pi^{-1}\circ\pi(i),\pi(i)} = \prod_{i=1}^n a_{i,\pi(i)}$. Dies zeigt $\det A = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^n a_{\pi(i),i} = \det A^t$. □

7.1.3 Kästchensatz

Satz. *Ist* A *von der Form* $A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right)$, *so gilt*

$$\det A = \det(B)\cdot\det(D).$$

Beweis. siehe Vorlesung. □

Folgerung. *Obere und untere Dreiecksmatrizen haben als Determinante das Produkt der Diagonaleinträge, d.h.* $\det A = a_{11}a_{22}\cdots a_{nn}$.

Beweis. Induktion nach n mit dem Kästchensatz, der im Induktionsschritt mit einer 1×1-Matrix B angewendet wird. □

7.1.4 Determinante und Gauß-Algorithmus

Die Regeln b), (i) und (vi) aus (7.1.1) erlauben es, die Matrix wie beim Gauß-Algorithmus auf eine Dreiecksform zu bringen, woraus sich dann die Determinante leicht ablesen lässt. Dabei sind sowohl Spalten- als auch Zeilentransformationen erlaubt (auch gemischt). Man beachte jedoch, dass der Gauß-Algorithmus im Allgemeinen nur über Körpern funktioniert.

Beispiel.

$$A = \begin{vmatrix} 3 & 0 & -2 \\ 6 & 0 & 1 \\ -9 & -2 & 5 \end{vmatrix} = 3\begin{vmatrix} 1 & 0 & -2 \\ 2 & 0 & 1 \\ -3 & -2 & 5 \end{vmatrix} = -3\begin{vmatrix} 1 & -2 & 0 \\ 2 & 1 & 0 \\ -3 & 5 & -2 \end{vmatrix} = -3\begin{vmatrix} 1 & 0 & 0 \\ 2 & 5 & 0 \\ -3 & -1 & -2 \end{vmatrix}$$
$$= (-3)\cdot 1\cdot 5\cdot(-2) = 30.$$

Folgerung. *Ist* $R = K$ *ein Körper, so gilt:* $\det A \neq 0 \Leftrightarrow \operatorname{Rg} A = n$. *Also:*

$$\mathrm{GL}_n(K) = \{A \in K^{n\times n} \mid \det A \neq 0\}.$$

Beweis. Sind A und A' Gauß-äquivalent, so ist $\det A' = \lambda \det A$ für ein $\lambda \in K\backslash\{0\}$ (siehe Regeln b), (i), (vi) aus (7.1.1)), also $\det A = 0 \Leftrightarrow \det A' = 0$. Wir können daher o.B.d.A. annehmen, dass A in Zeilenstufenform, also obere Dreiecksmatrix ist. Eine obere Dreiecksmatrix der Größe $n \times n$ hat offensichtlich genau dann Rang n (volle Stufenzahl), wenn alle Einträge entlang der Hauptdiagonalen ungleich 0 sind. Nach der Bemerkung ist das genau dann der Fall, wenn $\det A \neq 0$. (Man beachte, dass Körper nullteilerfrei sind.) □

7.1.5 Laplace-Entwicklung

Definition. Es seien $A \in R^{n\times n}$, $1 \leq i, j \leq n$. Wir betrachten die $(n-1) \times (n-1)$-Untermatrix M von A, die durch Streichen der i-ten Zeile und der j-ten Spalte entsteht, also

$$M := \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & & & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & & & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}.$$

Die Determinante $|M|$ dieser Untermatrix heißt *Minor von A zu ij*. Weiter wird der Ausdruck $\operatorname{cof}_{ij}(A) := (-1)^{i+j}|M|$ *Kofaktor von A zu ij* genannt.

Satz. *Für alle $1 \leq i, j \leq n$ gilt:*

$$\det A = \sum_{k=1}^{n} a_{ik} \operatorname{cof}_{ik}(A) = \sum_{k=1}^{n} a_{kj} \operatorname{cof}_{kj}(A).$$

Die erste Summe nennt man die Laplace-Entwicklung nach der i-ten Zeile, *die zweite Summe nennt man die* Laplace-Entwicklung nach der j-ten Spalte.

Beweis. Beweis mit Hilfe des Kästchensatzes (siehe Vorlesung). □

Beispiel.

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ -1 & -1 & 0 & 1 \\ 4 & 0 & 3 & -1 \\ 2 & 0 & -1 & 1 \end{vmatrix} = -2 \begin{vmatrix} -1 & 0 & 1 \\ 4 & 3 & -1 \\ 2 & -1 & 1 \end{vmatrix} + (-1) \begin{vmatrix} 1 & 3 & 4 \\ 4 & 3 & -1 \\ 2 & -1 & 1 \end{vmatrix}$$

$$= -2 \left(- \begin{vmatrix} 3 & -1 \\ -1 & 1 \end{vmatrix} + \begin{vmatrix} 4 & 3 \\ 2 & -1 \end{vmatrix} \right) - \left(\begin{vmatrix} 3 & -1 \\ -1 & 1 \end{vmatrix} - 3 \begin{vmatrix} 4 & -1 \\ 2 & 1 \end{vmatrix} + 4 \begin{vmatrix} 4 & 3 \\ 2 & -1 \end{vmatrix} \right)$$

Dabei wurde die 4×4-Matrix nach der 2. Spalte entwickelt, und die 3×3-Matrizen jeweils nach der 1. Zeile. Nun ist es sinnvoll, nach gleichen 2×2-Determinanten zu sortieren, bevor diese nach der Formel aus Beispiel (7.1.2) berechnet werden:

$$= \begin{vmatrix} 3 & -1 \\ -1 & 1 \end{vmatrix} - 6 \begin{vmatrix} 4 & 3 \\ 2 & -1 \end{vmatrix} + 3 \begin{vmatrix} 4 & -1 \\ 2 & 1 \end{vmatrix} = 2 - 6 \cdot (-10) + 3 \cdot 6 = 2 + 60 + 18 = 80.$$

7.1.6 Produktsatz

Satz. *Es gilt* $\det(AB) = \det A \cdot \det B$ *für alle* $A, B \in R^{n \times n}$.

Beweis. Wir sehen A als fest an und definieren die Abbildung

$$D : R^{n \times n} \to R, \quad B \mapsto \det(AB).$$

Man prüft leicht nach, dass D multilinear und alternierend ist. Das liegt daran, dass für die Spalten $s_1, \ldots, s_n$ von A gilt $D(s_1, \ldots, s_n) = \det(As_1, \ldots, As_n)$, dass det multilinear und alternierend ist, und dass die Matrixmultiplikation mit A ebenfalls linear ist (Rechnung als Übung). Nach (7.1) folgt

$$D(B) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} b_{\pi(i),i} D(E).$$

Nach der Leibniz-Formel für $\det B$ gilt somit $D(B) = \det B \cdot D(E)$. Wegen $D(B) = \det(AB)$ und $D(E) = \det(AE) = \det A$ ist das genau die Behauptung. □

Folgerung.

(i) *Ist* $A \in R^{n \times n}$ *invertierbar, so ist auch* $\det A \in R$ *invertierbar, und es gilt* $\det(A^{-1}) = (\det A)^{-1}$.

(ii) *Ähnliche Matrizen haben dieselbe Determinante.*

(iii) *Die Einschränkung der Abbildung* det *auf* $\mathrm{GL}_n(R)$ *ist ein Gruppenhomomorphismus*

$$\det : (\mathrm{GL}_n(R), \cdot) \to (R^\times, \cdot).$$

Da alle Abbildungsmatrizen eines Endomorphismus φ ähnlich zueinander sind (vgl. 6(6.5.6)), somit dieselbe Determinante haben, können wir diese als die *Determinante von* φ auffassen.

Definition. Für $\varphi \in \mathrm{End}(V)$ setzen wir $\det \varphi := \det M_\varphi^{\mathcal{B}}$, wobei $\mathcal{B}$ eine beliebige Basis von V ist.
(Die Definition ist unabhängig von der Wahl von $\mathcal{B}$.)

7.1.7 Cramer'sche Regel und Adjunktenformel

Satz. *Es sei $A \in \mathrm{GL}_n(R)$.*

a) Cramer'sche Regel:
Es sei $b \in R^n$ und $s_1, \ldots, s_n \in R^n$ bezeichne die Spalten von A. Die eindeutige Lösung von $Ax = b$ lautet

$$x_j := \frac{1}{\det A} \det(s_1, \ldots, s_{j-1}, b, s_{j+1}, \ldots, s_n).$$

b) Adjunktenformel:
Die Inverse von A lautet

$$A^{-1} = \frac{1}{\det A}(\mathrm{cof}_{ij}(A))^t.$$

Beweis. a) Da A invertierbar ist, existiert eine Lösung x von $Ax = b$, d.h. $b = \sum_{i=1}^n x_i s_i$. Es folgt

$$\begin{aligned}\det(s_1, \ldots, s_{j-1}, b, s_{j+1}, \ldots, s_n) &= \sum_{i=1}^n x_i \det(s_1, \ldots, s_{j-1}, s_i, s_{j+1}, \ldots, s_n) \\ &= x_j \det A.\end{aligned}$$

b) Der (i,j)-Eintrag von A^{-1} ist der i-te Eintrag x_i der Lösung von $Ax = e_j$. Nach a) lautet dieser: $x_i = \frac{1}{\det A} \det(s_1, \ldots, s_{i-1}, e_j, s_{i+1}, \ldots, s_n)$. Laplace-Entwicklung nach der i-ten Spalte ergibt:

$$x_i = \frac{1}{\det A}(0 + \ldots + 1 \cdot \mathrm{cof}_{ji}(A) + 0 + \ldots + 0) = \frac{\mathrm{cof}_{ji}(A)}{\det A}.$$

Also $A^{-1} = \frac{1}{\det A}(\mathrm{cof}_{ij}(A))^t$. □

Bemerkung a. Die transponierte Matrix der Kofaktoren $(\mathrm{cof}_{ji}(A))_{ij}$ wird *komplementäre Matrix* oder *Adjunkte* von A genannt, geschr. $\widetilde{A}$ oder $\mathrm{adj}(A)$.

Beispiel. siehe Vorlesung.

Bemerkung b. Falls die Matrix A ganzzahlige Einträge hat, dann liegt ein Vorteil der Cramer'schen Regel und der Adjunktenformel darin, dass die Nenner aller im Lösungsvektor bzw. in der Inversen auftretenden Brüche bereits in dem Term $\frac{1}{\det A}$ stecken. Die restliche Rechnung kommt ohne Brüche aus. Insbesondere sind alle Kofaktoren wieder ganzzahlig.

Übung. Für jedes $A \in R^{n \times n}$ gilt $A\widetilde{A} = \widetilde{A}A = (\det A)E_n$.

7.2 Eigenwerte und Eigenvektoren

In dem gesamten Abschnitt seien K ein Körper, $n \in \mathbb{N}$, $A \in K^{n\times n}$, V ein K-Vektorraum mit $0 < \dim V = n < \infty$ und $\varphi \in \mathrm{End}(V)$.

7.2.1 Das charakteristische Polynom

Definition a. Es sei $A = (a_{ij}) \in K^{n\times n}$. Man nennt

$$\det(X \cdot E_n - A) = \begin{vmatrix} X - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X - a_{22} & \cdots & -a_{2n} \\ \vdots & & \ddots & \\ -a_{n1} & \cdots & & X - a_{nn} \end{vmatrix} \in K[X]$$

das *charakteristische Polynom* von A, geschr. χ_A.

Beispiel.

(i) $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$. $\chi_A = X^3 - 6X^2 + 11X - 6 \in \mathbb{Q}[X]$.
(Rechnung siehe Vorlesung.)

(ii) $\chi_{E_n} = (X-1)^n = X^n - X^{n-1} + \ldots + (-1)^{n-1}X + (-1)^n$.
(Rechnung mit Kästchensatz).

Bemerkung.

(i) Das charakteristische Polynom ist normiert vom Grad n, d.h. hat die Form
$$\chi_A = X^n + c_{n-1}X^{n-1} + \ldots + c_1X + c_0.$$

(ii) Es gilt $c_{n-1} = -\mathrm{Spur}(A)$, wobei $\mathrm{Spur}(A) := a_{11} + \ldots + a_{nn}$.

(iii) Für die Polynomfunktion zu χ_A gilt:
$$\chi_A(\lambda) = \det(\lambda E - A) \quad \text{für alle } \lambda \in K.$$

(iv) Es gilt $c_0 = (-1)^n \det A$.

Beweis. (i) und (ii) ergeben sich aus der Leibniz-Formel, denn X^n und X^{n-1} können in der Leibniz-Formel nur für $\pi = \mathrm{id}$ entstehen.

(iii) folgt aus der Leibniz-Formel und der Tatsache, dass der Einsetzungshomomorphismus τ_λ ein Homomorphismus ist:

$$\chi_A(\lambda) = (\sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{\pi(i),i})(\lambda) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi)(\prod_{i=1}^{n} a_{\pi(i),i})(\lambda)$$
$$= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{\pi(i),i}(\lambda) = |\lambda E - A|.$$

(iv) folgt aus (iii), wenn man $\lambda = 0$ setzt. □

Folgerung. *Es gelten*

(i) $\chi_A = \chi_{A^t}$,

(ii) A, B *ähnlich* $\Rightarrow \chi_A = \chi_B$.

Beweis. (i) Nach Folgerung (7.1.2) gilt $\det A = \det A^t$. Wegen $(XE - A)^t = XE - A^t$ gilt also $\chi_a = \det(XE - A) = \det(XE - A^t) = \chi_{A^t}$.

(ii) Angenommen $A, B \in K^{n\times n}$ sind ähnlich, d.h. es gibt $T \in \mathrm{GL}_n(K)$ mit $B = T^{-1}AT$. Wir zeigen zunächst, dass dann auch $xE-B, xE-A \in K[X]^{n\times n}$ ähnlich sind:

$$xE - B = x(T^{-1}ET) - T^{-1}AT = T^{-1}xET - T^1AT$$
$$= T^{-1}(xET - AT) = T^{-1}(xE - A)T.$$

Nach Folgerung (7.1.6) ist somit $\chi_B = \det(xE-B) = \det(xE-A) = \chi_A$. □

Übung. Gilt auch die Umkehrung von Teil (ii) der Folgerung?

Wegen Teil (ii) der Folgerung ist folgende Definition unabhängig von der Wahl von $\mathcal{B}$.

Definition b. Dann heißt $\chi_\varphi := \chi_{M_\varphi^\mathcal{B}} \in K[X]$ das *charakteristische Polynom von* φ, wobei $\mathcal{B}$ eine beliebige Basis von V ist.

Beispiel. Die Drehung ρ_α von $\mathbb{R}^2$ um α im Uhrzeigersinn hat

$$\det \rho_\alpha = 1, \quad \chi_{\rho_\alpha} = X^2 - 2(\cos\alpha)X + 1.$$

(Rechnung siehe Vorlesung.)

7.2.2 Eigenwerte von Endomorphismen

Definition. Wir definieren für jedes $c \in K$ den Unterraum

$$V(c, \varphi) := \{v \in V \mid \varphi(v) = c \cdot v\} \leq V.$$

Wir nennen c einen *Eigenwert* von φ, wenn $V(c, \varphi) \neq \{\mathfrak{o}\}$. Ist c ein Eigenwert, so heißt $V(c, \varphi)$ der *Eigenraum von* φ. Die Vektoren $\mathfrak{o} \neq v \in V(c, \varphi)$ heißen *Eigenvektoren von* φ *zum Eigenwert* c.

Bemerkung.

(i) Es gilt $V(c, \varphi) = \mathrm{Kern}(\varphi - c \cdot \mathrm{id}) \leq V$.

(ii) Ein Eigenvektor von φ ist ein Vektor, dessen „Richtung“ sich unter der Abbildung nicht ändert.

(iii) Die Eigenvektoren von φ zum Eigenwert 1 sind genau die von $\mathfrak{o}$ verschiedenen Fixpunkte von φ. Demnach ist 1 genau dann ein Eigenwert von φ, wenn φ Fixpunkte $\neq \mathfrak{o}$ hat.

(iv) Die Eigenvektoren von φ zum Eigenwert 0 sind genau die von $\mathfrak{o}$ verschiedenen Elemente von $\mathrm{Kern}\,\varphi$. Demnach ist 0 ist genau dann ein Eigenwert von φ, wenn φ nicht-trivialen Kern hat.

Beispiel.

(i) Die Spiegelung des $\mathbb{R}^2$ an einer Ursprungsgeraden hat die Eigenwerte 1 und -1. Der Eigenraum zu 1 ist die Spiegelgerade, der Eigenraum zu -1 ist die Ursprungsgerade senkrecht zur Spiegelgeraden.

(ii) Die Drehung des $\mathbb{R}^2$ um einen Winkel, der kein Vielfaches von $180°$ ist, hat keine Eigenwerte.

Übung. Welche Eigenwerte haben Projektion und Scherung?

7.2.3 Eigenwerte von Matrizen

Definition. Wir definieren für jedes $c \in K$ den Unterraum

$$V(c, A) := V(c, \varphi_A) = \{x \in K^n \mid Ax = cx\} \leq K^n.$$

Wir nennen c einen *Eigenwert* von A, wenn $V(c, A) \neq \{0\}$. Ist c ein Eigenwert von A, so heißt $V(c, A)$ der *Eigenraum von* A *zum Eigenwert* c. Die Vektoren $0 \neq v \in V(c, A)$ heißen *Eigenvektoren von* A *zum Eigenwert* c.

Bemerkung.

(i) $V(c, A) = \mathbb{L}_0(A - cE) \leq K^n$.

(ii) Die Eigenvektoren zu 1 sind also gerade die nicht-trivialen Lösungen der Gleichung $Ax = x$, und die Eigenvektoren zu 0 sind die nicht-trivialen Elemente aus $\mathbb{L}_0(A)$.

(iii) c ist genau dann ein Eigenwert von A, wenn $\mathrm{Rg}(A - cE) < n$ bzw. wenn $\mathrm{Def}(A - cE) > 0$ bzw. wenn $\det(A - cE) \neq 0$.

Beispiel.

(i) $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^2$ hat keine Eigenwerte, weil $A - cE = \begin{pmatrix} -c & 1 \\ -1 & -c \end{pmatrix}$ die Zeilenstufenform $\begin{pmatrix} 1 & c \\ 0 & 1 + c^2 \end{pmatrix}$ hat, und somit für alle $c \in \mathbb{R}$ den Rang 2 besitzt. Geometrisch interpretiert beschreibt diese Matrix eine Drehung um 90° im Uhrzeigersinn.

Für die komplexen Zahlen $c = i$ und $c = -i$ ist $1 + c^2 = 0$, also hat die Matrix $A - cE$ den Rang 1. Über $\mathbb{C}$ besitzt A somit die Eigenwerte i und $-i$.

(ii) Eine obere Dreiecksmatrix $A = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} \in K^{n \times n}$ hat die Eigenwerte $a_{11}, \ldots, a_{nn}$, denn $A - cE = \begin{pmatrix} a_{11} - c & & * \\ & \ddots & \\ 0 & & a_{nn} - c \end{pmatrix}$ hat genau dann Determinante 0, wenn $c = a_{ii}$ ist für ein $1 \leq i \leq n$.

Die Methode, mittels Gauß-Algorithmus den Rang von $A - cE$ in Abhängigkeit vom Parameter c zu berechnen, kann bei grösseren Matrizen etwas umständlich werden. Eine Alternative bietet das charakteristische Polynom (siehe (7.2.5) unten).

7.2.4 Berechnung der Eigenräume

Bemerkung. Eigenräume von Endomorphismen und Matrizen hängen offensichtlich zusammen, und zwar über Koordinatenabbildungen. Ist $\mathcal{B}$ eine geordnete Basis von V, so gilt für jedes $\varphi \in \mathrm{End}_K(V)$:

$$\kappa_{\mathcal{B}}(V(c, \varphi)) = V(c, M_\varphi^{\mathcal{B}}).$$

Somit lässt sich die Berechnung von $V(c,\varphi)$ stets auf die Berechnung von $V(c, M_\varphi^{\mathcal{B}})$, also auf den Eigenraum einer Matrix, zurückführen. Insbesondere folgt, dass φ und $M_\varphi^{\mathcal{B}}$ diesselben Eigenwerte haben.

