**Vorlesungsskript**

# Lineare Algebra

PROF. DR. WALTER GUBLER

im Wintersemester 2010/2011 und Sommersemester 2011
an der Eberhard-Karls-Universität Tübingen

gesetzt von JULIEN SESSLER und TANJA PAPADOPOULOU mit LaTeX

Korrektur gelesen von CHRISTIAN POWER

Letzte Änderung: 12. Juli 2012

# Vorwort

Dieses Skript wurde während meiner Vorlesung *Lineare Algebra* im WS 10/11 und SS 11 an der Eberhard-Karls-Universtität Tübingen von Julien Sessler und Tanja Papadopoulou getippt und von Christian Power überarbeitet, denen ich dafür vielmals danke. Das Skript kann nur für die Hörer meiner Vorlesung von Nutzen sein. Wer sich sonst für Lineare Algebra interessiert, der sei auf die Literaturliste am Ende verwiesen. Mein Dank geht auch an Christian Christensen, der das Skript gründlich gelesen und mich auf viele Fehler hingewiesen hat. Wir sind dem Leser dankbar, wenn er gefundene Fehler an walter.gubler@mathematik.uni-regensburg.de meldet.

## Notation

Mit $\mathbb{N}$ bezeichnen wir die natürlichen Zahlen mit 0. Eine echte Inklusion von Mengen bezeichnen wir mit $A \subset B$, wenn Gleichheit zugelassen ist, dann benützen wir $A \subseteq B$. Die Gruppe der invertierbaren Elemente eines Ringes $R$ bezeichnen wir mit $R^*$.

# Inhaltsverzeichnis

# Teil I.

# Lineare Algebra I

Kapitel I.

# 1 Einführung

## 1. Logik

**1.1.** In der Logik gibt es Aussagen. Wir gehen davon aus, dass eine Aussage entweder wahr oder falsch ist (*Widerspruchsfreiheit der Mathematik*).

**1.2 Beispiel.** Aussage $A$: *Jede Primzahl ist ungerade.* Diese Aussage ist falsch, da 2 gerade und eine Primzahl ist $\Rightarrow$ Widerspruchsbeweis. Aussage $B$: *Jede ungerade Quadratzahl hat den Rest* 1 *bei Division durch* 8. Diese Aussage ist wahr (Beweis: siehe 1.9).

**1.3.** Die Mathematik ist aus **Axiomen** aufgebaut. Das sind Aussagen, die „von allen“ als richtig anerkannt werden, aber die nicht beweisbar sind.

Als Beispiel erwähnen wir das **Parallelenaxiom** aus der ebenen euklidischen Geometrie: „Durch jeden Punkt $P$ außerhalb einer Geraden $g$, gibt es genau eine Parallele zu $g$.“

**1.4.** Aus gegebenen Aussagen $A$ und $B$ kann man folgendermaßen neue Aussagen bilden:

- $\neg A$ (Negation von $A$, „nicht $A$“)
- $A \wedge B$ ($A$ und $B$)
- $A \vee B$ ($A$ oder $B$)
- $A \Longrightarrow B$ (wenn $A$ gilt, dann gilt auch $B$, „aus $A$ folgt $B$“)
- $A \Longleftrightarrow B$ ($A$ gilt genau dann, wenn $B$ gilt, „$A$ ist äquivalent zu $B$“)

**1.5.** Es ist wahr$=1$ und falsch$=0$.

| A | B | $\neg A$ | $A \wedge B$ | $A \vee B$ | $A \Longrightarrow B$ | $A \Longleftrightarrow B$ |
|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 | 1 | 1 | 0 |
| 0 | 0 | 1 | 0 | 0 | 1 | 1 |

**1.6.** Interessante Aussagen erhält man, in dem man die Quantoren „für alle“ (kurz: $\forall$) oder „es existiert“ (kurz: $\exists$) benutzt. Die Negation von $\exists$ ist $\nexists$ („es existiert kein“). „Es existiert genau ein“ kürzt man mit $\exists!$ ab.

**1.7 Beispiel.** Die Aussage $A$ aus Beispiel 1.2 kann man folgendermaßen formulieren: $\forall\ p$ Primzahl $\Rightarrow$ $p$ ungerade (Vorsicht: falsch). Beachte, dass die Umkehrung auch nicht gilt! Die Negation der Aussage $A$ lautet: $\exists\ p$ Primzahl mit $p$ gerade (wahr, $p = 2$).

**1.8.** Um in der Mathematik die Wahrheit einer Aussage zu prüfen, muss man einen **Beweis** führen. Dabei muss man die Aussage aus früher bewiesenen Aussagen und Axiomen herleiten (mit Hilfe von „$\Longrightarrow$“)

**1.9 Beispiel.** Wir beweisen Aussage $B$ von 1.2: *Jede ungerade Quadratzahl hat bei Division durch 8 den Rest 1.*

**Beweis:** Sei $q$ ungerade Quadratzahl. Da die Quadrate gerader Zahlen wieder gerade sind, muss $q = (2k+1)^2$ gelten, für eine natürliche Zahl $k$.

$$q = 4 \cdot k^2 + 4 \cdot k + 1 = 4 \cdot k \cdot (k+1) + 1$$

Von den benachbarten Zahlen $k$, $k+1$ ist genau eine gerade und deshalb ist $4 \cdot k \cdot (k+1)$ ein Vielfaches von 8. Das zeigt die Behauptung. □

**1.10.** Eine wichtige (wahre) Aussage wird als **Theorem** bezeichnet. Ein Zwischenergebnis, das für den Beweis eines Theorems benötigt wird, heißt **Lemma**. Eine **Proposition** ist eine naheliegende oder weniger wichtige Aussage. Eine **Behauptung** ist ebenfalls eine Aussage, die man beweisen muss. Ein **Korollar** ist eine Folgerung. Eine **Vermutung** ist eine Aussage, die man für wahr hält, aber bis jetzt nicht beweisen kann.

## 2. Elementare Zahlentheorie

**2.1.** Die natürlichen Zahlen werden durch die folgenden 5 *Peano-Axiome* charakterisiert:

**P1** 0 ist eine natürliche Zahl.

**P2** Für jede natürliche Zahl $n$ gibt es einen Nachfolger $n^*$.

**P3** 0 ist kein Nachfolger einer natürlichen Zahl.

**P4** Wenn zwei natürliche Zahlen denselben Nachfolger haben, dann sind sie gleich.

**P5** Es gilt das Prinzip der *vollständigen Induktion*.

**2.2.** Das **Prinzip der vollständigen Induktion** ist ein wichtiges Beweismittel. Es funktioniert folgendermaßen: Wir nehmen an, dass wir zu jeder natürlichen Zahl $n$ eine Aussage $A(n)$ haben. Weiter soll folgendes gelten:

- $A(0)$ sei wahr (**Induktionsanfang**).
- Wenn $A(n)$ für eine natürliche Zahl $n$ wahr ist, dann ist auch $A(n^*)$ wahr (**Induktionsschritt**).

Die vollständige Induktion besagt dann, dass jedes $A(n)$ wahr ist.

**2.3 Beispiel.** Wir schreiben dann $0, 1 := 0^*, 2 := 1^*, 3 := 2^*, \ldots$ Es kennzeichnet „:=“ hierbei eine Definition. **Behauptung:** *Wir erhalten mit dieser Liste alle natürlichen Zahlen.*

**Beweis** mit vollständiger Induktion. **Induktionsanfang** $n = 0$: $A(0)$ ist wahr, denn 0 ist die erste Zahl der Liste. **Induktionsschritt** $n \mapsto n+1$: Da $A(n)$ wahr ist, kommt $n$ in der Liste vor. Dann ist $n^*$ das nächste Element der Liste. Damit ist $A(n^*)$ wahr. Dies beweist den Induktionsschritt. Mit vollständiger Induktion folgt die Behauptung. □

**2.4 Bemerkung.** Wir sehen aus 2.3, dass wir die gewohnten natürlichen Zahlen erhalten. Weiter macht 2.3 die vollständige Induktion plausibel. Wenn $A(0)$ wahr ist, dann müssen auch

$$A(1) = A(0^*),\ A(2) = A(1^*),\ A(3) = A(2^*), \ldots$$

wahr sein. Man kann **P5** aber nicht aus den anderen Axiomen beweisen.

Mit Hilfe der vollständigen Induktion kann man $+, \cdot, \leq$ für die natürlichen Zahlen beweisen. Dies ist allerdings im Folgenden zu ausschweifend und wird als bereits bekannt vorausgesetzt.

**2.5.** Bei der vollständigen Induktion kann es passieren, dass man den Induktionsanfang für mehrere Zahlen prüfen muss, wie im folgenden Beweis. Weiter ist intuitiv klar, dass man im Induktionsschritt annehmen darf, dass die Behauptung richtig ist für alle Zahlen kleiner als $n$ (statt nur für $n-1$ wie in **P5**). Wir sprechen von **verbesserter vollständiger Induktion**. Für einen Beweis siehe [Gu, 1.1.12]

**2.6 Theorem (Division mit Rest).** *Für natürliche Zahlen $a$ und $b$, $b \neq 0$, gibt es eindeutige natürliche Zahlen $q$, $r$ mit*

$$a = b \cdot q + r \quad \wedge \quad 0 \leq r < b$$

*Dann heißt $r$ der **Rest** von $a$ bei der Division durch $b$.*

**Beweis:** Wir beweisen zuerst die *Existenz* mit vollständiger Induktion nach $a$ bei fixem $b$. *Induktionsanfang.* Wenn $0 \leq a < b$, dann gilt die Behauptung mit $q = 0$, $r = a < b$. *Induktionsschritt.* Wir dürfen $a \geq b$ wählen. Wir wollen die Aussage für dieses $a$ beweisen unter der Induktionsannahme, dass die Aussage richtig ist für alle $0 \leq a' < a$. Wegen $a \geq b$ folgt $a' := a - b \geq 0$. Weil $b > 0$, muss $a' < a$. Nach Induktionsannahme existiert ein $q'$ und ein $r'$ mit

$$a' = q' \cdot b + r' \quad \wedge \quad 0 \leq r' < b$$

Also folgt die Behauptung für $a$ durch $a = a' + b = q' \cdot b + r' + b = (q'+1) \cdot b + r'$. Dies zeigt den Induktionsschritt. Mit verbesserter vollständiger Induktion folgt die Behauptung uneingeschränkt. Um dies zu verdeutlichen, bemerken wir, dass die Behauptung für $a = 0, \ldots, b-1$ aus dem Induktionsanfang folgt und mit dem Induktionsschritt folgt sie sukzessive für $a = b$, $a = b+1, \ldots$ Dies zeigt die Existenz von $q$ und $r$.

*Eindeutigkeit.* Es gelte

$$a = q \cdot b + r \quad \wedge \quad a = q' \cdot b + r'$$

mit $0 \leq r, r' < b$. Wir dürfen $r \geq r'$ annehmen.

$$r - r' = (q' - q) \cdot b$$

Andererseits gilt $r - r' \leq r < b$ Es folgt

$$0 \leq (q' - q) \cdot b < b$$

Deshalb muss $q = q'$ gelten und damit folgt $r = r'$. Dies zeigt die Eindeutigkeit. □

**2.7 Definition.** Eine natürliche Zahl $b$ heißt **Teiler** der natürlichen Zahl $a$ genau dann, wenn es eine natürliche Zahl $c$ gibt mit $a = b \cdot c$. Wir nennen $a$ dann **Vielfaches** von $b$. Notation: $b \mid a$. Wenn $b$ kein Teiler von $a$ ist, dann schreiben wir $b \nmid a$.

**2.8 Definition.** Eine **Primzahl** ist eine natürliche Zahl $p$, die genau zwei Teiler hat.

**2.9 Bemerkung.** Die Teiler der Primzahl $p$ sind 1 und $p$. 1 ist *keine* Primzahl. Die ersten Primzahlen sind $2, 3, 5, 7, 11, 13, 17, 19, 23, \ldots$

**2.10 Proposition.** *Gegeben seien natürliche Zahlen $a, b, c, q$ mit $a = q \cdot b + c$. Dann sind die gemeinsamen Teiler von $a$ und $b$ dieselben, wie die gemeinsamen Teiler von $b$ und $c$.*

**Beweis:** Sei $d$ ein gemeinsamer Teiler von $a$ und $b$, d.h. $a = d \cdot e \wedge b = d \cdot f$. Wir müssen zeigen, dass $d$ ein Teiler von $c$ ist. Dies folgt aus

$$c = a - q \cdot b = d \cdot e - q \cdot d \cdot f = d \cdot (e - q \cdot f)$$

Umgekehrt kann man analog zeigen, dass jeder gemeinsame Teiler von $b$ und $c$ auch ein Teiler von $a$ ist. □

**2.11 Lemma (von Bezout).** *Seien $m$ und $n$ **teilerfremde** natürliche Zahlen, d.h. $m$ und $n$ haben nur 1 als gemeinsamen Teiler. Dann gibt es ganze Zahlen $a$ und $b$ mit*

$$m \cdot a + n \cdot b = 1, \quad a, b \in \mathbb{Z}.$$

**Beweis:** Wir dürfen $n \leq m$ annehmen. Unter dieser Voraussetzung beweisen wir die Behauptung mit vollständiger Induktion nach $n$. *Induktionsanfang* $n = 0$: Weil $m$ und $n$ nur 1 als Teiler haben, folgt $m = 1$. Das hat die Lösung $1 \cdot 1 + b \cdot 0 = 1$.

*Induktionsschritt:* Wir müssen die Behauptung für $n \geq 1$ zeigen unter der Induktionsannahme, dass die Behauptung stimmt für alle $n'$ mit $0 \leq n' < n$. Wir machen Division mit Rest (Theorem 2.6). Es gibt natürliche Zahlen $q, r$ mit $m = q \cdot n + r \wedge 0 \leq r < n$. Wir setzen $m' := n$ und $n' := r$. Nach Proposition 2.10 sind $m'$ und $n'$ teilerfremd. Nach Induktionsannahme gibt es ganze Zahlen $a'$ und $b'$ mit $a' \cdot n + b' \cdot r = 1$. Also folgt der Induktionsschritt aus

$$1 = a' \cdot n + b' \cdot r = a'n + b'(m - qn)$$

$$1 = (a' - b'q)n + b'm$$

Wenn wir $a := b'$ und $b := a' - b'q$ setzen, erhalten wir eine Lösung. Mit verbesserter vollständiger Induktion folgt die Behauptung uneingeschränkt. □

**2.12 Proposition.** *Eine natürliche Zahl $p \geq 2$ ist genau dann eine Primzahl, wenn folgende Aussage $A$ wahr ist: Wenn $p \mid mn$ für natürliche Zahlen $m$ und $n$, dann gilt $p \mid m \vee p \mid n$.*

**Beweis:** Wenn man die Äquivalenz von zwei Aussagen beweisen will, muss man zwei Richtungen beweisen.

$\Longrightarrow$: Es sei $p$ eine Primzahl. Wir wollen zeigen, dass $A$ wahr ist. Seien also $m$ und $n$ natürliche Zahlen mit $p \mid mn$. Wenn $p$ schon ein Teiler von $m$ ist, dann sind wir fertig. Also dürfen wir annehmen, dass $p \nmid m$. Weil $p$ eine Primzahl ist, hat $p$ nur 1 und $p$ als Teiler und damit ist 1 der einzige gemeinsame Teiler von $p$ und $m$. Nach dem Lemma von Bezout (siehe 2.11) gibt es ganze Zahlen $a$ und $b$ mit

$$a \cdot m + b \cdot p = 1$$

Es gilt

$$n = 1 \cdot n = (a \cdot m + bp)n = amn + bpn$$

Nach der Summenregel (Proposition 2.10) muss $p$ auch $n$ teilen, weil $p$ sowohl $m \cdot n$ wie auch $bpn$ teilt. Dies zeigt, dass $A$ wahr ist.

$\Longleftarrow$: Wir nehmen an, dass $A$ wahr ist und wir müssen zeigen, dass $p$ Primzahl ist. Sei dazu $d \mid p$, d.h. $p = k \cdot d$. Wegen $A$ folgt $p \mid k$ oder $p \mid d$. Wenn $p \mid k$ folgt wegen $k \mid p$ sofort $k = p$. Dann ist $d = 1$. Wenn $p \mid d$, folgt wegen $d \mid p$ wieder $p = d$. Als Teiler von $p$ haben wir nur 1 und $p$. Damit ist $p$ eine Primzahl. □

**2.13 Theorem (Fundamentalsatz der Zahlentheorie).** *Jede natürliche Zahl $n > 1$ lässt sich als Produkt von Primzahlen schreiben. Diese Faktorisierung $n = p_1 \cdot p_2 \cdots p_r$ ist bis auf die Reihenfolge eindeutig.*

**Beweis:** *Existenz* mit Induktion nach $n$. *Induktionsanfang* $n = 2$: Da $n = 2$ eine Primzahl ist, wählen wir $p_1 = 2$ und $n = p_1$ ist die gewünschte Faktorisierung.

*Induktionsschritt:* Es sei $n > 2$ und wir nehmen an, dass alle $n' < n$ eine Primfaktorisierung haben. Wir wollen zeigen, dass $n$ eine Primfaktorisierung hat. Falls $n$ eine Primzahl ist, dann sind wir wie beim Induktionsanfang fertig. Also dürfen wir annehmen, dass $n$ keine Primzahl ist, d.h.

$$n = n' \cdot n''$$

mit $1 < n' < n$, $1 < n'' < n$. Nach Induktionsannahme gibt es Primfaktorisierungen für

$$n' = p_1 \cdot p_2 \cdots p_r \quad \wedge \quad n'' = q_1 \cdot q_2 \cdots q_s.$$

Dann ist $n = n' \cdot n'' = (p_1 \cdot p_2 \cdots p_r) \cdot (q_1 \cdot q_2 \cdots q_s)$. Mit verbesserter vollständiger Induktion folgt die Existenz der Primfaktorzerlegung.

*Eindeutigkeit* der Primfaktorisierung mit Induktion nach $n$. *Induktionsanfang* $n = 2$. Es sei $2 = p_1 \cdots p_r$ eine andere Primfaktorisierung. Weil 2 eine Primzahl ist, folgt leicht aus Proposition 2.12, dass $2 \mid p_i$ für ein $p_i$. Weil $p_i$ eine Primzahl ist, muss $2 = p_i$ sein. Durch kürzen folgt aus $2 = p_1 \cdots p_r$ sofort, dass nur eine Primzahl vorkommen kann, d.h. $r = 1$ und $p_1 = 2$.

*Induktionsschritt.* Sei $n \geq 3$ und wir nehmen an, dass jede Zahl $n' < n$ eine eindeutige Primfaktorisierung hat. Seien $n = p_1 \cdots p_r = q_1 \cdots q_s$ zwei Primfaktorisierungen. Aus Übungsaufgabe 2.3, dass $p_r \mid q_i$ für ein $i$ und damit $p_r = q_i$. Durch Umnummerieren dürfen wir $q_i = q_s$ annehmen. Es gilt dann durch kürzen

$$n' := q_1 \cdots q_{s-1} = p_1 \cdots p_{r-1} < n.$$

Nach Induktionsannahme gilt $s - 1 = r - 1$ und $q_1, \ldots, q_{s-1}$ stimmt bis auf Reihenfolge mit $p_1 \cdots p_{r-1}$ überein. Wegen $p_r = q_s$ stimmt also die Behauptung auch für $n$. Damit ist der Induktionsschritt gezeigt.

Mit vollständige Induktion folgt die Behauptung für $n$. □

## 3. Mengenlehre

Wir geben hier eine intuitive Einführung in die Mengenlehre. Für einen axiomatischen Zugang verweisen wir auf die Literatur. Die Mengenlehre eignet sich hervorragend, mathematische Zusammenhänge kurz und präzise darzustellen.

**3.1.** Wir fassen eine **Menge** $X$ als Zusammenfassung gewisser Objekte auf, die wir **Elemente** von $X$ nennen. Formal schreiben wir $x \in X \iff$ „$x$ ist Element von $X$“. $x \notin X \iff$ „$x$ ist kein Element von $X$“.

**3.2.** Die **leere Menge** wird mit $\emptyset$ oder $\{\}$ bezeichnet. Sie hat überhaupt kein Element.

**3.3.** Zwei Mengen $X, Y$ sind **gleich**, wenn gilt $x \in X \iff x \in Y$. Also gilt

$$\{1,5,7,4,7,3\} = \{1,3,4,5,7\}$$

Hier werden mit den Mengenklammern $\{\}$ die Elemente der Menge umfasst. Auf Wiederholung und Reihenfolge wird bei Mengen keinen Wert gelegt! Wir bezeichnen mit $|X|$ die **Anzahl Elemente** von $X$. Im obigen Beispiel haben wir $|\{1,5,7,4,7,3\}| = 5$. Falls es unendlich viele Elemente in $X$ hat, schreiben wir $|X| := \infty$.

**3.4.** In der Mathematik sind die Zahlbereiche besonders wichtige Mengen, die in der Analysis eingeführt werden.

- $\mathbb{N} :=$ Menge der natürlichen Zahlen (mit 0)
- $\mathbb{Z} :=$ Menge der ganzen Zahlen
- $\mathbb{Q} :=$ Menge der rationalen Zahlen
- $\mathbb{R} :=$ Menge der reellen Zahlen

**3.5.** Eine **Teilmenge** $Y$ von der Menge $X$ ist durch $y \in Y \implies y \in X$ charakterisiert, Wir schreiben dann $Y \subseteq X$. Es gilt z.B.:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$$

wobei $\subset$ bedeutet, dass „$\subseteq$“ und $\neq$ gilt.

**3.6.** Oft definiert man Teilmengen durch Aussagen. Sei $X$ eine Menge und $A(x)$ eine Aussage, die für $x \in X$ entweder wahr oder falsch ist. Dann definieren wir $Y$ als diejenige Teilmenge von $X$, für die $A(x)$ wahr ist. Formal schreiben wir

$$Y := \{x \in X \mid A(x)\}$$

„Die Menge aller $x \in X$, für die $A(x)$ gilt." Zum Beispiel definiert

$$\{m \in \mathbb{N} \mid \exists n \in \mathbb{N}, m = 2n\}$$

die Menge der geraden Zahlen in $\mathbb{N}$.

**3.7.** Für Teilmengen $Y_1, Y_2$ von $X$ haben wir folgende Operatoren:

$$\begin{array}{lcll} Y_1 \cup Y_2 & := & \{x \in X \mid x \in Y_1 \vee x \in Y_2\} & \text{Vereinigung} \\ Y_1 \cap Y_2 & := & \{x \in X \mid x \in Y_1 \wedge x \in Y_2\} & \text{Durchschnitt} \\ Y_1 \setminus Y_2 & := & \{x \in Y_1 \mid x \notin Y_2\} & \text{Differenz} \end{array}$$

Wenn $Y_1 \cap Y_2 = \emptyset$, dann heißen $Y_1$ und $Y_2$ **disjunkt**.

**3.8.** Oft haben wir mehrere oder sogar unendlich viele Mengen im Spiel. Formal bedeutet das, dass man eine **Indexmenge** $I$ hat und für jedes $i \in I$ hat man eine Menge $X_i$. Wir sprechen dann von einer **Familie** $(X_i)_{i \in I}$ von Mengen $X_i$.

**3.9.** Wenn die Familie $(Y_i)_{i \in I}$ aus Teilmengen $Y_i$, von einer einzigen Menge $X$ besteht (also $Y_i \subseteq X$), dann definieren wir die **Vereinigung**

$$\bigcup_{i \in I} Y_i := \{x \in X \mid \exists i \in I, x \in Y_i\}$$

Bei paarweise disjunkten Teilmengen schreiben wir: $\dot{\bigcup}_{i \in I} Y_i$. Den **Durchschnitt** schreiben wir

$$\bigcap_{i \in I} Y_i := \{x \in X \mid \forall i \in I \Longrightarrow x \in Y_i\}$$

Oft besteht die Indexmenge aus $I = \{1, \ldots, n\}$ bzw. $I = \mathbb{N}$, dann setzen wir

$$\bigcup_{i=1}^{n} Y_i := Y_1 \cup \cdots \cup Y_n := \bigcup_{i \in \{1, \ldots, n\}} Y_i$$

bzw.

$$\bigcup_{i=0}^{\infty} Y_i := \bigcup_{i \in \mathbb{N}} Y_i$$

Analoges gilt für den Durchschnitt.

**3.10.** Wir definieren das **kartesische Produkt** $X_1 \times X_2$ der Mengen $X_1$ und $X_2$ als die Menge der Paare $(x_1, x_2)$ mit $x_1 \in X_1$ und $x_2 \in X_2$. Dabei gilt $(x_1, x_2) = (x_1', x_2') :\Longleftrightarrow x_1 = x_1' \wedge x_2 = x_2'$.

**3.11 Beispiel.** Nehme $X_1 := X_2 := \mathbb{R}$, dann erhalten wir $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$. Allgemein setzen wir $X^2 := X \times X$.

**3.12.** Wir verallgemeinern das **kartesische Produkt** auf Familien $(X_i)_{i \in I}$ von beliebigen Mengen $X_i$ durch

$$\prod_{i \in I} X_i := \{(x_i)_{i \in I} \mid \forall i \in I \Longrightarrow x_i \in X_i\}$$

Dabei sind die Elemente des kartesischen Produkts Familien $(x_i)_{i \in I}$ von Elementen $x_i \in X_i$. Es gilt $(x_i)_{i \in I} = (x_i')_{i \in I} \Longleftrightarrow (\forall i \in I \Longrightarrow x_i = x_i')$.

**3.13 Beispiel.** In der Praxis hat die Familie oft die Form $X_1, \ldots, X_n$ und dann setzen wir

$$X_1 \times \cdots \times X_n = \prod_{i \in \{1,\ldots,n\}} X_i$$

Die Elemente dieses kartesischen Produktes sind **n-Tupel** $(x_1, \ldots, x_n)$ mit $x_i \in X_i$. Wenn sogar $X_1 := \cdots := X_n := X$, dann setzen wir $X^n := X_1 \times \cdots \times X_n$. Ein wichtiges Beispiel ist

$$\mathbb{R}^n := \{(\alpha_1, \ldots, \alpha_n) \mid \alpha_1, \ldots, \alpha_n \in \mathbb{R}\}$$

**3.14.** Eine **Abbildung** $f : X \longrightarrow Y$ zwischen Mengen $X$ und $Y$ ist eine Vorschrift, die jedem $x \in X$ genau ein $y \in Y$ zuordnet. Wir schreiben dafür $y = f(x)$ und oft auch $x \longmapsto f(x)$.

$X$ heißt der **Definitionsbereich** und $Y$ heißt der **Wertebereich** von $f$. Weiter ist

$$f(X) := \{f(x) \mid x \in X\}$$

das **Bild** von $f$. Die Abbildung heißt **injektiv**, wenn

$$\big(x_1, x_2 \in X, f(x_1) = f(x_2)\big) \Longrightarrow x_1 = x_2$$

Die Abbildung heißt **surjektiv**, wenn $Y = f(X)$ gilt. Eine Abbildung heißt **bijektiv**, wenn sie injektiv und surjektiv ist.

**3.15.** Die **identische Abbildung** der Menge $X$ ist definiert durch

$$\mathrm{id}_X : X \longrightarrow X,\ x \longmapsto x$$

**3.16.** Zwei beliebige Abbildungen kann man **nicht** verknüpfen. Das geht nur, wenn der Wertebereich $Y$ der zweiten Abbildung $f : X \longrightarrow Y$ gleich dem Definitionsbereich der ersten Abbildung $g : Y \longrightarrow Z$ ist. Die **Verknüpfung** $g \circ f$ ist dann definiert durch

$$g \circ f : X \longrightarrow Z,\ x \longmapsto g\big(f(x)\big)$$

**3.17.** Für eine Abbildung $f : X \longrightarrow Y$ und eine Teilmenge $W \subseteq Y$ definieren wir das **Urbild** von $W$ als

$$f^{-1}(W) := \{x \in X \mid f(x) \in W\}$$

**Warnung:** Dies ist eine Menge und sie hat nichts mit der Umkehrfunktion zu tun, die im Allgemeinen nicht zu existieren braucht.

**3.18.** Die **Umkehrfunktion** existiert nur für eine **bijektive** Abbildungen $f \colon X \to Y$ und ist die Abbildung, die jedem $y \in Y$, das eindeutige $x \in X$ zuordnet, mit $f(x) = y$. Die Existenz von $x \in X$ mit $f(x) = y$ folgt aus der Surjektivität von $f$ und die Eindeutigkeit von $x$ folgt aus der Injektivität. Die Umkehrfunktion wird oft mit $f^{-1}$ bezeichnet, weil nach Konstruktion

$$f^{-1} \circ f = \mathrm{id}_X \quad \wedge \quad f \circ f^{-1} = \mathrm{id}_Y$$

gilt.

**3.19.** Eine **Relation** auf der Menge $X$ ist eine Teilmenge $R \subseteq X \times X$. Wir sagen, dass $x_1$ **in Relation** zu $x_2$ steht, wenn $(x_1, x_2) \in R$. Wir schreiben dann $x \sim_R y$ oder einfach $x \sim y$ und vergessen $R$ wieder.

- $\sim$ heißt **reflexiv**, falls $x \sim x \ \forall x \in X$
- $\sim$ heißt **symmetrisch**, falls $x \sim y \Longrightarrow y \sim x$
- $\sim$ heißt **antisymmetrisch**, falls $(x \sim y \wedge y \sim x) \Longrightarrow x = y$

- $\sim$ heißt **transitiv**, falls $(x \sim y \wedge y \sim z) \Longrightarrow x \sim z$

**3.20.** Eine Relation $\sim$, die reflexiv, symmetrisch und transitiv ist, heißt Äquivalenzrelation. Dann heißen $x, y \in X$ mit $x \sim y$ **äquivalent**. Wir nennen

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

die **Äquivalenzklasse** von $x$.

**3.21 Proposition.** *Es sei $\sim$ eine Äquivalenzrelation auf $X$. Für $x, y \in X$ sind folgende Aussagen äquivalent*

*$A_1$) $x \sim y$;*

*$A_2$) $[x] = [y]$;*

*$A_3$) $[x] \cap [y] \neq \emptyset$.*

*Insbesondere ist $X$ disjunkte Vereinigung der Äquivalenzklassen zu $\sim$.*

**Beweis:** Übungsblatt 4, Aufgabe 1. $\square$

**3.22 Beispiel.**

$$X = \{\triangle \subseteq \mathbb{R}^2 \mid \triangle \text{ Dreieck}\}, \ \triangle \sim \triangle' :\Longleftrightarrow \triangle \text{ ähnlich zu } \triangle'.$$

Wieder bedeutet $:\Longleftrightarrow$, dass wir $\sim$ durch diese Eigenschaft definieren. Dann ist $\sim$ eine Äquivalenzrelation. Auch die Kongruenz von Figuren ist eine Äquivalenzrelation.

**3.23 Definition.** Eine **partielle Ordnung** $\leq$ auf einer Menge $X$ ist eine Relation auf $X$, die reflexiv, antisymmetrisch und transitiv ist. Wir schreiben $x < y$ für $x \leq y \wedge x \neq y$. Weiter sei $x \geq y :\Longleftrightarrow y \leq x$. Eine **Totalordnung** ist eine partielle Ordnung, für die

$$x \leq y \quad \vee \quad y \leq x$$

gilt.

**3.24 Beispiel.** Die übliche Ordnung $\leq$ auf $\mathbb{R}$ ist eine Totalordnung. Die **Potenzmenge** $\operatorname{Pot}(X) := \{A \mid A \subseteq X\}$ ist partiell geordnet durch $\subseteq$, aber keine Totalordnung.

**3.25.** Ein Element $x$ von $X$ heißt **maximal**, wenn

$$y \geq x \Longrightarrow y = x$$

gilt auf $X$. Weiter heißt $x$ eine **obere Schranke** von $Y \subseteq X$, wenn

$$\forall y \in Y \Longrightarrow y \leq x.$$

Eine obere Schranke von $X$ selber heißt **größtes Element** von $X$ und ist wegen der Antisymmetrie eindeutig.

**3.26.** Es ist klar, dass das größte Element von $X$ maximal ist. Es kann aber passieren, dass $X$ kein größtes Element hat aber immerhin maximale Elemente. Zum Beispiel hat $X = \{Y \subseteq \mathbb{N} \mid |Y| \leq 5\}$ alle 5-elementige Teilmengen $Y$ als maximale Elemente, aber kein größtes Element bzgl. $\subseteq$. $\mathbb{R}$ hat überhaupt keine maximalen Elemente.

**3.27.** Streng genommen müsste man die Mengenlehre axiomatisch aufziehen. Dabei muss man aufpassen, dass man mit Hilfe der Logik nicht zu „große“ Mengen macht. Zum Beispiel ist die „Menge“ aller Mengen keine Menge mehr, sondern man nennt das eine **Klasse**. Die Klassen sind eine Verallgemeinerung von Mengen, auf die wir hier nicht näher eingehen.

Damit umgeht man zum Beispiel das **Paradoxon von Russell**. Sei $C := \{X \mid X \text{ Menge } \wedge X \notin X\}$. Wäre $C$ eine Menge, dann wäre sowohl die Aussage

$$C \in C$$

wie auch ihre Negation

$$C \notin C$$

falsch und die **Widerspruchsfreiheit** der Mathematik wäre widerlegt. Bei vernünftigen mathematischen Problemen tauchen keine solcher logischen Schwierigkeiten auf. Zum Beispiel sind die Potenzmenge und kartesische Produkte von Mengen wieder Mengen.

**3.28.** In der modernen Mathematik braucht man auch das **Auswahlaxiom**, das besagt, dass es zu jeder Menge $F$ von nichtleeren Mengen eine Funktion $f$ gibt, die jedem $X \in F$ ein Element $f(X) \in X$ zuordnet. Man kann zeigen, dass dies äquivalent ist zu folgendem Axiom, was sich in der Praxis als nützlich erwiesen hat.

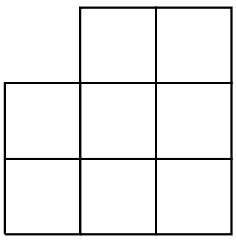
**3.29 Axiom (Zornsches Lemma).** Ist $X$ eine Menge mit einer partiellen Ordnung $\leq$ und besitzt jede total geordnete Teilmenge von $X$ eine obere Schranke in $X$, dann hat $X$ ein maximales Element.

**3.30.** Wie gesagt kann man das Zornsche Lemma nicht beweisen, außer man setzt das Auswahlaxiom voraus. Deshalb ist der Name Lemma unglücklich gewählt, aber nicht mehr zu ändern. Der axiomatische Aufbau der Mathematik geht auf David Hilbert zurück (Hilbert-Programm, 1921). Kurt Gödel hat 1931 gezeigt, dass man die Widerspruchsfreiheit der Mathematik nicht zeigen kann und dass es immer Aussagen gibt, die man nicht beweisen kann. Man nennt sie **unentscheidbar**.

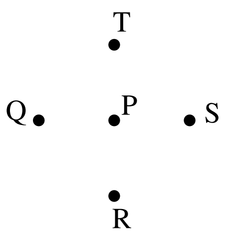
## 4. Lineare Gleichungssysteme

Wir geben eine Einführung in reelle lineare Gleichungssysteme.

**4.1 Beispiel.** Wie betrachten eine Platte, die aus Quadraten zusammengesetzt ist.

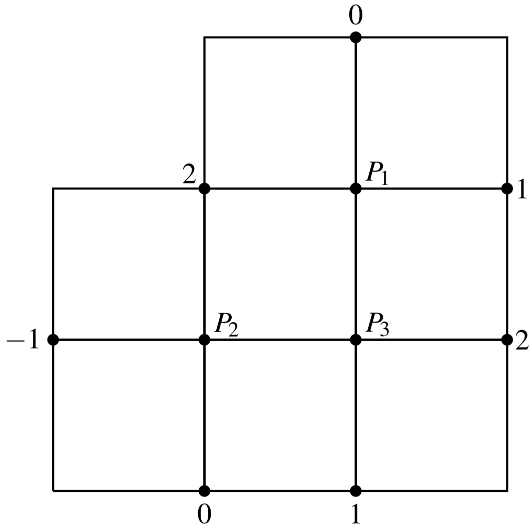


Wir geben uns eine Temperaturverteilung $u$ auf der Platte vor und wollen aus der Temperaturverteilung am Rand die Temperatur im Innern bestimmen. Dabei wollen wir die Temperatur nur in den Gitterpunkten bestimmen. Aus der Physik folgt




$$u(P) = \frac{1}{4}\big(u(Q) + u(R) + u(S) + u(T)\big) \tag{I.1}$$

Das gilt nur näherungsweise, das kümmert uns aber nicht! Wir gehen nur von folgenden Randwerten aus:




Gesucht: $x_1 = u(P_1), x_2 = u(P_2), x_3 = u(P_3)$.

$$\begin{aligned}
&\text{(I.1) in } P_1: \qquad x_1 = \frac{1}{4}(x_3 + 1 + 0 + 2)\\
&\text{(I.1) in } P_2: \qquad x_2 = \frac{1}{4}(x_3 + 2 - 1 + 0)\\
&\text{(I.1) in } P_3: \qquad x_3 = \frac{1}{4}(x_1 + x_2 + 1 + 2)
\end{aligned}$$

Wir erhalten das inhomogene lineare Gleichungssystem:

$$\begin{array}{cccc}
x_1 & & -\frac{1}{4}x_3 & = \frac{3}{4}\\
& x_2 & -\frac{1}{4}x_3 & = \frac{1}{4}\\
-\frac{1}{4}x_1 & -\frac{1}{4}x_2 & +x_3 & = \frac{3}{4}
\end{array}$$

**4.2.** Ein lineares Gleichungssystem hat die Form

$$\begin{array}{c}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1\\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2\\
\dots\dots\dots\dots\dots\dots\dots\dots\\
a_{m1}x_1 + \ a_{m2} \ + \cdots + a_{mn}x_n = b_m
\end{array} \tag{I.2}$$

für gegebene reelle Zahlen $a_{ij}$ und $b_k$. Weiter nennt man $x_1, \dots, x_n$ **Unbekannte** oder **Variable**. Weiter **Homogen** heißt (I.2) ein lineares Gleichungssystem, wenn alle $b_i = 0$ sind Dabei ist (I.2) ein **inhomogenes** lineares Gleichungssystem, wenn ein $b_i \neq 0$.

**4.3 Proposition.** *Seien* $\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n$ *und* $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \in \mathbb{R}^n$ *Lösungen eines homogenen linearen Gleichungssystems. Dann sind auch* $\begin{pmatrix} \alpha_1+\beta_1 \\ \vdots \\ \alpha_n+\beta_n \end{pmatrix}$ *und* $\begin{pmatrix} \lambda\cdot\alpha_1 \\ \vdots \\ \lambda\cdot\alpha_n \end{pmatrix}$ *Lösungen des homogenen linearen Gleichungssystems für alle* $\lambda \in \mathbb{R}$.

**Beweis:** Weil es Lösungen sind, gilt

$$a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1n}\alpha_n = 0 \quad \text{und} \quad a_{11}\beta_1 + a_{12}\beta_2 + \cdots + a_{1n}\beta_n = 0.$$

Addiert man diese beiden Gleichungen, dann folgt

$$a_{11}(\alpha_1+\beta_1)+a_{12}(\alpha_2+\beta_2)+\cdots+a_{1n}(\alpha_n+\beta_n)=0.$$

Also erfüllt $\begin{pmatrix}\alpha_1+\beta_1\\ \vdots\\ \alpha_n+\beta_n\end{pmatrix}$ die erste Gleichung. Analog für die anderen Gleichungen. Damit folgt die erste Behauptung. Durch Multiplizieren der Gleichung mit $\lambda$ folgt die zweite Behauptung. □

**4.4 Proposition.** *Wir nehmen an, dass das inhomogene Gleichungssystem* (I.2) *eine Lösung* $\begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix} \in \mathbb{R}^n$ *hat. Wir nennen sie die* ***spezielle Lösung****. Dann sind die folgenden Aussagen äquivalent:*

1. $\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}$ *ist Lösung des zugehörigen homogenen linearen Gleichungssystems.*

2. $\begin{pmatrix}\alpha_1+\gamma_1\\ \vdots\\ \alpha_n+\gamma_n\end{pmatrix}$ *ist Lösung von* (I.2).

Die allgemeine Lösung eines inhomogenen linearen Gleichungssystems ergibt sich aus Summe der allgemeinen Lösung des zugehörigen homogenen linearen Gleichungssystems und einer speziellen Lösung des inhomogenen Systems.

**Beweis:** Siehe Übungsblatt 3, Aufgabe 2. □

**4.5 Illustration.** Wir betrachten folgendes einfache Beispiel:

$$\begin{array}{rrcr} x_1 & -x_2 & = & 1\\ -2x_1 & +2x_2 & = & -2 \end{array} \tag{I.3}$$

Spezielle Lösung von (I.3): $x_1=1, x_2=0$. Allgemeine Lösung des zugehörigen homogenen linearen Gleichungssystems:

$$\begin{array}{rrcr} x_1 & -x_2 & = & 0\\ -2x_1 & +2x_2 & = & 0 \end{array} \tag{I.4}$$

ist gleich $x_1=\lambda, x_2=\lambda$ $(\forall\lambda\in\mathbb{R})$. Damit ist die allgemeine Lösung von (I.3) = spezielle Lösung von (I.3) + allgemeine Lösung von (I.4) = $\binom{1}{0}+\binom{\lambda}{\lambda}=\binom{1+\lambda}{\lambda}$.

**4.6 Bemerkung.** Ein homogenes lineares Gleichungssystem hat immer $\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}\in\mathbb{R}^n$ als Lösung. Es gibt inhomogene lineare Gleichungssysteme, die keine Lösung besitzen. Zum Beispiel:

$$\begin{array}{rrcr} x_1 & -x_2 & = & 0\\ -2x_1 & +2x_2 & = & 1 \end{array}$$

**4.7.** Folgende Operationen ändern die Lösungsmenge von (I.2) nicht und sind deshalb zulässig.

a) Vertauschen von Gleichungen,

b) Multiplizieren einer Gleichung mit $\lambda\neq 0$,

c) Zu einer Gleichung ein Vielfaches einer **anderen** Gleichung addieren.

Die Beweise sind einfach, weil man kann die Operationen wieder rückgängig machen und ändern deshalb die Lösungsmenge nicht (Äquivalenzumformung). Mit diesen Operationen können wir jedes lineare Gleichungssystem mit folgendem Algorithmus lösen.

**4.8 Beispiel.** Mit dem **Gauß-Algorithmus** kann man jedes lineare Gleichungssystem systematisch lösen. Wir führen das an unserem einleitenden Beispiel 4.1 vor:

$$x_1 \qquad -\frac{1}{4}x_3 = \frac{3}{4} \tag{I.5}$$

$$x_2 - \frac{1}{4}x_3 = \frac{1}{4} \tag{I.6}$$

$$-\frac{1}{4}x_1 - \frac{1}{4}x_2 + \quad x_3 = \frac{3}{4} \tag{I.7}$$

**1.Schritt** Wir vertauschen die Gleichungen bis die oberste Gleichung ein $x_1$ enthält. Hier nicht nötig!
**2.Schritt** Durch ein Addieren eines geeigneten Vielfaches der oberen Gleichung zu den anderen Gleichungen bringt man dort die Unbekannte $x_1$ weg. Wir machen deshalb (I.7')=(I.7)+$\frac{1}{4}$(I.5)].

$$x_1 \qquad -\frac{1}{4}x_3 = \frac{3}{4} \tag{I.5}$$

$$x_2 - \frac{1}{4}x_3 = \frac{1}{4} \tag{I.6}$$

$$-\frac{1}{4}x_2 + \frac{15}{16}x_3 = \frac{15}{16} \tag{I.7'}$$

Beachte, dass dies eine Äquivalenzumformung aus (I.7) war, denn (I.7')=(I.7)+$\frac{1}{4}$(I.5). Ab jetzt lassen wir die erste Gleichung unverändert und beginnen wieder mit dem 1. Schritt für $x_2$ für die restlichen Gleichungen. Der 1. Schritt ist wieder unnötig und der 2. Schritt legt (I.7)"= (I.7)' + $\frac{1}{4}$(I.6) nahe.

$$x_1 \qquad -\frac{1}{4}x_3 = \frac{3}{4} \tag{I.5}$$

$$x_2 - \frac{1}{4}x_3 = \frac{1}{4} \tag{I.6}$$

$$\frac{7}{8}x_3 = 1 \tag{I.7''}$$

Jetzt müssen wir (I.5) und (I.6) unverändert lassen und beginnen wieder mit dem 1.Schritt für die restlichen Gleichungen. Das bringt hier nichts mehr, weil wir nur noch eine Gleichung übrig ist. Wir können jetzt das Gleichungssystem von unten nach oben lösen.

$$\text{(I.7'')} \Longrightarrow x_3 = \frac{8}{7}, \quad \text{(I.6)} \Longrightarrow x_2 = \frac{15}{28}, \quad \text{(I.5)} \Longrightarrow x_1 = \frac{29}{28}$$

Dies gibt die Wärmeverteilung auf der Platte in Beispiel 4.1.