Beweis. Da $\kappa_{\mathcal{B}}$ ein Isomorphismus ist, und unter Benutzung von $\kappa_{\mathcal{B}}(\varphi(v)) = M_\varphi^{\mathcal{B}} \cdot \kappa_{\mathcal{B}}(v)$, gilt:

$$\begin{aligned} v \in V(c,\varphi) &\Leftrightarrow \varphi(v) = cv \Leftrightarrow \kappa_{\mathcal{B}}(\varphi(v)) = \kappa_{\mathcal{B}}(cv) \\ &\Leftrightarrow M_\varphi^{\mathcal{B}} \cdot \kappa_{\mathcal{B}}(v) = c\kappa_{\mathcal{B}}(v) \Leftrightarrow \kappa_{\mathcal{B}}(v) \in V(c, M_\varphi^{\mathcal{B}}). \end{aligned}$$

□

Wir berechnen die Eigenräume in Beispielen, deren Eigenwerte wir schon kennen.

Beispiel.

(i) Es sei σ_0 die Spielgung von $\mathbb{R}^2$ an der e_1-Achse, von der wir wissen, dass sie die Eigenwerte 1 und -1 besitzt. Die Abbildungsmatrix von σ_0 bzgl. der Standardbasis lautet $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Wir berechnen $V(-1,\sigma_0) = V(-1,A)$. Aus

$$A - (-1)E = A + E = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

liest man $V(-1,A) = \mathbb{L}_0(A-(-1)E) = \langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle = \langle e_2 \rangle$ ab. Wie vermutet, ergibt sich als Eigenraum zum Eigenwert 1 also die Ursprungsgerade senkrecht zur Spiegelgeraden.

(ii) $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{C}^{2\times 2}$ hat die Eigenwerte i und $-i$.

$V(i,A)$:

$$\begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix}, \text{ also } V(i,A) = \langle \begin{pmatrix} i \\ -1 \end{pmatrix} \rangle.$$

$V(-i,A)$:

$$\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix}, \text{ also } V(-i,A) = \langle \begin{pmatrix} i \\ 1 \end{pmatrix} \rangle.$$

(iii) Es sei σ_α die Spielgung von $\mathbb{R}^2$ an der um α gedrehten e_1-Achse. Nach Beispiel 6.5.3 hat σ_α bzgl. der Standardbasis die Matrix

$$S_\alpha = \begin{pmatrix} \cos^2\alpha - \sin^2\alpha & 2\sin\alpha\cos\alpha \\ 2\sin\alpha\cos\alpha & \sin^2\alpha - \cos^2\alpha \end{pmatrix}.$$

Als Übung berechne man den Eigenraum zum Eigenwert 1. (Wir wissen bereits, dass die Spiegelgerade, also die um α gedrehte e_1-Achse, herauskommen muss.)

7.2.5 Eigenwerte als Nullstellen von χ

Satz. *Die Eigenwerte von φ bzw. A sind genau die Nullstellen des charakteristischen Polynoms χ_φ bzw. χ_A.*

Beweis. Für A: Genau dann ist c Eigenwert von A, wenn $\mathbb{L}_0(A - cE)$ nichttrivial ist, d.h. genau dann, wenn $\det(A - cE) = 0$ bzw. $\det(cE - A) = 0$. Nach Bemerkung (7.2.1) ist $\det(cE - A) = \chi_A(c)$. Somit ist (i) gezeigt.

Für φ: Setze $A = M_\varphi^{\mathcal{B}}$ für eine beliebige Basis $\mathcal{B}$. Per Definition ist $\chi_\varphi = \chi_A$. Nach Bemerkung (7.2.4) haben φ und A diesselben Eigenwerte. Damit ist alles gezeigt. □

Folgerung.

(i) Es gibt höchstens n verschiedene Eigenwerte von A und von φ.

(ii) A und A^t haben diesselben Eigenwerte.

(iii) Ähnliche Matrizen haben gleiche Eigenwerte.

Beweis. Das charakteristische Polynom hat Grad n und damit höchstens n verschiedene Nullstellen. A und A^t haben dasselbe charakteristische Polynom. Das gleiche trifft auf ähnliche Matrizen zu. □

Übung. Haben A, A^t und zu A ähnliche Matrizen auch dieselben Eigenvektoren wie A? Wenn nicht, finde man ein Gegenbeispiel.

Beispiel. Wir verifizieren den Satz anhand der bisherigen Beispiele aus (7.2.3) und (7.2.4):

(i) $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $\chi_A = X^2 + 1$. Dieses Polynom hat keine Nullstellen in $\mathbb{R}$, zerfällt aber über $\mathbb{C}$ in $\chi_A = (X + i)(X - i)$.

(ii) Die Spiegelungsmatrix $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ hat $\chi_A = (X-1)(X+1)$. Dieses Polynom hat die Nullstellen ± 1.

(iii) $A = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$, $\chi_A = (X - a_{11}) \cdots (X - a_{nn})$. Dieses Polynom hat die Nullstellen $a_{11}, \ldots, a_{nn}$.

(iv) $\chi_{E_n} = (X-1)^n$.

(v) Als Übung berechne man das charakteristische Polynom der Matrix

$$S_\alpha = \begin{pmatrix} \cos^2\alpha - \sin^2\alpha & 2\sin\alpha\cos\alpha \\ 2\sin\alpha\cos\alpha & \sin^2\alpha - \cos^2\alpha \end{pmatrix}$$

aus Beispiel (7.2.4). (Es muss $(X-1)(X+1)$ herauskommen, weil S_α eine Spiegelung beschreibt, und somit ± 1 die einzigen Eigenwerte sind.)

7.2.6 Vielfachheit von Eigenwerten

Definition. Es sei c ein Eigenwert von A bzw. von φ. Die Vielfachheit von c als Nullstelle von χ_A bzw. χ_φ wird *(algebraische) Vielfachheit* von c genannt, geschr. $a_c(A)$ bzw. $a_c(\varphi)$.

Die Dimension von $V(c, A)$ bzw. $V(c, \varphi)$ wird *geometrische Vielfachheit* von c genannt, geschr. $g_c(A)$ bzw. $g_c(\varphi)$.

Beispiel a. Wir betrachten die Spiegelung des $\mathbb{R}^3$ an der e_1-e_2-Ebene. Der Eigenraum zum Eigenwert 1 ist die e_1-e_2-Ebene, also 2-dimensional. Somit hat der Eigenwert 1 die geometrische Vielfachheit 2.

Die Abbildungsmatrix bzgl. (e_1, e_2, e_3) lautet $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Also ist $\chi = (x+1)(x-1)^2$, und die algebraische Vielfachheit von 1 ist ebenfalls 2.

Die Aussage aus Folgerung (7.2.5) gilt auch, wenn man jeden Eigenwert mit seiner Vielfachheit zählt:

Folgerung. *Es gilt stets $\sum_c a_c(\varphi) \leq n$ bzw. $\sum_c a_c(A) \leq n$, wobei die Summe über alle Eigenwerte c gebildet wird.*

Beweis. Zählt man die Nullstellen mit ihrer Vielfachheit, so hat ein Polynom vom Grad n höchstens n Nullstellen. □

Beispiel b. Wir berechnen alle Eigenwerte, Eigenräume und Vielfachheiten der Eigenwerte von $A = \begin{pmatrix} -3 & 0 & 0 \\ 2 & -3 & 1 \\ 10 & 0 & 2 \end{pmatrix}$ bzw. $A = \begin{pmatrix} -3 & 0 & 0 \\ 1 & -3 & 1 \\ 10 & 0 & 2 \end{pmatrix}$.
(Rechnung siehe Vorlesung.)

Satz. *Es gilt stets $g_c(\varphi) \leq a_c(\varphi)$ bzw. $g_c(A) \leq a_c(A)$.*

Beweis. Wir zeigen die Aussage für φ (die Aussage für A folgt daraus, indem man φ_A betrachtet). Es sei c ein Eigenwert von φ, und $g = g_c(\varphi)$ sei seine geometrische Vielfachheit. Wir wählen eine geordnete Basis $(v_1, \ldots, v_g)$ von $V(c, \varphi)$ und ergänzen diese zu einer Basis $\mathcal{B} := (v_1, \ldots, v_n)$ von V. Dann hat $M := M_\varphi^{\mathcal{B}}$ die Form

$$M = \left(\begin{array}{ccc|c} c & & 0 & \\ & \ddots & & * \\ 0 & & c & \\ \hline & 0 & & B \end{array}\right), \text{ und } XE - M = \left(\begin{array}{ccc|c} X-c & & 0 & \\ & \ddots & & * \\ 0 & & X-c & \\ \hline & 0 & & XE-B \end{array}\right),$$

wobei oben links jeweils ein $g \times g$-Block steht. Nach dem Kästchensatz folgt $\chi_\varphi = \det(XE - M) = (x-c)^g \det(XE - B) = (x-c)^g \chi_B$. Nach Definition der algebraischen Vielfachheit ist somit $a_c(\varphi) \geq g$. □

Übung. Haben A, A^t und zu A ähnliche Matrizen dieselben geometrischen Vielfachheiten?

7.2.7 Spiegelungen

Definition. Es sei $1 \neq -1$ in K. Ein Endomorphismus $\varphi \in \mathrm{End}(V)$ heißt eine *Spiegelung*, falls gelten:

(i) 1 und -1 sind Eigenwerte von φ, und

(ii) $g_1(\varphi) = n - 1$.

(φ ist Spiegelung an $V(1, \varphi)$, der sogenannten *Spiegelungshyperebene*.)

Bemerkung. Es sei φ eine Spiegelung. Wähle $0 \neq v_1 \in V(-1, \varphi)$ und wähle eine geordnete Basis $(v_2, ..., v_n)$ von $V(1, \varphi)$. Wegen $\varphi(v_1) = -v_1 \neq v_1$ ist $v_1 \notin V(1, \varphi)$. Daraus folgt, dass $\mathcal{B} := (v_1, ..., v_n)$ linear unabhängig, wegen $\dim V = n$ also sogar Basis von V ist. Da alle Basisvektoren Eigenvektoren

sind, ist $M_\varphi^{\mathcal{B}}$ eine Diagonalmatrix, nämlich

$$M_\varphi = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}.$$

7.3 Diagonalisierbarkeit

In dem gesamten Abschnitt seien K ein Körper, $n \in \mathbb{N}$, $A \in K^{n\times n}$, V ein K-Vektorraum mit $0 < \dim V = n < \infty$ und $\varphi \in \mathrm{End}(V)$.

7.3.1 Diagonalisierbare Endomorphismen und Matrizen

Bemerkung a. Eine Basis $\mathcal{A}$ von V besteht genau dann aus Eigenvektoren von A, wenn $M_\varphi^{\mathcal{A}}$ eine Diagonalmatrix ist.

Definition.

(i) φ heißt *diagonalisierbar*, wenn eine Basis von V existiert, die aus Eigenvektoren von φ besteht.

(ii) A heißt *diagonalisierbar*, wenn A ähnlich zu einer Diagonalmatrix ist. Ist etwa $T^{-1}AT$ eine Diagonalmatrix mit $T \in \mathrm{GL}_n(K)$, so sagt man A *wird durch* T *diagonalisiert.*

Satz. *Für jede Basis* $\mathcal{B}$ *von* V *gilt:*

$$\varphi \textit{ diagonalisierbar} \Leftrightarrow M_\varphi^{\mathcal{B}} \textit{ diagonalisierbar.}$$

Beweis. $\Rightarrow$: Sei φ diagonalisierbar. Wähle eine Basis $\mathcal{A}$ von V aus Eigenvektoren und setze $T := {}^{\mathcal{B}}T^{\mathcal{A}}$. Dann ist $T^{-1}M_\varphi^{\mathcal{B}}T = M_\varphi^{\mathcal{A}}$ eine Diagonalmatrix, also ist $M_\varphi^{\mathcal{B}}$ diagonalisierbar.

$\Leftarrow$: Sei $M_\varphi^{\mathcal{B}}$ durch $T \in \mathrm{GL}_n(K)$ diagonalisierbar, d.h. $T^{-1}M_\varphi^{\mathcal{B}}T$ sei eine Diagonalmatrix. Nach Folgerung 6.5.4 gibt es eine Basis $\mathcal{A}$ von V mit ${}^{\mathcal{B}}T^{\mathcal{A}} = T$. Dann ist $M_\varphi^{\mathcal{A}} = T^{-1}M_\varphi^{\mathcal{B}}T$ eine Diagonalmatrix, also ist φ diagonalisierbar. □

Bemerkung b. A ist genau dann diagonalisierbar, wenn φ_A diagonalisierbar ist, also genau dann, wenn eine Basis von K^n aus Eigenvektoren von A existiert.

Ist $\mathcal{A} = (v_1, \ldots, v_n)$ eine solche Basis aus Eigenvektoren von A, so wird A durch $T := {}^{\mathcal{E}}T^{\mathcal{A}}$ diagonalisiert (die Spalten von T lauten $v_1, \ldots, v_n$). Genauer: $T^{-1}AT = \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix}$, wobei v_i Eigenvektor zum Eigenwert c ist.

Beweis. Dies folgt aus dem Satz, indem man $\varphi = \varphi_A$ und $\mathcal{B} = \mathcal{E}$ wählt. Da v_i Eigenvektor von A zum Eigenwert c_i ist, hat $M^{\mathcal{A}}_{\varphi_A} = T^{-1}AT$ die besagte Diagonalform. □

Beispiel a. Ist $A = \begin{pmatrix} -3 & 0 & 0 \\ 2 & -3 & 1 \\ 10 & 0 & 2 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$ diagonalisierbar?

Aus Beispiel (7.2.6) sind die Eigenvektoren $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$ zum Eigenwert -3 und $\begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix}$ zum Eigenwert 2 bekannt. Da diese drei Vektoren linear unabhängig sind, also eine Basis bilden, wird A durch die Matrix $T = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 1 & 0 \\ 5 & 0 & 2 \end{pmatrix}$ diagonalisiert: $T^{-1}AT = \begin{pmatrix} 2 & & \\ & -3 & \\ & & -3 \end{pmatrix}$.

Beispiel.

(i) $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in K^{2\times 2}$ nicht diagonalisierbar.

(ii) $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in K^{2\times 2}$ nicht diagonalisierbar für $K = \mathbb{R}$, aber diagonalisierbar für $K = \mathbb{C}$.

Anschaulich beschreibt (i) eine *Scherung* und (ii) eine Drehung um 90°. Für beides gibt es über $\mathbb{R}$ keine Basis aus Eigenvektoren.

Beweis. (i) Wegen $\chi_A = (X-1)^2$ lautet der einzig mögliche Eigenwert 1. Wäre A diagonalisierbar, so gäbe es ein $T \in \mathrm{GL}_2(K)$, mit $T^{-1}AT = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2$. Daraus folgt aber der Widerspruch $A = E_2$.

(ii) Nach Beispiel (7.2.4) hat A überhaupt keine Eigenvektoren über $\mathbb{R}$, also kann auch keine Basis aus Eigenvektoren existieren. Über $\mathbb{C}$ existieren

aber die beiden linear unabhängigen Eigenvektoren $\begin{pmatrix} i \\ -1 \end{pmatrix}$ zu i und $\begin{pmatrix} i \\ 1 \end{pmatrix}$ zu $-i$. Somit gilt $T^{-1}AT = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ für $T = \begin{pmatrix} i & i \\ -1 & 1 \end{pmatrix}$. □

7.3.2 Kriterien

Wir beschränken uns auf Matrizen; für Endomorphismen geht alles analog.

Erinnerung an die Vielfachheiten von Eigenwerten. Für jeden Eigenwert c gilt:

$$1 \leq g_c(A) \leq a_c(A) \leq n.$$

Ist l die Anzahl der verschiedenen Eigenwerte von A, so folgt:

$$l \leq \sum_c g_c(A) \leq \sum_c a_c(A) \leq n. \tag{7.2}$$

Satz. *Folgende Aussagen sind äquivalent:*

(i) *A ist diagonalisierbar.*

(ii) $\sum_c g_c(A) = n$.

(iii) *χ_A zerfällt vollständig in Linearfaktoren und für jeden Eigenwert c von A ist $g_c(A) = a_c(A)$.*

Beweis. (ii)⇔(iii): Wegen (7.2) ist (ii) äquivalent zu $\sum_c g_c(A) = \sum_c a_c(A)$ und $\sum_c a_c(A) = n$. Ersteres ist äquivalent dazu, dass $g_c(A) = a_c(A)$ für jeden Eigenwert c gilt. Letzteres ist äquivalent dazu, dass χ_A vollständig in Linearfaktoren zerfällt.

(i)⇒(ii): Sei A diagonalisierbar. Die Basisvektoren einer Basis aus Eigenvektoren stammen aus den Eigenräumen. Jeder Eigenraum liefert höchstens $g_c(A)$ linear unabhängige Vektoren. Damit folgt $n \leq \sum_c g_c(A)$ und wegen (7.2) auch die Gleichheit.

(ii)⇒ (i): Sei $\sum_c g_c(A) = n$. Wähle zu jedem Eigenwert c eine Basis B_c des Eigenraums $V(c, A)$. Dann ist $|B_c| = g_c(A)$. Nach folgendem Lemma ist $B := \cup_c B_c$ linear unabhängig und $|B| = \sum_c g_c(A) = n$, also B eine Basis aus Eigenvektoren von A. □

Übung.

(i) Eigenräume zu paarweise verschiedenen Eigenwerten sind *disjunkt*, d.h. der Schnitt ist gleich $\{0\}$.

(ii) Eigenvektoren zu paarweise verschiedenen Eigenwerten sind linear unabhängig.

Lemma. *Es seien $c_1, \ldots, c_l$ paarweise verschiedene Eigenwerte von A. Seien $B_i \subseteq V(c_i, A)$ linear unabhängig, $i = 1, \ldots, l$. Dann sind die B_i paarweise disjunkt und $B_1 \cup \ldots \cup B_l$ ist linear unabhängig.*

Beweis. Nach Teil (i) der Übung sind die B_i paarweise disjunkt. Angenommen $\sum_{j=1}^{m} \lambda_j v_j = 0$ ist eine lineare Abhängigkeit in B, d.h. $v_1, \ldots, v_m \in B$ paarweise verschieden, $m \geq 1$, und $\lambda_1, \ldots, \lambda_m \in K \setminus \{0\}$. Durch Zusammenfassen der Summanden aus jeweils demselben Eigenraum $V(c_i, A)$ bekommen wir eine Summe $w_1 + \ldots + w_l = 0$ mit $w_i \in V(c_i, A)$. Es sind nicht alle $w_i = 0$, denn sonst wäre für dasjenige i_0 mit $v_1 \in B_{i_0}$ die Gleichungen $w_{i_0} = 0$ eine lineare Abhängigkeit in B_{i_0}. Somit ist $\{w_1, \ldots, w_l\}$ linear abhängig, im Widerspruch zu Teil (ii) der Übung. Also ist die Annahme falsch, d.h. B ist linear unabhängig. □

Beispiel. Wir betrachten die Matrizen aus den Beispielen (7.3.1) erneut.

(i) $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in K^{2\times 2}$ ist nicht diagonalisierbar, obwohl $\chi_A = (X-1)^2$ vollständig zerfällt. Man rechnet leicht nach, dass $g_1(A) = 1 < 2 = a_1(A)$ ist.

Die Tatsache, dass χ_A vollständig in Linearfaktoren zerfällt, ist also allein nicht hinreichend für die Diagonalisierbarkeit von A.

(ii) $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ ist nicht diagonalisierbar, weil $\chi_A = X^2 + 1$ nicht vollständig zerfällt.

(iii) Für $A = \begin{pmatrix} -3 & 0 & 0 \\ 2 & -3 & 1 \\ 10 & 0 & 2 \end{pmatrix}$ und $B = \begin{pmatrix} -3 & 0 & 0 \\ 1 & -3 & 1 \\ 10 & 0 & 2 \end{pmatrix}$ gilt $\chi_A = \chi_B = (X-2)(X+3)^2$. Nach Beispiel (7.2.6) ist $g_2(A) + g_{-3}(A) = 1 + 2 = 3$ und $g_2(B) + g_{-3}(B) = 1 + 1 = 2$. Somit ist A diagonalisierbar und B nicht.