**4.9 Beispiel.**

$$\begin{array}{rrrrrcl} & x_2 & -x_3 & +x_4 & +x_5 & = & 0 \\ 2x_1 & -x_2 & +x_3 & +x_4 & -x_5 & = & 0 \\ 4x_1 & +x_2 & -x_3 & +x_4 & +x_5 & = & 1 \\ 2x_1 & +x_2 & -x_3 & +x_4 & +2x_5 & = & 0 \end{array} \tag{I.8}$$

Wir wollen dieses inhomogene lineare Gleichungssystem lösen. Mathematiker sind faul, deshalb lassen sie die Variablen weg und schreiben nur das folgene Schema:

$$\left.\begin{array}{rrrrr|l} 0 & 1 & -1 & 1 & 1 & 0 \\ 2 & -1 & 1 & 1 & -1 & 0 \\ 4 & 1 & -1 & 1 & 1 & 1 \\ 2 & 1 & -1 & 1 & 2 & 0 \end{array}\right.$$

Man muss sich die Unbekannte $x_i$ immer bei der $i$-ten Spalte dazu denken und das $=$ statt dem senkrechten Strich. **1. Schritt:** Vertauschen der 1. und 2. Gleichung. **2. Schritt:** $x_1$ eliminieren.

$$\begin{array}{rrrrr|l} 2 & -1 & 1 & 1 & -1 & 0 \;\; |\cdot(-2) \quad |\cdot(-1) \\ 0 & 1 & -1 & 1 & 1 & 0 \\ 4 & 1 & -1 & 1 & 1 & 1 \;\; + \\ 2 & 1 & -1 & 1 & 2 & 0 \;\; + \end{array} \quad \underset{\rightsquigarrow}{\text{2.Schritt}} \quad \begin{array}{rrrrr|l} 2 & -1 & 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & 1 & 1 & 0 \\ 0 & 3 & -3 & -1 & 3 & 1 \\ 0 & 2 & -2 & 0 & 3 & 0 \end{array}$$

Ab jetzt bleibt die erste Zeile unverändert und der Prozess wird mit den restlichen Zeilen für $x_2$ wiederholt.

$$\left.\begin{array}{ccccc|l}
2 & -1 & 1 & 1 & -1 & 0 \\
0 & 1 & -1 & 1 & 1 & 0 \;|\cdot(-3) \quad |\cdot(-2) \\
0 & 3 & -3 & -1 & 3 & 1 \; + \\
0 & 2 & -2 & 0 & 3 & 0 \; +
\end{array}\right. \rightsquigarrow
\begin{array}{ccccc|c}
2 & -1 & 1 & 1 & -1 & 0 \\
0 & 1 & -1 & 1 & 1 & 0 \\
0 & 0 & 0 & -4 & 0 & 1 \\
0 & 0 & 0 & -2 & 1 & 0
\end{array}$$

Da $x_3$ in Zeile 3 und 4 nicht vorkommen, eliminiert man gleich $x_4$.

$$\begin{array}{ccccc|l}
2 & -1 & 1 & 1 & -1 & 0 \\
0 & 1 & -1 & 1 & 1 & 0 \\
0 & 0 & 0 & -4 & 0 & 1 \;|\cdot\left(-\frac{1}{2}\right) \\
0 & 0 & 0 & -2 & 1 & 0 \; +
\end{array} \rightsquigarrow
\begin{array}{ccccc|c}
2 & -1 & 1 & 1 & -1 & 0 \\
0 & 1 & -1 & 1 & 1 & 0 \\
0 & 0 & 0 & -4 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & -\frac{1}{2}
\end{array}$$

Wenn wir so eine **Zeilenstufenform** erreicht haben, ist der Gauß-Algorithmus fertig und wir finden die Lösung wieder von unten nach oben: $x_5 = -\frac{1}{2}$, $x_4 = -\frac{1}{4}$, $x_3 = x_3$, $x_2 = x_3 + \frac{3}{4}$, $x_1 = \frac{1}{4}$ allgemeine Lösung:

$$\begin{pmatrix} \frac{1}{4} \\ \lambda + \frac{3}{4} \\ \lambda \\ -\frac{1}{4} \\ -\frac{1}{2} \end{pmatrix} \quad \text{mit } \lambda \in \mathbb{R}.$$

Kapitel II.

# 2 Algebraische Grundlagen

## 1. Gruppen

Gruppen treten in der Mathematik und den Naturwissenschaften da auf, wo es Symmetrien gibt. In diesem Abschnitt wollen wir die grundlegenden Definitionen kennenlernen. Mehr wird in der Algebra I geboten.

**1.1.** Eine **Gruppe** ist eine Menge $G$ versehen mit einer inneren Verknüpfung $G \times G \longrightarrow G$, $(a,b) \longmapsto a \cdot b$, die folgenden Axiomen genügt:

(G1) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ für alle $a,b,c \in G$. (Assoziativität)

(G2) $\exists e \in G$ so, dass $\forall a \in G \Longrightarrow a \cdot e = e \cdot a = a$. (neutrales Element)

(G3) $\forall a \in G \Longrightarrow \exists a^{-1} \in G$ mit $a \cdot a^{-1} = a^{-1} \cdot a = e$. (inverses Element)

Die Gruppe $G$ heißt **kommutativ** (oder **abelsch**), falls

(G4) $\forall a,b \in G \Longrightarrow a \cdot b = b \cdot a$

Bei einer abelschen Gruppe benutzt man oft $+$ statt $\cdot$ für die Verknüpfung. Das Neutralelement einer additiv geschriebenen abelschen Gruppe heißt Null und wird mit 0 bezeichnet. Die Inverse von $a$ wird mit $-a$ statt mit $a^{-1}$ notiert. Weiter benutzen wir $a - b := a + (-b)$.

**1.2.** Eine Gruppe hat die folgenden Eigenschaften

i) Das neutrale Element $e$ einer Gruppe ist eindeutig bestimmt.

ii) Das inverse Element zu $a \in G$ ist eindeutig bestimmt.

iii) $(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$ für alle $a,b \in G$.

iv) Für $a,b \in G$ hat die Gleichung $a \cdot x = b$ eine eindeutige Lösung in $G$. Die Gleichung $y \cdot a = b$ hat eine eindeutige Lösung in $G$. Es gilt $x = a^{-1} \cdot b$ und $y = b \cdot a^{-1}$.

**Beweis:** i) Angenommen, es gibt noch ein weiteres neutrales Element $e' \in G$ mit $e' \cdot a = a \cdot e' = a$ für alle $a \in G$.

$$e' \underset{\text{G2 für e}}{=} e \cdot e' \underset{\text{G2 für e'}}{=} e$$

ii) Angenommen $a'$ sei ein weiteres inverses Element zu $a \in G$

$$a^{-1} = a^{-1} \cdot e = a^{-1} \cdot (a \cdot a') = (a^{-1} \cdot a) \cdot a' = e \cdot a' = a'$$

iii) $(a \cdot b)^{-1} = (a \cdot b)^{-1}(a \cdot a^{-1}) = (a \cdot b)^{-1} \cdot a(b \cdot b^{-1})a^{-1} = (ab)^{-1}(ab)(b^{-1}a^{-1}) = e \cdot (b^{-1}a^{-1}) = b^{-1}a^{-1}$.

iv) Wird als Übung überlassen.

□

**1.3.** In der Mathematik definiert man **Homomorphismen** als strukturerhaltende Abbildungen. Hier in der Gruppentheorie geschieht das folgendermaßen:

Eine Abbildung $\varphi\colon G_1 \longrightarrow G_2$ zwischen zwei Gruppen $(G_1, \cdot)$ und $(G_2, *)$ heißt genau dann **Gruppenhomomorphismus**, wenn

$$\varphi(a \cdot b) = \varphi(a) * \varphi(b)$$

für alle $a, b \in G_1$. Der Einfachheit halber wollen wir ab jetzt ebenfalls $\cdot$ für die Verknüpfung auf $G_2$ schreiben (statt $*$).

**1.4 Definition.** Eine Teilmenge $H$ von $G$ heißt **Untergruppe** von $G$, wenn folgende Axiome erfüllt sind.

(U1) $a, b \in H \Longrightarrow a \cdot b \in H$ (abgeschlossen unter $\cdot$)

(U2) $e \in H$

(U3) $a \in H \Longrightarrow a^{-1} \in H$

Es ist leicht zu sehen, dass dann $(H, \cdot)$ eine Gruppe ist.

**1.5.** Für Gruppenhomomorphismen $\varphi\colon G_1 \longrightarrow G_2$ und $\psi\colon G_2 \longrightarrow G_3$ gelten die folgenden Eigenschaften:

i) $\varphi(e_1) = e_2$ ($e_i$ Neutralelement von $G_i$)

ii) $\varphi(a^{-1}) = \varphi(a)^{-1}$ für alle $a \in G$.

iii) $\psi \circ \varphi$ ist ein Gruppenhomomorphismus.

**Beweis:** i) Es gilt für $g_1 \in G_1$:

$$\varphi(g_1) = \varphi(e_1 \cdot g_1) = \varphi(e_1) \cdot \varphi(g_1)$$

Betrachte die Gleichung $ya = b$ mit $b := \varphi(g_1)$ und $a := \varphi(g_1)$. Offenbar hat diese Gleichung die Lösungen $y = \varphi(e_1)$ und $y = e_2$. Nach 1.2 iv) folgt $\varphi(e_1) = e_2$.

ii) Für $g_1 \in G_1$ gilt

$$e_2 = \varphi(e_1) = \varphi(g_1 \cdot g_1^{-1}) = \varphi(g_1) \cdot \varphi(g_1^{-1})$$

Betrachte die Gleichung $a \cdot x = b$ mit $a := \varphi(g_1)$ und $b := e_2$. Die hat die Lösungen $x = \varphi(g_1^{-1})$ und $x = \varphi(g_1)^{-1}$ nach (G3). Also folgt wieder mit 1.2 iv) $\varphi(g_1)^{-1} = \varphi(g_1^{-1})$.

iii) Für $a, b \in G_1$ gilt

$$(\psi \circ \varphi)(a \cdot b) = \psi\big(\varphi(a \cdot b)\big) = \psi\big(\varphi(a) \cdot \varphi(b)\big) = \psi\big(\varphi(a)\big) \cdot \psi\big(\varphi(b)\big) \qquad \square$$

**1.6.** Für einen Gruppenhomomorphismus $\varphi\colon G_1 \longrightarrow G_2$ definieren wir den **Kern** als

$$\ker(\varphi) := \varphi^{-1}(e_2) = \{a \in G_1 \mid \varphi(a) = e_2\}.$$

**1.7 Proposition.** *Sei $\varphi\colon G_1 \longrightarrow G_2$ ein Gruppenhomomorphismus.*

*a) $Ker(\varphi)$ ist eine Untergruppe von $G_1$*

*b) $\varphi(G_1)$ ist eine Untergruppe von $G_2$*

*c) $\varphi$ injektiv $\Longleftrightarrow Ker(\varphi) = \{e_1\}$.*

**Beweis:** Aufgabe 4.3 $\square$

**1.8 Beispiel.** a) $(\mathbb{N},+)$ ist keine Gruppe. Es gibt zwar ein neutrales Element 0, aber die Elemente $n>0$ haben kein Inverses!

b) $\mathbb{Z},\mathbb{Q},\mathbb{R}$ sind Gruppen bezüglich $+$

c) $\mathbb{Z},\mathbb{Q},\mathbb{R}$ sind keine Gruppen bezüglich $\cdot$. Es gibt zwar wieder ein neutrales Element 1 bezüglich $\cdot$, aber 2 hat kein multiplikativ Inverses in $\mathbb{Z}$. Also ist $\mathbb{Z}$ keine Gruppe bezüglich $\cdot$. Aber 0 hat weder in $\mathbb{Z}$, noch in $\mathbb{Q}$ oder $\mathbb{R}$ eine multiplikativ Inverse und damit sind auch $\mathbb{Q}$ und $\mathbb{R}$ keine Gruppen bezüglich $\cdot$.

**1.9.** Eine Menge $M$ mit einer inneren Verknüpfung $\cdot$ heißt genau dann **Monoid**, wenn sie assoziativ ist und ein neutrales Element besitzt. Zum Beispiel sind $\mathbb{Z},\mathbb{Q},\mathbb{R}$ und $\mathbb{C}$ bezüglich der Multiplikation Monoide. Im folgenden definieren wir nun die Menge der invertierbaren Elemente im Monoid $M$ durch

$$M^* := \{a\in M \mid \exists a^{-1}\in M, a\cdot a^{-1} = a^{-1}\cdot a = e\}$$

Es ist leicht zu sehen, dass $M^*$ bezüglich $\cdot$ eine Gruppe bildet. Zum Beispiel gilt

$$(\mathbb{Z},\cdot)^* = \{1,-1\}, (\mathbb{Q},\cdot)^* = \mathbb{Q}\setminus\{0\}, (\mathbb{R},\cdot)^* = \mathbb{R}\setminus\{0\}$$

**1.10 Beispiel.** Sei $X$ eine Menge und

$$M(X) := \{f\colon X\longrightarrow X \mid f \text{ ist Abbildung}\}$$

Dann bildet $M(X)$ ein Monoid bezüglich $\circ$, d.h. der Verknüpfung von Selbstabbildungen. Verwenden wir nun 1.9, dann bildet $M(X)^*$ eine Gruppe bezüglich $\circ$. Aus unseren Überlegungen in der Mengenlehre folgt leicht für $f\in M(X)$: $f$ invertierbar in $M(X) \iff f$ hat eine Umkehrabbildung $\iff f$ bijektiv. Somit ist $S(X) := \{f\colon X\longrightarrow X \mid f \text{ bijektiv}\}$ eine Gruppe bezüglich $\circ$. Sie heißt die **symmetrische Gruppe** auf $X$.

**1.11.** Wir betrachten eine Gruppe $G$ und eine Untergruppe $H$. Wir nehmen an, dass die Gruppe abelsch ist (d.h. kommutativ) und wir schreiben deshalb für die Verknüpfung $+$ statt $\cdot$. Für $g_1,g_2\in G$ definieren wir

$$g_1\equiv g_2 \pmod H \quad :\iff \quad g_1-g_2\in H.$$

Wir sagen, dass $g_1$ **kongruent** zu $g_2$ ist **modulo H**.

**1.12 Proposition.** *Die Kongruenz modulo $H$ ist eine Äquivalenzrelation.*

**Beweis:** Es gilt $g-g = g+(-g) = 0\in H$ und damit ist $g\equiv g \pmod H$ für alle $g\in G$. Somit ist $\equiv$ *reflexiv. Symmetrisch:* Es sei $g_1\equiv g_2 \pmod H$. Also gilt $g_1-g_2\in H$. Aus (U3) folgt $-(g_1-g_2)\in H$. Also gilt $g_2\equiv g_1 \pmod H$. *Transitiv:* Es seien $g_1\equiv g_2 \pmod H$ und $g_2\equiv g_3 \pmod H$. $\Longrightarrow g_1-g_2\in H\wedge g_2-g_3\in H$. $\Longrightarrow g_1-g_3 = (g_1-g_2)+(g_2-g_3)\in H$. Somit gilt $g_1\equiv g_3 \pmod H$ und $\equiv$ ist transitiv. $\square$

**1.13.** Wir bezeichnen den Raum der Äquivalenzklassen der Kongruenz modulo $H$ mit

$$G/H = \{[g]\mid g\in G\}.$$

Hier ist wie gewohnt $[g]$ die Äquivalenzklasse von $g$.

**1.14 Proposition.** *$G/H$ ist eine abelsche Gruppe bezüglich der Verknüpfung*

$$+ : (G/H)\times(G/H)\to G/H, \quad ([g_1],[g_2])\mapsto [g_1]+[g_2] := [g_1+g_2]$$

**Beweis:** Wir müssen zuerst zeigen, dass diese Verknüpfung wohldefiniert ist, d.h. wir müssen zeigen, dass die Definition von $[g_1]+[g_2] := [g_1+g_2]$ unabhängig ist von der Wahl der Repräsentanten $g_i$ aus der gegebenen Äquivalenzklasse $[g_i]$. Wir zeigen zuerst, dass die Definition unabhängig von der Wahl der Repräsentanten von $g_1$ ist. Sei also $g_1' \in [g_1]$, d.h. $g_1 \equiv g_1' \pmod{H}$.

$$\Longrightarrow g_1 - g_1' \in H \Longrightarrow (g_1+g_2)-(g_1'+g_2) \in H.$$

Daraus folgt $g_1+g_2 \equiv g_1'+g_2 \pmod{H}$, d.h. $[g_1+g_2]=[g_1'+g_2]$. Analog oder mit der Kommutativität zeigt man, dass die Definition auch unabhängig von der Wahl von $g_2$ ist. Die Gruppenaxiome vererben sich repräsentantenweise auf $G/H$, z.B. ist $[0]$ das Neutralelement, weil $[g]+[0]=[g+0]=[g]$. □

**1.15 Proposition.** *Die Quotientenabbildung*

$$\pi\colon G \longrightarrow G/H, \quad g \longmapsto [g]$$

*ist ein surjektiver Gruppenhomomorphismus mit* $\ker(\pi)=H$.

**Beweis:** Es seien $g_1, g_2 \in G$, dann gilt:

$$\pi(g_1+g_2)=[g_1+g_2]=[g_1]+[g_2]=\pi(g_1)+\pi(g_2)$$

Damit ist $\pi$ ein Gruppenhomomorphismus. Wir nehmen ein Element aus $G/H$, d.h. eine Äquivalenzklasse $[g]$ für ein $g \in G$. Dann gilt $\pi(g)=[g]$. Also ist $\pi$ surjektiv.

$$g \in \ker(\pi) \Longleftrightarrow \pi(g)=[0] \Longleftrightarrow [g]=[0] \Longleftrightarrow g \equiv 0 \pmod{H} \Longleftrightarrow g = g-0 \in H$$

Also gilt $H=\ker(\pi)$. □

**1.16 Beispiel.** $(\mathbb{Z},+)$ ist eine abelsche Gruppe. Für jedes $m \geq 2$ bilden die Vielfachen von $m$ eine Untergruppe

$$m\mathbb{Z} := \{m \cdot k \mid k \in \mathbb{Z}\}$$

von $\mathbb{Z}$. In diesem Fall hat die Äquivalenzrelation aus 1.11 die Form

$$n_1 \equiv n_2 \pmod{m} \Longleftrightarrow n_1 - n_2 \in m\mathbb{Z} \Longleftrightarrow m \mid n_1 - n_2$$

Wir sagen dann, dass $n_1$ **kongruent zu** $n_2$ ist **modulo** $m$. Wir nennen die Äquivalenzklassen in diesem Fall **Restklassen** (modulo $m$).

**1.17 Proposition.** $\mathbb{Z}/m\mathbb{Z}$ *ist eine abelsche Gruppe bestehend aus den m verschiedenen Restklassen* $[0],[1],\ldots,[m-1]$.

**Beweis:** Aus Proposition 1.14 wissen wir, dass $\mathbb{Z}/m\mathbb{Z}$ eine abelsche Gruppe bezüglich $+$ ist. Ein beliebiges Element aus $\mathbb{Z}/m\mathbb{Z}$ ist eine Restklasse $[n]$ mit $n \in \mathbb{Z}$. Mit der Division mit Rest (Aufgabe 2.4) folgt $n = qm+r$ mit $r \in \{0,\ldots,m-1\}$. $\Longrightarrow n-r=qm$. Also gilt $n \equiv r \pmod{m}$, d.h. $[n]=[r]$. Dies zeigt, dass $\mathbb{Z}/m\mathbb{Z} = \big\{[0],[1],\ldots,[m-1]\big\}$. Es bleibt zu zeigen, dass die Restklassen $[0],[1],\ldots,[m-1]$ paarweise verschieden sind. Es sei also $[n_1]=[n_2]$ mit $0 \leq n_1 \leq n_2 \leq m-1$. Es folgt, dass $n_1 \equiv n_2 \pmod{m} \Longrightarrow m \mid n_2-n_1$ und $0 \leq n_2-n_1 \leq m-1$. Es muss also $n_2-n_1=0$, d.h. $n_1=n_2$ gilt wie gewünscht. □

**1.18 Beispiel.** Wir berechnen die **Verknüpfungstafel** der Gruppe $\mathbb{Z}/4\mathbb{Z}$:

| $+$ | [0] | [1] | [2] | [3] |
|---|---|---|---|---|
| [0] | [0] | [1] | [2] | [3] |
| [1] | [1] | [2] | [3] | [0] |
| [2 | [2] | [3] | [0] | [1] |
| [3] | [3] | [0] | [1] | [2] |

**1.19 Beispiel.** Wann treffen sich der Minuten- und der Stundenzeiger? Wir beginnen um Mitternacht mit 0 zu zählen. Dann sei $x \in \mathbb{R}$ die Anzahl Stunden. Während der Stundenzeiger $\frac{x}{12}$ des Uhrkreises durchlaufen hat, überdeckte der Minutenzeiger schon den $x$-ten Teil des Uhrkreises. $x \in \mathbb{R}$ ist Treffpunkt, genau dann wenn $x \equiv \frac{x}{12} \pmod{\mathbb{Z}} \iff \frac{11}{12}x \equiv 0 \pmod{\mathbb{Z}} \iff x = \frac{12}{11}k$ mit $k \in \mathbb{Z}$. Also treffen sich die Zeiger nach $\frac{12}{11}$ Stunden, $\frac{24}{11}$ Stunden, $\ldots$, $\frac{120}{11}$ Stunden, $\frac{132}{11}$ Stunden=12 Stunden. Diese Rechnung fand in der Gruppe $\mathbb{R}/\mathbb{Z}$ statt.

## 2. Ringe und Körper

In der Schule lernt man schon früh den Ring $\mathbb{Z}$ kennen. Als weiteres Leitbeispiel werden wir im nächsten Abschnitt den Ring der Polynome kennenlernen. In diesem Abschnitt werden wir Ringe und die für die lineare Algebra grundlegenden Körper studieren. Körper sind kommutative Ringe, in denen man dividieren kann. Sie werden die Rolle der Zahlbereiche übernehmen.