Insbesondere sind A und B nicht ähnlich, haben aber gleiches charakteristisches Polynom.

7.3.3 Ein hinreichendes Kriterium

Folgerung. *Wenn χ_A vollständig in paarweise verschiedene Linearfaktoren zerfällt, dann ist A diagonalisierbar.*

Beweis. In diesem Fall ist $a_c(A) = 1$ für alle Eigenwerte, also offensichtlich $g_c(A) = a_c(A)$. Die Aussage ergibt sich also aus Satz (7.3.2). □

Beispiel a. $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 2 & -1 \end{pmatrix}$ hat $\chi_A = (X^2 - 2)(X + 1)$. Über $\mathbb{Q}$ zerfällt dieses Polynom nicht vollständig in Linearfaktoren ($\sqrt{2} \notin \mathbb{Q}$), somit ist A nicht diagonalisierbar über $\mathbb{Q}$.

Über $\mathbb{R}$ dagegen zerfällt das Polynom vollständig in paarweise verschiedene Linearfaktoren ($\chi_A = (X - \sqrt{2})(X + \sqrt{2})(X + 1)$), somit ist A diagonalisierbar über $\mathbb{R}$.

Beispiel b. $A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \in K^{2\times 2}$ hat $\chi_A = (X - 1)(X + 1)$.

Ist $1 \neq -1$ (z.B. für $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ oder $\mathbb{F}_p$ mit $p \neq 2$), so zerfällt also χ_A in paarweise verschiedene Linearfaktoren, also ist A diagonalisierbar.

Ist dagegen $1 = -1$ (z.B. in $K = \mathbb{F}_2, F_4, \ldots$), so ist das nicht der Fall. Die Matrix ist auch nicht diagonalisierbar, wie bereits in Beispiel (7.3.1) gezeigt wurde.

7.3.4 Triangulierbarkeit

Definition.

(i) φ heißt *triangulierbar*, wenn eine Basis von V existiert, bzgl. der die Abbildungsmatrix von φ eine obere Dreiecksmatrix ist.

(ii) A heißt *triangulierbar*, wenn A ähnlich zu einer oberen Dreiecksmatrix ist.
Ist etwa $T^{-1}AT$ eine obere Dreiecksmatrix mit $T \in \mathrm{GL}_n(K)$, so sagt man *A wird durch T trianguliert.*

Bemerkung. Für jede beliebige Basis $\mathcal{A}$ von V gilt:

$$\varphi \text{ triangulierbar } \Leftrightarrow M_\varphi^{\mathcal{A}} \text{ triangulierbar.}$$

Satz. *A ist genau dann triangulierbar, wenn χ_A vollständig in Linearfaktoren zerfällt.*

Beweis. Sei A triangulierbar, also ähnlich zu einer oberen Dreiecksmatrix D mit den Diagonaleinträgen $d_1, \ldots, d_n$. Laut Kästchensatz ist $\chi_A = \chi_D = (X - d_1) \cdots (X - d_n)$, wobei $d_1, \ldots, d_n$ die Diagonaleinträge von D sind.

Wir zeigen die Umkehrung mittels Induktion nach n. Der Induktionsanfang $n = 1$ ist trivial. Sei nun $n > 1$ und die Aussage für $n - 1$ bereits bewiesen. Es zerfalle χ_A vollständig, etwa $\chi_A = (X - c_1) \ldots (X - c_n)$.

Wegen $\chi_A(c_1) = 0$ ist c_1 ein Eigenwert von A. Wähle einen Eigenvektor $v_1 \in K^n$ von A zum Eigenwert c_1 und ergänze diesen zu einer geordneten Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von K^n. Dann hat $M_{\varphi_A}^{\mathcal{B}}$ die Form $\left(\begin{array}{c|c} c_1 & * \\ \hline 0 & D \end{array}\right) \in K^{n\times n}$ mit $D \in K^{(n-1)\times(n-1)}$. Also (Kästchensatz):

$$(X - c_1)(X - c_2)\cdots(X - c_n) = \chi_A = \chi_{M_{\varphi_A}^{\mathcal{B}}} = (X - c_1)\chi_D.$$

Per Kürzungsregel (im nullteilerfreien Polynomring $K[X]$) folgt $\chi_D = (X - c_2)\cdots(X - c_n)$, d.h. χ_D zerfällt vollständig in Linearfaktoren. Nach Induktionsvoraussetzung gibt es $S \in \mathrm{GL}_{n-1}(K)$ so, dass $S^{-1}DS$ eine obere Dreiecksmatrix ist. Setze $T := \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S \end{array}\right) \in K^{n\times n}$. Dann ist $\det T = \det S \neq 0$ (Kästchensatz), d.h. $T \in \mathrm{GL}_n(K)$. Weiter ist $T^{-1} = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S^{-1} \end{array}\right)$ und

$$\begin{aligned} T^{-1}M_{\varphi_A}^{\mathcal{B}}T &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S^{-1} \end{array}\right)\left(\begin{array}{c|c} c_1 & * \\ \hline 0 & D \end{array}\right)\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S \end{array}\right) \\ &= \left(\begin{array}{c|c} c_1 & * \\ \hline 0 & SD \end{array}\right)\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S \end{array}\right) = \left(\begin{array}{c|c} c_1 & * \\ \hline 0 & S^{-1}DS \end{array}\right) \end{aligned}$$

Also ist $M_{\varphi_A}^{\mathcal{B}}$, und somit A, triangulierbar. □

Folgerung. *Über $\mathbb{C}$ ist jede quadratische Matrix triangulierbar.*

Beweis. Der Fundamentalsatz der Algebra. □

7.3.5 Begleitmatrix

Sei $f = X^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0 \in K[X]$ ein normiertes Polynom.

Definition. Die *Begleitmatrix* von f ist definiert als

$$C(f) := \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

Satz. $\chi_{C(f)} = f$.

Beweis. Übung. □

7.3.6 Anwendung: lineare rekursive Folgen

Definition. Eine Folge (a_n) in K, die definiert ist durch eine Rekursionsgleichung

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k}$$

sowie durch die Anfangswerte $a_0, \ldots, a_{k-1}$, heißt *lineare rekursive Folge*. Das normierte Polynom

$$f := X^k - c_1 X^{k-1} - \ldots - c_{k-1} X - c_k$$

heißt *charakteristisches Polynom* der Folge (a_n).

Bemerkung. Sei (a_n) eine lineare rekursive Folge mit charakteristischem Polynom $f = X^k - c_1 X^{k-1} - \ldots - c_{k-1} X - c_k$.

(i) Es gilt für alle $n \geq k+1$:

$$\begin{pmatrix} a_{n-k+1} \\ \vdots \\ a_n \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ c_k & c_{k-1} & \cdots & c_2 & c_1 \end{pmatrix}}_{=:C} \begin{pmatrix} a_{n-k} \\ \vdots \\ a_{n-1} \end{pmatrix} = C^{n-k+1} \begin{pmatrix} a_0 \\ \vdots \\ a_{k-1} \end{pmatrix},$$

bzw. für alle $n \in \mathbb{N}_0$:

$$\begin{pmatrix} a_n \\ \vdots \\ a_{n+k-1} \end{pmatrix} = C^n \begin{pmatrix} a_0 \\ \vdots \\ a_{k-1} \end{pmatrix}.$$

(ii) $C = C(f)^t$ und $\chi_C = f$.

(iii) Ist C diagonalisierbar, etwa $T^{-1}CT = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_k \end{pmatrix} =: D$ mit $T \in \mathrm{GL}_n(K)$, so folgt

$$C^n = (TDT^{-1})^n = TD^nT^{-1} = T \begin{pmatrix} d_1^n & & \\ & \ddots & \\ & & d_k^n \end{pmatrix} T^{-1}.$$

Daraus ergibt sich eine geschlossene Formel für a_n.

Beispiel (Die Fibonacci-Folge)**.** Wir betrachten die Rekursionsgleichung $a_n = a_{n-1} + a_{n-2}$ mit den Anfangsgliedern $a_0 = 0, a_1 = 1$. Das charakteristische Polynom lautet $f = X^2 - X - 1$. Für alle $n \in \mathbb{N}_0$ gilt

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = C^n \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Die Nullstellen von f (Eigenwerte von C) lauten $c_1 = \frac{1+\sqrt{5}}{2}$ und $c_2 = \frac{1-\sqrt{5}}{2}$. Eigenvektoren zu c_i berechnet man so:

$$C - c_i E = \begin{pmatrix} -c_i & 1 \\ 1 & 1-c_i \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1-c_i \\ 0 & c_i - c_i^2 + 1 \end{pmatrix} = \begin{pmatrix} 1 & 1-c_i \\ 0 & 0 \end{pmatrix},$$

also $v_i = \begin{pmatrix} c_i - 1 \\ 1 \end{pmatrix}$. Somit ist

$$D := T^{-1}CT = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} \text{ für } T = \begin{pmatrix} c_1 - 1 & c_2 - 1 \\ 1 & 1 \end{pmatrix}.$$

Man berechnet a_n als den $(1,2)$-Eintrag von $C^n = TD^nT^{-1}$:

$$\begin{aligned} TD^n &= \begin{pmatrix} c_1^n(c_1-1) & c_2^n(c_2-1) \\ c_1^n & c_2^n \end{pmatrix} \\ \det T &= (c_1-1) - (c_2-1) = c_1 - c_2 = \sqrt{5} \\ T^{-1} &= \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 1-c_2 \\ -1 & c_1 - 1 \end{pmatrix} \\ a_n &= \frac{1}{\sqrt{5}} \left(c_1^n(c_1-1)(1-c_2) + c_2^n(c_2-1)(c_1-1)\right) \end{aligned}$$

Wegen $-1 = f(1) = (1-c_1)(1-c_2)$ folgt die Formel

$$a_n = \frac{1}{\sqrt{5}}(c_1^n - c_2^n) = \frac{1}{\sqrt{5}} \left((\frac{1+\sqrt{5}}{2})^n - (\frac{1-\sqrt{5}}{2})^n \right).$$

7.4 Der PageRank-Algorithmus

7.4.1 Einleitung und Idee

Gegeben seien n Webseiten $S_1, \ldots, S_n$, die sich untereinander verlinken, wobei wir keine Links einer Seite auf sich selbst zulassen. Wir veranschaulichen die Situation durch einen gerichteten Graphen ohne Schleifen mit Knotenmenge $\{S_1, \ldots, S_n\}$ und Kantenmenge $\{S_j \to S_i \,|\, S_j \text{ verlinkt auf } S_i\}$. Es sei

n_j die Anzahl der Kanten, die von S_j ausgehen. Wir nehmen an, dass $n_j \geq 1$ für jedes j (Bemerkung 7.4.4 unten erklärt, wie man auf diese Voraussetzung verzichten kann). Definiere die *Link-Matrix* $L = (l_{ij})_{ij} \in \mathbb{R}^{n\times n}$ durch

$$l_{ij} := \begin{cases} \frac{1}{n_j} & \text{falls } S_j \text{ auf } S_i \text{ verlinkt und } i \neq j, \\ 0 & \text{sonst.} \end{cases}$$

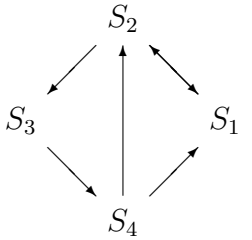
Die j-te Spalte von L enthält die Links, die von der Seite S_j ausgehen, und die Spaltensumme ist $\sum_{i=1}^{n} l_{ij} = n_j \cdot \frac{1}{n_j} = 1$ für alle j. In der Praxis ist L *dünn besetzt*, d.h. enthält viele Nullen.

Idee. Die Seiten stimmen selbst über ihre Wichtigkeit ab.

1. Ansatz: Jede Seite hat eine Stimme, die sie gleichmässig auf diejenigen Seiten verteilt, die von ihr verlinkt werden. Das *Gewicht* (= Wichtigkeit) der Seite S_i ergibt sich dann als die Zeilensumme $x_i := \sum_{j=1}^{n} l_{ij}$. Problem: "Unwichtige Seiten", die sich gegenseitig verlinken, werden wichtig.

2. Ansatz: Wie 1., aber jede Seite hat genau so viele Stimmen, wie ihrem Gewicht entspricht. Das Gewicht von S_i lautet dann $x_i := \sum_{j=1}^{n} l_{ij}x_j$. Problem: Wie wird die Selbst-Reflexivität aufgelöst? Lösung: Der *Gewichtsvektor* $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ muss die Bedingung $x = Lx$ erfüllen, ist also Eigenvektor von L zum Eigenwert 1.

Beispiel.



$$L = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad x = \begin{pmatrix} 3 \\ 4 \\ 2 \\ 2 \end{pmatrix} = Lx$$

Frage. Gibt es immer einen Eigenvektor zum Eigenwert 1? Gibt es einen Eigenvektor mit Einträgen ≥ 0? Ist er eindeutig?

7.4.2 Markov-Matrizen

Definition. Eine beliebige reelle Matrix $M \in \mathbb{R}^{m\times n}$ heißt *positiv* (bzw. *negativ, nicht-negativ, nicht-positiv*), geschrieben $M > 0$ (bzw. $M < 0$, $M \geq 0$,

$M \leq 0$), wenn alle Einträge > 0 (bzw. < 0, ≥ 0, ≤ 0) sind. Wir definieren $l(M)$ als die Summe aller Einträge.

Eine quadratische nicht-negative reelle Matrix M heißt *Markov-Matrix*, wenn $l(s) = 1$ für jede Spalte s von M.

Satz. *Jede Markov-Matrix $M \in \mathbb{R}^{n\times n}$ hat einen nicht-negativen Eigenvektor zum Eigenwert* 1.

Beweis. Die Matrix M^t hat Zeilensummen gleich 1, also ist $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ ein Eigenvektor zum Eigenwert 1. Nach Folgerung 7.2.5 hat damit auch M den Eigenwert 1.

Sei nun $M = (a_{ij})_{ij}$ und $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ ein Eigenvektor von A zum Eigenwert 1. Wegen $x \neq 0$ ist mindestens ein $x_i \neq 0$. Wir nehmen an, dass ein $x_i > 0$ (sonst ersetze x durch $-x$).

Durch Permutieren der Basisvektoren der Standardbasis können wir erreichen, dass $x_1, \ldots, x_r \leq 0$ und $x_{r+1}, \ldots, x_n > 0$ mit $0 \leq r \leq n-1$. (Eine Permutation der Standardbasis ist ein Basiswechsel, der auf A die Anwendung derselben Permutation auf die Spalten und ihrer Inversen auf die Zeilen bewirkt. So entsteht wieder eine Markov-Matrix.) Wir zerlegen nun M und x in Blöcke $M = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right), x = \begin{pmatrix} y \\ \hline z \end{pmatrix}$ derart, dass A quadratisch ist und A und y genau r Zeilen haben. Dann sind $A, B, C, D \geq 0, y \leq 0$ und $z > 0$. Aus $Mx = x$ folgt $Ay + Bz = y$, also $l(y) = l(Ay + Bz) = l(Ay) + l(Bz)$. Offensichtlich gilt

$$l(\begin{pmatrix} A \\ \hline C \end{pmatrix} y) = l(\begin{pmatrix} Ay \\ \hline Cy \end{pmatrix}) = l(Ay) + l(Cy).$$

Wegen $\sum_{i=1}^{n} a_{ij} = 1$ für alle $j = 1, \ldots, n$ ist andererseits

$$l(\begin{pmatrix} A \\ \hline C \end{pmatrix} y) = \sum_{i=1}^{n}(\sum_{j=1}^{r} a_{ij}x_j) = \sum_{j=1}^{r}(\sum_{i=1}^{n} a_{ij})x_j = \sum_{j=1}^{r} x_j = l(y).$$

Durch Gleichsetzen von $l(y)$ folgt $l(Cy) = l(Bz)$. Wegen $Cy \leq 0$ und $Bz \geq 0$ ist das nur möglich, wenn $Cy = 0$ und $Bz = 0$. Man rechnet leicht nach, dass dann mit $\begin{pmatrix} y \\ \hline z \end{pmatrix}$ auch $\begin{pmatrix} -y \\ \hline z \end{pmatrix} \geq 0$ ein Eigenvektor zum Eigenwert 1 ist. □

Beispiel. Die Matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ zeigt, dass der nicht-negative Eigenvektor x zum Eigenwert 1 nicht eindeutig sein muss, denn $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ haben beide diese Eigenschaft.

Das Phänomen tritt z.B. auf, wenn der "Link-Graph" nicht zusammenhängend ist. Dann hat die Link-Matrix (bei geeigneter Nummerierung der Seiten) eine Blockstruktur der Form

$$\begin{pmatrix} L_1 & & 0 \\ & \ddots & \\ 0 & & L_k \end{pmatrix}$$

7.4.3 Markov-Prozesse

Es sei M eine $n \times n$ Markov-Matrix, $v \in \mathbb{R}^n$.

Bemerkung. Für alle $i \geq 0$ gilt:

(i) $v \geq 0 \Rightarrow M^i v \geq 0$

(ii) $l(M^i v) = l(v)$

Beweis. Als Übung (die Argumente sind dieselben wie im Beweis von Satz 7.4.2). □

Satz. *Konvergiert die Folge $v, Mv, M^2v, \ldots$ in $\mathbb{R}^n$ gegen x, so ist x ein Eigenvektor von M zum Eigenwert 1.*

Beweisskizze. Mit Konvergenz in $\mathbb{R}^n$ ist "komponentenweise Konvergenz" gemeint (das wird in der mehrdimensionalen Analysis genau definiert). Konvergiert $M^i v \to x$, so auch die Teilfolge $M^{i+1}v \to x$. Andererseits konvergiert $M^{i+1}v = M(M^i v) \to Mx$ (hier wird benötigt, dass $v \mapsto Mv$ eine stetige Abbildung ist). Folglich $Mx = x$, d.h. x ist Eigenwert zu 1. □

Definition. Ist $M \in \mathbb{R}^{n \times n}$ eine Markov-Matrix und $v \in \mathbb{R}^n$, so wird die Folge $v, Mv, M^2v, \ldots$ in $\mathbb{R}^n$ *Markov-Prozess* mit *Anfangswert* v genannt.

Folgerung. *Konvergiert der Markov-Prozess $v, Mv, M^2v, \ldots \longrightarrow x$, so ist x ein nicht-negativer Eigenvektor von M zum Eigenwert 1 mit $l(x) = 1$.*

Beweisskizze. Aufgrund des Satzes ist x Eigenwert zu 1. Die Aussagen $x \geq 0$ und $l(x) = 1$ folgen aus der Tatsache, dass $M^i v \geq 0$ und $l(M^i v) = 1$ für alle $i \geq 0$ (siehe Bemerkung). □

Beispiel a. Es sei L die Link-Matrix aus Beispiel 7.4.1. Dann ist

$$L^i v \overset{i\to\infty}{\longrightarrow} \frac{1}{11}\begin{pmatrix}3\\4\\2\\2\end{pmatrix} \text{ z.B. für } v = \begin{pmatrix}\frac{1}{4}\\\frac{1}{4}\\\frac{1}{4}\\\frac{1}{4}\end{pmatrix}.$$

Die Interpretation ist die eines "Zufallssurfer", der sich zu Beginn mit Wahrscheinlichkeit v_i auf der Seite S_i aufhält, und dann in jedem Schritt zufällig mit gleicher Wahrscheinlichkeit einen beliebigen Link auf S_i verfolgt. Nach gewisser Zeit hält er sich mit Wahrscheinlichkeit x_i auf der Seite S_i auf.

Es gilt sogar

$$L^i \overset{i\to\infty}{\longrightarrow} \frac{1}{11}\begin{pmatrix}3&3&3&3\\4&4&4&4\\2&2&2&2\\2&2&2&2\end{pmatrix}, \text{ also } L^i v \overset{i\to\infty}{\longrightarrow} \frac{1}{11}\begin{pmatrix}3\\4\\2\\2\end{pmatrix}$$

für alle Anfangswerte v.

Übung. Was passiert in dem Modell des Zufallssurfers, wenn man zulässt, dass eine Webseite keine herausführenden Links besitzt ($n_j = 0$)?