**2.1.** Ein **Ring** ist eine Menge $R$ mit zwei inneren Verknüpfungen $+, \cdot$ so, dass $(R,+)$ eine abelsche Gruppe und $(R,\cdot)$ ein Monoid ist. Weiter verlangen wir, dass die Distributivgesetze gelten, d.h.

(D1) $a \cdot (b+c) = a \cdot b + a \cdot c$

(D2) $(b+c) \cdot a = b \cdot a + c \cdot a$

für alle $a,b,c \in R$. *Beachte*, dass wir bei einem Monoid insbesondere ein neutrales Element bzgl. $\cdot$ verlangen, das wir bei Ringen immer mit 1 bezeichnen werden.

**2.2.** Ein Ring heißt **kommutativ**, wenn er bezüglich der Multiplikation kommutativ ist.

**2.3.** Das Neutralelement bzgl. $+$ wird wie gewohnt mit 0 bezeichnet. Wir wollen das Inverse von $a \in R$ mit $-a$ bezeichnen (bzgl. der Addition). Weiter ist die Subtraktion $-$ definiert als

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

Für $a,b,c \in R$ gelten dann folgende Rechenregeln, die in Aufgabe 5.5 bewiesen werden:

i) $a \cdot 0 = 0 = 0 \cdot a$

ii) Das Einselement ist eindeutig festgelegt. Es gilt $1 = 0$ genau dann, wenn $R = \{0\}$.

iii) $-a = (-1) \cdot a$

iv) $a \cdot (b-c) = a \cdot b - a \cdot c \wedge (b-c) \cdot a = b \cdot a - c \cdot a$

**2.4 Definition.** Ein **Körper** ist ein kommutativer Ring $K$ so, dass $K \setminus \{0\}$ eine Gruppe ist. Insbesondere gilt $0 \neq 1$.

**2.5 Beispiel.** $\mathbb{Z}$ ist ein kommutativer Ring, aber kein Körper, weil $m \geq 2$ kein multiplikatives Inverses hat. Als Körper kommen uns $\mathbb{Q}$ und $\mathbb{R}$ in den Sinn.

**2.6 Proposition.** *Für $m \geq 2$ ist $\mathbb{Z}/m\mathbb{Z}$ ein kommutativer Ring bzgl. den Verknüpfungen*

$$[n_1] + [n_2] := [n_1 + n_2], \quad [n_1] \cdot [n_2] := [n_1 \cdot n_2]$$

**Beweis:** Es bleibt nur zu zeigen, dass die Multiplikation wohldefiniert ist, da wir $+$ schon früher betrachtet haben und die Rechenregeln sich vererben. Das lassen wir als Übung. $\square$

**2.7.** Für einen Ring $R$ bezeichnen wir mit

$$R^* := \{a \in R \mid \exists a^{-1} \in R, a \cdot a^{-1} = a^{-1} \cdot a = 1\}$$

die Menge der invertierbaren Elemente aus $R$. Da $(R,\cdot)$ ein Monoid ist, folgt aus 1.9, dass $(R^*,\cdot)$ eine Gruppe ist.

**2.8 Proposition.** *Sei $m \geq 2$ und $n \in \mathbb{N}$. Dann gilt:* $[n] \in (\mathbb{Z}/m\mathbb{Z})^* \Longleftrightarrow \mathrm{ggT}(m,n) = 1$.

**Beweis:** $[n] \in (\mathbb{Z}/m\mathbb{Z})^* \Longleftrightarrow \exists x \in \mathbb{Z}, [x] \cdot [n] = [1] \Longleftrightarrow \exists x, y \in \mathbb{Z}, x \cdot n - 1 = y \cdot m$

„$\Longleftarrow$“: Wir nehmen $\mathrm{ggT}(m,n) = 1$ an. Nach dem Lemma von Bezout hat die diophantische Gleichung $xn - ym = 1$ eine Lösung $(x,y) \in \mathbb{Z}^2$ $\big($da $\mathrm{ggT}(n,m) = 1\big)$. Nach dem Beginn des Beweises folgt $[n] \in (\mathbb{Z}/m\mathbb{Z})^*$.

„$\Longrightarrow$“: Sei $[n] \in (\mathbb{Z}/m\mathbb{Z})^*$. Wir bezeichnen den $\mathrm{ggT}(m,n)$ mit $d$. Nach dem Beginn des Beweises gilt $\exists x, y \in \mathbb{Z}$ mit $xn - ym = 1$. Nach Proposition I.2.10 folgt aus $d \mid n$ und $d \mid m$ auch $d \mid xn - ym = 1$. Also folgt $d = 1$. □

**2.9 Beispiel.** $(\mathbb{Z}/9\mathbb{Z})^* = \{[1],[2],[4],[5],[7],[8]\}$. Dies geschieht am Besten mit Proposition 2.8. Wir berechnen jetzt noch $[5]^{-1} = [2]$, denn $[5] \cdot [2] = [10] = [1]$.

**2.10 Korollar.** *Sei $m \geq 2$, dann ist $\mathbb{Z}/m\mathbb{Z}$ genau dann ein Körper, wenn $m$ eine Primzahl ist.*

**Beweis:** Nach 2.6 und 1.17 ist $\mathbb{Z}/m\mathbb{Z} = \big\{[0], \ldots, [m-1]\big\}$ ein kommutativer Ring. Es folgt:

$$\begin{aligned}\mathbb{Z}/m\mathbb{Z} \text{ Körper} &\Longleftrightarrow (\mathbb{Z}/m\mathbb{Z})^* = \big\{[1], \ldots, [m-1]\big\} \\ &\overset{2.2.8}{\Longleftrightarrow} \big[\forall n \in \{1, \ldots, m-1\} \Longrightarrow \mathrm{ggT}(n,m) = 1\big] \\ &\Longleftrightarrow m \text{ Primzahl}\end{aligned}$$

□

**2.11 Proposition.** *In einem Körper $K$ gelten die folgenden Eigenschaften:*

*i)* $a \cdot b = 0 \Longrightarrow a = 0 \vee b = 0$ *(nullteilerfrei)*

*ii)* $\big[a \cdot b = a \cdot c \wedge a \neq 0\big] \Longrightarrow b = c$ *(Kürzungsregel)*

**Beweis:**

i) Es sei $a \cdot b = 0$. Wenn $a = 0$ ist, dann sind wir fertig. Also dürfen wir oBdA $a \neq 0$.

$$\Longrightarrow 0 = a^{-1} \cdot 0 = a^{-1}(ab) = b$$

Dies zeigt i).

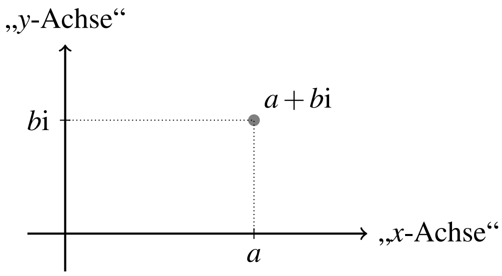
ii) Aus $ab = ac$ folgt $0 = ab - ac = a(b-c)$. Weil $a \neq 0$ folgt mit i) $b - c = 0$. □

**2.12 Beispiel.** In einem Ring müssen diese Eigenschaften nicht gelten. Zum Beispiel ist der kommutative Ring $\mathbb{Z}/m\mathbb{Z}$ genau dann nullteilerfrei, wenn $m$ eine Primzahl ist.

„$\Longleftarrow$“: Wenn $m$ eine Primzahl ist, dann ist $\mathbb{Z}/m\mathbb{Z}$ ein Körper nach Korollar 2.10 und somit nullteilerfrei nach Proposition 2.10.

„$\Longrightarrow$“: Wenn $m$ keine Primzahl ist, dann gilt $m = n_1 \cdot n_2$ mit $1 < n_1 < m$ und $1 < n_2 < m$. Also gilt $[n_1] \cdot [n_2] = [n_1 \cdot n_2] = [m] = [0] \in \mathbb{Z}/m\mathbb{Z}$. Also sind $[n_1]$ und $[n_2]$ Nullteiler und es gilt offenbar $[n_1] \neq [0], [n_2] \neq [0]$. Dies beweist die Negation von $\Longleftarrow$ und damit $\Longrightarrow$. □

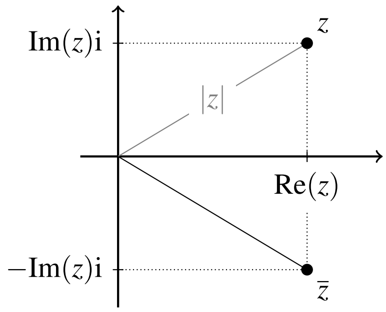
**2.13.** Wir wollen noch auf die Zahlbereichserweiterung $\mathbb{R} \subset \mathbb{C}$ eingehen. Als Menge gilt $\mathbb{C} = \mathbb{R}^2$, aber wir schreiben **komplexe Zahlen** als $a + b \cdot \mathrm{i}$ statt $(a,b) \in \mathbb{R}^2$. Dabei ist $i$ ein formales Symbol, das der Rechenregel $\mathrm{i}^2 = -1$ genügen soll. Die komplexen Zahlen veranschaulicht man sich in der **Gauß'schen Zahlenebene**:



Die reellen Zahlen sind dann auf der $x$-Achse, wenn man $a \in \mathbb{R}$ mit $a + 0 \cdot \mathrm{i}$ identifiziert. Die Zahlen auf der $y$-Achse, außer 0, nennt man rein **imaginär**. Für eine komplexe Zahl $z = a + b \cdot \mathrm{i}$ heißt $a$ **Realteil** von $z$. Er wird mit $\mathrm{Re}(z)$ bezeichnet. Der **Imaginärteil** $\mathrm{Im}(z)$ ist als $b$ definiert. Weiter ist

$$|z| := \sqrt{a^2 + b^2}$$

der **Betrag** von $z$. Wir nennen $\bar{z} := a - bi$ die **komplex-Konjugierte** von $z$.




Wir definieren die Addition von komplexen Zahlen $z = a + b\mathrm{i}$ und $w = c + d\mathrm{i}$ durch

$$z + w := (a + c) + (b + d) \cdot \mathrm{i}$$

analog zur Vektoraddition im $\mathbb{R}^2$. Die Multiplikation dieser Zahlen ist definiert durch

$$z \cdot w := (a \cdot c - b \cdot d) + (a \cdot d + b \cdot c) \cdot \mathrm{i}$$

Der Vorteil der Schreibweise $z = a + b \cdot \mathrm{i}$ wird hier deutlich, weil man sich die Multiplikation dadurch merken kann, dass sie für reelle Zahlen wie üblich funktioniert, dass die üblichen Rechenregeln gelten und dass $\mathrm{i}^2 = -1$ gilt. Zum Beispiel gilt:

$$(2 + 3\mathrm{i})(1 - \mathrm{i}) = 2 \cdot 1 - 2 \cdot \mathrm{i} + 3 \cdot \mathrm{i} - 3\mathrm{i}^2 = 2 - 2\mathrm{i} + 3\mathrm{i} + 3 = 5 + \mathrm{i}$$

**2.14 Satz.** $(\mathbb{C}, +, \cdot)$ *ist ein Körper.*

**Beweis:** Die Körperaxiome sind einfach zu prüfen. Wir prüfen hier nur, dass jede komplexe Zahl $z = a + b\mathrm{i} \neq 0$ eine multiplikative Inverse hat. Es gilt

$$z \cdot \bar{z} = (a + b\mathrm{i})(a - b\mathrm{i}) = a^2 - (b\mathrm{i})^2 = a^2 + b^2 = |z|^2 \tag{II.1}$$

Wegen $z \neq 0$ folgt $|z| > 0$ und mit (II.1) folgt, dass $|z|^{-2} \cdot \bar{z}$ die multiplikative Inverse zu $z$ ist. □

## 3. Polynome

Wir beleuchten Polynome in diesem Abschnitt vom Standpunkt der Algebra aus.

Es sei dabei $K$ immer ein Körper.

**3.1.** Wenn $(a_i)_{i \in I}$ eine endliche Familie von Elementen $a_i \in K$ ist, dann bedeutet $\sum_{i \in I} a_i$, dass wir die Familie aufsummieren. Durch Umnummerieren können wir $I = \{1, \ldots, n\}$ annehmen. Dann schreiben wir

$$\sum_{i \in I} a_i := \sum_{i=1}^{n} a_i = a_1 + \ldots + a_n$$

Wenn $I = \emptyset$, dann setzen wir $\sum\limits_{i \in I} a_i := 0$ Beachte, dass es oben nicht auf die Wahl der Nummerierung ankommt wegen der Assoziativität und der Kommutativität von $+$.

**3.2 Proposition.** *Es sei* $(a_i)_{i \in I}, (b_i)_{i \in I}$ *und* $(c_j)_{j \in J}$ *endliche Familien aus K und* $d \in K$. *Dann gilt* *a)* $\sum_{i\in I} da_i = d\sum_{i\in I} a_i$,

*b)* $\sum_{i\in I}(a_i+b_i) = \sum_{i\in I} a_i + \sum_{i\in I} b_i$,

*c)* $(\sum_{i\in I} a_i)\cdot(\sum_{j\in J} c_j) = \sum_{i\in I}(\sum_{j\in J} a_i c_j) = \sum_{(i,j)\in I\times J} a_i c_j = \sum_{j\in J}(\sum_{i\in I} a_i c_j)$,

*d) falls $I$ die disjunkte Vereinigung von $I'$ und $I''$ ist, dann gilt:* $\sum_{i\in I} a_i = \sum_{i\in I'} a_i + \sum_{i\in I''} a_i$.

**Beweis:** Aufgabe 6.1. □

**3.3 Definition.** Ein **Polynom** mit **Koeffizienten** $a_i \in K$ ist ein formaler Ausdruck

$$p(x) := a_m x^m + a_{m-1}x^{m-1} + \ldots + a_1 x + a_0$$

Dabei ist $x$ ein formales Symbol, das wir **Variable** nennen. Dabei sind zwei Polynome $p(x)$ und $q(x)$ genau dann gleich, wenn die Koeffizienten $\neq 0$ gleich sind, d.h. $a_i \neq 0 \Longleftrightarrow b_i \neq 0$ und $a_i = b_i$ für alle diese Koeffizienten. Der **höchste Koeffizient** ist das $a_j$ und mit maximalem $j$ mit $a_j \neq 0$. Dieses $j$ heißt dann **Grad** von $p(x)$. Wir bezeichnen den Grad mit $\deg g(x)$. Konvention: $\deg(0) = -\infty$. Die Menge aller Polynome wird mit $K[x]$ bezeichnet.

**3.4 Definition.** Wir definieren die **Addition** der Polynome

$$p(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0 \quad \text{und} \quad q(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_0$$

durch

$$p(x) + q(x) := (a_k + b_k)x^k + (a_{k-1} + b_{k-1})x^{k-1} + \cdots + (a_0 + b_0),$$

wobei $k := \max\{m,n\}$ und wir Koeffizienten $a_i, b_j$ mit $i > m$ bzw. $j > n$ als 0 setzen.

**3.5 Definition.** Wir definieren die **Multiplikation** der obigen Polynome durch

$$p(x)\cdot q(x) := c_{m+n}x^{m+n} + c_{m+n-1}x^{m+n-1} + \cdots + c_0$$

mit

$$c_k := a_k b_0 + a_{k-1}b_1 + \cdots + a_1 b_{k-1} + a_0 b_k = \sum_{i+j=k} a_i b_j.$$

Wieder werden alle nicht definierten $a_i, b_j$ als 0 gesetzt.

**3.6 Proposition.** *Mit diesen Verknüpfungen ist $K[x]$ ein kommutativer Ring.*

**Beweis:** Es ist leicht zu sehen, dass $K[x]$ eine abelsche Gruppe bzgl. $+$ bildet. Das Neutralelement bzgl. $+$ ist das Nullpolynom 0. Ebenfalls ist klar, dass $K[x]$ bzgl. der Multiplikation kommutativ ist und das Neutralelement $p(x) = 1$ hat. Wir werden in Aufgabe 6.2 sehen, dass die Multiplikation assoziativ ist. Also ist $(K[x], \cdot\,)$ ein kommutatives Monoid.

Es bleibt die Distributivität $p(x)\cdot\big(q(x) + r(x)\big) = p(x)\cdot q(x) + p(x)\cdot r(x)$ für Polynome $p(x), q(x), r(x) \in K[x]$ zu beweisen. Es sei

$$p(x) = a_m x^m + \cdots + a_0$$

und durch Aufstocken mit Nullkoeffizienten dürfen wir

$$q(x) = b_n x^n + \cdots + b_0, \qquad r(x) = c_n x^n + \cdots + c_0.$$

annehmen. Der $k$-te Koeffizient von $p(x)(q(x) + r(x))$ ist

$$\sum_{i+j=k} a_i(b_j + c_j)$$

Der $k$-te Koeffizient von $p(x)q(x)+p(x)r(x)$ ist

$$\sum_{i+j=k} a_i b_j + \sum_{i+j=k} a_i c_j = \sum_{i+j=k} a_i b_j + a_i c_j$$

Aufgrund der Distributivität von $K$ folgt $p(x)(q(x)+r(x)) = p(x)q(x)+p(x)r(x)$. $\square$

**3.7.** In ein Polynom $p(x) = a_m x^m + \cdots + a_0$ kann man Werte $\alpha \in K$ einsetzen und erhält

$$p(\alpha) = a_m \alpha^m + \cdots + a_0 \in K$$

Weiter heißt $\alpha \in K$ **Nullstelle** von $p(x) :\Longleftrightarrow p(\alpha) = 0$

**3.8 Proposition.** *Für $p(x), q(x) \in K[x]$ und $\alpha \in K$ gilt*

$$(p+q)(\alpha) = p(\alpha)+q(\alpha), \quad (p\cdot q)(\alpha) = p(\alpha)\cdot q(\alpha)$$

**Beweis:** Das erste Gesetz folgt sofort aus Proposition 3.2 b). Wir beweisen das zweite Gesetz:

$$\begin{aligned} p(\alpha)\cdot q(\alpha) &= \Big(\sum_{i=0}^{m} a_i \alpha^i\Big)\cdot\Big(\sum_{j=0}^{n} b_j \alpha^j\Big) \\ &\underset{3.2\ c)}{=} \sum_{(i,j)\in\{0,\ldots,m\}\times\{0,\ldots,n\}} a_i b_j \alpha^{i+j} \end{aligned}$$

Wir schreiben $\{0,\ldots,m\}\times\{0,\ldots,n\} = \bigcup_{k=0}^{m+n}\big\{(i,j)\in\{0,\ldots,m\}\times\{0,\ldots,n\} \mid i+j=k\big\}$ als Vereinigung paarweise disjunkter Teilmengen. Deshalb folgt aus 3.2 d).

$$\begin{aligned} p(\alpha)\cdot q(\alpha) &= \sum_{(i,j)\in\{0,\ldots,m\}\times\{0,\ldots,n\}} a_i b_j \alpha^{i+j} \\ &= \sum_{k=0}^{m+n}\sum_{i+j=k} a_i b_j \alpha^{i+j} = \sum_{k=0}^{m+n}\sum_{i+j=k} a_i b_j \alpha^k \\ &\underset{3.2\ a)}{=} \sum_{k=0}^{m+n}\left(\sum_{i+j=k} a_i b_j\right)\alpha^k = (p\cdot q)(\alpha) \end{aligned}$$ $\square$

Im folgenden soll die Rechenregel $-\infty + m = -\infty$ gelten für alle $m \in \mathbb{N}\cup\{-\infty\}$.

**3.9 Proposition.** *Für Polynome $p(x), q(x) \in K[x]$ gilt:*

*a)* $\deg\big(p(x)+q(x)\big) \le \max\big\{\deg\big(p(x)\big), \deg\big(q(x)\big)\big\}$. *Weiter gilt „=“, falls* $\deg\big(p(x)\big) \ne \deg\big(q(x)\big)$.

*b)* $\deg\big(p(x)\cdot q(x)\big) = \deg\big(p(x)\big) + \deg\big(q(x)\big)$.