Beispiel b. Angenommen zu n_1 vorhandenen Webseiten erzeugen wir n_2 neue Webseiten, die sich gegenseitig verlinken, auf die aber sonst kein Link führt. Insgesamt gibt es also $n = n_1 + n_2$ Webseiten. In der Praxis ist die Anzahl der neuen Webseiten immer klein im Verhältnis zur Gesamtzahl, d.h. $\frac{n_1}{n} \approx 1$ und $\frac{n_2}{n} \approx 0$. Der "Link-Graph" zerfällt dann in zwei Zusammenhangskomponenten mit n_1 bzw. n_2 vielen Knoten. Entsprechend hat die Link-Matrix L die Blockform $\begin{pmatrix}L_1 & 0\\0 & L_2\end{pmatrix}$, wobei L_1 ein $n_1 \times n_1$-Block ist und L_2 der $n_2 \times n_2$-Block

$$L_2 = \frac{1}{n_2 - 1}\begin{pmatrix}0&1&\cdots&1\\1&0&&1\\\vdots&&&\vdots\\1&\cdots&1&0\end{pmatrix}.$$

Sowohl L_1 als auch L_2 ist wieder eine Markov-Matrix. Aufgrund der Blockstruktur ist der Eigenvektor zu 1 nicht eindeutig. Z.B. gibt es einen Eigenvektor zu 1 der Form $x = \begin{pmatrix}0\\x_2\end{pmatrix}$ mit x_2 Eigenvektor von L_2. Als Gewichtsvektor interpretiert würde dieser nur den neuen Webseiten Bedeutung zumessen.

Das kann nicht passieren, wenn man den Markov-Prozess mit "gleichverteiltem" Anfangsvektor $v = \frac{1}{n}\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ bildet. Angenommen $v, Lv, L^2v, \ldots$ konvergiert gegen $x \in \mathbb{R}^n$. Zerlegt man $v = \begin{pmatrix} v_1 \\ \hline v_2 \end{pmatrix}$ und $x = \begin{pmatrix} x_1 \\ \hline x_2 \end{pmatrix}$ mit $v_1, x_1 \in \mathbb{R}^{n_1}, v_2, x_2 \in \mathbb{R}^{n_2}$, so ist $v_2 = \frac{1}{n}\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ und $v_2, L_2v_2, L_2^2v_2, \ldots$ konvergiert gegen x_2. Da L_2 Markov-Matrix ist, folgt nach der Bemerkung $l(x_2) = l(v_2) = \frac{n_2}{n} \approx 0$. D.h. die neuen Webseiten sind selbst in ihrer Gesamtheit gemäß des Gewichtsvektors x unbedeutend.

7.4.4 Positive Markov-Matrizen

Satz. *Ist M eine positive Markov-Matrix M, so ist*

(i) jeder Eigenvektor zu 1 entweder positiv oder negativ,

(ii) der Eigenraum zu 1 ein-dimensional,

(iii) der Eigenvektor x zu 1 mit $l(x) = 1$ eindeutig bestimmt und positiv.

Beweis. Wir führen den Beweis zunächst genau wie im Satz 7.4.2 bis zu der Stelle $Bz = 0$. Wegen $z > 0$ folgt daraus $B = 0$. Eine positive Matrix enthält keine Nullmatrix als echte Teilmatrix, daher ist $r = 0$. Das bedeutet $x_1, \ldots, x_n > 0$ bzw. $x > 0$. Wir haben damit (i) gezeigt (denn x wurde als beliebiger Eigenvektor zu 1 angenommen und dann entweder x oder $-x$ betrachtet). Die untenstehende Übung zeigt, dass daraus (ii) folgt. Somit gibt es genau einen Eigenvektor x zu 1 mit $l(x) = 1$. Nach (i) ist $x > 0$. □

Beispiel. Die Matrix $\begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$ ist eine Markov-Matrix, die keinen positiven Eigenvektor zum Eigenwert 1 besitzt.

Übung. Jeder 2-dimensionale Unterraum von $\mathbb{R}^n$ enthält Vektoren die weder positiv noch negativ sind.

Bemerkung. Ohne Beweis sei folgende Tatsache erwähnt: Ist M eine positive Markov-Matrix, so konvergiert der Markov-Prozess M^iv für jeden Anfangswert v. Aus Folgerung 7.4.3 ergibt sich, dass der Grenzwert x dann der eindeutige Vektor aus Teil (iii) des obigen Satzes ist. Da dies auch für die

Anfangswerte e_i gilt, konvergiert M^i folglich gegen die Matrix, deren Spalten alle identisch x sind (vgl. Beispiel 7.4.3a).

In der Anwendung macht es Sinn, ein kleines $\alpha > 0$ zu wählen und in der Link-Matrix zu allen Einträgen $\frac{\alpha}{n}$ zu addieren (und danach die Spalten wieder auf Spaltensumme 1 zu "normieren"). Die Interpretation von α ist die Wahrscheinlichkeit, dass der Zufallssurfer keinem vorhandenen Link folgt, sondern auf eine beliebige andere Seite wechselt. Die Link-Matrix ist dann stets positiv (und der Link-Graph zusammenhängend). Dadurch konvergiert der Markov-Prozess zu einem eindeutigen Grenzwert unabhängig vom Anfangswert. Ausserdem kann man auf die Voraussetzung $n_j > 0$ für alle j verzichten.

7.5 Satz von Cayley-Hamilton

Es sei V in diesem Paragrafen ein endlich-dimensionaler K-Vektorraum mit $0 < \dim_K V = n < \infty$. Weiter seien $\varphi \in \mathrm{End}(V)$ und $A \in K^{n\times n}$.

7.5.1 Einsetzungshomomorphismus

Definition. Die Abbildungen

$$\tau_A : K[X] \to K^{n\times n}, \quad A \mapsto f(A),$$
$$\tau_\varphi : K[X] \to \mathrm{End}(V), \quad \varphi \mapsto f(\varphi)$$

werden jeweils *Einsetzungshomomorphismus* genannt.

Bemerkung. Die Einsetzungshomomorphismen sind sowohl Ring- als auch Vektorraum-Homomorphismen.

Frage. Was ist $\chi_A(A)$ und $\chi_\varphi(\varphi)$?

Beispiel.

(i) $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathbb{Q}^{2\times 2}$, $f = X^2 - 5X - 2 \in \mathbb{Q}[X]$. Es gilt $A^0 = E_2$, also

$$f(A) = A^2 - 5A - 2E_2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} - 5\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

(ii) $\varphi : \mathbb{Q}^2 \to \mathbb{Q}^2, e_1 \mapsto e_1 + e_2, e_2 \mapsto 2e_1 + e_2$, $f = X^2 - 2X - 1 \in \mathbb{Q}[X]$. Es gilt $\varphi^0 = \mathrm{id}_{\mathbb{Q}^2} =: \mathrm{id}$, also $f(\varphi) = \varphi^2 - 2\varphi - \mathrm{id}$. Wir berechnen

$$f(\varphi)(e_1) = \ldots = 0,$$
$$f(\varphi)(e_2) = \ldots = 0,$$

also ist $f(\varphi) = 0$ (Nullabbildung).

7.5.2 Invariante Unterräume

Definition. Ein Unterraum $U \leq V$ heißt *invariant unter* φ bzw. *φ-invariant*, wenn $\varphi(U) \subseteq U$.

Beispiel.

(i) Die Drehung des $\mathbb{R}^3$ an der e_3-Achse hat die invarianten Unterräume $\langle e_3 \rangle$ (die Drehachse) und $\langle e_1, e_2 \rangle$ (die Drehebene). Die Abbildungsmatrix bzgl. (e_1, e_2, e_3) hat die Form

$$\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & R_\alpha \end{array}\right),$$

wobei R_α die übliche 2×2-Drehmatrix ist (vgl. Beispiel 6.5.1).

(ii) Es sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine geordnete Basis von V. Zerlege $M = M_\varphi^{\mathcal{B}}$ in Blöcke $M = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$, so dass A ein $r \times r$-Block ist. Genau dann ist $U := \langle v_1, \ldots, v_r \rangle$ φ-invariant, wenn $C = 0$. In diesem Fall ist $A = M_{\varphi|_U}^{(v_1,\ldots,v_r)}$. Entsprechend ist $\langle v_{r+1}, \ldots, v_n \rangle$ genau dann φ-invariant, wenn $B = 0$.

(iii) $\{\mathfrak{o}\}$ und V sind stets φ-invariant, für jedes φ.

(iv) Sei $\mathfrak{o} \neq v \in V$ beliebig und sei $r \in \mathbb{N}$ maximal mit $(v, \varphi(v), \varphi^2(v), \ldots, \varphi^{r-1}(v))$ linear unabhängig. Dann besitzt $\varphi^r(v)$ eine Darstellung $\varphi^r(v) = \sum_{i=0}^{r-1} \lambda_i \varphi^i(v)$ (vgl. Lemma 6.2.4). Folglich ist

$$U_v := \langle v, \varphi(v), \ldots, \varphi^{r-1}(v) \rangle$$

ein φ-invarianter Unterraum. Es hat U_v die Dimension r und $\mathcal{B} = (v, \varphi(v), \varphi^2(v), \ldots, \varphi^{r-1}(v))$ ist eine geordnete Basis.

Es ist U_v der kleinste φ-invariante Unterraum, der v enthält. Im Allgemeinen kann $U_v = V$ sein. Falls v ein Eigenvektor ist, so ist $U_v = \langle v \rangle$.

(v) Es sei $f \in K[X]$. Dann ist $\operatorname{Kern} f(\varphi)$ ein φ-invarianter Unterraum (denn $f(\varphi)(v) = \mathfrak{o} \Rightarrow f(\varphi)(\varphi(v)) = \varphi(f(\varphi)(v)) = \varphi(\mathfrak{o}) = \mathfrak{o}$), z.B.:
für $f = c \in K \setminus \{0\}$ ist $\operatorname{Kern} f(\varphi) = \{\mathfrak{o}\}$,
für $f = X - c$ ist $\operatorname{Kern} f(\varphi) = \operatorname{Kern}(\varphi - c \cdot \mathrm{id}) = V(c, \varphi)$.

Frage. Gibt es nicht-triviale φ-invariante Unterräume und wie findet man sie (von möglichst kleiner Dimension)?

Lemma. *a) Für jeden φ-invarianten Unterraum $U \leq V$ gilt $\chi_{\varphi|_U} | \chi_\varphi$.*
b) Für $f, g \in K[X]$ gilt: $f|g \Rightarrow \operatorname{Kern} f(\varphi) \leq \operatorname{Kern} g(\varphi)$.

Beweis. a) Wähle eine geordnete Basis von U und ergänze diese zu einer Basis von V. Dann folgt die Aussage aus Beispiel (ii) und dem Kästchensatz für charakteristische Polynome.

b) Sei $h \in K[X]$ mit $g = h \cdot f$ und $v \in \operatorname{Kern} f(\varphi)$. Dann ist $g(\varphi) = (h \cdot f)(\varphi) = h(\varphi) \circ f(\varphi)$. Also $g(\varphi)(v) = h(\varphi)(f(\varphi)(v)) = h(\varphi)(\mathfrak{o}) = \mathfrak{o}$. □

7.5.3 Satz von Cayley-Hamilton

Wir bestimmen nun $\operatorname{Kern} \chi_\varphi(\varphi)$.

Lemma. *Für jedes $\mathfrak{o} \neq v \in V$ ist $\chi_{\varphi|_{U_v}}(\varphi)(v) = \mathfrak{o}$.*

Insbesondere folgt $v \in U_v \leq \operatorname{Kern} \chi_{\varphi|_{U_v}}(\varphi)$.

Beweis. Es seien alle Notationen wie in Beispiel 7.5.2iv. Die Abbildungsmatrix von $\varphi|_{U_v}$ bzgl. $\mathcal{B}$ hat die Form

$$\begin{pmatrix} 0 & & \cdots & 0 & \lambda_0 \\ 1 & 0 & \cdots & 0 & \lambda_1 \\ 0 & \ddots & \ddots & & \vdots \\ 0 & & 1 & 0 & \lambda_{r-2} \\ 0 & \cdots & 0 & 1 & \lambda_{r-1} \end{pmatrix} \in K^{r \times r}.$$

Da dies eine Begleitmatrix ist gilt nach Satz 7.3.5:

$$\chi_{\varphi|_{U_v}} = X^r - \lambda_{r-1} X^{r-1} - \ldots - \lambda_1 X - \lambda_0.$$

Einsetzen von φ liefert

$$(\chi_{\varphi|_{U_v}})(\varphi) = \varphi^r - \lambda_{r-1} \varphi^{r-1} - \ldots - \lambda_1 \varphi - \lambda_0.$$

Damit ist klar, dass v im Kern von $\chi_{\varphi|_{U_v}}(\varphi)$ liegt:

$$\begin{aligned} (\chi_{\varphi|_{U_v}})(\varphi)(v) &= (\varphi^r - \lambda_{r-1} \varphi^{r-1} - \ldots - \lambda_1 \varphi - \lambda_0)(v) \\ &= \varphi^r(v) - \lambda_{r-1} \varphi^{r-1}(v) - \ldots - \lambda_1 \varphi(v) - \lambda_0 v = \mathfrak{o}. \end{aligned}$$

□

Satz. *Es gilt stets $\chi_\varphi(\varphi) = 0$ bzw. $\chi_A(A) = 0$.*

In anderen Worten lautet die Aussage: Kern $\chi_\varphi(\varphi) = V$.

Beweis. Wir beweisen nur die Aussage für φ; für A folgt sie durch Übergang zu φ_A. Sei $\mathfrak{o} \neq v \in V$ beliebig. Nach obigem Lemma sowie Lemma 7.5.2 gilt: $v \in \operatorname{Kern} \chi_{\varphi|U_v}(\varphi) \leq \operatorname{Kern} \chi_\varphi(\varphi)$. □

Beispiel. $V = \mathbb{R}^3, \varphi =$ Drehung um 90° um e_3-Achse. [siehe Vorlesung]

Bemerkung a. $\dim U_v \leq \deg f$ für alle $v \in \operatorname{Kern} f(\varphi)$.

Beweis. Ist $f = a_r X^r + \ldots + a_1 X + a_0$ mit $a_r \neq 0$ und $f(\varphi)(v) = a_r \varphi^r(v) + \ldots + a_1 \varphi(v) + a_0 v = \mathfrak{o}$, so ist $(v, \varphi(v), \ldots, \varphi^r(v))$ linear abhängig, also $\dim U_v \leq r = \deg f$. □

Folgerung. *Sei $\chi_\varphi = f_1 \cdots f_r$ mit $f_i \in K[X], \deg f_i \geq 1$. Sei $m = \max\{\deg f_i \mid i = 1, \ldots, r\}$. Dann existiert ein φ-invarianter Unterraum U mit $0 < \dim U \leq m$.*

Beweis. Für $r = 1$ ist $m = \deg f_1 = \deg \chi_\varphi = n$ und V selbst ein φ-invarianter Unterraum der Dimension n. Wir zeigen die Behauptung für $r = 2$ (allgemein führe man Induktion nach n). Sei $\chi_\varphi = f \cdot g$. Nach Cayley-Hamilton ist $\chi_\varphi(\varphi) = f(\varphi) \circ g(\varphi) = 0$. Wähle ein beliebiges $\mathfrak{o} \neq v \in V$. Es gilt $f(\varphi)(g(\varphi)(v)) = \mathfrak{o}$. Falls $g(\varphi)(v) = \mathfrak{o}$, so ist $0 < \dim U_v \leq \deg g \leq m$ nach Bemerkung a. Falls $v' := g(\varphi)(v) \neq \mathfrak{o}$ so ist $f(\varphi)(v') = \mathfrak{o}$, also $0 < \dim U_{v'} \leq \deg f \leq m$ nach Bemerkung a. □

Übung (Berechnung von A^{-1}). Es sei $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathbb{Q}^{2\times 2}$. Man berechne A^{-1} mit Hilfe des Satzes von Cayley-Hamilton.

Wir schliessen mit einer allgemeinen Bemerkung zu invarianten Unterräumen.

Bemerkung b.

(i) Gibt es einen φ-invarianten Unterraum U, so besitzt χ_φ einen Teiler vom Grad $\dim U$ (dieser ist $\chi_{\varphi|U}$).

(ii) Ist f ein Teiler von χ_φ, so ist Kern $f(\varphi)$ ein invarianter Unterraum der Dimension $\geq \deg f$. Es stellt sich heraus (Satz 7.5.4b unten), dass die Dimensionen von Kern $f(\varphi)$, wobei f die Teiler von χ_φ durchläuft, bereits φ wesentlich charakterisieren.

(iii) Nicht jeder φ-invariante Unterraum U hat die Form Kern $f(\varphi)$ für ein Polynom f.

Beweis. (i) wurde in Lemma 7.5.2 gezeigt.
Für lineare Polynome $f = X - c$ kennen wir die Aussage (ii) bereits, denn Kern $f(\varphi)$ ist dann der Eigenraum zu c und seine Dimension ist die geometrische Vielfachheit von c. Bekanntlich ist die geometrische Vielfachheit $\geq 1 = \deg f$ (kann aber auch > 1 sein). Für beliebiges f ergibt sich der Beweis erst im Rahmen der Normalformtheorie in der Linearen Algebra II.
(iii) sieht man schon am Beispiel der Identiät $\varphi = \mathrm{id}_V$, denn dafür ist Kern $f(\varphi)$ stets $\{\mathfrak{o}\}$ oder ganz V, während es φ-invariante Unterräume jeder Dimension $\leq \dim V$ gibt. □

7.5.4 Ausblick: Normalformen

In der Linearen Algebra II werden *Normalformen* von quadratischen Matrizen diskutiert, die eine solche Matrix bis auf Ähnlichkeit charakterisieren. Wir geben hier eine Zusammenfassung der Resultate ohne Beweise.

Bemerkung. Für ähnliche Matrizen $A, B \in K^{n\times n}$ stimmen überein:

(i) Determinante und Spur,

(ii) das charakteristische Polynom,

(iii) die Eigenwerte mit algebraischer Vielfachheit,

(iv) die geometrischen Vielfachheiten, also der Defekt von $A-cE$ und $B-cE$ wobei c ein beliebiger Eigenwert c von A ist,

(v) der Defekt von $f(A)$ und $f(B)$ wobei f ein beliebiger Teiler von χ_A ist,

(vi) der Rang von $f(A)$ und $f(B)$ wobei f ein beliebiger Teiler von χ_A ist.

Übung. Man zeige die letzten beiden Teile der Bermerkung.

Definition. Eine quadratische Matrix A heißt *zerlegbar*, wenn A ähnlich zu einer Matrix in Blockform $\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$ mit einer $r \times r$-Matrix A und $r > 0$ ist. Anderenfalls heißt A *unzerlegbar*.

Satz a (Allgemeine Normalform)**.** *Es seien $A, B \in K^{n\times n}$.*

(i) A ist ähnlich zu einer Matrix der Form $\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}$ mit unzerlegbaren quadratischen Blöcken $A_1, \ldots, A_r$.

(ii) Jede unzerlegbare Matrix A ist ähnlich zu einer Begleitmatrix (zu dem Polynom χ_A).

Eine zu A ähnliche Matrix der Form $\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}$ *mit Begleitmatrizen $A_1, \ldots, A_r$ wird* allgemeine Normalform *von A genannt.*

(iii) Die allgemeine Normalform von A ist bis auf Reihenfolge der A_i eindeutig durch A bestimmt.

Satz b. *Für $A, B \in K^{n\times n}$ sind äquivalent:*

(i) A und B sind ähnlich.

(ii) A und B haben gleiche allgemeine Normalformen.

(iii) $\chi_A = \chi_B$ und $\operatorname{Rg} f(A) = \operatorname{Rg} f(B)$ für jeden Teiler f von χ_A.

Beispiel. Die Matrizen $\begin{pmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}$ und $\begin{pmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & \\ & & & 1 \end{pmatrix}$ haben gleiche geometrische Vielfachheiten, sind aber nicht ähnlich, weil sie das Kriterium (iii) aus Satz b für $f = (X-1)^2$ nicht erfüllen.

Übung. Man folgere aus Satz b, dass A und A^t ähnlich sind.