**Beweis:** Wenn $p(x) = 0$ oder $q(x) = 0$, dann sind die Behauptungen trivial. Also dürfen wir annehmen, dass $p(x), q(x)$ nicht das Nullpolynom sind. Also haben wir

$$p(x) = a_m x^m + \cdots + a_0, \quad a_m \ne 0$$

und

$$q(x) = b_n x^n + \cdots + b_0, \quad b_n \ne 0$$

Wegen der Kommutativität dürfen wir $m \ge n$ annehmen. Falls $m > n$, dann gilt

$$p(x)+q(x) = a_m x^m + \text{ kleinere Grade.}$$

In diesem Fall gilt

$$\deg(p(x)+q(x)) = m = \deg(p(x)) = \max\{\deg(p(x)), \deg(q(x))\}$$

Falls $m = n$, dann gilt
$$p(x) + q(x) = (a_m + b_m)x^m + \text{ kleinere Grade.}$$
Es folgt
$$\deg(p(x) + q(x)) \leq m = \max\{\deg(p(x)), \deg(q(x))\}$$
Um b) zu zeigen, bemerkt man
$$p(x) \cdot q(x) = \sum_{k=0}^{m+n} c_k x^k$$
mit
$$c_k = \sum_{i+j=k} a_i b_j$$
Insbesondere gilt
$$c_{m+n} = a_m b_n \neq 0$$
nach Proposition 2.11 (Körper ist nullteilerfrei). Es folgt $\deg(p(x) \cdot q(x)) = m + n$ wie gewünscht. Dies zeigt b). $\square$

**3.10 Korollar.** *Im Ring $K[x]$ gilt die Kürzungsregel*
$$p(x)q(x) = p(x)r(x) \ \wedge \ p(x) \neq 0 \Longrightarrow q(x) = r(x)$$
*und er ist nullteilerfrei:*
$$p(x)q(x) = 0 \Longrightarrow p(x) = 0 \ \vee \ q(x) = 0.$$

**Beweis:** Wir zeigen zuerst die Nullteilerfreiheit. Sei $p(x)q(x) = 0$ und $p(x) \neq 0$, $q(x) \neq 0$. Ziel: Widerspruch! (Indirekter Beweis). Wegen
$$\deg\big(p(x)q(x)\big) = \deg\big(p(x)\big) + \deg\big(q(x)\big) \neq -\infty$$
und somit $p(x)q(x) \neq 0 \Longrightarrow$ Widerspruch! Aus der Nullteilerfreiheit folgt immer die Kürzungsregel:
$$p(x)q(x) = p(x)r(x) \ \wedge \ p(x) \neq 0$$
$$\Longrightarrow p(x)q(x) - p(x)r(x) = p(x)\big(q(x) - r(x)\big) = 0$$
Weil $K[x]$ wie oben gesehen nullteilerfrei ist und $p(x) \neq 0$, muss $q(x) - r(x) = 0$ gelten, d.h. $q(x) = r(x)$. $\square$

**3.11 Theorem (Polynomdivision).** *Für $a(x), b(x) \in K[x]$ mit $b(x) \neq 0$ gibt es genau ein $q(x), r(x) \in K[x]$ mit*
$$a(x) = q(x)b(x) + r(x) \quad \wedge \quad \deg\big(r(x)\big) < \deg\big(b(x)\big)$$

**Beweis:** Vollständiger Induktion nach $d = \deg(a(x))$ analog zu Theorem 2.6. Der Beweis wird in Aufgabe 6.3 gemacht. $\square$

**3.12 Korollar (Abspalten einer Nullstelle).** *Sei $\alpha \in K$ eine Nullstelle von $p(x) \in K[x]$. Dann $\exists!\ q(x) \in K[x]$ mit*
$$p(x) = (x - \alpha)q(x)$$

**Beweis:** Durch Polynomdivision (Theorem 3.11) $\exists q(x), r(x) \in K[x]$ mit
$$p(x) = q(x)(x - \alpha) + r(x) \quad \wedge \quad \deg(r(x)) < 1$$
Also gilt $r(x) = c \in K$. Durch Einsetzen von $\alpha$ folgt
$$0 = p(\alpha) = q(\alpha)(\alpha - \alpha) + c = 0$$
Also gilt $c = 0$ und es folgt die Existenz. Die Eindeutigkeit folgt entweder auch aus der Polynomdivision (Theorem 3.11) oder aus der Kürzungsregel 3.10. $\square$

**3.13 Warnung.** Wir haben die Polynome als formale Summen von Termen der Form $a_n x^n$ definiert. In der Schule definiert man Polynome oft als die entsprechenden Funktionen. Hier soll darauf hingewiesen werden, dass Polynome nicht mit den entsprechenden Polynomfunktionen

$$K \to K, \quad \alpha \mapsto a_m \alpha^m + \cdots + a_0$$

identifiziert werden dürfen. Wir betrachten folgendes Beispiel im Fall $K := \mathbb{Z}/2\mathbb{Z} = \{\overline{0}, \overline{1}\}$. Die Polynome

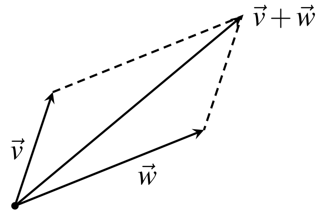
$$p(x) := \overline{1}x^2 = x^2 \in K[x] \quad \text{und} \quad q(x) := \overline{1}x = x \in K[x]$$

sind verschieden, aber die Funktionen $K \to K$, $\alpha \mapsto p(\alpha)$, $\alpha \mapsto q(\alpha)$ sind gleich.

# 3 Kapitel III.
# Vektorräume

## 1. Grundbegriffe

In der Schule haben wir in der Geometrie die 3-dimensionale Vektorrechnung gelernt. Wesentlich ist, dass man Vektoren mit reellen Zahlen strecken kann und mit Hilfe der Parallelogrammkonstruktion addieren kann.




In der analytischen Geometrie lernt man, dass diese Vektorrechnung durch $\mathbb{R}^3$ und den beiden Operationen

$$\lambda\cdot\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}=\begin{pmatrix}\lambda v_1\\\lambda v_2\\\lambda v_3\end{pmatrix},\qquad\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}+\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}=\begin{pmatrix}v_1+w_1\\v_2+w_2\\v_3+w_3\end{pmatrix}$$

beschrieben werden kann. Wir nennen $\lambda\cdot v$ die skalare Multiplikation des Vektors

$$\vec{v}=\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}$$

mit dem Skalar $\lambda\in\mathbb{R}$. In diesem Abschnitt bezeichnen wir mit $K$ immer einen Körper. Analog kann man die Addition auf $K^n$ erklären und ebenso eine skalare Multiplikation von Elementen aus $K^n$ mit einem Skalar $\lambda\in K$. Wir werden zuerst weitere Beispiele mit Addition und skalarer Multiplikation kennen lernen und dann solche Situationen axiomatisch als Vektorräume definieren. Die Vektorräume bilden die Grundbausteine für die lineare Algebra. Man darf sie nicht mit Ringen verwechseln, bei denen die Multiplikation eine innere Verknüpfung ist.

**1.1 Beispiel.** Wir betrachten ein homogenes lineares Gleichungssystem.

$$\begin{aligned}a_{11}x_1+\cdots+a_{1n}x_n&=0\\ \cdots\cdots\cdots&\cdots\cdots\\ a_{m1}x_1+\cdots+a_{mn}x_n&=0\end{aligned}\tag{III.1}$$

Die in §I.4 bewiesenen Aussagen für lineare Gleichungssysteme stimmen nicht nur für den Körper $\mathbb{R}$, sondern auch für beliebige Körper. Wir nehmen also $a_{ij}\in K$ an und suchen Lösungen in $K^n$. Nach Theorem I.4.3 ist der Lösungsraum von (III.1) eine Teilmenge des $K^n$, der abgeschlossen ist unter $+$ und skalarer Multiplikation.

**1.2 Beispiel.** Es sei $V_n:=\left\{p(x)\in K[x]\mid\deg\left(p(x)\right)\le n\right\}$. Nach Proposition II.3.9 ist $V_n$ abgeschlossen unter $+$ und skalarer Multiplikation mit Skalaren aus $K$.

**1.3 Beispiel.** Es sei $X$ eine beliebige Menge und $V:=\{f\colon X\to K\mid f\text{ Funktion}\}$. Dann kann man leicht zeigen, dass $V$ ein Ring ist bezüglich den Verknüpfungen

$$(f+g)\colon X\to K,\ x\mapsto f(x)+g(x)\qquad(f\cdot g)\colon X\to K,\ x\mapsto f(x)\cdot g(x)$$

für $f,g \in V$. Man braucht dazu nur die Ringaxiome von $K$. Wir definieren eine skalare Multiplikation von $f \in V$ mit $\lambda \in K$ durch

$$\lambda \cdot f \colon X \to K, \quad x \mapsto \lambda \cdot f(x)$$

Wir axiomatisieren diese Beispiele durch:

**1.4 Definition.** Ein **K-Vektorraum** ist eine Menge $V$ mit einer **Addition** $+\colon V \times V \to V$ und einer **skalaren Multiplikation** $K \times V \to V$, $(\lambda, v) \mapsto \lambda \cdot v$, die folgenden Axiomen genügt für alle $\lambda, \mu \in K$ und $v, w \in V$:

A) $(V,+)$ ist eine abelsche Gruppe

B) Es gelten die Distributivgesetze: $(\lambda+\mu)\cdot v = \lambda \cdot v + \mu \cdot v \wedge \lambda \cdot (v+w) = \lambda \cdot v + \lambda \cdot w$

C) $\lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v$

D) $1 \cdot v = v$

**1.5.** Wir nennen die Elemente eines Vektorraumes **Vektoren**. Wir verzichten aber auf die Schreibweise $\vec{v}$, die wir der Schule überlassen und notieren Vektoren einfach mit $v$ oder $w$. Es ist leicht zu sehen, dass die zuvor betrachteten Beispiele Vektorräume sind.

**1.6 Beispiel.** Das trivialste Beispiel eines $K$-Vektorraumes ist der **Nullraum** $\{0\}$.

**1.7 Beispiel.** Man kann $K$ selber als Vektorraum auffassen, in dem man die Multiplikation als skalare Multiplikation auffasst.

**1.8 Proposition.** *Für $\lambda \in K$ und $v$ aus dem $K$-Vektorraum $V$ gilt:*

*a)* $\lambda \cdot 0_V = 0_V$

*b)* $0_K \cdot v = 0_V$

*c)* $(-\lambda)\cdot v = \lambda \cdot (-v) = -(\lambda \cdot v)$

*d)* $\lambda \cdot v = 0_V \Longrightarrow \lambda = 0_K \vee v = 0_V$

**Beweis:** Aufgabe 7.1. □

**1.9 Definition.** Eine **lineare Abbildung** $\varphi\colon V \to W$ zwischen den $K$-Vektorräumen $V$ und $W$ ist eine Abbildung so, dass $\varphi(v+w) = \varphi(v) + \varphi(w)$, $\lambda \cdot \varphi(v) = \varphi(\lambda \cdot v)$ für alle $v, w \in V$ und $\lambda \in K$.

**1.10 Beispiel.** Analog zu Beispiel 1.2 ist $K[x]$ ein $K$-Vektorraum. Wir wählen $\alpha \in K$ fest. Dann ist

$$\varphi_\alpha \colon K[x] \to K, \quad p(x) \mapsto p(\alpha)$$

nach Proposition II.3.8 eine lineare Abbildung.

**1.11 Beispiel.** Sei $[a,b]$ ein abgeschlossenes Intervall in $\mathbb{R}$. Dann ist der Raum

$$C\big([a,b]\big) := \{f\colon [a,b] \to \mathbb{R} \mid f \text{ stetige Funktion}\}$$

ein reeller Vektorraum analog zu Beispiel 1.3. Aus der Analysis wissen wir, dass die folgende Abbildung

$$\varphi\colon C\big([a,b]\big) \to \mathbb{R}, \quad f \mapsto \int_a^b f(t)\mathrm{d}t$$

eine lineare Abbildung ist, d.h. es gilt $\forall f,g \in C\big([a,b]\big)$, $\forall \lambda \in \mathbb{R}$

$$\int_a^b (f+g)(t)\mathrm{d}t = \int_a^b f(t)\mathrm{d}t + \int_a^b g(t)\mathrm{d}t, \quad \lambda \cdot \int_a^b f(t)\mathrm{d}t = \int_a^b \big(\lambda \cdot f(t)\big)\mathrm{d}t.$$

**1.12 Beispiel.** Sei $C^1\big(]a,b[\big)$ der Raum der einmal stetig differenzierbaren reellen Funktionen auf dem offenen Intervall $]a,b[$. Wie $C\big([a,b]\big)$ bildet auch $C^1\big(]a,b[\big)$ einen reeller Vektorraum mit den in Beispiel 1.3 definierten Operationen.

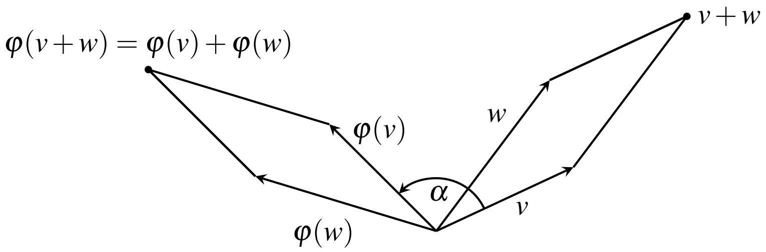
Aus der Analysis wissen wir, dass

$$\psi\colon C^1\big(]a,b[\big) \to C\big(]a,b[\big), f \mapsto f'$$

eine lineare Abbildung ist, d.h. es gilt $\forall f,g \in C\big([a,b]\big)$, $\forall \lambda \in \mathbb{R}$

$$(f+g)' = f'+g' \quad \wedge \quad (\lambda f)' = \lambda \cdot f'$$

**1.13 Beispiel.** Die **Drehung** der Ebene um einen gegebenen Winkel $\alpha$ induziert eine lineare Abbildung $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$.




Begründung: Das Parallelogramm dreht sich mit. Streckung dreht sich auch mit, d.h. $\varphi(\lambda v) = \lambda\varphi(v)$.

**1.14 Beispiel.** Die Abbildung $\varphi\colon \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ ist nicht linear. $\varphi\big((-1)\cdot x\big) \neq (-1)\cdot\varphi(x)$ z.B. für $x = 1$

**1.15 Proposition.** *Es sei $\varphi\colon V \to W$ eine lineare Abbildung von $K$-Vektorräumen. Dann gilt:*

*a)* $\varphi(0) = 0 \quad \wedge \quad \varphi(-v) = -\varphi(v)$

*b) Wenn $\psi\colon W \to U$ eine weitere $K$-lineare Abbildung ist, dann ist auch $\psi\circ\varphi\colon V \to U$ eine $K$-lineare Abbildung.*

**Beweis:** Beachte, dass jede lineare Abbildung insbesondere ein Gruppenhomomorphismus ist bzgl. $+$. Die entsprechenden Eigenschaften für Gruppenhomomorphismen wurden schon in Proposition II.1.5 bewiesen. Also folgt a) und

$$(\psi\circ\varphi)(v_1+v_2) = (\psi\circ\varphi)(v_1)+(\psi\circ\varphi)(v_2).$$

Für $\lambda \in K$ und $v \in V$ gilt

$$(\psi\circ\varphi)(\lambda\cdot v) = \psi\big(\varphi(\lambda\cdot v)\big) = \psi\big(\lambda\cdot\varphi(v)\big) = \lambda\cdot\psi\big(\varphi(v)\big) = \lambda(\psi\circ\varphi)(v)$$

Also folgt auch b). □

**1.16 Definition.** Eine $K$-lineare Abbildung $\varphi\colon V \to W$ heißt **Isomorphismus**, wenn es eine $K$-lineare Abbildung $\psi\colon W \to V$ gibt mit:

$$\psi\circ\varphi = \mathrm{id}_V \quad \wedge \quad \varphi\circ\psi = \mathrm{id}_W \tag{III.2}$$

**1.17 Proposition.** *Für eine $K$-lineare Abbildung $\varphi\colon V \to W$ gilt: $\varphi$ Isomorphismus $\Longleftrightarrow$ $\varphi$ bijektiv.*

**Beweis:** „$\Longrightarrow$“ Es sei $\varphi$ Isomorphismus. Dann ist die lineare Abbildung $\psi\colon W \to V$ aus der Definition 1.16 offenbar die Umkehrfunktion von $\varphi$ und damit ist $\varphi$ bijektiv (siehe I.3.18).

„$\Longleftarrow$“ Es sei also $\varphi$ eine bijektive lineare Abbildung. Nach I.3.18 gibt es eine Umkehrabbildung $\psi\colon W \to V$, d.h. (III.2) gilt. Wir müssen zeigen, dass $\psi$ linear ist. Seien $w_1, w_2 \in W$, $\lambda \in K$. Weil $\varphi$ surjektiv ist, $\exists v_i \in V$ mit $\varphi(v_i) = w_i$. Es folgt

$$v_i = (\psi\circ\varphi)(v_i) = \psi(w_i)$$

Weiter gilt

$$\psi(\lambda w_1) = \psi\big(\lambda\varphi(v_1)\big) = \psi\Big(\varphi(\lambda(v_1))\Big) = \lambda v_1 = \lambda\psi(w_1)$$

wie gewünscht. Weiter haben wir

$$\psi(w_1 + w_2) = \psi\big(\varphi(v_1) + \varphi(v_2)\big) = \psi\big(\varphi(v_1 + v_2)\big) = v_1 + v_2 = \psi(w_1) + \psi(w_2)$$

Also ist $\psi$ linear. Es folgt, dass $\varphi$ ein Isomorphismus ist. □

**1.18.** Nach Proposition 1.17 könnte man einen Isomorphismus direkt als bijektive lineare Abbildung definieren. Der Vorteil der Definition 1.16 ist, dass sie sich problemlos auf andere mathematische Strukturen übertragen lässt.

**1.19.** Alle bisher betrachteten linearen Abbildungen waren **keine** Isomorphismen (mit Ausnahme von 1.13). Um doch noch ein Beispiel zu geben, sei $W := \{p(x) \in K[x] \mid \deg\big(p(x)\big) \leq n\}$ der Vektorraum aus Beispiel 1.2 und $V := K^{n+1}$. Die lineare Abbildung

$$\varphi\colon V \to W, \quad (a_1, \dots, a_{n+1}) \mapsto a_1x^n + a_2x^{n-1} + \dots + a_nx + a_{n+1},$$

ist offenbar bijektiv und damit nach Proposition 1.17 ein Isomorphismus.

## 2. Kartesische Produkte, Unterräume und Summen

In diesem Abschnitt sehen wir einige Grundkonstruktionen, die es erlauben, aus gegebenen Vektorräumen neue Vektorräume zu konstruieren. Weiter gehen wir näher auf die Begründung ein, wieso die Beispiele aus III.1 Vektorräume sind.

Wie immer bezeichnet $K$ einen Körper.

**2.1.** Sei $(V_i)_{i\in I}$ eine Familie von $K$-Vektorräumen $V_i$. Dann wird das kartesische Produkt $\prod_{i\in I} V_i$ zu einem $K$-Vektorraum durch

$$(v_i)_{i\in I} + (w_i)_{i\in I} := (v_i + w_i)_{i\in I}$$

und

$$\lambda \cdot (v_i)_{i\in I} := (\lambda \cdot v_i)_{i\in I}$$

Die Vektorraumaxiome sind einfach zu prüfen, z.B. gilt:

$$\begin{aligned}\lambda \cdot ((v_i)_{i\in I} + (w_i)_{i\in I}) &= \lambda \cdot (v_i + w_i)_{i\in I} = \big(\lambda \cdot (v_i + w_i)\big)_{i\in I} \\ &= (\lambda \cdot v_i + \lambda \cdot w_i)_{i\in I} = (\lambda \cdot v_i)_{i\in I} + (\lambda \cdot w_i)_{i\in I} \\ &= \lambda \cdot (v_i)_{i\in I} + \lambda \cdot (w_i)_{i\in I}.\end{aligned}$$

Dies zeigt eine Form der Distributivität. Die anderen Axiome werden in Aufgabe 8.1 gezeigt.

**2.2 Beispiel.** Als Spezialfall betrachten wir $I = \{1, \dots, n\}$ und $V_1 = V_2 = \dots = V_n = K$, dann erhalten wir den bekannten Vektorraum $K^n$.

**2.3 Definition.** Eine Teilmenge $U$ des $K$-Vektorraumes $V$ heißt **Unterraum** von $V$ genau dann, wenn folgende 3 Axiome erfüllt sind:

U1) $u_1, u_2 \in U \Longrightarrow u_1 + u_2 \in U$

U2) $[\lambda \in K \,\wedge\, u \in U] \Longrightarrow \lambda u \in U$

U3) $0 \in U$

Das letzte Axiom brauchen wir nur, um die leere Menge als Unterraum auszuschließen.

**2.4 Proposition.** *Für eine Teilmenge $U$ des $K$-Vektorraumes $V$ sind folgende Aussagen äquivalent:*

*a) $U$ ist ein Unterraum.*

*b) Die Addition und die skalare Multiplikation von $V$ machen aus $U$ einen $K$-Vektorraum.*

**Beweis:** a) $\Longrightarrow$ b) Sei $U$ ein Unterraum von $V$. Nach U1) ist $+$ eine innere Verknüpfung auf $U$ und nach U3) gilt $0 \in U$. Weiter folgt aus U2):

$$u \in U \Longrightarrow -u = (-1) \cdot u \underset{U2)}{\in} U$$

Aus II.1.5 folgt, dass $(U,+)$ eine Untergruppe von $(V,+)$ ist. Also ist $(U,+)$ eine abelsche Gruppe. Weiter folgt aus U2), dass $U$ unter skalarer Multiplikation abgeschlossen ist. Die anderen Vektorraumaxiome für $U$ vererben sich von $V$. Das zeigt b). b) $\Longrightarrow$ a) trivial. $\square$

**2.5 Beispiel.** Der Lösungsraum des homogenen Gleichungssystems in Beispiel 1.1 ist ein Unterraum von $K^n$ (siehe Beispiel 1.1). Also folgt aus Proposition 2.4, dass der Lösungsraum ein $K$-Vektorraum ist.

**2.6 Beispiel.** Sei $V$ der reelle Vektorraum der reellen Funktionen auf dem offenen Intervall $]a,b[\subseteq \mathbb{R}$. Dann ist $C\big(]a,b[\big)$ ein Unterraum von $V$. Weiter ist $C^1\big(]a,b[\big)$ ein Unterraum von $C\big(]a,b[\big)$ und damit auch von $V$.

**2.7 Proposition.** *Sei $\varphi\colon V \to W$ eine lineare Abbildung von $K$-Vektorräumen. Dann ist $\ker(\varphi)$ ein Unterraum von $V$ und $\varphi(V)$ ist ein Unterraum von $W$.*

**Beweis:** Nach Proposition II.1.7 sind der Kern und das Bild Untergruppen, weil $\varphi$ bzgl. $+$ ein Gruppenhomomorphismus ist. Es bleibt U2) nachzurechnen. Sei $v \in \ker(\varphi), \lambda \in K$, Dann gilt:

$$\varphi(\lambda v) = \lambda \varphi(v) = \lambda 0 = 0$$

Analog für das Bild. $\square$

**2.8 Proposition.** *Sei $(U_i)_{i\in I}$ eine Familie von Unterräumen des $K$-Vektorraumes $V$. Dann ist auch $\bigcap_{i\in I} U_i$ ein Unterraum von $V$.*

**Beweis:** Sei $U := \bigcap_{i\in I} U_i$. Weil $0 \in U_i$ gilt $\forall i \in I$ nach U3) für $U_i$, folgt auch $0 \in \bigcap_{i\in I} U_i$. Also gilt U3) auch für $U$. Sei $u_1, u_2 \in U$, dann gilt $u_1, u_2 \in U_i\ \forall i \in I$. Nach U1) für $U_i$ gilt $u_1 + u_2 \in U_i\ \forall i \in I$. Dann gilt $u_1 + u_2 \in \bigcap_{i\in I} U_i = U$, also folgt U1) für $U$. Sei $u \in U$, $\lambda \in K \Longrightarrow [\lambda u \in U_i \forall i \in I]$. damit auch $\lambda u \in \bigcap_{i\in I} U_i = U$, d.h. U2) für $U$. $\square$

**2.9 Proposition.** *Es seine $U_1, U_2$ Unterräume des $K$-Vektorraumes $V$. Dann ist*

$$\{u_1 + u_2 \mid u_1 \in U_1,\, u_2 \in U_2\}$$

*ein Unterraum von $V$.*

**Beweis:** Wir müssen die Unterraumaxiome prüfen für $U := \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$. Seien $u, u' \in U$. Dann gilt $u = u_1 + u_2$, $u' = u_1' + u_2'$ mit $u_1, u_1' \in U_1$, $u_2, u_2' \in U_2$. Wegen

$$u + u' = (u_1 + u_2) + (u_1' + u_2') = (u_1 + u_1') + (u_2 + u_2')$$

folgt $u + u' \in U$. Also gilt U1). Sei $u \in U$, $\lambda \in K$. Dann gilt $u = u_1 + u_2$ wie oben und

$$\lambda u = \lambda(u_1 + u_2) = \lambda u_1 + \lambda u_2$$

Also folgt $\lambda u \in U$ und damit gilt U2). Da $0 = 0 + 0$ ist, gilt auch U3). $\square$

**2.10 Definition.** Der Unterraum $\{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$ aus Proposition 2.9 heißt die **Summe der Unterräume $\mathbf{U_1}$ und $\mathbf{U_2}$**. Es wird mit $U_1 + U_2$ bezeichnet.