Satz c (Jordan'sche Normalform)**.** *Es sei $A \in K^{n\times n}$ und χ_A zerfalle vollständig in Linearfaktoren (z.B. $K = \mathbb{C}$). Ist A unzerlegbar, so sind alle Eigenwerte von A identisch, etwa gleich $\lambda \in K$, und A ist ähnlich zu dem* Jordan-Block $J_n(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$. *Eine zu A ähnliche Matrix der Form* $\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}$ *mit Jordan-Blöcken $A_1, \ldots, A_r$ wird* Jordan'sche Normalform *von A genannt. Die Jordan'sche Normalform von A ist bis auf Reihenfolger der Jordan-Blöcke eindeutig durch A bestimmt.*

Als Anwendung der Jordan'schen Normalform kann man zeigen:

Folgerung a. *Jede homogene lineare Differentialgleichung über $\mathbb{C}$ mit gegebenen Anfangswerten hat eine geschlossene Lösung.*

Folgerung b. *Jede linear rekursive Folge (a_n) über $\mathbb{C}$ besitzt eine geschlossene Formel für a_n.*

Beweisskizze. Es sei $f = X^k - c_1X^{k-1} - \ldots - c_{k-1}X - c_k$ das charakteristische Polynom der Folge (a_n) und C die Transponierte der Begleitmatrix von f (vgl. § 7.3.6). Dann existiert eine Jordan'sche Normalform D von C. Es reicht zu zeigen, dass es geschlossene Formeln für die Einträge von D^n gibt. Wir können o.B.d.A. annehmen, dass D ein einzelner Jordan-Block $J_k(\lambda)$ ist. Dann kann man zeigen:

$$D^n = \begin{pmatrix} \binom{n}{0}\lambda^n & \binom{n}{1}\lambda^{n-1} & \cdots & \binom{n}{k}\lambda^{n-k} \\ 0 & \binom{n}{0}\lambda^n & \binom{n}{1}\lambda^{n-1} & \vdots \\ 0 & 0 & \ddots & \binom{n}{1}\lambda^{n-1} \\ 0 & 0 & 0 & \binom{n}{0}\lambda^n \end{pmatrix}$$

Für $D = J_4(\lambda)$ ist z.B.

$$D = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix} \quad D^2 = \begin{pmatrix} \lambda^2 & 2\lambda & 1 & 0 \\ 0 & \lambda^2 & 2\lambda & 1 \\ 0 & 0 & \lambda^2 & 2\lambda \\ 0 & 0 & 0 & \lambda^2 \end{pmatrix}$$

$$D^3 = \begin{pmatrix} \lambda^3 & 3\lambda^2 & 3\lambda & 1 \\ 0 & \lambda^3 & 3\lambda^2 & 3\lambda \\ 0 & 0 & \lambda^3 & 3\lambda^2 \\ 0 & 0 & 0 & \lambda^3 \end{pmatrix} \quad D^4 = \begin{pmatrix} \lambda^4 & 4\lambda^3 & 6\lambda^2 & 4\lambda \\ 0 & \lambda^4 & 4\lambda^3 & 6\lambda^2 \\ 0 & 0 & \lambda^4 & 4\lambda^3 \\ 0 & 0 & 0 & \lambda^4 \end{pmatrix}$$

□

Kapitel 8

Euklidische Vektorräume

In diesem Kapitel sei stets $K = \mathbb{R}$ und V sei ein $\mathbb{R}$-Vektorraum.

8.1 Euklidische Vektorräume

8.1.1 Skalarprodukte

Definition. Eine Abbildung $\langle , \rangle : V \times V \to \mathbb{R}$ heißt *Skalarprodukt* auf V, wenn für alle $\lambda, \mu \in \mathbb{R}$ und $v, w \in V$ gelten:

(S1) $\langle v, \lambda w_1 + \mu w_2 \rangle = \lambda \langle v, w_1 \rangle + \mu \langle v, w_2 \rangle$,

(S2) $\langle v, w \rangle = \langle w, v \rangle$,

(S3) $\langle v, v \rangle > 0$ für alle $v \neq \mathfrak{o}$.

Ist auf V ein Skalarprodukt $\langle , \rangle$ definiert und ist V endlich-dimensional, so heißt V, genauer $(V, \langle , \rangle)$, ein *euklidischer Vektorraum*.

Bemerkung. Aus der Definition eines Skalarproduktes folgt sofort:

(i) $\langle \lambda v_1 + \mu v_2, w \rangle = \lambda \langle v_1, w \rangle + \mu \langle v_2, w \rangle$.

(ii) $\langle v, \mathfrak{o} \rangle = \langle \mathfrak{o}, v \rangle = 0$.

(iii) $\langle v, v \rangle \geq 0$, und $\langle v, v \rangle = 0 \Leftrightarrow v = \mathfrak{o}$.

Man sagt, ein Skalarprodukt ist eine *positiv definite, symmetrische Bilinearform*.

Beispiel.

(i) Standard-Skalarprodukt auf $\mathbb{R}^n$:

$$\langle \begin{pmatrix} a_! \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} b_! \\ \vdots \\ b_n \end{pmatrix} \rangle = \sum_{i=1}^{n} a_i b_i \in \mathbb{R}.$$

Man nennt $\mathbb{R}^n$, ausgestattet mit dem Standardskalarprodukt, den *n-dimensionalen euklidischen Raum*. Das Standardskalarprodukt lässt sich als Matrixprodukt einer Zeile mit einer Spalte schreiben:

$$\langle x, y \rangle = x^t \cdot y.$$

Ist umgekehrt $A \in \mathbb{R}^{m \times n}$ mit Zeilen $z_1, \ldots, z_m$ und $x \in \mathbb{R}^n$, so gilt

$$A \cdot x = \begin{pmatrix} \langle z_1^t, x \rangle \\ \vdots \\ \langle z_m^t, x \rangle \end{pmatrix}.$$

(ii) Auf dem $\mathbb{R}$-Vektorraum $C^0([0,1])$ ist ein Skalarprodukt definiert durch

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt.$$

(iii) Ist $\langle , \rangle$ ein Skalarprodukt auf V und $\varphi \in \mathrm{End}(V)$ ein Isomorphismus, d.h. $\varphi \, \mathrm{Aut}(V)$, so wird durch

$$\langle v, w \rangle^\varphi := \langle \varphi(v), \varphi(w) \rangle$$

ein neues Skalarprodukt $\langle , \rangle^\varphi$ definiert.

(iv) Unterräume euklidischer Vektorräume sind bzgl. der Einschränkung des Skalarproduktes wieder euklidische Vektorräume.

Beweis. Die Nachweise, dass es sich um Skalarprodukte handelt, sind eine leichte Übungsaufgabe. □

8.1.2 Die Norm (Länge)

Es sei $\langle , \rangle$ ein Skalarprodukt auf V.

Definition. Die *Norm* oder *Länge* von $v \in V$ ist definiert als

$$\|v\| := \sqrt{\langle v, v \rangle}.$$

Wir sagen v ist *normiert*, wenn $\|v\| = 1$.

Bemerkung.

(i) $\langle v, v\rangle = \|v\|^2$.

(ii) $\|v\| \geq 0$, und $\|v\| = 0 \Leftrightarrow v = \mathfrak{o}$.

(iii) $\|\lambda v\| = |\lambda| \cdot \|v\|$. Insbesondere ist $\frac{v}{\|v\|}$ stets normiert.

Beispiel. (i) Die Länge eines Vektor $v = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ bzgl. des Standard-Skalarproduktes ist $\|v\| = \sqrt{a^2 + b^2}$.

(ii) Die Länge von $f \in C^0([0,1])$ bzgl. des Skalarproduktes aus Beispiel 8.1.1 ist $\|f\| = \int_0^1 f^2(t)dt$.

8.1.3 Cauchy-Schwarz'sche Ungleichung

Satz (Cauchy-Schwarz'sche Ungleichung)**.** *Für alle $v, w \in V$ gilt*

$$|\langle v, w\rangle| \leq \|v\| \cdot \|w\|.$$

Weiter ist $|\langle v, w\rangle| = \|v\| \cdot \|w\|$ genau dann, wenn (v, w) linear abhängig ist.

Beweis. Die Ungleichung ist äquivalent zu $\langle v, w\rangle^2 \leq \langle v, v\rangle\langle w, w\rangle$. Wir zeigen $\langle v, v\rangle\langle w, w\rangle - \langle v, w\rangle^2 \geq 0$. Mit $\lambda = \frac{\langle v,w\rangle}{\langle w,w\rangle}$ gilt:

$$\begin{aligned}
\langle v, v\rangle\langle w, w\rangle - \langle v, w\rangle^2 &= \langle w, w\rangle \left(\langle v, v\rangle - \frac{\langle v, w\rangle^2}{\langle w, w\rangle}\right) \\
&= \langle w, w\rangle \left(\langle v, v\rangle - 2\frac{\langle v, w\rangle^2}{\langle w, w\rangle} + \frac{\langle v, w\rangle^2}{\langle w, w\rangle}\right) \\
&= \langle w, w\rangle \left(\langle v, v\rangle - 2\lambda\langle v, w\rangle + \lambda^2\langle w, w\rangle\right) \\
&= \langle w, w\rangle\langle v - \lambda w, v - \lambda w\rangle \geq 0.
\end{aligned}$$

□

Folgerung. *Für alle $v, w \in V$ gilt*

(i) (Dreiecksungleichung) $\|v + w\| \leq \|v\| + \|w\|$.

(ii) (umgekehrte Dreiecksungleichung) $|\|v\| - \|w\|| \leq \|v - w\|$.

(iii) (Polarisationsformel) $\langle v, w\rangle = \frac{1}{2}(\|v + w\|^2 - \|v\|^2 - \|w\|^2)$.

(iv) (Parallelogramm-Identität) $\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2$.

Beweis. Übung. □

Übung. Gibt es ein Skalarprodukt auf $\mathbb{R}^n$ derart, dass für alle $x \in \mathbb{R}^n$ gilt: $\|x\| = \sum_{i=1}^n |x_i|$ wobei $x^t = (x_1, \ldots, x_n)$?

8.1.4 Winkel

Es sei $\langle , \rangle$ ein Skalarprodukt auf V. Nach der Cauchy-Schwarz'schen Ungleichung ist für alle $v, w \neq \mathfrak{o}$ stets $\frac{|\langle v,w\rangle|}{\|v\|\|w\|} \leq 1$, d.h.

$$-1 \leq \frac{\langle v, w\rangle}{\|v\|\|w\|} \leq 1.$$

Da $\cos : [0, \pi] \to [-1, 1]$ bijektiv ist, gibt es ein eindeutiges $\alpha \in [0, \pi]$ mit $\cos \alpha = \frac{\langle v,w\rangle}{\|v\|\|w\|}$.

Definition. Der *Winkel* zwischen $v, w \in V\backslash\{\mathfrak{o}\}$, geschr. $\angle(v, w)$, ist das eindeutige $\alpha \in [0, \pi]$ mit $\cos \alpha = \frac{\langle v,w\rangle}{\|v\|\|w\|}$.

Zwei beliebige Vektoren $v, w \in V$ heißen *orthogonal*, geschr. $v \perp w$, wenn $\langle v, w\rangle = 0$.

Beispiel a. Im 2-dimensionalen euklidischen Raum $\mathbb{R}^2$ (mit Standard-Skalarprodukt) gilt für jeden normierten Vektor $v = \begin{pmatrix} a \\ b \end{pmatrix}$:

$$\cos(\angle(v, e_1)) = \frac{\langle v, e_1\rangle}{\|v\|\|e_1\|} = a.$$

Die Definition des Winkels stimmt also mit der geometrischen Interpretation überein.

Beispiel b. Wir betrachten den $\mathbb{R}$-Vektorraum $V = C^0([-\pi, \pi])$ mit Skalarprodukt

$$\langle f, g\rangle := \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)dx.$$

Es gelten

$$\begin{aligned}
\|\sin\| &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin^2(x)dx = \frac{1}{\pi}[\frac{x}{2} - \frac{\sin(2x)}{4}]_{-\pi}^{\pi} = 1,\\
\|\cos\| &= \frac{1}{\pi}\int_{-\pi}^{\pi} \cos^2(x)dx = \frac{1}{\pi}[\frac{x}{2} + \frac{\sin(2x)}{4}]_{-\pi}^{\pi} = 1,\\
\langle \sin, \cos\rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(x)\cos(x)dx = \frac{1}{\pi}[-\frac{1}{2}\cos^2 x]_{-\pi}^{\pi} = 0.
\end{aligned}$$

D.h. sin und cos sind normiert und orthogonal zueinander. (Dass das letzte Integral 0 ist, sieht man ohne Rechnung schon daran, dass $\sin(x)\cos(x)$ eine ungerade Funktion ist.)

Definiere nun $s_t \in V$, $t \in \mathbb{R}$, durch $s_t(x) := \sin(x-t)$, d.h. s_t ist eine Phasenverschiebung des Sinus um den Winkel t. Z.B. ist $s_0 = \sin, s_\pi = -\sin, s_{\pi/2} = -\cos, s_{-\pi/2} = \cos$. Wir berechnen $\angle(s_0, s_t)$. Mit dem trigonometrischen Additionstheorem $\sin(x-t) = \sin x \cos t - \cos x \sin t$ ergibt sich

$$\begin{aligned}\int \sin x \sin(x-t)dx &= \int (\cos t \sin^2 x - \sin t \sin x \cos x)dx \\ &= \cos t(\frac{x}{2} - \frac{\sin(2x)}{4}) + \sin t \frac{\cos^2 x}{2} + C,\end{aligned}$$

also

$$\langle s_0, s_t\rangle = \frac{1}{\pi}[\cos t(\frac{x}{2} - \frac{\sin(2x)}{4}) + \sin t \frac{\cos^2 x}{2}]_{-\pi}^{\pi} = \cos t.$$

Also gilt $\angle(s_0, s_t) = t$, d.h. der Winkel stimmt mit der Phasenverschiebung überein.

Bemerkung.

(i) $\mathfrak{o} \perp v$ für alle $v \in V$.

(ii) (v, w) linear abhängig $\Leftrightarrow \angle(v, w) = 0$ oder π.

(iii) (Pythagoras) Ist $v \perp w$, so gilt

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Beweis. (i) $\langle \mathfrak{o}, v\rangle = 0$ für alle $v \in V$.

(ii) Der zweite Teil der Aussage von Satz (8.1.2).

(iii) Die Polarisationsformel.

□

8.2 Orthogonalität

Es sei V in diesem Abschnitt ein euklidischer Vektorraum.

8.2.1 Orthogonalräume

Definition. Es seien $v \in V, M \subseteq V$ und $U \leq V$.

(i) v heißt *orthogonal* zu M, geschr. $v \perp M$, wenn $\langle v, w\rangle = 0$ für alle $w \in M$.

(ii) Der *Orthogonalraum* zu M ist definiert als

$$M^\perp := \{v \in V \mid v \perp M\} \subseteq V.$$

(iii) Eine Zerlegung $v = v_0 + v_1$ mit $v_0 \in U$ und $v_1 \in U^\perp$ heißt *Orthogonalzerlegung* von v bzgl. U.

Bemerkung. Es seien $v \in V, M \subseteq V$ und $U \leq V$.

(i) $M^\perp$ ist ein Unterraum von V.

(ii) $M^\perp = \langle M \rangle^\perp$.

(iii) $M \cap M^\perp \subseteq \{\mathfrak{o}\}$.

(iv) Existiert eine Orthogonalzerlegung von v bzgl. U, so ist diese eindeutig.

Beweis. Übung. □

Beispiel.

(i) Im euklidischen Raum $\mathbb{R}^3$ ist $\{v\}^\perp$ für jedes $v \neq 0$ einen Ebene.

(ii) Im euklidischen Raum $\mathbb{R}^n$ ist für jedes $A \in \mathbb{R}^{n \times l}$: $\mathrm{SR}(A)^\perp = \mathbb{L}_0(A^t)$.

Beweis. (ii) $\mathrm{SR}(A)^\perp = \{s_1, \ldots, s_l\}^\perp = \{x \in \mathbb{R}^n \mid s_i^t \cdot x = 0\} = \mathbb{L}_0(A^t) \leq \mathbb{R}^n$. □

8.2.2 Orthogonalsysteme

Definition. Ein Tupel $(v_1, \ldots, v_r)$ mit $v_i \in V$ und $v_i \neq \mathfrak{o}$ heißt *Orthogonalsystem*, wenn $v_1, \ldots, v_r$ paarweise orthogonal sind, d.h. $v_i \perp v_j$ falls $i \neq j$. Sind $v_1, \ldots, v_r$ zusätzlich normiert, so nennen wir $(v_1, \ldots, v_r)$ ein *Orthonormalsystem*.

Ist $(v_1, \ldots, v_r)$ eine Basis von V, so sprechen wir auch von *Orthogonalbasen* bzw. *Orthonormalbasen*.

Bemerkung. Ein Tupel $\mathcal{B} = (v_1, \ldots, v_r)$ mit $v_i \in V$ ist genau dann ein Orthonormalsystem, wenn

$$\langle v_i, v_j \rangle = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases}$$

Lemma. *Ist $(v_1, \ldots, v_r)$ ein Orthogonalsystem und $v = \sum_{i=1}^r \lambda_i v_i$, so gilt für jedes $1 \leq j \leq r$:*

$$\langle v, v_j \rangle = \lambda_j \langle v_j, v_j \rangle.$$

Beweis. $\langle v, v_j\rangle = \langle \sum_{i=1}^r \lambda_i v_i, v_j\rangle = \sum_{i=1}^r \lambda_i \langle v_i, v_j\rangle = \lambda_j \langle v_j, v_j\rangle$. □

Satz. *Sei $U \leq V$ und $(u_1, \ldots, u_r)$ eine Orthogonalbasis von U. Dann gibt es für jedes $v \in V$ die Orthogonalzerlegung $v = v_0 + v_1$ und es gilt:*

$$v_0 = \sum_{i=1}^{r} \frac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle} \cdot u_i. \tag{8.1}$$

Beweis. Es sei v_0 wie in (8.1), $v_1 := v - v_0$. Offensichtlich ist $v_0 \in U$. Für jedes $1 \leq j \leq r$ ist nach dem Lemma, angewendet auf v_0:

$$\langle v_0, u_j\rangle = \frac{\langle v, u_j\rangle}{\langle u_j, u_j\rangle}\langle u_j, u_j\rangle = \langle v, u_j\rangle,$$

also

$$\langle v_1, u_j\rangle = \langle v - v_0, u_j\rangle = \langle v, u_j\rangle - \langle v_0, u_j\rangle = 0.$$

Das bedeutet $v_1 \in \{u_1, \ldots, u_r\}^\perp = \{u_1, \ldots, u_r\}^\perp = \langle u_1, \ldots, u_r\rangle^\perp = U^\perp$. □

Beispiel. Sei $u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, v = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Dann ist (u_1, u_2) eine Orthogonalbasis der Ebene $U = \langle u_1, u_2\rangle$. Wir berechnen die Orthogonalzerlegung von v bzgl. U:

$$v_0 = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle}u_1 + \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle}u_2 = \frac{1}{2} \cdot u_1 + 0 \cdot u_2 = \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \end{pmatrix}.$$

$$v_1 = v - v_0 = \begin{pmatrix} -1/2 \\ 0 \\ 1/2 \end{pmatrix}.$$

Übung. Man zeige (evtl. mit Hilfe des Lemmas), dass jedes Orthogonalsystem linear unabhängig ist.

8.2.3 Das Gram-Schmidt-Verfahren

Satz. *Jeder endlich-dimensionale euklidische Vektorraum besitzt eine Orthogonalbasis.*

Beweis. Induktion nach $n = \dim V$. Falls $n = 1$, so ist (v) für jedes $v \neq \mathbf{o}$ eine Orthogonalbasis von V. Sei nun $n > 1$. Wähle einen $n-1$-dimensionalen Unterraum $U \leq V$. Nach Induktionsvoraussetzung hat U eine Orthogonalbasis $(u_1, \ldots, u_{n-1})$. Wähle $v \in V \setminus U$ beliebig. Nach Satz 8.2.2 gibt es die

Orthogonalzerlegung $v = v_0 + v_1$ bzgl. U. Setze $\mathcal{B} := (u_1, \ldots, u_{n-1}, v_1)$. Wegen $v_1 \notin U$ (sonst wäre $v \in U$) ist $\mathcal{B}$ Basis von V. Wegen $v_1 \in U^\perp$ ist $\mathcal{B}$ ein Orthonormalsystem. □

Der Beweis des Satzes lässt sich sofort in folgenden rekursiven Algorithmus übersetzen:

Algorithmus (Orthogonalisierungsverfahren von Gram-Schmidt)**.** *Es sei U ein endlich-dimensionaler Unterraum von V. Die in der Abbildung dargestellte Prozedur* GRAM-SCHMIDT *berechnet zu jeder Basis $(v_1, \ldots, v_r)$ von U eine Orthogonalbasis $(w_1, \ldots, w_r)$ von U.*

GRAM-SCHMIDT$(v_1, \ldots, v_r)$
1 **if** $r = 1$
2 **then return** v_1
3 $w_1, \ldots, w_{r-1} \leftarrow$ GRAM-SCHMIDT$(v_1, \ldots, v_{r-1})$
4 $v_{r0} \leftarrow \sum_{i=1}^{r-1} \frac{\langle v_r, w_i \rangle}{\langle w_i, w_i \rangle} w_i$
5 $w_r \leftarrow v_r - v_{r0}$
6 **return** $w_1, \ldots, w_r$

Abbildung 8.1: Prozedur Gram-Schmidt

Beispiel. $V = \mathbb{R}^4$, $U = \langle v_1, v_2, v_3 \rangle$ mit

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

Wir berechnen mit Gram-Schmidt eine Orthogonalbasis von U:

1. $w_1 = v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

2. $v_{20} = \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = \frac{-1}{1} w_1 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}, w_2 = v_2 - v_{20} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}$.

3. $v_{30} = \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 = \frac{0}{1} w_1 + \frac{-5}{5} w_2 = \begin{pmatrix} -1 \\ 0 \\ -2 \\ 0 \end{pmatrix}, w_3 = v_3 - v_{30} =$

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

Es ist (w_1, w_2, w_3) eine Orthogonalbasis von U.

Übung. Was passiert, wenn man das Gram-Schmidt-Verfahren nicht mit einer Basis, sondern nur mit einem Erzeugendensystem von U beginnt?

8.2.4 Die Orthogonalprojektion

Definition. Ein $\pi \in \mathrm{End}(V)$ heißt *Projektion* auf $U \leq V$, wenn $\mathrm{Bild}(\pi) = U$ und $\pi \circ \pi = \pi$. Wir sprechen von einer *Orthogonal-Projektion* (OP), wenn $\mathrm{Kern}(\pi) = U^\perp$.

Satz. *Es sei U ein endlich-dimensionaler Unterraum von V. Dann gibt es genau eine Orthogonal-Projektion π_U auf U.*

Beweis. Eindeutigkeit: Es sei π eine Orthogonalprojektion auf U und $v \in V$. Setzt man $v_0 := \pi(v) \in U$ und $v_1 := v - \pi(v)$, so gilt $\pi(v_1) = \pi(v) - \pi(\pi(v)) = \mathfrak{o}$, also $v_1 \in U^\perp$. D.h. $v = v_0 + v_1$ ist eine Orthogonalzerlegung von v bzgl. U. Nach Übung 8.2.1 ist v_0 eindeutig durch v bestimmt.

Existenz: Nach Satz 8.2.3 hat U eine Orthogonalbasis. Nach Satz 8.2.2 hat jedes $v \in V$ eine Orthogonalzerlegung $v = v_0 + v_1$ bzgl. U. Definiere eine Abbildung $\pi : V \to U$ durch $\pi(v) := v_0$. Als Übung zeige man, dass π linear ist. □

Bemerkung. Ist $(u_1, \ldots, u_r)$ eine Orthogonalbasis von U, so ist $\pi_U(v)$ durch die Formel (8.1) für v_0 gegeben, die wir auch *Projektionsformel* nennen. Ist $(u_1, \ldots, u_r)$ sogar eine Orthonormalbasis von U, so vereinfacht sich die Projektions-Formel zu

$$v_0 = \sum_{i=1}^{r} \langle v, u_i \rangle \cdot u_i.$$

Beispiel. Es sei $U \leq \mathbb{R}^4$ wie in Beispiel 8.2.3. Wir verwenden die dort berechnete Orthogonalbasis (w_1, w_2, w_3) von U, um $\pi_U(v)$ für folgendes v zu berechnen:

$$v = \begin{pmatrix} 2 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \pi_U(v) = \sum_{i=1}^{3} \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle} w_i = \frac{-1}{1} w_1 + \frac{5}{5} w_2 + \frac{1}{1} w_3 = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \end{pmatrix}.$$

8.2.5 Die Dimensionsformel

Es sei $U \leq V$ und $\dim V = n < \infty$.

Satz. $\dim U + \dim U^\perp = \dim V$.

Beweis. Dies folgt aus der Dimensionsformel für die lineare Abbildung π_U (siehe Folgerung 6.5.2), denn $\text{Bild}(\pi_U) = U$ und $\text{Kern}(\pi_U) = U^\perp$. □

Übung. Man beweise den Satz ohne Verwendung der Dimensionsformel für lineare Abbildungen. Hinweis: Führe eine Basisergänzung mit dem Gram-Schmidt verfahren durch.

Folgerung. $(U^\perp)^\perp = U$.

Beweis. Offensichtlich ist $U \subseteq (U^\perp)^\perp$. Nach dem Satz gilt ausserdem $\dim(U^\perp)^\perp = n - \dim U^\perp = n - (n - \dim U) = \dim U$. Damit folgt die Gleichheit. □

Beispiel. Betrachte $\mathbb{R}^n$ und $\mathbb{R}^m$ jeweils mit dem Standard-Skalarprodukt. Für jedes $A \in \mathbb{R}^{m\times n}$ gelten:

(i) $\text{Kern}\,\varphi_{A^t} = \mathbb{L}_0(A^t) = \text{SR}(A)^\perp = (\text{Bild}\,\varphi_A)^\perp \leq \mathbb{R}^m$.

(ii) $\text{Bild}\,\varphi_{A^t} = \text{SR}(A^t) = \mathbb{L}_0(A)^\perp = (\text{Kern}\,\varphi_A)^\perp \leq \mathbb{R}^n$,

Beweis. Die beiden "äusseren" Gleichheitszeichen sind jeweils klar aus der Definition von φ_A.. Die Gleichung $\text{SR}(A)^\perp = \mathbb{L}_0(A^t)$ wurde bereits in Beispiel (8.2.1) gezeigt. Dieselbe Aussage für A^t statt A lautet $\text{SR}(A^t)^\perp = \mathbb{L}_0(A) \leq \mathbb{R}^n$. Mit der Folgerung ergibt sich $\mathbb{L}_0(A)^\perp = (\text{SR}(A^t)^\perp)^\perp = \text{SR}(A^t)$. □

8.2.6 Die Orthogonalentwicklung

Satz. *Es sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Orthogonalbasis von V und $v \in V$. Dann:*

$$v = \sum_{i=1}^{n} \frac{\langle v, v_i\rangle}{\langle v_i, v_i\rangle} \cdot v_i, \quad \textit{d.h. } \kappa_{\mathcal{B}}(v) = \begin{pmatrix} \frac{\langle v,v_1\rangle}{\langle v_1,v_1\rangle} \\ \vdots \\ \frac{\langle v,v_n\rangle}{\langle v_n,v_n\rangle} \end{pmatrix}.$$

Im Fall einer Orthonormalbasis gilt:

$$v = \sum_{i=1}^{n} \langle v, v_i\rangle \cdot v_i, \quad \kappa_{\mathcal{B}}(v) = \begin{pmatrix} \langle v, v_1\rangle \\ \vdots \\ \langle v, v_n\rangle \end{pmatrix}.$$

Beweis. Betrachtet man die Orthogonalprojektion von V auf sich selbst, so ist $\pi_V(v) = v$ und die Aussage ergibt sich aus (8.1). □

Beispiel. Betrache die Vektoren

$$v_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3.$$

Dann ist $\mathcal{B} = (v_1, v_2, v_3)$ eine Orthonormalbasis von $\mathbb{R}^3$ und die Orthogonalentwicklung von v nach $\mathcal{B}$ lautet:

$$v = \sqrt{\frac{2}{3}} e_1 + \sqrt{2} e_2 - \sqrt{\frac{1}{3}} e_3.$$

Beweis. Man rechnet nach, dass

$$\langle v_i, v_j \rangle = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j \end{cases}$$

für $i, j \in \{1, 2, 3\}$. D.h. (v_1, v_2, v_3) ist ein ONS, also nach Satz 8.2.2 auch linear unabhängig, also eine ONB. Die Koordinaten bzgl. dieser Basis berechnet man als

$$\langle v, v_1 \rangle = \frac{1 - 1 + 2}{\sqrt{6}} = \sqrt{\frac{2}{3}}, \langle v, v_2 \rangle = \frac{1 + 1 + 0}{\sqrt{2}} = \sqrt{2}, \langle v, v_3 \rangle = \frac{1 - 1 - 1}{\sqrt{3}} = -\sqrt{\frac{1}{3}}.$$

□

8.3 Approximation

In einem euklidischen Vektorraum V seien $M \subset V$ und $v \in V$ gegeben. Es wird ein Element $x \in M$ gesucht, dass eine „beste Näherung“ an v darstellt.

8.3.1 Winkelapproximation

Vorausgesetzt, dass v und alle $x \in M$ normiert sind, kann man nach dem von v und x eingeschlossenen Winkel approximieren. Nach der Cauchy-Schwarz'schen Ungleichung gilt für normierte Vektoren $-1 \leq \langle v, x \rangle \leq 1$, wobei $\langle v, x \rangle = 1$ genau dann, wenn $v = w$, $\langle v, x \rangle = 0$ genau dann, wenn $v \perp w$, und $\langle v, x \rangle = -1$ genau dann, wenn $v = -w$. Für eine beste Approximation ist daher $\langle v, x \rangle$ zu maximieren.

Beispiel a. Gegeben seien n Dokumente $D_1, \ldots, D_n$ und m Terme $T_1, \ldots, T_m$. Wir definieren zu Dokument j den Vektor

$$d'_j = \begin{pmatrix} d'_{1j} \\ \vdots \\ d'_{mj} \end{pmatrix} \in \mathbb{R}^m, \text{ wobei } d'_{ij} = \begin{cases} 1 & \text{falls } T_i \text{ in } D_j \text{ vorkommt,} \\ 0 & \text{sonst} \end{cases}$$

und normieren zu $d_j := \frac{d'_j}{\|d'_j\|}$.

Aus einer Suchanfrage nach den Termen $T_{i_1}, \ldots, T_{i_l}$ wird entsprechend ein Suchvektor

$$q' = \begin{pmatrix} q'_1 \\ \vdots \\ q'_m \end{pmatrix} \in \mathbb{R}^m, \text{ wobei } q'_i = \begin{cases} 1 & \text{falls } i \in \{i_1, \ldots, i_l\}, \\ 0 & \text{sonst} \end{cases}$$

gebildet und zu $q := \frac{q'}{\|q'\|}$ normiert.

Das am besten zur Suchanfrage passende Dokument ist dann D_j für dasjenige j, für das $\langle q, d_j \rangle$ maximal wird.

Man beachte, dass man für das Standardskalarprodukt alle $\langle q, d_j \rangle$ durch eine einzelne Matrixmultiplikation errechnen kann. Schreibt man $d_1, \ldots, d_n$ in die Spalten einer Matrix D, so ist D eine Markov-Matrix und es gilt

$$q^t \cdot D = (\langle q, d_1 \rangle, \ldots, \langle q, d_n \rangle) \in \mathbb{R}^{1 \times n}.$$

Beispiel b. Hat man analoge Audiosignale statt Textdokumenten, so kann man diese als Vektoren $d'_j \in C^0([0,1])$ auffassen (z.B. bei Abspiel-Länge 1s). Unter Verwendung des Skalarproduktes $\langle f, g \rangle = \int_0^1 f(x)g(x)dx$ könnte man analog zum vorherigen Beispiel verfahren.

8.3.2 Abstandsapproximation

Im Allgemeinen, d.h. wenn die Vektoren nicht normiert sind, versucht man den Abstand $\|v - x\|$ für $x \in M$ zu minimieren.

Definition. Man nennt $d(v, M) := \inf\{\|v - x\| \mid x \in M\}$ den *Abstand* von v zu M.

Falls M ein Unterraum ist, so wird die beste Approximation gerade von der Projektion geliefert:

Satz. *Es seien $U \leq V$ und $v \in V$. Dann gilt für alle $u \in U$:*

$$\|v - u\| \geq \|v - \pi_U(v)\|.$$

Insbesondere ist $d(v, U) = \|v - \pi_U(v)\|$.

Beweis. Setze $u_0 := \pi_U(v)$. Dann ist $v - u_0 \in U^\perp$. Für jedes $u \in U$ gilt somit $(v - u_0) \perp (u_0 - u)$, also nach Pythagoras:

$$\|v - u\|^2 = \|v - u_0\|^2 + \|u_0 - u\|^2 \geq \|v - u_0\|^2.$$

Daraus folgt die Behauptung. □

Bemerkung. Die Wahl des Skalarproduktes bestimmt die Definition des Abstandes und legt damit das Kriterium fest, nach welchem approximiert wird. Verschiedene Skalarprodukte liefern im Allgemeinen verschiedene beste Approximationen.

Beispiel. Ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine Orthogonalbasis von V und $U = \langle v_1, \ldots, v_r \rangle$ so lässt sich die π_U in Koordinaten bzgl. $\mathcal{B}$ ohne Rechnung sofort angeben:

$$\kappa_{\mathcal{B}}(v) = (x_1, \ldots, x_n)^t \Rightarrow \kappa_{\mathcal{B}}(\pi_U(v)) = (x_1, \ldots, x_r, 0, \ldots, 0)^t.$$

Übung. Wie lautet die Abbildungsmatrix von π_U bzgl. der angegebenen Basis in obigem Beispiel?

8.3.3 Datenkompression

Die Abstandsapproximationen durch Elemente eines möglichst kleinen Unterraums U kann zur verlustbehafteten Datenkompression verwendet werden. Sei dazu eine Orthongonalbasis $\mathcal{B} = (v_1, \ldots, v_n)$ von V gewählt. Um einen gegebenen "Datensatz" $v \in V$ mit "Genauigkeit" $1 \leq r \leq n$ zu komprimieren, wird $\kappa_{\mathcal{B}}(v) = (x_1, \ldots, x_n)^t$ berechnet und nur das Tupel $(x_1, \ldots, x_r)$ gespeichert bzw. übertragen. Gemäß Beispiel 8.3.2 definiert dieses Tupel die beste Approximation aus dem Unterraum $U_r := \langle v_1, \ldots, v_r \rangle$ an v. Bei gewählter Genauigkeit r ergibt sich die Kompressionsrate $\frac{r}{n}$. Die Schwierigkeit liegt darin, die Orthogonalbasis $\mathcal{B}$ so zu wählen, dass sich die für die gegebene Anwendung signifikante Information in den ersten Koordinaten sammelt und die Information aus den letzten Koordinaten verzichtbar ist.

Beispiel a. Wir fassen Binärzahlen der Länge n bit als Vektoren aus $\mathbb{R}^n$ auf, etwa die Zahl $1011\ldots0$ als $(1, 0, 1, 1, \ldots, 0)^t$. Die "least significant bits" stehen in der Binärzahl rechts, die "most significant bits" links. Betrachtet man das Standard-Skalarprodukt und wählt als Orthogonalbasis die Standardbasis, so läuft das beschriebene Verfahren darauf hinaus, nur die r "most significant bits" zu speichern.

Beispiel b. Bei der *diskreten Kosinustransformation* wird die Orthogonalbasis $\mathcal{B} = (v_0, \ldots, v_{n-1})$ von $\mathbb{R}^n$ bzgl. Standard-Skalarprodukt gewählt, die

definiert ist durch:

$$v_i := \begin{pmatrix} \cos\left(\frac{1}{2n} \cdot i\pi\right) \\ \cos\left(\frac{3}{2n} \cdot i\pi\right) \\ \vdots \\ \cos\left(\frac{2n-1}{2n} \cdot i\pi\right) \end{pmatrix}$$

Bezeichnet $T_n := {}^{\mathcal{E}}T^{\mathcal{B}}$ die zugehörige Basiswechselmatrix, so ist z.B.

$$T_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad T_3 = \begin{pmatrix} 1 & \sqrt{3}/2 & 1/2 \\ 1 & 0 & -1 \\ 1 & -\sqrt{3}/2 & 1/2 \end{pmatrix}.$$

Unter Verwendung geeigneter trigonometrischer Identitäten kann man nachrechnen, dass $\mathcal{B}$ tatsächlich eine Orthogonalbasis ist ($\langle v_i, v_j \rangle = 0$ für alle $i \neq j$) und ausserdem

$$\|v_i\| = \begin{cases} \sqrt{n} & \text{falls } i = 0, \\ \sqrt{\frac{n}{2}} & \text{falls } i > 0. \end{cases}$$

Die Idee, die hinter der diskreten Kosinustransformation steckt, wird ersichtlich, wenn man folgende kontinuierliche Variante betrachtet.

Beispiel c. Es sei $V = C^0([0,1])$ mit dem Skalarprodukt $\langle f, g \rangle = \int_0^\pi f(x)g(x)dx$. Für jedes $n \in \mathbb{N}$ ist ein Orthogonalsystem $\mathcal{B}_n = (f_0, \ldots, f_{n-1})$ definiert durch $f_i(x) := \cos(ix)$. Unter Verwendung geeigneter trigonometrischer Identitäten kann man nachrechnen, dass $\mathcal{B}$ tatsächlich ein Orthogonalsystem ist ($\langle f_i, f_j \rangle = 0$ für alle $i \neq j$) und ausserdem

$$\|f_i\| = \begin{cases} \sqrt{\pi} & \text{falls } i = 0, \\ \sqrt{\frac{\pi}{2}} & \text{falls } i > 0. \end{cases}$$

Die Projektion auf die Unterräume $U_n = \langle f_0, \ldots, f_{n-1} \rangle$ bedeutet, eine gegebene Funktion f durch eine Überlagerung (Linearkombination) von Kosinusfunktion verschiedener Frequenzen zu approximieren.

Bemerkung a. Man sagt, bei der Kosinustransformation wird vom "Zeitraum" in den "Frequenzraum" transformiert. (besser wäre: von der "Zeitbasis" in die "Frequenzbasis"). Verschiedene Modifikationen der diskreten Kosinustransformation werden bei Audio-Codecs verwendet. Eine zweidimensionale Variante ist die Grundlage des JPEG-Verfahrens.

8.4 Positiv definite Matrizen

Es sei V in diesem Abschnitt ein $\mathbb{R}$-Vektorraum mit $0 < n = \dim V < \infty$.

8.4.1 Die Gram-Matrix

Definition. Es sei $\langle , \rangle$ ein Skalarprodukt auf V. Für jede geordnete Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V definieren wir die *Gram-Matrix* von $\langle , \rangle$ bzgl. $\mathcal{B}$ als

$$G^{\mathcal{B}}(\langle , \rangle) := (\langle v_i, v_j \rangle)_{ij} \in \mathbb{R}^{n \times n}.$$

Beispiel. Wir betrachten $\mathbb{R}^n$ mit dem Standard-Skalarprodukt $\langle , \rangle$.

(i) Es sei $\mathcal{B} = (s_1, \ldots, s_r)$ mit $s_i \in \mathbb{R}^n$. Da $\langle s_i, s_j \rangle = s_i^t \cdot s_j$ für alle $1 \leq i, j \leq r$, ergibt sich
$$G^{\mathcal{B}}(\langle , \rangle) = A^t A,$$
wobei A die Matrix mit Spalten $s_1, \ldots, s_r$ ist, also $A = {}^{\mathcal{E}}T^{\mathcal{B}}$.

(ii) Bezüglich der Standardbasis $\mathcal{E}$ ist $G^{\mathcal{E}}(\langle , \rangle) = E_n$.