**2.11 Proposition.** *Es seien $U_1, U_2$ Unterräume des $K$-Vektorraumes $V$ und $U := U_1 + U_2$. Dann sind folgende Bedingungen äquivalent.*

*a)* $U_1 \cap U_2 = \{0\}$.

*b)* $\forall u \in U \Longrightarrow \exists!(u_1, u_2) \in U_1 \times U_2$ *mit* $u = u_1 + u_2$

**Beweis:** a) $\Longrightarrow$ b): Sei $u \in U$. Aus der Annahme $U = U_1 + U_2$ folgt die Existenz von $u_1 \in U_1$, $u_2 \in U_2$ mit $u = u_1 + u_2$. Es bleibt zu zeigen, dass so eine Darstellung eindeutig ist. Sei $u = u_1' + u_2'$ eine weiter Darstellung.

$$\Longrightarrow u_1 + u_2 = u_1' + u_2' \Longrightarrow u_1 - u_1' = u_2' - u_2$$

Die linke Seite ist in $U_1$ und die rechte Seite in $U_2$. Also gilt

$$u_1 - u_1' = u_2' - u_2 \in U_1 \cap U_2 = \{0\}$$

Also folgt $u_1 = u_1'$ und $u_2 = u_2'$ und damit die Eindeutigkeit. Dies zeigt b).

b) $\Longrightarrow$ a): wir müssen $U_1 \cap U_2 = \{0\}$ zeigen. Sei $u \in U_1 \cap U_2$. Dann gilt

$$u = u + 0 = 0 + u$$

Dies sind zwei Darstellungen der Form $u = u_1 + u_2$ mit $u_i \in U_i$. Aufgrund der Eindeutigkeit in b) folgt $u = 0$. Also gilt $U_1 \cap U_2 = \{0\}$. Dies zeigt a). □

**2.12 Definition.** Wenn eine der beiden äquivalenten Bedingungen in Proposition 2.11 erfüllt ist, dann heißt $U$ die **direkte Summe** der Unterräume $U_1$ und $U_2$. Wir schreiben dann $U = U_1 \oplus U_2$.

**2.13 Beispiel.** Es seien $0 \leq k < n$ natürliche Zahlen und $V := \{p(x) \in K[x] \mid \deg(p) \leq n\}$. Wir haben in Beispiel 1.2 gesehen, dass $V$ ein $K$-Vektorraum ist. Dann sind

$$U_1 := \Big\{ \sum_{j=0}^{k} a_j x^j \mid a_j \in K \Big\}$$

und

$$U_2 := \Big\{ \sum_{j=k+1}^{n} a_j x^j \mid a_j \in K \Big\}$$

und

$$U_3 := \Big\{ \sum_{j=k}^{n} a_j x^j \mid a_j \in K \Big\}$$

Unterräume von $V$, wie man leicht nachrechnen kann. Offenbar gilt

$$U_1 + U_2 = U_1 + U_3 = V$$

Wegen $U_1 \cap U_2 = \{0\}$, folgt mit Proposition 2.11 $V = U_1 \oplus U_2$. Wegen $U_1 \cap U_3 = \{\lambda x^k \mid \lambda \in K\} \neq \{0\}$, folgt mit Proposition 2.11, dass $U_1 + U_3 = V$ **keine** direkte Summe ist.

**2.14.** Es seien $U_1, \ldots, U_r$ Unterräume des $K$-Vektorraumes $V$. Wenn jeder Vektor $v \in V$ genau eine Darstellung $v = u_1 + \cdots + u_r$ mit $u_i \in U_i$ hat, dann heißt $V$ die **direkte Summe** der Unterräume $U_1, \ldots, U_r$. Wir schreiben dann

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_r$$

Für $r > 2$ ist das hier nicht mehr ganz so einfach wie in Proposition 2.11. Dies werden wir in Aufgabe 8.2 untersuchen. In Beispiel 2.13 gilt, dass $V$ die direkte Summe der $n+1$ Unterräume $K \cdot x^k$ ist, wobei $k = 0, \ldots, n$. Es gilt also

$$V = \bigoplus_{i=0}^{n} K \cdot x^k.$$

## 3. Quotientenräume und Dualraum

In diesen Abschnitt werden wir sehen, dass man auch Kongruenzen in einem Vektorraum modulo einem Unterraum betrachten kann. Diese Konstruktion ist genau so wichtig wie $\mathbb{Z}/m\mathbb{Z}$. Weiter werden wir sehen, dass die Menge aller linearen Abbildungen zwischen zwei Vektorräumen wieder ein Vektorraum bildet. Im Spezialfall des Wertebereichs gleich $K$ sprechen wir von Dualraum, dessen Konstruktion in vielen Bereichen der Mathematik wichtig ist.

In diesem Abschnitt bezeichnet $K$ ein Körper, $V, V_1, V_2$ sind $K$-Vektorräume und $W$ ist ein Unterraum von $V$.

**3.1.** Wir können $W$ als Untergruppe der abelschen Gruppe $(V,+)$ betrachten und erhalten die Kongruenz modulo $W$ wie in II.1.11, d.h. für $v_1, v_2 \in V$ gilt

$$v_1 \equiv v_2 \pmod{W} :\Longleftrightarrow v_1 - v_2 \in W.$$

Die Äquivalenzklassen bezeichnen wir mit $[v]$. Wir nennen die Menge der Äquivalenzklasse $\{[v] \mid v \in V\}$ den **Quotientenraum** $V/W$.

**3.2 Proposition.** *Der Quotientenraum $V/W$ ist ein $K$-Vektorraum bzgl. den Operationen*

$$[v_1] + [v_2] := [v_1 + v_2], \quad \lambda[v] := [\lambda v].$$

**Beweis:** Nach Proposition II.1.14 ist $+$ wohldefiniert auf $V/W$ und macht aus $V/W$ eine abelsche Gruppe. Für $\lambda \in K$ und $v_1 \equiv v_2 \pmod{W}$ gilt

$$v_1 \equiv v_2 \pmod{W} \Longrightarrow v_1 - v_2 \in W \Longrightarrow \lambda v_1 - \lambda v_2 = \lambda(v_1 - v_2) \in W$$

und damit folgt $\lambda v_1 \equiv \lambda v_2 \pmod{W}$. Somit ist die Definition $\lambda[v] := [\lambda v]$ unabhängig von der Wahl des Repräsentanten $v \in V$, d.h. wohldefiniert. Die anderen Vektorraumaxiome vererben sich von $V$ auf $V/W$, z.B. gilt der zweite Teil von B) für $V/W$ wegen

$$(\lambda + \mu)[v] \underset{\text{Def.}}{=} \big[(\lambda + \mu)v\big] \underset{\text{B) für } V}{=} [\lambda v + \mu v] \underset{\text{Def. } +,\cdot}{=} \lambda[v] + \mu[v]. \qquad \square$$

**3.3 Proposition.** *Die kanonische Abbildung $\pi \colon V \to V/W, v \mapsto [v]$ hat folgende Eigenschaften:*

*a)* $\pi$ *ist eine surjektive lineare Abbildung.*

*b)* $\ker(\pi) = W$.

**Beweis:** Dies folgt sofort aus Proposition II.1.15 und Proposition 3.2. $\square$

**3.4 Beispiel.** Sei $V = U \oplus W$ für Untervektorräume $U, W$. Dann haben wir einen Isomorphismus $\varphi \colon U \xrightarrow{\sim} V/W$, $u \mapsto [u]$.

⌜Nach Proposition 3.3 ist $\varphi$ linear und $\ker(\varphi) = U \cap W = \{0\}$ wegen der direkten Summe. Also ist $\varphi$ injektiv nach Proposition II.1.7. Um surjektiv zu zeigen, sei $[v] \in V/W$. Es existiert genau ein $u \in U, w \in W$ mit $v = u + w$. $\Longrightarrow$ $[v] = [u]$ und somit $\varphi(u) = [w]$.⌟

**3.5 Theorem (Homomorphiesatz).** *Sei $\varphi \colon V_1 \to V_2$ eine lineare Abbildung von $K$-Vektorräumen und $W_1$ ein Unterraum von $V_1$ mit $W_1 \subseteq \ker(\varphi)$. Dann gibt es genau eine lineare Abbildung $\overline{\varphi} \colon V_1/W_1 \to V_2$ mit $\overline{\varphi}\big([v_1]\big) = \varphi(v_1)$ für alle $v_1 \in V_1$.*

**Beweis:** Sei $[v_1]$ ein beliebiges Element aus dem Quotientenraum $V_1/W_1$. Wir definieren $\overline{\varphi}\big([v_1]\big) := \varphi(v_1)$ und müssen zeigen, dass dies nicht von der Wahl des Repräsentanten $v_1 \in V_1$ abhängt. Für $v_1' \in V_1$ gilt

$$v_1 \equiv v_1' \pmod{W_1} \Longrightarrow v_1 - v_1' \in W_1 \subseteq \ker(\varphi)$$

und damit folgt aus der Linearität $\varphi(v_1) = \varphi(v_1')$ wie gewünscht. Die Linearität von $\overline{\varphi}$ ergibt sich aus

$$\begin{aligned}\overline{\varphi}\big([v_1]+[v_1']\big) &\underset{\text{Def. }+}{=} \overline{\varphi}\big([v_1+v_1']\big) \underset{\text{Def. }\overline{\varphi}}{=} \varphi(v_1+v_1') \\ &\underset{\varphi\text{ linear}}{=} \varphi(v_1)+\varphi(v_1') \underset{\text{Def. }\overline{\varphi}}{=} \overline{\varphi}\big([v_1]\big)+\overline{\varphi}\big([v_1']\big)\end{aligned}$$

und

$$\begin{aligned}\overline{\varphi}\big(\lambda[v_1]\big) &= \overline{\varphi}\big([\lambda v_1]\big) = \varphi(\lambda v_1) = \lambda\varphi(v_1) \\ &= \lambda\overline{\varphi}\big([v_1]\big).\end{aligned}$$

*Eindeutigkeit*: Es ist also $\overline{\varphi}\colon V_1/W_1 \to V_2$ linear mit: $\forall v \in V$ gilt $\varphi(v) = \overline{\varphi}\big([v]\big)$. Ist dann $[v] \in V_1/W_1$ ein Vektor und $v \in \pi^{-1}\big([v]\big)$ ein (beliebiges) Urbild (wobei $\pi\colon V_1 \to V_1/W_1$ die kanonische Projektion aus Proposition 3.3 ist), so folgt $\overline{\varphi}\big([v]\big) = \varphi(v)$. Damit ist $\overline{\varphi}\big([v]\big)$ eindeutig durch $[v]$ bestimmt. □

**3.6 Theorem (Isomorphiesatz).** *Sei $\varphi\colon V_1 \to V_2$ eine lineare Abbildung. Dann gilt*

$$\overline{\varphi}\colon V_1/\ker(\varphi) \xrightarrow{\sim} \varphi(V_1), \quad [v_1] \mapsto \varphi(v_1).$$

**Beweis:** Nach Theorem 3.5 erhalten wir eine Abbildung $\overline{\varphi}\colon V_1/\ker(\varphi) \to \varphi(V_1)$, wenn wir noch den Wertebereich $V_2$ durch das Bild $\varphi(V_1)$ ersetzen. Es bleibt zu zeigen, dass $\overline{\varphi}$ bijektiv ist nach Proposition 1.17. Wegen

$$[v_1] \in \ker(\overline{\varphi}) \overset{\text{Def. }\overline{\varphi}}{\Longleftrightarrow} \varphi(v_1) = 0 \Longleftrightarrow v_1 \in \ker(\varphi) \Longleftrightarrow [v_1] = 0$$

folgt $\ker(\overline{\varphi}) = \{0\}$, d.h. $\overline{\varphi}$ ist injektiv nach Proposition II.1.7.

Sei $v_2 \in \varphi(V_1)$. Dann $\exists v_1 \in V_1$ mit $v_2 = \varphi(v_1)$. Nach Definition von $\overline{\varphi}$ gilt $v_2 = \overline{\varphi}\big([v_1]\big)$. Damit ist $\overline{\varphi}$ auch surjektiv. □

**3.7.** Für $K$-Vektorräume $V_1, V_2$ sei

$$\mathrm{Hom}_K(V_1,V_2) := \{\varphi\colon V_1 \to V_2 \mid \varphi\, K\text{-lineare Abbildung}\}.$$

**3.8 Proposition.** *Der Raum* $\mathrm{Hom}_K(V_1,V_2)$ *wird zu einem $K$-Vektorraum bezüglich der Addition*

$$\varphi+\psi\colon V_1 \to V_2, \quad v_1 \mapsto (\varphi+\psi)(v_1) := \varphi(v_1)+\psi(v_1)$$

*und der skalaren Multiplikation*

$$\lambda\cdot\varphi\colon V_1 \to V_2, \quad v_1 \mapsto (\lambda\varphi)(v_1) := \lambda\varphi(v_1)$$

*mit $\lambda \in K$.*

**Beweis:** Wir zeigen exemplarisch den ersten Teil von B), die anderen Axiome sehen wir in den Übungen: Für $v_1 \in V_1$ gilt

$$\begin{aligned}\big(\lambda\cdot(\varphi+\psi)\big)(v_1) &\underset{\text{Def. }\cdot}{=} \lambda\cdot\big((\varphi+\psi)(v_1)\big) \underset{\text{Def. }+}{=} \lambda\big(\varphi(v_1)+\psi(v_1)\big) \\ &\underset{\text{B) für }V_2}{=} \lambda\cdot\varphi(v_1)+\lambda\cdot\psi(v_1) \underset{\text{Def. }\cdot}{=} (\lambda\cdot\varphi)(v_1)+(\lambda\cdot\psi)(v_1) \\ &= (\lambda\cdot\varphi+\lambda\cdot\psi)(v_1)\end{aligned}$$

und damit sind $\lambda\cdot(\varphi+\psi)$ und $\lambda\cdot\varphi+\lambda\cdot\psi$ als Abbildungen gleich. □

**3.9 Definition.** Für einen Vektorraum $V$ definieren wir den **Dualraum** $V^*$ als $V^* := \mathrm{Hom}_K(V,K)$. Die Elemente aus $V^*$ sind lineare Abbildungen $f\colon V \to K$, die wir **lineare Funktionale** auf $V$ nennen.

**3.10 Korollar.** *$V^*$ ist mit den Operationen*

$$f+g\colon V \to K, v \mapsto f(v)+g(v) \quad \textit{und} \quad \lambda \cdot f\colon V \to K, v \mapsto \lambda \cdot f(v)$$

*für $\lambda \in K$ ein $K$-Vektorraum.*

**Beweis:** Nach Beispiel 1.7 ist $K$ selber ein $K$-Vektorraum und deshalb folgt die Behauptung aus Proposition 3.8. □

**3.11 Proposition.** *Die Abbildung*

$$\langle\ ,\ \rangle\colon V \times V^* \to K, \quad (v,f) \mapsto f(v)$$

*ist **bilinear**, d.h. sie ist linear im ersten Argument*

$$\langle v_1+v_2, f\rangle = \langle v_1, f\rangle + \langle v_2, f\rangle \quad \wedge \quad \langle \lambda v, f\rangle = \lambda \langle v, f\rangle$$

*und linear im zweiten Argument*

$$\langle v, f+g\rangle = \langle v, f\rangle + \langle v, g\rangle \quad \wedge \quad \langle v, \lambda f\rangle = \lambda \langle v, f\rangle.$$

**Beweis:** Die Linearität im ersten Argument folgt aus der Linearität von $f$ und die Linearität im zweiten Argument folgt aus der Definition von $+$ und der skalaren Multiplikation linearer Funktionale. □

**3.12.** Wir können die Bilinearform $\langle\ ,\ \rangle$ auf $V \times V^*$ benutzen um jedem $v \in V$ eine Funktion $\langle v, \cdot\rangle$ auf $V^*$ zuzuordnen. Dabei bedeutet der Punkt $\cdot$, dass wir $v$ als fest betrachten und die Funktion $f \mapsto \langle v, f\rangle \in K$ meinen. Wegen der Linearität im zweiten Argument ist $\langle v, \cdot\rangle$ ein lineares Funktional auf $V^*$. In anderen Worten gehört $\langle v, \cdot\rangle$ zum **Bidualraum** $V^{**} := (V^*)^*$

**3.13 Proposition.** *Die kanonische Abbildung $V \to V^{**}$, $v \mapsto \langle v, \cdot\rangle$ ist linear.*

**Beweis:** Die Linearität folgt aus der Linearität im ersten Argument in Proposition 3.11. □

**3.14 Bemerkung.** Wir werden später sehen, dass diese lineare Abbildung immer injektiv ist. Falls $V$ endlich dimensional ist, werden wir sogar zeigen, dass sie ein Isomorphismus ist, d.h. es gibt **Bidualität**. Das bedeutet, dass es einen kanonischen Isomorphismus von $V$ auf den Bidualraum gibt, mit dem man $V$ mit $V^{**}$ identifizieren kann.

**3.15 Proposition.** *Sei $\varphi\colon V_1 \to V_2$ eine lineare Abbildung von $K$-Vektorräumen. Dann existiert genau eine Abbildung $\varphi^*\colon V_2^* \to V_1^*$ mit*

$$\langle \varphi(v_1), f_2\rangle = \langle v_1, \varphi^*(f_2)\rangle$$

*für alle $v_1 \in V_1$, $f_2 \in V_2^*$.*

**Beweis:** Die obige Identität ist äquivalent zu

$$f_2 \circ \varphi(v_1) = f_2\big(\varphi(v_1)\big) = \big(\varphi^*(f_2)\big)(v_1),$$

deshalb setzen wir

$$\varphi^*(f_2) := f_2 \circ \varphi.$$

Dies zeigt die Existenz und die Eindeutigkeit folgt aus der Konstruktion. □

**3.16 Definition.** Diese Abbildung $\varphi^*\colon V_2^* \to V_1^*$ aus Proposition 3.15 heißt die **duale Abbildung** zu $\varphi$.

**3.17 Proposition.** *Die duale Abbildung $\varphi^*$ ist linear.*

**Beweis:** Für $f_2, g_2 \in V_2^*$ und $v_1 \in V_1$ gilt

$$\langle \varphi(v_1), f_2 \rangle = \langle v_1, \varphi^*(f_2) \rangle \quad \wedge \quad \langle \varphi(v_1), g_2 \rangle = \langle v_1, \varphi^*(g_2) \rangle.$$

Mit der Linearität im zweiten Argument folgt für jedes $\lambda \in K$:

$$\langle \varphi(v_1), \lambda f_2 + g_2 \rangle = \langle v_1, \lambda \varphi^*(f_2) + \varphi^*(g_2) \rangle.$$

Andererseits ist die linke Seite gleich $\langle v_1, \varphi^*(\lambda f_2 + g_2) \rangle$ nach Proposition 3.15. Weil $v_1 \in V_1$ beliebig war, folgt $\varphi^*(\lambda f_2 + g_2) = \lambda \varphi^*(f_2) + \varphi^*(g_2)$ nach Definition der Paarung. $\square$

# 4 Kapitel IV.
# Dimensionstheorie

In diesem Kapitel werden wir alle Vektorräume bis auf Isomorphie klassifizieren.

## 1. Linearkombinationen

In diesem Abschnitt betrachten wir alle Möglichkeiten, wie man aus gegebenen Vektoren eines Vektorraumes neue Vektoren konstruieren kann. Damit kann man den kleinsten Unterraum beschreiben, der eine gegebene Teilmenge $S$ enthält. Diesen Unterraum werden wir die lineare Hülle von $S$ nennen.

In diesem Abschnitt sei $K$ ein Körper und $V$ ein Vektorraum.

**1.1.** Wir beginnen mit Vektoren $v_1, v_2, \dots, v_n \in V$, die nicht notwendigerweise verschieden sein müssen. Ein Vektor der Form $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$ mit $\lambda_1, \dots, \lambda_n \in K$ heißt **Linearkombination** von $v_1, \dots, v_n$.

**1.2 Beispiel.** Die Menge der Linearkombinationen der Vektoren $v_1 = \left(\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right)$ und $v_2 = \left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right)$ bilden die Ebene durch $v_1$ und $v_2$.

**1.3 Proposition.** *Für jedes $S \subseteq V$ existiert ein kleinster Unterraum $W$ von $V$ mit der Eigenschaft $S \subseteq W$.*

**Beweis:** Es gibt offensichtlich mindestens einen Unterraum $W$ von $V$ mit $S \subseteq W$, denn wir können einfach $W := V$ wählen. Wegen Proposition III.2.8 ist

$$U := \bigcap_{\substack{W \text{ Unterraum,} \\ S \subseteq W}} W$$

ein Unterraum von $V$. Offenbar gilt $S \subseteq U$. Also ist $U$ auch ein Unterraum mit $S \subseteq U$ und andererseits gilt nach Konstruktion, dass $U \subseteq W$ für alle Unterräume $S \subseteq W$. Also ist $U$ der kleinste solche Unterraum. $\square$

**1.4 Definition.** Der kleinste Unterraum von $V$, der $S$ enthält, heißt **lineare Hülle** von $S$. Wir bezeichnen ihn mit $\langle S \rangle$. Wenn $S = \emptyset$, dann gilt $\langle \emptyset \rangle = \{0\}$. Falls $S \neq \emptyset$, dann haben wir folgende Beschreibung:

**1.5 Proposition.** *Die lineare Hülle von $S$ ist gleich der Menge der Linearkombinationen, die man aus endlich vielen Elementen aus $S$ bilden kann.*

**Beweis:** Die Menge $M$

$$M := \{\lambda_1 v_1 + \cdots + \lambda_n v_n \mid v_1, \dots, v_n \in S, \lambda_1, \dots, \lambda_n \in K, n \in \mathbb{N}\}$$

bildet einen Unterraum von $V$:

- die Summe von zwei solchen Linearkombinationen ist offenbar wieder in $M$.
- ein skalares Vielfaches einer Linearkombination aus $M$ ist aufgrund der Distributivität wieder eine solche Linearkombination.
- $0 = 0 \cdot v$ für irgendein $v \in S \Longrightarrow 0 \in M$.