(iii) Auf $\mathbb{R}^2$ bzgl. $\mathcal{B} = (\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \end{pmatrix})$ ist z.B.

$$G^{\mathcal{B}}(\langle , \rangle) = \begin{pmatrix} 1 & -1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 3 \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ -3 & 9 \end{pmatrix}.$$

Bemerkung. Es seien $\langle , \rangle$ ein Skalarprodukt auf V, $\mathcal{B}$ eine geordnete Basis von V und $G \in \mathbb{R}^{n \times n}$.

(i) $\mathcal{B}$ Orthogonalbasis $\Leftrightarrow G^{\mathcal{B}}(\langle , \rangle)$ Diagonalmatrix.

(ii) $\mathcal{B}$ Orthonormalbasis $\Leftrightarrow G^{\mathcal{B}}(\langle , \rangle) = E_n$.

(iii) $G = G^{\mathcal{B}}(\langle , \rangle) \Leftrightarrow$

$$\langle v, w \rangle = \kappa_{\mathcal{B}}(v)^t \cdot G \cdot \kappa_{\mathcal{B}}(w) \text{ für alle } v, w \in V. \tag{8.2}$$

(iv) $\mathcal{B}$ Orthonormalbasis $\Rightarrow \langle v, w \rangle = \kappa_{\mathcal{B}}(v)^t \cdot \kappa_{\mathcal{B}}(w)$ für alle $v, w \in V$.

Nach (iii) ist das Skalarprodukt schon eindeutig durch die Gram-Matrix definiert. Nach (iv) verhält sich jedes Skalarprodukt wie das Standardskalarprodukt, wenn man zu den Koordinatenvektoren bzgl. einer Orthonormalbasis übergeht. Insbesondere ist das Skalarprodukt schon eindeutig durch die Angabe einer Orthonormalbasis definiert.

Beweis. Beweis als Übung. □

Frage.

1. Wie bekommt man alle Skalarprodukte auf V?
2. Welche quadratischen Matrizen treten als Gram-Matrizen auf, d.h. für welche quadratischen Matrizen wird durch die Formel (8.2) ein Skalarprodukt definiert?
3. Wir verhält sich die Gram-Matrix unter Basiswechsel?
4. Wir verhalten sich die Gram-Matrizen verschiedener Skalarprodukte zueinander?

8.4.2 Charakterisierung durch Orthonormalbasen

Wir widmen uns der Frage, wie man alle Skalarprodukt auf V bekommt.

Definition. Es sei $\mathcal{B}$ eine beliebige geordnete Basis von V. Wir bezeichnen mit $\langle , \rangle_{\mathcal{B}}$ das Skalarprodukt, dass definiert ist durch

$$\langle v, w\rangle_{\mathcal{B}} := \kappa_{\mathcal{B}}(v)^t \cdot \kappa_{\mathcal{B}}(w)$$

für alle $v, w \in V$. Das Standard-Skalarprodukt auf $\mathbb{R}^n$ wird demnach mit $\langle , \rangle_{\mathcal{E}}$ bezeichnet.

Übung. Man zeige, dass es sich bei $\langle , \rangle_{\mathcal{B}}$ um ein Skalarprodukt mit Orthonormalbasis $\mathcal{B}$ handelt. Man vergleiche dieses Skalarprodukt mit Beispiel 8.1.1iii.

Satz. *Die Zuordnung*

$$\{\textit{geordnete Basen von } V\} \to \{\textit{Skalarprodukte von } V\},$$
$$\mathcal{B} \mapsto \langle v, w\rangle_{\mathcal{B}}$$

ist surjektiv und die Faser zu einem Skalarprodukt $\langle , \rangle$ besteht aus allen Orthonormalbasen bzgl. $\langle , \rangle$.

Beweis. Sei $\langle , \rangle$ ein beliebiges Skalarprodukt auf V. Nach Teil (iv) von Bemerkung 8.4.1 ist jede Orthonormalbasis bzgl. $\langle , \rangle$ ein Urbild von $\langle , \rangle$. Ist umgekehrt $\mathcal{B}$ ein Urbild von $\langle , \rangle$, so ist $\mathcal{B}$ nach obiger Übung eine Orthonormalbasis von $\langle , \rangle_{\mathcal{B}} = \langle , \rangle$. Somit besteht die Faser zu $\langle , \rangle$ genau aus den Orthonormalbasen bzgl. $\langle , \rangle$. Da es stets eine Orthonormalbasis gibt (Gram-Schmidt) sind die Fasern nicht-leer, d.h. die Zuordnung ist surjektiv. □

Folgerung. *Es seien $\langle , \rangle$ ein beliebiges Skalarprodukt auf V und $\mathcal{B}'$ eine Orthonormalbasis von V bzgl. $\langle , \rangle$. Für jede geordnete Basis $\mathcal{B}$ ist $G^{\mathcal{B}}(\langle , \rangle) = T^t T$ wobei $T = {}^{\mathcal{B}'}T^{\mathcal{B}}$. Weiter gilt*

$$\{G^{\mathcal{B}}(\langle , \rangle) \mid \mathcal{B} \textit{ geordnete Basis von } V\} = \{T^t T \mid T \in \mathrm{GL}_n(\mathbb{R})\}.$$

Beweis. Es sei bemerkt, dass die erste Aussage für den Fall das Standard-Skalarproduktes bereits aus Beispiel 8.4.1i bekannt ist. Allgemein gilt nach Bemerkung 8.4.1iv

$$\langle v, w\rangle = \kappa_{\mathcal{B}'}(v)^t \cdot \kappa_{\mathcal{B}'}(w) = (T\kappa_{\mathcal{B}}(v))^t \cdot (T\kappa_{\mathcal{B}}(w)) = \kappa_{\mathcal{B}}(v)^t \cdot (T^tT) \cdot \kappa_{\mathcal{B}}(w).$$

Also nach Bemerkung 8.4.1iii: $G^{\mathcal{B}}(\langle , \rangle_{\mathcal{B}}) = T^tT$.

Die Inklusion $\subseteq$ ist damit trivial. Sei umgekehrt $T \in \mathrm{GL}_n(\mathbb{R})$ gegeben. Dann gibt es eine Basis $\mathcal{B}$ mit $T = {}^{\mathcal{B}'}T^{\mathcal{B}}$ (Folgerung 6.5.4). Wegen $G^{\mathcal{B}}(\langle , \rangle_{\mathcal{B}}) = T^tT$ ist damit auch $\supseteq$ gezeigt. □

8.4.3 Basiswechselsatz

Satz. *Seien $\langle , \rangle$ ein beliebiges Skalarprodukt auf V und $\mathcal{B}_1, \mathcal{B}_2$ zwei beliebige geordnete Basen von V. Dann gilt $G^{\mathcal{B}_2}(\langle , \rangle) = S^t \cdot G^{\mathcal{B}_1}(\langle , \rangle) \cdot S$ mit $S = {}^{\mathcal{B}_1}T^{\mathcal{B}_2}$.*

Beweis. Als Übung, entweder durch zweimalige Anwendung von Folgerung 8.4.2 oder direkt aus der Definition der Gram-Matrix. □

Übung. Wir definieren eine Relation $\sim$ auf $\mathbb{R}^{n\times n}$ durch $A \sim B$, falls $T \in \mathrm{GL}_n(K)$ existiert mit $T^tAT = B$. Zeige, dass $\sim$ eine Äquivalenzrelation ist.

Übung. Man zeige durch eine Abwandlung des Gauß-Algorithmus und mit Hilfe des Basiswechselsatzes für Gram-Matrizen, dass jeder endlich-dimensionale euklidische Vektorraum eine Orthogonalbasis besitzt.

8.4.4 Positiv definite Matrizen

Definition. Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt *positiv definit*, wenn sie symmetrisch ist (d.h. $A = A^t$) und wenn für alle $0 \neq x \in \mathbb{R}^n$ gilt:

$$x^tAx > 0.$$

Satz. *Es seien $\mathcal{B}$ eine geordnete Basis von V und $G \in \mathbb{R}^{n\times n}$. Durch die Formel (8.2) wird genau dann ein Skalarprodukt auf V definiert, wenn G positiv definit ist.*

Beweis. Die Gram-Matrix eines Skalarproduktes ist stets positiv definit, weil sie offensichtlich symmetrisch ist und weil die Formel (8.2) gilt und jedes $x \in \mathbb{R}^n$ als Koordinatenvektor vorkommt. Umgekehrt rechnet man leicht nach, dass (8.2) ein Skalarprodukt definiert, wenn G positiv definit ist. □

Folgerung a. *Für jede Basis $\mathcal{B}$ von V ist folgende Zuordnung eine Bijektion:*

$$\{\textit{Skalarprodukte auf } V\} \to \{\textit{positiv definite } n \times n\textit{-Matrizen}\},$$
$$\langle , \rangle \mapsto G^{\mathcal{B}}(\langle , \rangle)$$

Das Urbild einer positiv definiten $n \times n$-Matrix G ist durch (8.2) gegeben.

Beweis. Der obige Satz besagt, dass es diese Zuordnung gibt. Weiter besagt er, dass für gegebenes positiv definites $G \in \mathbb{R}^{n\times n}$ durch (8.2) ein Skalarprodukt definiert wird. Nach Bemerkung 8.4.1iii ist dieses Skalarprodukt das eindeutige Urbild von G. Also ist die Zuordnung bijektiv. □

Beispiel.

(i) E_n ist positiv definit. Das zugeordnete Skalarprodukt lautet

$$\langle v, w\rangle_{\mathcal{B}} := \kappa_{\mathcal{B}}(v)^t \cdot \kappa_{\mathcal{B}}(w)$$

und besitzt $\mathcal{B}$ als Orthonormalbasis.

(ii) Eine Diagonalmatrix D ist genau dann positiv definit, wenn alle Diagonaleinträge > 0 sind. Das liegt daran, dass $e_i D e_i$ gerade der i-te Diagonaleintrag ist.

(iii) $A = \begin{pmatrix} 4 & -2 \\ -2 & 3 \end{pmatrix}$ ist positiv definit, denn $\begin{pmatrix} a & b \end{pmatrix} A \begin{pmatrix} a \\ b \end{pmatrix} = 4a^2 - 4ab + 3b^2 = (2a - b)^2 + 2b^2 > 0$ falls $a \neq 0$ oder $b \neq 0$.

(iv) $A = \begin{pmatrix} 4 & -2 \\ -2 & -1 \end{pmatrix}$ ist nicht positiv definit, denn $\begin{pmatrix} 0 & 1 \end{pmatrix} A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -1$.

(v) $A = \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}$ ist nicht positiv definit, denn $\begin{pmatrix} 1 & 2 \end{pmatrix} A \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 0$.

Folgerung b.

(i) $\{\textit{positiv definite } n \times n\textit{-Matrizen}\} = \{S^t S \mid S \in \mathrm{GL}_n(\mathbb{R})\}$.

(ii) Für $A \in \mathbb{R}^{n\times n}$ und $T \in \mathrm{GL}_n(K)$ gilt: A positiv definit $\Leftrightarrow$ $T^t A T$ positiv definit.

Beweis. (i) ergibt sich aus Folgerung a oben sowie Folgerung 8.4.2: jede positiv definite Matrix ist eine Gram-Matrix und deshalb von der Form $T^t T$ mit $T \in \mathrm{GL}_n(\mathbb{R})$; umgekehrt ist $T^t T$ für jedes $T \in \mathrm{GL}_n(\mathbb{R})$ eine Gram-Matrix und deshalb positiv definit.

(ii) Sei A positiv definit. Nach (i) ist dann $A = S^t S$ mit $S \in \mathrm{GL}_n(K)$, also auch $T^t A T = T^t S^t S T = (ST)^t (ST)$ positiv definit. Die Richtung $\Rightarrow$ folgt wegen $(T^{-1})^t (T^t A T) T^{-1} = A$. □

Frage. Wie stellt man systematisch fest, ob eine symmetrische Matrix positiv definit ist?

8.4.5 Der Spektralsatz

Es sei $\langle , \rangle$ ein Skalarprodukt auf V,

Definition.

(i) $\varphi \in \mathrm{End}(V)$ heißt *selbstadjungiert*, wenn für alle $v, w \in V$ gilt:

$$\langle \varphi(v), w \rangle = \langle v, \varphi(w) \rangle.$$

(ii) $A \in \mathbb{R}^{n \times n}$ heißt *symmetrisch*, wenn $A^t = A$.

Satz (Spektralsatz)**.**

(i) Zu jedem selbstadjungierten $\varphi \in \mathrm{End}(V)$ gibt es eine Eigenvektorbasis von V, die gleichzeitig Orthonormalbasis bzgl. $\langle , \rangle$ ist.

(ii) Zu jeder reellen symmetrischen Matrix gibt es eine Eigenvektorbasis von $\mathbb{R}^n$, die gleichzeitig Orthonormalbasis bzgl. $\langle , \rangle_{\mathcal{E}}$ ist.

Folgerung. *Es sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Dann ist A genau dann positiv definit, wenn alle Eigenwerte von A positiv sind.*

Beweis. Es sei $\mathcal{B}$ eine orthonormale Eigenvektorbasis zu A (existiert nach Spektralsatz). Setze $T = {}^{\mathcal{E}}T^{\mathcal{B}}$. Nach Beispiel 8.4.1i und Bemerkung 8.4.1ii haben wir $T^tT = G^{\mathcal{B}} = E_n$, also $T^{-1} = T^t$. Da $\mathcal{B}$ Eigenvektorbasis ist, ist $D := T^tAT = T^{-1}AT$ eine Diagonalmatrix, deren Diagonaleinträge genau die Eigenwerten von A sind. Nach Teil (ii) von Folgerung 8.4.4b ist A genau dann positiv definit, wenn D positiv definit ist. Nach Beispiel 8.4.4ii also genau dann, wenn alle Eigenwerte von A positiv sind. □

Übung. Man überprüfe dieses Kriterium für die positive Definitheit anhand von Beispiel 8.4.4.

Wir widmen uns nun dem Beweis des Spektralsatzes.

Bemerkung a. Es seien $\mathcal{B}$ eine Orthonormalbasis von V, $\varphi \in \mathrm{End}(V)$ und $U \leq V$.

(i) φ selbstadjungiert $\Leftrightarrow M_{\varphi}^{\mathcal{B}}$ symmetrisch.

(ii) φ selbstadjungiert, U φ-invariant $\Rightarrow U^{\perp}$ φ-invariant.

Beweis. Übung. □

Lemma. *Ist $A \in \mathbb{R}^{n\times n}$ symmetrisch, so zerfällt χ_A über $\mathbb{R}$ vollständig in Linearfaktoren.*

Beweis. Nach dem Fundamentalsatz der Algebra zerfällt χ_A über $\mathbb{C}$ vollständig in Linearfaktoren. Zu zeigen bleibt: alle Eigenwerte von A sind reell. Sei $c \in \mathbb{C}$ Eigenwert von A, etwa $Av = cv, v \in \mathbb{C}^n \setminus \{0\}$. Bezeichne mit $\overline{A}, \overline{v}, \ldots$ die komplexe Konjugation aller Einträge. Da A reell ist, gilt $\overline{A} = A$.
1. $\overline{v}$ ist Eigenvektor zum Eigenwert $\overline{c}$: $A\overline{v} = \overline{A}\overline{v} = \overline{Av} = \overline{cv} = \overline{c}\overline{v}$.
2. $z = a + bi \in \mathbb{C} \setminus \{0\} \Rightarrow z\overline{z} = a^2 + b^2 > 0$.
3. $v^t\overline{v} > 0$: $v = (x_1, \ldots, x_n)^t$ nicht alle $x_i = 0 \Rightarrow v^t\overline{v} = \sum_{i=1}^n x_i\overline{x_i} > 0$.
4. $c(v^t\overline{v}) = (cv)^t\overline{v} = (Av)^t\overline{v} = v^tA^t\overline{v} = v^tA\overline{v}$.
5. $\overline{c}(v^t\overline{v}) = v^t(\overline{cv}) = v^tA\overline{v}$.
Es folgt $c = \overline{c}$, d.h. c ist reell. □

Beweis des Spektralsatzes. Es sei $\varphi \in \mathrm{End}(V)$ selbstadjungiert. (Statt $A \in \mathbb{R}^{n\times n}$ betrachte φ_A.) Wir führen Induktion nach $n = \dim V$. Der Fall $n = 1$ ist klar, denn jeder normierte Vektor stellt eine Orthonormalbasis aus Eigenvektoren dar. Sei nun $n > 1$. Nach dem Lemma gibt es einen reellen Eigenwert c. Sei v ein normierter Eigenvektor zu c. Ergänze v zu einer Orthonormalbasis $\mathcal{B} = (v, v_2, \ldots, v_n)$ von V (Gram-Schmidt). Betrachte $U = \langle v\rangle$ und $U^\perp = \langle v_2, \ldots, v_n\rangle$. Da φ selbstadjungiert ist, ist mit U auch $U^\perp$ φ-invariant. Offensichtlich ist $\varphi|_{U^\perp}$ wieder selbstadjungiert. Nach Induktionsvoraussetzung hat $U^\perp$ eine Orthonormalbasis aus Eigenvektoren. Zusammen mit v ergibt das die gesuchte Basis. □

Bemerkung b. Man kann sogar die Umkehrung des Spektralsatzes zeigen: gibt es zu $A \in \mathbb{R}^{n\times n}$ eine orthonormale Eigenvektorbasis, so ist A symmetrisch.

8.5 Orthogonale Abbildungen

Es seien V, W zwei $\mathbb{R}$-Vektorräume mit Skalarprodukten $\langle , \rangle_V$ und $\langle , \rangle_W$.

8.5.1 Orthogonale Homomorphismen

Es sei $\varphi \in \mathrm{Hom}(V, W)$.

Bemerkung. Es sei B eine Basis von V. Folgende Aussagen sind äquivalent:

(i) $\langle \varphi(v), \varphi(w)\rangle_W = \langle v, w\rangle_V$ für alle $v, w \in V$.

(ii) $\langle \varphi(v), \varphi(w) \rangle_W = \langle v, w \rangle_V$ für alle $v, w \in B$.

(iii) $\|\varphi(v)\| = \|v\|$ für alle $v \in V$.

(iv) $v \in V$ normiert $\Rightarrow \varphi(v)$ normiert.

In diesem Fall gelten insbesondere:

(iv) φ injektiv,

(v) $\angle(\varphi(v), \varphi(w)) = \angle(v, w)$ für alle $v, w \in V$,

(vi) $v \perp w \Rightarrow \varphi(v) \perp \varphi(w)$ für alle $v, w \in V$,

(vii) $\varphi(U^\perp) \subseteq \varphi(U)^\perp$ für alle $U \leq V$,

(viii) $\mathcal{B}$ Orthonormalsystem in $V \Rightarrow \varphi(\mathcal{B})$ Orthonormalsystem,

(ix) c Eigenwert von $\varphi \Rightarrow c \in \{\pm 1\}$.

Beweis. Übung. □

Definition. Ein $\varphi \in \mathrm{Hom}(V, W)$ heißt *orthogonal*, wenn die Bedingungen aus der Bemerkung gelten. Die Menge aller orthogonalen Automorphismen von V wird mit $O(V)$ bezeichnet.

Beispiel.

(i) Der Endomorphismus $-\mathrm{id}_V$ ist orthogonal, denn

$$\langle -v, -w \rangle = (-1)^2 \langle v, w \rangle = \langle v, w \rangle.$$

(ii) Ist φ bijektiv und orthogonal, so ist auch φ^{-1} orthogonal. (leichte Übung)

(iii) Es sei $\dim V = n < \infty$ und $\mathcal{B}$ eine Orthonormalbasis von V. Betrachtet man $\mathbb{R}^n$ mit dem Standard-Skalarprodukt, so ist $\kappa_{\mathcal{B}} : V \to \mathbb{R}^n$ orthogonal.

(iv) Orthogonalprojektionen sind im Allgemeinen nicht injektiv, also nicht orthogonal.

(v) Eine Spiegelung φ im Sinne von Definition 7.2.7 ist im Allgemeinen nicht orthogonal (Bild siehe Vorlesung), sondern nur dann, wenn $V(1, \varphi) \perp V(-1, \varphi)$. In diesem Abschnitt meinen wir mit *Spiegelung* stets eine orthogonale Spiegelung.

(vi) Jede Drehung des $\mathbb{R}^2$ um den Ursprung ist orthogonal, weil sie längenerhaltend ist.

Übung a. Man zeige: $O(V)$ ist eine Untergruppe von $\mathrm{Aut}(V)$.

Übung b. Es seien $\mathcal{A}, \mathcal{B}$ Orthonormalbasen von V, W. Weiter sei $\psi \in \mathrm{Hom}(W, V)$ die (eindeutig bestimmte) lineare Abbildung mit ${}^{\mathcal{A}}M^{\mathcal{B}}_{\psi} = ({}^{\mathcal{B}}M^{\mathcal{A}}_{\varphi})^t$. Man zeige: $\mathrm{Bild}\,\psi = (\mathrm{Kern}\,\varphi)^{\perp}$ und $\mathrm{Kern}\,\psi = (\mathrm{Bild}\,\varphi)^{\perp}$.
Hinweis: Beispiel 8.2.5.

Beweis. Es ist $\mathrm{Bild}\,\psi = \kappa_{\mathcal{B}}^{-1}(\mathrm{SR}(A^t)) = \kappa_{\mathcal{B}}^{-1}(\mathbb{L}_0(A)^{\perp}) = \kappa_{\mathcal{B}}^{-1}(\mathbb{L}_0(A))^{\perp} = \mathrm{Kern}(\varphi)^{\perp}$. □

8.5.2 Orthogonale Matrizen

Bemerkung a. Für $A \in \mathbb{R}^{n\times n}$ sind folgende Aussagen äquivalent:

(i) $A^t A = E_n$.

(ii) $A \in \mathrm{GL}_n(\mathbb{R})$ und $A^{-1} = A^t$.

(iii) Die Spalten von A bilden eine Orthonormalbasis von $\mathbb{R}^n$ bzgl. $\langle , \rangle_{\mathcal{E}}$.

(iv) Die Zeilen von A bilden eine Orthonormalbasis von $\mathbb{R}^n$ bzgl. $\langle , \rangle_{\mathcal{E}}$.

In diesem Fall gelten insbesondere:

(v) $|\det A| = 1$.

Beweis. Übung. □

Definition. $A \in \mathbb{R}^{n\times n}$ heißt *orthogonal*, wenn die Bedingungen aus Bemerkung a gelten. Die Menge aller orthogonalen $n \times n$-Matrizen

$$O(n) := \{A \in \mathbb{R}^{n\times n} \mid A^t A = E_n\}$$

wird *orthogonale Gruppe* genannt.

Übung. Man zeige $O(n) \leq \mathrm{GL}_n(\mathbb{R})$ mit $O(n)^t = O(n)$.

Satz. *Es sei $0 < n = \dim V < \infty$ und $\mathcal{B}$ sei eine Orthonormalbasis von V. Für $\varphi \in \mathrm{End}(V)$ sind äquivalent:*

(i) $\varphi \in O(V)$.

(ii) $\varphi(\mathcal{B})$ *Orthonormalbasis von* V.

(iii) $M_\varphi^{\mathcal{B}} \in O(n)$.

In diesem Fall gelten insbesondere:

(iv) $|\det \varphi| = 1$,

(v) $\varphi \in \mathrm{Aut}(V)$,

(vi) $U \leq V$ *φ-invariant* $\Rightarrow U^\perp$ *φ-invariant.*

Beweis. Es gilt: φ Automorphismus $\Leftrightarrow \varphi(\mathcal{B})$ Basis $\Leftrightarrow M_\varphi^{\mathcal{B}}$ invertierbar. Wir können also voraussetzen, dass dies der Fall ist. Die Äquivalenz von (i) und (ii) folgt aus Teil (ii) von Bemerkung 8.5.1. Da $\mathcal{B}$ Orthonormalbasis bzgl. $\langle , \rangle$ ist, gilt $\langle , \rangle = \langle , \rangle_{\mathcal{B}}$. Nach Definition von $\langle , \rangle_{\mathcal{B}}$ ist $\varphi(\mathcal{B})$ ONB bzgl. $\langle , \rangle_{\mathcal{B}}$ genau dann, wenn $\kappa_{\mathcal{B}}(\mathcal{B})$ Orthonormalbasis von $\mathbb{R}^n$ bzgl. $\langle , \rangle_{\mathcal{E}}$ ist. Da $\kappa_{\mathcal{B}}(\mathcal{B})$ gerade aus den Spalten von $M_\varphi^{\mathcal{B}}$ besteht, folgt die Äquivalenz von (ii) und (iii) aus Bemerkung aiii. Der Rest ist Übung. □

Beispiel.

(i) $A \in O(1) \Leftrightarrow A = (1)$ oder $A = (-1)$.

(ii) Die orthogonalen 1×1-Matrizen sind genau (1) und (-1).

(iii) Eine Diagonalmatrix ist genau dann orthogonal, wenn alle Diagonaleinträge gleich ± 1 sind.

(iv) Die 2×2-Matrizen R_α (Drehung) und S_α (Spiegelung) aus Beispiel 6.5.1 sind orthogonal. Z.B.

$$R_{\pi/4} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in O(2).$$

(v) Die Matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (Scherung) hat $\det A = 1$, ist aber nicht orthogonal.

Bemerkung b. Es seien $0 < n = \dim V < \infty$ und $\mathcal{B}, \mathcal{B}'$ zwei Orthonormalbasen von V. Dann ist ${}^{\mathcal{B}}T^{\mathcal{B}'} \in O(n)$.
Als Folge davon liesst sich der Spektralsatz so: Jede symmetrische reelle Matrix $A \in \mathbb{R}^{n \times n}$ ist *orthogonal diagonalisierbar*, d.h. es gibt ein $T \in O(n)$ derart, dass $T^{-1}AT = T^tAT$ eine Diagonalmatrix ist.

Beweis. Wähle $\varphi \in \mathrm{Aut}(V)$ mit $\varphi(\mathcal{B}) = \mathcal{B}'$. Dann ist ${}^{\mathcal{B}}T^{\mathcal{B}'} = M_\varphi^{\mathcal{B}} \in O(n)$ aufgrund des Satzes. Nach dem Spektralsatz gibt es eine Orthonormalbasis $\mathcal{B}$ von $\mathbb{R}^n$ bzgl. $\langle , \rangle$, die aus Eigenvektoren von A besteht. Folglich ist $T := {}^{\mathcal{E}}T^{\mathcal{B}} \in O(n)$ und $T^{-1}AT$ eine Diagonalmatrix. □

Übung. Zeige, dass jede orthogonal diagonalisierbare Matrix reelle Matrix symmetrisch ist.

8.5.3 $O(2)$

Satz. *Jede Matrix* $A \in O(2)$ *hat die Form* $A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = R_\alpha$ *oder* $A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} = S_{\alpha/2}$ *mit* $\alpha \in (-\pi, \pi]$.

Bemerkung. Wir betrachten $\mathbb{R}^2$ bzgl. $\langle , \rangle_\mathcal{E}$.

(i) Die normierten Vektoren aus $\mathbb{R}^2$ sind genau die Vektoren der Form $\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$ mit $\alpha \in [-\pi, \pi]$.

(ii) Zu jedem normierten $v = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 \backslash \{0\}$ gibt es genau zwei normierte $w \in \mathbb{R}^2$ mit $v \perp w$; diese lauten $w = \begin{pmatrix} -b \\ a \end{pmatrix}$ und $w = \begin{pmatrix} b \\ -a \end{pmatrix}$.

Beweis. (i) Sei $v = \begin{pmatrix} a \\ b \end{pmatrix}$ normiert, d.h. $a^2 + b^2 = 1$. Sei $\alpha = \angle(v, e_1)$, d.h. $\alpha \in [0, \pi]$ mit $\cos\alpha = \langle v, e_1 \rangle = a$. Wegen $\sin^2 = 1 - \cos^2$ (Analysis) folgt $\sin^2\alpha = 1 - a^2 = b^2$, also $\sin\alpha = \pm b$. Falls $b < 0$, dann ersetze α durch $-\alpha$. So bekommen wir $\alpha \in (-\pi, \pi]$ mit $\cos\alpha = a$ und $\sin\alpha = b$. Umgekehrt ist jeder Vektor dieser Form normiert wegen $\sin^2 + \cos^2 = 1$.

(ii) Wegen $\dim\langle v \rangle = 1$ ist $\dim\langle v \rangle^\perp = 2 - 1 = 1$ und jeder 1-dimensionale $\mathbb{R}$-Vektorraum enthält genau 2 normierte Vektoren. Die angegebenen w sind verschieden und erfüllen offenbar $v \perp w$. □

Beweis des Satzes. Es seien $s_1, s_2 \in \mathbb{R}^2$ die Spalten von $A \in O(2)$. Dann ist $\|s_1\| = \|s_2\| = 1$ und $s_1 \perp s_2$. Nach Teil (i) der Bemerkung gibt es $\alpha \in (-\pi, \pi]$ mit $s_1 = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$. Nach Teil (ii) der Bemerkung folgt $s_2 = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$ oder $s_2 = \begin{pmatrix} \sin\alpha \\ -\cos\alpha \end{pmatrix}$. □

Folgerung (Übersicht über die orthogonalen Endomorphismen von $\mathbb{R}^2$)**.**

Abbildung	Drehung um $\alpha \in (-\pi, \pi) \setminus \{0\}$	Drehung um π	Spiegelung
Matrix	$R_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$	$R_\pi = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$S_\alpha = \begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}$
Determinante	1	1	-1
Spur	$2\cos\alpha$	-2	0
char. Polynom	$X^2 - 2\cos(\alpha)X + 1$	$(X+1)^2$	$(X+1)(X-1)$
Eigenwerte	keine	-1,-1	-1,1
Eigenvektorbasis	keine	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}, \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$
Diagonalform	nicht diag.bar	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
Orthogonalität	eigentlich	eigentlich	uneigentlich
selbstadjungiert	nein	ja	ja

Beispiel. Welche Art von Abbildung beschreibt $A = \frac{1}{5}\begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix}$?

1. $A^tA = E_2 \Rightarrow A \in O(2)$.
2. $\det A = -1 \Rightarrow A$ Spiegelung.

Wie lautet die Spiegelachse?

Die Spiegelachse ist gerade der Eigenraum $V(1, A) = \mathbb{L}_0(A - E) = \langle \begin{pmatrix} 3 \\ 1 \end{pmatrix} \rangle$.

Übung. Bei einer Spiegelung $A \in O(2)$ ist die Spiegelachse gleich $\langle v + Av \rangle$ für jedes $v \in \mathbb{R}^2 \setminus \{0\}$. Man prüfe die Aussage an dem obigen Beispiel.

8.5.4 $SO(n)$

Es sei $0 < \dim V = n < \infty$.

Definition. Eine Matrix $A \in O(n)$ bzw. ein Endomorphismus $\varphi \in O(V)$ heißt *eigentlich orthogonal*, wenn die Determinante 1 ist, und *uneigentlich orthogonal*, wenn die Determinante -1 ist.

Die Menge aller eigentlich orthogonalen Matrizen bzw. Endomorphismen

$$SO(n) := \{A \in O(n) \mid \det A = 1\},$$
$$SO(V) := \{\varphi \in O(V) \mid \det \varphi = 1\}$$

nennen wir die *spezielle orthogonale Gruppe*.

Bemerkung.

(i) $SO(n) \leq O(n) \leq \mathrm{GL}_n(\mathbb{R})$ und $SO(V) \leq O(V) \leq \mathrm{End}(V)$. Die uneigentlich orthogonalen Matrizen bzw. Endomorphismen bilden keine Gruppe. Aufgrund der Multiplikativität der Determinante gilt

$$\begin{aligned} \text{eigentlich} \circ \text{eigentlich} &= \text{eigentlich}, \\ \text{eigentlich} \circ \text{uneigentlich} &= \text{uneigentlich}, \\ \text{uneigentlich} \circ \text{uneigentlich} &= \text{eigentlich}. \end{aligned}$$

(ii) Betrachte $V = \mathbb{R}^2$ mit $\langle , \rangle_{\mathcal{E}}$. Laut Folgerung 8.5.3 sind die eigentlich orthogonalen Endomorphismen genau die Drehungen und die uneigentlich orthogonalen Endomorphismen genau die (orthogonalen) Spiegelungen. (Das gilt nicht für $V = \mathbb{R}^n$ mit $n \geq 3$, vgl. Beispiel (8.5.5) unten). Ausserdem:

$$\begin{aligned} \text{Drehung} \circ \text{Spiegelung} &= \text{Spiegelung}, \\ \text{Spiegelung} \circ \text{Spiegelung} &= \text{Drehung}. \end{aligned}$$

Jede Drehung ist das Produkt von zwei Spiegelungen, denn

$$R_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = S_{\alpha/2} \cdot S_0.$$

8.5.5 $O(n)$

Es sei $0 < \dim V = n < \infty$.

Satz. *Es sei $\varphi \in \mathrm{End}(V)$ orthogonal. Dann gibt es eine Orthonormalbasis $\mathcal{B}$ von V so, dass*

$$M_\varphi^{\mathcal{B}} = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & A_1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & A_k \end{pmatrix},$$

wobei $A_k \in SO(2)$ mit $A_k \neq \pm E_2$. Lässt man $A_k = -E_2$ zu, so bekommt

man die Form

$$M_\varphi^{\mathcal{B}} = \begin{pmatrix} \pm 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & A_1 & & \\ & & & & & \ddots & \\ & & & & & & A_l \end{pmatrix},$$

in der $\det\varphi$ *gleich dem Eintrag oben links ist.*

Beweis. Wir führen Induktion nach $n = \dim V$. Der Fall $n = 1$ ist klar, da es nur die orthogonalen Matrizen (-1) und (1) gibt. Sei nun $n > 1$. Nach Folgerung 2.6.6 (aus dem Fundamentalsatz der Algebra) zerfällt χ_φ in Faktoren vom Grad ≤ 2. Nach Folgerung 7.5.3 (aus dem Satz von Cayley-Hamilton) gibt es einen nicht-trivialen φ-invarianten Unterraum U der Dimension ≤ 2. Wähle ein soches U von minimaler Dimension. Nach Satz 8.5.2 ist auch $U^\perp$ φ-invariant. Wähle eine beliebige Orthonormalbasis $\mathcal{B}_1$ von U. Im Fall $\dim U = 1$ ist $A := M_{\varphi|_U}^{\mathcal{B}_1}$ eine orthogonale 1×1-Matrix, also (1) oder (-1). Im Fall $\dim U = 2$ ist A eine orthogonale 2×2-Matrix, also eine der Matrizen aus der Tabelle aus 8.5.3. Da $\varphi|_U$ keinen 1-dimensionalen φ-invarianten Unterraum enthält, ist $A \in SO(2)$ und $A \neq -E_2$. Nach Induktionsvoraussetzung können wir eine Orthonormalbasis $\mathcal{B}_2$ von $U^\perp$ so wählen, dass $B := M_{\varphi|_{U^\perp}}^{\mathcal{B}_2}$ eine Form wie in der Aussage hat. Hängt man $\mathcal{B}_1$ und $\mathcal{B}_2$ aneinander, so bekommt man die gesuchte Basis $\mathcal{B}$ (bis auf Reihenfolge), denn $M_\varphi^{\mathcal{B}} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$. □

Beispiel $(n = 3)$**.** Es sei $A \in O(3)$. Nach dem Satz gibt es eine Orthonormalbasis (v_1, v_2, v_3) des $\mathbb{R}^3$, bzgl. der φ_A eine Matrix der Form

$$\begin{pmatrix} 1 & \\ & A \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} -1 & \\ & A \end{pmatrix} \quad \text{mit } A \in SO(2) \text{ hat.}$$

(Hier ist auch $A = E_2$ zugelassen.) Der erste Fall ist der eigentlich orthogonale, der zweite der uneigentlich orthogonale. Da A eine Drehmatrix ist, beschreibt $\begin{pmatrix} 1 & \\ & A \end{pmatrix}$ eine Drehung um die v_1-Achse. Aufgrund der Zerlegung $\begin{pmatrix} -1 & \\ & A \end{pmatrix} = \begin{pmatrix} -1 & \\ & E_2 \end{pmatrix} \cdot \begin{pmatrix} 1 & \\ & A \end{pmatrix}$ beschreibt $\begin{pmatrix} 1 & \\ & A \end{pmatrix}$ eine Drehung um die v_1-Achse mit anschliessender Spiegelung an der $\langle v_2, v_3 \rangle$-Ebene beschreibt.

An diesem Beispiel sieht man, dass es uneigentlich orthogonale Matrizen $\begin{pmatrix} -1 & \\ & A \end{pmatrix}$ gibt, die keine Spiegelung beschreiben, nämlich wenn A nicht die Einheitsmatrix ist (Drehspiegelung).

Folgerung. *Im $\mathbb{R}^3$ ist*

(i) jeder eigentlich orthogonale Endomorphismus eine Drehung,

(ii) jeder eigentlich orthogonale Endomorphismus das Produkt von zwei Spiegelungen,

(iii) jeder uneigentlich orthogonale Endomorphismus das Produkt von drei Spiegelungen.

Beweis. Nach Bemerkung (8.5.4) ist jedes $A \in SO(2)$ das Produkt von zwei Spiegelmatrizen, d.h. $A = SS'$ mit $S, S' \in O(2)$ und $\det S, S' = -1$. Dann ist

$$\begin{pmatrix} 1 & \\ & A \end{pmatrix} = \begin{pmatrix} 1 & \\ & S \end{pmatrix} \cdot \begin{pmatrix} 1 & \\ & S' \end{pmatrix},$$
$$\begin{pmatrix} -1 & \\ & A \end{pmatrix} = \begin{pmatrix} -1 & \\ & E_2 \end{pmatrix} \cdot \begin{pmatrix} 1 & \\ & S \end{pmatrix} \cdot \begin{pmatrix} 1 & \\ & S' \end{pmatrix},$$

wobei alle Matrizen auf der rechten Seite Spiegelungen sind. □

Beispiel. Welche Art von Abbildung beschreibt $A = \frac{1}{11}\begin{pmatrix} 2 & -6 & 9 \\ 6 & -7 & -6 \\ 9 & 6 & 2 \end{pmatrix}$?

1. $A^t A = E_3 \Rightarrow A \in O(3)$.
2. $\det A = 1 \Rightarrow A$ Drehung.

Wie lauten Drehachse, Drehebene und Drehwinkel?

Die Drehachse ist gerade der Eigenraum $V(1, A) = \mathbb{L}_0(A - E) = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \rangle$.

Die Drehebene ist der Orthogonalraum $V(1, A)^\perp = \langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \rangle$.

Fügt man Orthonormalbasen von Drehachse und Drehebene zusammen, bekommt man z.B. die Orthonormalbasis $\mathcal{B} = \left(\frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right)$ von $\mathbb{R}^3$. Mit dem Basiswechselsatz berechnet man

$$M_\varphi^{\mathcal{B}} = \frac{1}{11}\begin{pmatrix} 11 & 0 & 0 \\ 0 & -7 & 6\sqrt{2} \\ 0 & -6\sqrt{2} & -7 \end{pmatrix}.$$

Für den Drehwinkel α gilt $\cos\alpha = -\frac{7}{11}$, d.h. $\alpha \approx 0.72\pi$.

Übung. Wie kann man den Drehwinkel im obigen Beispiel berechnen, ohne vorher $M_\varphi^{\mathcal{B}}$ bestimmen zu müssen?

Literaturverzeichnis

[1] M. Aigner. *Diskrete Mathematik.* Vieweg, 2004.

[2] H. Anton. *Lineare Algebra.* Spektrum, 1995.

[3] A. Beutelspacher. *Lineare Algebra.* Vieweg, 2003.

[4] G. Fischer. *Lineare Algebra.* Vieweg, 2005.

[5] S. Teschl G. Teschl. *Mathematik für Informatiker, Band 1.* Springer, 2007.

[6] K. Jänich. *Lineare Algebra.* Springer, 2003.

[7] A. Steger. *Diskrete Strukturen.* Springer, 2001.

[8] K. Meyberg und P. Vachenauer. *Höhere Mathematik.* Springer, 2001.