Dies zeigt, dass $M$ ein Unterraum ist. Wegen $s = 1 \cdot s$ für $s \in S$ gilt $S \subseteq M$. Nach Definition der linearen Hülle folgt $\langle S \rangle \subseteq M$. Weil ein Unterraum abgeschlossen ist unter $+$ und skalaren Multiplikation, muss er jede Linearkombination aus seinen Elementen enthalten. Aus $S \subseteq \langle S \rangle$ folgt damit $M \subseteq \langle S \rangle$. Also gilt $M = \langle S \rangle$ wie gewünscht. $\square$

**1.6 Proposition.** *Seien $V, V'$ $K$-Vektorräume und $\varphi \colon V \to V'$, $\varphi' \colon V \to V'$ zwei lineare Abbildungen. Weiter nehmen wir an, dass die Einschränkungen von $\varphi$ und $\varphi'$ auf $S \subseteq V$ übereinstimmen. Dann stimmen $\varphi$ und $\varphi'$ auch auf der linearen Hülle von $S$ überein.*

**1.Beweis:** Sei $v \in \langle S \rangle$. Nach Proposition 1.5 gilt $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$ mit $v_i \in S$, $\lambda_i \in K$. Es folgt aus der Linearität von $\varphi$ und $\varphi'$

$$\varphi(v) = \varphi(\sum_{i=1}^{n} \lambda_i v_i) = \sum_{i=1}^{n} \lambda_i \varphi(v_i) = \sum_{i=1}^{n} \lambda_i \varphi'(v_i) = \varphi'(\sum_{i=1}^{n} \lambda_i v_i) = \varphi'(v)$$

wie gewünscht. $\square$

**2.Beweis:** Die Menge

$$\{v \in V \mid \varphi(v) = \varphi'(v)\} = \ker(\varphi - \varphi')$$

ist ein Unterraum von $V$ (Proposition III.2.7), der $S$ enthält. Nach Definition der linearen Hülle enthält obige Menge die lineare Hülle $\langle S \rangle$, was zu zeigen war. $\square$

**1.7 Definition.** Die Vektoren $v_1, \ldots, v_n \in V$ heißen **linear abhängig**, wenn es $\lambda_1, \ldots, \lambda_n \in K$ gibt, die nicht alle 0 sind und für die

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$

gilt. Wenn $v_1, \ldots, v_n$ nicht linear abhängig sind, dann heißen sie **linear unabhängig**, d.h. aus

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$

mit $\lambda_i \in K$ folgt immer $\lambda_i = \cdots = \lambda_n = 0$. Allgemein definiert man eine Familie $(v_i)_{i \in I}$ mit $v_i \in V$ als

- linear abhängig, wenn es eine endliche Teilfamilie gibt, die linear abhängig ist im obigen Sinn
- linear unabhängig, wenn sie nicht linear abhängig ist, d.h. jede endliche Teilfamilie ist linear unabhängig im obigen Sinn.

**1.8 Proposition.** *Es seien $v_1, \ldots, v_n \in V$ mit $n \geq 2$. Dann sind folgende Aussagen äquivalent:*

*a) $v_1, \ldots, v_n$ sind linear abhängig.*

*b) Einer der Vektoren $v_1, \ldots, v_n$ lässt sich als Linearkombination der anderen Vektoren schreiben.*

**Beweis:** „ $a \Longrightarrow b$ “: Wenn $v_1, \ldots, v_n$ linear abhängig sind, dann $\exists \lambda_1, \ldots, \lambda_n \in K$, nicht alle 0, mit

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$

Es gibt also ein $i$ mit $\lambda_i \neq 0$ und es folgt

$$\lambda_i v_i = -\sum_{j \neq i} \lambda_j v_j$$

und damit

$$v_i = -\sum_{j \neq i} \frac{\lambda_j}{\lambda_i} v_j$$

wie gewünscht. Dies zeigt b) aus a).

„ $b \Longrightarrow a$ “: $\exists i \in \{1,\dots,n\}$ mit

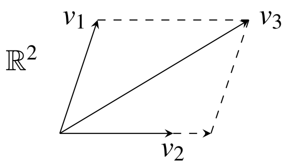
$$v_i = \sum_{j \neq i} \lambda_j v_j$$

Wenn wir $\lambda_i := -1$ setzen, dann folgt

$$0 = \sum_{j=1}^{n} \lambda_j v_j$$

Also sind $v_1,\dots,v_n$ linear abhängig, d.h. a) gilt. □

**1.9 Beispiel.** Wir betrachten die Vektoren $v_1, v_2, v_3$ aus dem beiliegenden Bild. Sie sind in $\mathbb{R}^2$ linear abhängig, da wir $v_3$ als Linearkombination von $v_1$ und $v_2$ darstellen können. Wir nehmen nun $v_4$ im Raum $\mathbb{R}^3 \setminus (xy\text{-Ebene})$ hinzu und betrachten die obigen $v_1, v_2, v_3$ als Vektoren in der $xy$-Ebene. Dann gilt




- $v_1, v_2, v_3, v_4$ sind linear abhängig (da schon $v_1, v_2, v_3$ linear abhängig war).
- $v_1, v_2, v_4$ sind linear unabhängig, denn sonst wären die drei Vektoren nach Proposition 1.8 in einer Ebene.

**1.10 Definition.** Eine Familie $(v_i)_{i\in I}$ von Vektoren $v_i \in V$ heißt ein **Erzeugendensystem** von $V$ genau dann, wenn ihre lineare Hülle gleich $V$ ist.

**1.11.** Nach Proposition 1.5 ist dies äquivalent dazu, dass jeder Vektor aus $V$ sich als Linearkombination einer endlichen Teilfamilie von $(v_i)_{i\in I}$ schreiben lässt.

Eine Teilmenge $S \subseteq V$ heißt **Erzeugendensystem** von $V :\Longleftrightarrow \langle S \rangle = V$.

**1.12 Beispiel.** $S \subseteq \{v_1, v_2, v_3, v_4\}$ aus Beispiel 1.9 bildet genau dann ein Erzeugendensystem von $\mathbb{R}^3$, wenn $|S| \geq 3$ und $v_4 \in S$.

## 2. Basis

Die Basis ist ein zentraler Begriff in der Theorie der Vektorräume, weil man damit Koordinaten einführen kann, mit denen man rechnen kann. In diesem Abschnitt werden wir die grundlegenden Eigenschaften von Basen studieren.

Dabei bezeichnet $K$ einen Körper und $V$ einen $K$-Vektorraum.

**2.1 Definition.** Eine **Basis** von $V$ ist ein linear unabhängiges Erzeugendensystem $(b_i)_{i\in I}$ von $V$. Konvention: leere Familie ist Basis des Nullraums.

**2.2 Beispiel.** $\mathbb{R}^3$ besitzt die Standardbasis

$$\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$$

Es gibt noch viele andere Basen wie z.B. $v_1, v_2, v_4$ oder $v_1, v_3, v_4$ oder $v_2, v_3, v_4$ in Beispiel 1.12 bzw. 1.9.

**2.3 Beispiel.** Auch der Vektorraum $K[x]$ der Polynome hat eine Standardbasis: $1, x, x^2, x^3, \dots, x^n, \dots$ aufgrund der Definition der Polynome.

**2.4 Theorem.** *Für eine Familie $(b_i)_{i\in I}$ von Vektoren $b_i \in V$ sind folgende Aussagen äquivalent.*

*a) $(b_i)_{i\in I}$ ist eine Basis*

*b) $(b_i)_{i\in I}$ ist eine maximale linear unabhängige Familie, d.h.*

$$\forall v \in V \Longrightarrow v, (b_i)_{i\in I} \text{ ist linear abhängig}$$

*c) $(b_i)_{i\in I}$ linear unabhängig und $\exists$ Erzeugendensystem $(w_j)_{j\in J}$ von $V$ so, dass*

$$\forall j \in J \Longrightarrow w_j, (b_i)_{i\in I} \text{ ist linear abhängig}$$

*d) $(b_i)_{i\in I}$ ist ein minimales Erzeugendensystem, d.h.*

$$\forall k \in I \Longrightarrow (b_i)_{i\in I\setminus\{k\}} \text{ ist kein Erzeugendensystem}$$

**Beweis:** Wir zeigen zuerst „ a) $\Longrightarrow$ b) “. Sei $(b_i)_{i\in I}$ eine Basis von $V$. Sei $v \in V$. Zu zeigen: $v, (b_i)_{i\in I}$ ist linear abhängig. Weil $(b_i)_{i\in I}$ ein Erzeugendensystem ist, können wir $v$ als Linearkombination von endlich vielen $b_i$ schreiben. Nach Proposition 1.8 ist dann $v, (b_i)_{i\in I}$ linear abhängig. Also gilt b).

„ b) $\Longrightarrow$ c) “. Wir wählen als Erzeugendensystem die Familie $(v)_{v\in V}$ (die Menge $V$ erzeugt natürlich den Vektorraum $V$). Dann folgt c) sofort aus b).

„ c) $\Longrightarrow$ d) “. Zuerst zeigen wir, dass $(b_i)_{i\in I}$ ein Erzeugendensystem von $V$ ist. Wir nehmen dabei an, dass c) gilt für das Erzeugendensystem $(w_j)_{j\in J}$. Wir benutzen folgende Aussage aus Aufgabe 9.2: „Wenn jedes Element aus einem Erzeugendensystem sich als Linearkombination von endlich vielen Vektoren aus einer gegebenen Familie schreiben lässt, dann ist diese Familie ebenfalls ein Erzeugendensystem.“ Also genügt es hier zu zeigen, dass jedes $w_j$ eine Linearkombination von endlich vielen $b_i$ ist. Nach c) gilt

$$\lambda w_j + \sum_{i\in I_0} \lambda_i b_i = 0 \tag{IV.1}$$

für ein endliches $I_0 \subseteq I$ und $\lambda, (\lambda_i)_{i\in I_0}$ aus $K$, die nicht alle 0 sind. Weil die $b_i$ nach Voraussetzung c) linear unabhängig sind, muss $\lambda \neq 0$ gelten in (IV.1). $\Longrightarrow w_j = -\sum_{i\in I_0} \frac{\lambda_i}{\lambda} b_i$. Wir müssen noch zeigen, dass dieses Erzeugendensystem $(b_i)_{i\in I}$ minimal ist. Sei dazu $k \in I$. Wir müssen zeigen, dass $(b_i)_{i\in I\setminus\{k\}}$ kein Erzeugendensystem ist. Wir machen einen indirekten Beweis, nehmen also an, dass $(b_i)_{i\in I\setminus\{k\}}$ ein Erzeugendensystem ist. Wir werden dazu einen Widerspruch herleiten.

Nach Annahme lässt sich also $b_k$ als Linearkombination von endlich vielen $b_i$ schreiben mit $i \neq k$. Nach Proposition 1.8 ist dann die ganze Familie $(b_i)_{i\in I}$ linear abhängig. Dies widerspricht der Annahme c). Also ist $(b_i)_{i\in I\setminus\{k\}}$ *kein* Erzeugendensystem, wie gewünscht. Dies zeigt d).

"d) $\Longrightarrow$ a)". Es sei $(b_i)_{i\in I}$ ein minimales Erzeugendensystem. Um a) zu zeigen, müssen wir nur die lineare Unabhängigkeit von $(b_i)_{i\in I}$ zeigen. Wir argumentieren wieder mit einem indirekten Beweis und nehmen an, dass $(b_i)_{i\in I}$ linear abhängig ist, d.h.

$$\sum_{i\in I_0} \lambda_i b_i = 0$$

für endlichen $I_0 \subseteq I$ und $\lambda_i \in K\ \forall i \in I_0$, $\lambda_k \neq 0$ für ein $k \in I_0$. Wie in (IV.1) folgt, dass $b_k$ eine Linearkombination der $(b_i)_{i\in I_0\setminus\{k\}}$ ist. Nach Aufgabe 9.2 ist auch $(b_i)_{i\in I\setminus\{k\}}$ ein Erzeugendensystem von $V$. Dies widerspricht der Annahme d), dass das Erzeugendensystem $(b_i)_{i\in I}$ minimal ist. Also ist $(b_i)_{i\in I}$ linear unabhängig und damit wie gesehen eine Basis. Dies zeigt a). □

**2.5 Theorem (Basisergänzungssatz).** *Seien $v_1, \ldots, v_n$ linear unabhängige Vektoren aus $V$ und $w_1$, $\ldots, w_m$ ein Erzeugendensystem von $V$. Dann gibt es $1 \leq j_1 < \ldots < j_r \leq m$ so, dass $v_1, \ldots, v_n, w_{j_1}$, $\ldots, w_{j_r}$ eine Basis von $V$ ist.*

**Beweis:** Wir machen den Beweis im Hinblick auf die spätere Verallgemeinerung für beliebige Familien. Wir betrachten alle Möglichkeiten $1 \leq j_1 < \ldots < j_r \leq m$ so, dass

$$v_1, \ldots, v_n, w_{j_1}, \ldots, w_{j_r} \tag{IV.2}$$

linear unabhängig bleiben. Diese Möglichkeiten werden partiell geordnet durch

$$\left[(1 \leq j_1 < \ldots < j_r \leq m) \leq (1 \leq k_1 < \ldots < k_s \leq m)\right] :\Longleftrightarrow \{j_1, \ldots, j_r\} \subseteq \{k_1, \ldots, k_s\}$$

Wir lassen auch die leere Ergänzung zu in (IV.2), d.h. $r = 0$. Weil die $v_1, \ldots, v_n$ linear unabhängig sind, haben wir damit mindestens eine Möglichkeit. Weil es nur endlich viele Möglichkeiten in (IV.2) gibt, muss es eine maximale Möglichkeit geben (z.B. wenn $r$ maximal ist). Wir nehmen an, dass (IV.2) so eine maximale Möglichkeit ist. Wir behaupten, dass damit (IV.2) eine Basis ist. Dies folgt sofort aus Theorem 2.4 weil c) aufgrund der Maximalität von (IV.2) erfüllt ist. □

**2.6 Korollar.** *Wenn $V$ ein Erzeugendensystem $w_1, \ldots w_m$ hat, dann bildet ein geeigneter Teil $w_{j_1}, \ldots,$ $w_{j_r}$ eine Basis von $V$.*

**Beweis:** Dies folgt sofort aus Theorem 2.5, wenn wir für $(v_i)_{i \in I}$ für leere Familie nehmen, d.h. wenn kein $v_i$ auftritt. □

**2.7.** Das Korollar 2.6 zeigt, dass Vektorräume mit einem endlichen Erzeugendensystem eine Basis haben. Für beliebige Vektorräume kann man dies auch zeigen unter Benutzung des Zornschen Lemmas.

**2.8 Theorem (allgemeiner Basisergänzungssatz).** *Sei $(v_i)_{i \in I}$ eine linear unabhängige Familie aus $V$ und sei $(w_j)_{j \in J}$ ein Erzeugendensystem von $V$. Dann $\exists S \subseteq J$ so, dass $(v_i)_{i \in I}, (w_j)_{j \in S}$ eine Basis von $V$ ist.*

**Beweis:** Analog zum Beweis von Theorem 2.5 betrachten wir alle $S \subseteq J$ so, dass

$$(v_i)_{i \in I}, (w_j)_{j \in S} \tag{IV.3}$$

linear unabhängig bleibt. Sei $X$ die Menge aller solcher $S$. Wir definieren eine partielle Ordnung

$$S \leq S' :\Longleftrightarrow S \subseteq S'$$

auf $X$. Da $\emptyset \in X \Longrightarrow X \neq \emptyset$. Wir behaupten, dass es ein maximales $S \in X$ gibt. Wenn dies stimmt, dann kann man wie im Beweis von Theorem 2.5 zeigen, dass das zum maximalen $S$ gehörige (IV.3) eine Basis ist.

Die Existenz eines maximalen $S \in X$ folgt aus dem Zornschen Lemma (siehe Axiom I.3.29). Dazu müssen wir zeigen, dass jede total geordnete Menge $Y \subseteq X$ eine obere Schranke in $X$ hat. Wir werden zeigen, dass

$$T := \bigcup_{S \in Y} S$$

eine obere Schranke von $Y$ ist. Offenbar gilt $T \subseteq J$ und $S \subseteq T \; \forall S \in Y$. Es bleibt zu zeigen, dass

$$(v_i)_{i \in I}, (w_j)_{j \in T} \tag{IV.4}$$

linear unabhängig bleibt. Es sei also

$$\sum_i \lambda_i v_i + \sum_j \lambda'_j w_j = 0 \tag{IV.5}$$

für endlich viele $i \in I$ und endlich viele $j \in T$ mit $\lambda_i, \lambda'_j \in K$. Jedes solcher $j$ gehört zu einem $S_j \in Y$. Weil $Y$ total geordnet ist, muss es ein größtes $S_j$ geben, da nur endlich viele $j$ in (IV.5) auftreten. Wir bezeichnen dieses größte $S_j$ mit $S$. Damit liegen alle Indizes $j$ aus (IV.5) in $S$. Weil (IV.3) für dieses $S$ linear unabhängig ist, folgt $\lambda_i = \lambda'_j = 0$ für alle $i$ und $j$ aus (IV.4). Damit ist (IV.4) linear unabhängig, d.h. $T$ ist obere Schranke von $Y$ in $X$. □

**2.9 Korollar.** *Jeder Vektorraum hat eine Basis.*

**Beweis:** Analog zu Korollar 2.6. □

**2.10 Korollar.** *Sei $U$ ein Unterraum von $V$. Dann gibt es einen Unterraum $W$ von $V$ so, dass $V = U \oplus W$.*

**Beweis:** Nach Korollar 2.9 hat $U$ eine Basis $(u_i)_{i\in I}$. Nach Theorem 2.8 können wir $(u_i)_{i\in I}$ ergänzen zu einer Basis $(u_i)_{i\in I\dot\cup J}$ von $V$. Sei $W$ die lineare Hülle der Familie $(u_j)_{j\in J}$. Nach Konstruktion gilt $V = U+W$, dann sei $v \in V$, dann gilt

$$v = \lambda_{i_1}u_{i_1} + \ldots + \lambda_{i_m}u_{i_m} + \lambda_{i_{m+1}}u_{i_{m+1}} + \ldots \lambda_{i_r}u_{i_r},$$

wobei $i_1,\ldots i_m \in I$, $i_{m+1},\ldots,i_r \in J$ sein sollen. Also gilt $v = u+w$ mit

$$u := \lambda_{i_1}u_{i_1} + \ldots + \lambda_{i_m}u_{i_m} \text{ und } w := \lambda_{i_{m+1}}u_{i_{m+1}} + \ldots + \lambda_{i_r}u_{i_r}.$$

Weil $u \in U, w \in W$ folgt $V = U+W$. Es bleibt $U \cap W = \{0\}$ zu zeigen. Sei dazu $v \in U \cap W$. Somit ist $v$ eine Linearkombination von endlich vielen $u_i$, $i \in I$ und Linearkombination von endlich vielen $u_j$, $j \in J$.

$$\Longrightarrow \sum_{i\in I_0} \lambda_i u_i = v = \sum_{j\in J_0} \lambda_j u_j$$

für endliche $I_0 \subseteq I$ und $J_0 \subseteq J$. Da $(u_k)_{k\in I\dot\cup J}$ eine Basis ist, folgt sofort $\lambda_i = 0$, $\lambda_j = 0\ \forall i \in I_0$, $\forall j \in J_0$. Also gilt $v = 0$ wie gewünscht. Es folgt $U \cap W = \{0\}$ und damit auch $V = U \oplus W$. □

**2.11 Korollar.** *Sei $U$ ein Unterraum von $V$ und $f \in U^*$. Dann existiert $g \in V^*$ mit $g|_U = f$.*

**Beweis:** Nach Korollar 2.10 gibt es einen Unterraum $W$ von $V$ so, dass $V = U \oplus W$. Für jedes $v \in V$ $\exists!(u,w) \in U \times W$ mit $v = u+w$. Wir setzen $\pi(v) := u$. Das ergibt eine Abbildung $\pi\colon V \to U$. Aufgrund der Eindeutigkeit der Darstellung folgt leicht, dass $\pi$ linear ist. Dies kann auch aus Aufgabe 8.3 gewonnen werden. Offensichtlich gilt $\pi(u) = u$ für alle $u \in U$ (denn $u = u+0$ mit $u \in U, 0 \in W$). Nach Proposition III.1.15 ist $g\colon f \circ \pi$ ein lineares Funktional auf $V$. Für $u \in U$ gilt

$$g(u) = f\big(\pi(u)\big) = f(u)$$

und damit folgt die Behauptung. □

**2.12 Korollar.** *Die kanonische Abbildung $V \to V^{**} := (V^*)^*$, $v \mapsto \langle v,\cdot\rangle$ aus Proposition III.3.13 ist eine injektive lineare Abbildung.*

**Beweis:** Wir haben schon in Proposition III.3.13 gesehen, dass die kanonische Abbildung $V \to V^{**}$ linear ist. Um Injektivität zu zeigen, müssen wir für $v \in V \setminus \{0\}$ zeigen, dass $\langle v,\cdot\rangle$ nicht Null ist in $V^{**}$. Beachte, dass $v$ eine Basis ist der linearen Hülle $U := \langle v\rangle = \{\lambda v \mid \lambda \in K\}$. Wir definieren nun $f \in U^*$ durch $f(\lambda v) := \lambda$. Nach Korollar 2.11 lässt sich $f$ zu einem linearen Funktional $g \in V^*$ fortsetzen. Es gilt

$$\langle v,g\rangle \overset{\text{Def. von } \langle\cdot,\cdot\rangle}{=} g(v) \overset{\text{da } g|_U = f}{=} f(v) = 1$$

Also ist $\langle v,\cdot\rangle \neq 0 \in V^{**}$ wie gewünscht. □

Im Folgenden werden wir uns der Einfachheit halber auf den Fall der Vektorräume mit endlicher Basis beschränken. Für Verallgemeinerungen verweisen wir auf die Literatur.

**2.13 Theorem.** *Für Vektoren $b_1,\ldots,b_n \in V$ sind folgende Aussagen äquivalent.*

*a)* $b_1,\ldots,b_n$ *ist eine Basis von* $V$

*b)* $\forall v \in V \Longrightarrow \exists!\lambda_1,\ldots,\lambda_n \in K$ *mit* $v = \lambda_1 b_1 + \ldots + \lambda_n b_n$

**Beweis:** „a) $\Longrightarrow$ b)“: Es sei $b_1,\dots,b_n$ eine Basis. Weil $b_1,\dots,b_n$ damit ein Erzeugendensystem ist, gibt es $\lambda_1,\dots,\lambda_n$ mit $v=\lambda_1 b_1+\ldots+\lambda_n b_n$. Wenn $b=\lambda_1' b_1+\ldots+\lambda_n' b_n$ eine weitere solche Linearkombination ist, dann folgt

$$(\lambda_1'-\lambda_1)b_1+\ldots+(\lambda_n'-\lambda_n)b_n=b-b=0$$

Aus der linearen Unabhängigkeit folgt $\lambda_1=\lambda_1',\dots,\lambda_n=\lambda_n'$ und damit die Eindeutigkeit.

„b) $\Longrightarrow$ a)“: Aus der Existenz in b) folgt sofort, dass die Vektoren ein Erzeugendensystem von $V$ bilden. Aus der Eindeutigkeit in b) für $v=0$ folgt sofort die lineare Unabhängigkeit von $b_1,\dots,b_n$. Also ist es eine Basis. □

**2.14.** Wenn $b_1,\dots,b_n$ eine Basis von $V$ ist, dann nennen wir die eindeutigen Zahlen $\lambda_1,\dots,\lambda_n$ aus Theorem 2.13 b) die **Koordinaten** von $v$ bzgl. $b_1,\dots,b_n$. Weiter heißt $\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix}\in K^n$ der **Koordinatenvektor** von $V$ bzgl. $b_1,\dots,b_n$.

**2.15 Proposition.** *Wenn $V$ die Basis $b_1,\dots,b_n$ hat, dann ist die Abbildung $\varphi\colon K^n\to V$, $\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix}\mapsto \lambda_1 b_1+\ldots+\lambda_n b_n$ ein Isomorphismus, dessen Umkehrabbildung ein Vektor $v\in V$ seinen Koordinatenvektoren bzgl. $b_1,\dots,b_n$ zuordnet.*

**Beweis:** Es ist leicht zu sehen, dass $\varphi$ eine lineare Abbildung ist. Sei $\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix},\begin{pmatrix}\mu_1\\ \vdots\\ \mu_n\end{pmatrix}\in K^n$. Dann gilt

$$\begin{aligned}\varphi(\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix}+\begin{pmatrix}\mu_1\\ \vdots\\ \mu_n\end{pmatrix})&=\varphi(\begin{pmatrix}\lambda_1+\mu_1\\ \vdots\\ \lambda_n+\mu_1\end{pmatrix})=(\lambda_1+\mu_1)b_1+\ldots+(\lambda_n+\mu_n)b_n\\ &=(\lambda_1 b_1+\ldots\lambda_n b_n)+(\mu_1 b_1+\ldots+\mu_n b_n)=\varphi(\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix})+\varphi(\begin{pmatrix}\mu_1\\ \vdots\\ \mu_n\end{pmatrix})\end{aligned}$$

Analog gilt

$$\varphi(\lambda\cdot\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix})=\lambda\cdot\varphi(\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_n\end{pmatrix})$$

für $\lambda\in K$. Also ist $\varphi$ linear. Sei $\psi\colon V\to K^n$ dadurch gegeben, dass wir $v\in V$ den Koordinatenvektor bzgl. $b_1,\dots,b_n$ zuordnen. Nach Konstruktion gilt $\psi\circ\varphi=\mathrm{id}_{K^n}$ und $\varphi\circ\psi=\mathrm{id}_V$, d.h. $\varphi$ ist bijektiv. Nach Proposition III.1.17 ist $\varphi$ ein Isomorphismus. □

**2.16 Bemerkung.** Die obige Proposition zeigt, dass Vektorräume mit $n$ Basisvektoren isomorph zu $K^n$ sind. Es ist aber zu beachten, dass der Isomorphismus *nicht* eindeutig ist, weil er von der Wahl der Basis abhängt.

Zum Schluss wollen wir noch zeigen, wie man in der Praxis entscheidet, ob gegebene Vektoren in $K^n$ linear unabhängig sind, ein Erzeugendensystem oder gar eine Basis bilden.

**2.17 Beispiel.** Wir betrachten die Vektoren

$$v_1=\begin{pmatrix}0\\2\\4\\2\end{pmatrix},v_2=\begin{pmatrix}1\\-1\\1\\1\end{pmatrix},v_3=\begin{pmatrix}-1\\1\\-1\\-1\end{pmatrix},v_4=\begin{pmatrix}1\\1\\1\\1\end{pmatrix},v_5=\begin{pmatrix}1\\-1\\1\\2\end{pmatrix}$$

aus $\mathbb{R}^4$ und wir wollen obige Frage beantworten.

Die Vektoren sind genau dann linear unabhängig, wenn das homogene lineare Gleichungssystem

$$x_1 \cdot v_1 + x_2 \cdot v_2 + x_3 \cdot v_3 + x_4 \cdot v_4 + x_5 \cdot v_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

nur die triviale Lösung $x_1 = \ldots = x_5 = 0$ hat. Schematisch schreiben wir das Gleichungssystem als

$$\begin{array}{rrrrr|l} 0 & 1 & -1 & 1 & 1 & 0 \\ 2 & -1 & 1 & 1 & -1 & 0 \\ 4 & 1 & -1 & 1 & 1 & 0 \\ 2 & 1 & -1 & 1 & 2 & 0 \end{array}$$

In Beispiel I.4.9 haben wir mit dem Gaußalgorithmus das äquivalente Gleichungssystem in Zeilenstufenform gefunden:

$$\begin{array}{rrrrr|l} 2 & -1 & 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}$$

Den Lösungsraum findet man durch Rückwärtslösen und ist gleich

$$\mathbb{L} = \{ \begin{pmatrix} 0 \\ \lambda \\ \lambda \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^5 \mid \lambda \in \mathbb{R} \}$$

Damit sind $v_1, v_2, v_3, v_4, v_5$ linear abhängig. Dafür ist $v_1, v_2, v_4, v_5$ linear unabhängig, weil es nur die triviale Lösung mit $x_3 = 0$ gibt (siehe $\mathbb{L}$). Analog ist $v_1, v_3, v_4, v_5$ linear unabhängig, aber alle anderen 4 Vektoren sind linear abhängig.

Die Vektoren $v_1, \ldots, v_5$ bilden ein Erzeugendensystem, genau dann, wenn das inhomogene lineare Gleichungssystem

$$\begin{array}{rrrrr|l} 0 & 1 & -1 & 1 & 1 & y_1 \\ 2 & -1 & 1 & 1 & -1 & y_2 \\ 4 & 1 & -1 & 1 & 1 & y_3 \\ 2 & 1 & -1 & 1 & 2 & y_4 \end{array}$$

für alle $y_1, y_2, y_3, y_4 \in \mathbb{R}$ eine Lösung hat. Bei den Zeilenoperationen in I.4.7 passiert $y_i \to \lambda y_i$, $y_i \leftrightarrow y_j$, $y_i \mapsto y_i + \lambda y_j$ mit $\lambda \in \mathbb{R} \setminus 0$ und $i \neq j$. Dies sind alle bijektive lineare Abbildungen $\mathbb{R}^4 \to \mathbb{R}^4$, weil man die wieder rückgängig machen kann. Also ist $v_1, \ldots, v_5$ genau dann ein Erzeugendensystem, wenn

$$\begin{array}{rrrrr|l} 2 & -1 & 1 & 1 & -1 & z_1 \\ 0 & 1 & -1 & 1 & 1 & z_2 \\ 0 & 0 & 0 & -4 & 0 & z_3 \\ 0 & 0 & 0 & 0 & 1 & z_4 \end{array}$$

für alle $z_1, z_2, z_3, z_4 \in \mathbb{R}$ eine Lösung hat. Dies hat aber offenbar immer eine Lösung, weil die Zeilenstufenform bis ganz nach unten geht. Also haben wir ein Erzeugendensystem. Wir erhalten sogar immer eine Lösung mit $x_3 = 0$, d.h. auch $v_1, v_2, v_4, v_5$ sind ein Erzeugendensystem und aus unseren früheren Betrachtungen folgt, dass $v_1, v_2, v_4, v_5$ eine Basis ist. Analog ist $v_1, v_3, v_4, v_5$ eine Basis. Wenn $x_4 = 0$ (bzw. $x_5 = 0$, bzw. $x_1 = 0$) ist, dann gibt es offenbar nicht immer eine Lösung, und so ist $v_1, v_2, v_3, v_5$ (bzw. $v_1, v_2, v_3, v_4$, bzw. $v_2, v_3, v_4, v_5$) kein Erzeugendensystem und damit keine Basis.

Mit dem Argument kann man zeigen, dass $k$ Vektoren in $K^n$ genau dann linear unabhängig sind, wenn sie Zeilenstufenform

$$
\begin{array}{c} k\left\{\vphantom{\begin{array}{c}1\\1\\1\end{array}}\right. \\ \vphantom{\begin{array}{c}1\\1\end{array}} \end{array}
\underbrace{\begin{array}{|ccc|}\hline 1 & & \\ 0 & 1 & \\ \vdots & 0 & 1 \\ \vdots & & 0 \\ 0 & \cdots & 0 \\ \hline \end{array}}_{k}
$$

haben. Sie bilden genau dann ein Erzeugendensystem, wenn die Zeilenstufenform ganz nach unten geht. Damit bilden sie genau dann einen Basis, wenn die Zeilenstufenform

$$
\begin{array}{|cccc|}\hline 1 & * & * & * \\ 0 & \ddots & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & 1 \\ \hline \end{array}
$$

vorliegt. Insbesondere bestehen Basen von $K^n$ aus genau $n$ Vektoren. Dies werden wir im nächsten Abschnitt vertiefen.

## 3. Dimension

Wie schon im letzten Abschnitt angedeutet, haben zwei Basen eines Vektorraums immer dieselbe Länge. Diese Länge nennt man die Dimension des Vektorraums. Sie klassifiziert die endlich erzeugten Vektorräume bis auf Isomorphie. Die Dimension ist also eine der wichtigsten Begriffe in der linearen Algebra.

In diesem Abschnitt sei $V$ immer ein Vektorraum über dem Körper $K$.

**3.1 Lemma (Austauschlemma).** *Sei $b_1, \ldots, b_n$ eine Basis von $V$ und sei $v$ eine Linearkombination*

$$v = \lambda_1 b_1 + \ldots + \lambda_n b_n \tag{IV.6}$$

*mit $\lambda_1 \neq 0$. Dann ist auch $v, b_2, \ldots, b_n$ eine Basis von $V$.*

**Beweis:** Wir zeigen zuerst linear unabhängig. Seien $\mu_1, \ldots, \mu_n \in K$ mit

$$\mu_1 v + \mu_2 b_2 + \ldots + \mu_n b_n$$

Dann liefert Einsetzen von (IV.6):

$$\mu_1 \lambda_1 b_1 + (\mu_1 \lambda_2 + \mu_2) b_2 + \ldots + (\mu_1 \lambda_n + \mu_n) b_n = 0$$

Aus der linearen Unabhängigkeit von $b_1, \ldots, b_n$ und aus $\lambda_1 \neq 0$ folgt sukzessive:

$$\mu_1 = 0,\ \mu_2 = 0,\ \ldots,\ \mu_{n-1} = 0,\ \mu_n = 0$$

Somit ist $v, b_2, \ldots, b_n$ linear unabhängig.

Es bleibt zu zeigen, dass $v, b_2, \ldots, b_n$ ein Erzeugendensystem von $V$ ist. Nach Aufgabe 9.2 genügt es zu zeigen, dass $b_1$ eine Linearkombination von $v, b_2, \ldots, b_n$ ist, da $b_1, \ldots, b_n$ ein Erzeugendensystem bildet und $b_2, \ldots, b_n$ trivialerweise Linearkombinationen von $v, b_2, \ldots, b_n$ sind. Dies folgt aus (IV.6) wegen

$$\begin{aligned} b_1 &= \lambda_1^{-1}(v - \lambda_2 b_2 - \ldots - \lambda_n b_n) \\ &= \lambda_1^{-1} v - \lambda_1^{-1}\lambda_2 b_2 - \ldots - \lambda_1^{-1}\lambda_n b_n \end{aligned}$$

□

**3.2 Theorem (Austauschsatz von Steinitz).** *Es sei $b_1, \dots, b_n$ eine Basis und $(w_j)_{j\in J}$ ein Erzeugendensystem von $V$. Für $k \in \{0, \dots, n\}$ gibt es eine k-elementige Teilmenge $S$ von $J$ so, dass $(w_j)_{j\in S}, b_{k+1}, \dots, b_n$ eine Basis von $V$ ist.*

**Beweis** mit Induktion nach $k$: Induktionsanfang für $k = 0$ ist trivial, $S = \emptyset$.

Induktionsschritt $(k-1) \mapsto k$: Es sei $k \geq 1$ und die Behauptung sei richtig für $k-1$. Also gibt es eine $(k-1)$-elementige Teilmenge $T$ von $J$ so, dass $(w_j)_{j\in T}, b_k, \dots, b_n$ eine Basis von $V$ ist. Für $\iota \in J$ können wir $w_\iota$ als Linearkombination dieser Basis schreiben. Dabei muss es ein $\iota \in J \setminus T$ geben so, dass

$$w_\iota = \sum_{j\in T} \mu_j w_j + \sum_{i=k}^{n} \lambda_i b_i$$

mit $\mu_j, \lambda_i \in K$ und $\lambda_k \neq 0$. Denn sonst wäre schon $(w_j)_{j\in T}, b_{k+1}, \dots, b_n$ ein Erzeugendensystem nach Aufgabe 9.2 und dies widerspricht der Tatsache, dass jede Basis ein minimales Erzeugendensystem ist nach Theorem 2.4. Also gibt es so ein $\iota$ und mit dem Austauschlemma kann man $b_k$ mit $w_\iota$ tauschen. Wir setzen $S := T \cup \{\iota\}$ und damit folgt der Induktionsschritt. □

**3.3 Theorem.** *In einem endlich erzeugten Vektorraum gibt es eine endliche Basis und alle Basen von $V$ haben dieselbe Länge.*

**Beweis:** Nach Korollar 2.6 hat $V$ eine endliche Basis $b_1, \dots, b_n$. Sei $(w_j)_{j\in J}$ eine beliebige Basis. Nach dem Austauschsatz von Steinitz gibt es eine $n$-elementige Teilmenge $S$ von $J$ so, dass $(w_j)_{j\in S}$ eine Basis von $V$ bilden. Das geht nur, wenn $S = J$ ist. □

**3.4 Definition.** Die Länge einer endlichen Basis von $V$ nennen wir die **Dimension** von $V$ und bezeichnen sie mit $\dim(V)$. Wenn $V$ keine endliche Basis haben sollte, dann heißt $V$ **unendlich-dimensional** und wir setzen $\dim(V) := \infty$.

**3.5 Beispiel.** $K^n$ hat die Standardbasis $e_i = (0, \dots, 0, 1, 0, \dots, 0)$, $i = 1, \dots, n$ und damit die Dimension $n$.

**3.6 Beispiel.** $K[x]$ hat die Standardbasis $(x^n)_{n\in\mathbb{N}}$ und ist somit unendlich-dimensional.

Der Vollständigkeit halber wollen wir den Steinitz'schen Austauschsatz für beliebige Basen verallgemeinern. Das wird später nicht benötigt.

**3.7 Theorem.** *Es sei $(b_i)_{i\in I}$ eine Basis und $(w_j)_{j\in J}$ ein Erzeugendensystem von $V$. Dann gibt es für jedes $T \subseteq I$ eine injektive Abbildung $\pi\colon T \to J$ so, dass $(w_j)_{j\in\pi(T)}, (b_i)_{i\in I\setminus T}$ eine Basis von $V$ ist.*

**Beweis:** Wir skizzieren den Beweis nur, die Einzelheiten überlassen wir dem Leser. Wir betrachten

$$X := \{\pi : S \to J \mid S \subseteq T, \pi \text{ injektiv und } (w_j)_{j\in\pi(S)}, (b_i)_{I\setminus S} \text{ Basis von } V\}.$$

Wir ordnen die Menge $X$ partiell, in dem wir für $(\pi : S \to J), (\pi : S' \to J) \in X$ definieren, dass $\pi \leq \pi'$ genau dann gelten soll, wenn $S \subseteq S'$ und $\pi'|_S = \pi$ gilt. Man zeigt dann mit dem Zornschen Lemma, dass $X$ ein maximales Element $\pi : S \to J$ besitzt. Mit dem Austauschlemma (Lemma 3.1) folgt dann, dass $S = T$ gelten muss. □

**3.8 Theorem.** *Jeder Vektorraum hat eine Basis $(b_i)_{i\in I}$. Für jede weitere Basis $(b'_j)_{j\in J}$ gibt es eine Bijektion $\pi\colon I \to J$.*

**Beweis:** Analog wie Theorem 3.3. □

**3.9.** Zwei Menge $I$ und $J$ heißen dann **gleichmächtig**, wenn es eine Bijektion $\pi\colon I \to J$ gibt.

**3.10 Theorem.** *Zwei Vektorräume sind genau dann isomorph, wenn sie gleichmächtige Basen haben.*

**Beweis:** „$\Longrightarrow$“ Sei $\varphi: V \to V'$ ein Isomorphismus von Vektorräumen und sei $(b_i)_{i\in I}$ eine Basis von $V$. Dann folgt sofort, dass $(\varphi(b_i))_{i\in I}$ eine Basis von $V$ ist und die ist offenbar gleichmächtig.

„$\Longleftarrow$“ Es seien $V, V'$ Vektorräume mit gleichmächtigen Basen $(b_i)_{i\in I}$ bzw. $(b'_j)_{j\in J}$. Also gibt es eine Bijektion $\pi: I \to J$ und nach Umnummerierung der Basis $(b'_j)_{j\in J}$ dürfen wir $I = J$ annehmen. Nun hat eine Basis folgende wichtige Eigenschaft: Man kann sich beliebige Basisbilder in einem anderen Vektorraum aussuchen und dann gibt es genau eine lineare Abbildung, die dies fortsetzt und die die vorgegebenen Basisbilder hat (siehe Proposition V1.1). Es gibt also genau eine lineare Abbildung $\varphi: V \to V'$ mit $\varphi(b_i) = b'_i \; \forall i \in I$. Analog gibt es genau eine lineare Abbildung $\psi: V' \to V$ mit $\psi(b'_i) = b_i \; \forall i \in I$. Dann folgt

$$\begin{aligned}(\varphi \circ \psi)(b'_i) &= \varphi(b_i) = b'_i \; \forall i \in I \\ (\psi \circ \varphi)(b_i) &= \psi(b'_i) = b_i \; \forall i \in I\end{aligned}$$

Wieder mit Proposition V.1.1 folgt $\varphi \circ \psi = \mathrm{id}_V$, $\psi \circ \varphi = \mathrm{id}_V$. Also ist $\varphi$ ein Isomorphismus von $V$ auf $V'$ wie gewünscht. □

**3.11 Proposition.** *Es sei $V$ ein Vektorraum der Dimension $n < \infty$. Dann bilden $n$ linear unabhängige Vektoren immer eine Basis von $V$.*

**Beweis:** Seien $b_1, \ldots, b_n$ linear unabhängig. Nach dem Basisergänzungssatz 2.5 können wir $b_1, \ldots, b_n$ zu einer Basis $b_1, \ldots, b_m$ ergänzen mit $m \geq n$. Nach Theorem 3.3 gilt $m = n$. □

**3.12 Korollar.** *Sei $W$ ein Unterraum von $V$. Dann gilt*

*a)* $\dim(W) \leq \dim(V)$

*b)* $\dim(W) = \dim(V) < \infty \Longrightarrow W = V$

**Beweis:** a) folgt unmittelbar aus dem Basisergänzungssatz. b) aus Proposition 3.11. □

**3.13.** Die folgenden Dimensionssätze sind eigentlich nur interessant für endlich-dimensionale Vektorräume. Wenn man aber die in der Analysis übliche Arithmetik $n + \infty = \infty$ für $n \in \mathbb{N}$ oder $n = \infty$ benutzt, dann stimmen diese Aussagen auch für unendlich-dimensionale Vektorräume. Man muss nur aufpassen, dass man keine Subtraktionen benutzt, weil $\infty - \infty$ nicht definiert ist.

**3.14 Proposition.** *Seien $V_1, \ldots V_r$ Vektorräume über dem Körper $K$. Dann gilt*

$$\dim(V_1 \times \ldots \times V_r) = \dim(V_1) + \ldots + \dim(V_r).$$

**Beweis:** Für $i \in \{1, \ldots, r\}$ sei $(b_j^{(i)})_{j\in J_i}$ eine Basis in $V_i$. Dann ist

$$((0, \ldots, 0, \underbrace{b_j^{(i)}}_{\text{i-te Komponente}}, 0 \ldots, 0))_{\substack{i\in\{1,\ldots,r\}, \\ j\in J_i}} \in V_1 \times \ldots \times V_r$$

eine Basis von $V_1 \times \ldots \times V_r$, wobei $i \in \{1, \ldots, r\}$ und $j \in J_i$. Das beweist man leicht durch komponentenweises rechnen. □

**3.15 Korollar.** *Seien $U_1, \ldots, U_r$ Unterräume von $V$ mit $V = U_1 \oplus \ldots \oplus U_r$. Dann gilt*

$$\dim(V) = \dim(U_1) + \ldots + \dim(U_r).$$

**Beweis:** Nach Aufgabe 8.2 ist $V$ isomorph zu $U_1 \times \ldots \times U_r$ und damit folgt die Behauptung aus Proposition 3.14. □

**3.16 Korollar.** *Sei $U$ ein Unterraum von $V$. Dann gilt*

$$\dim(V) = \dim(U) + \dim(V/U).$$

**Beweis:** Nach Korollar 2.10 gibt es einen Unterraum $W$ von $V$ so, dass $V = U \oplus W$. Nach Beispiel III.3.4 gilt $W = V/U$. Mit Korollar 3.15 folgt

$$\dim(V) = \dim(U) + \dim(W) = \dim(U) + \dim(V/U). \qquad \square$$

Nun können wir die **Dimensionsformel für lineare Abbildungen** beweisen:

**3.17 Korollar.** *Sei $\varphi\colon V_1 \to V_2$ lineare Abbildung von $K$-Vektorräumen. Dann gilt*

$$\dim(V_1) = \dim(\ker(\varphi)) + \dim(\varphi(V_1))$$

**Beweis:** Nach dem Isomorphiesatz III.3.6 gilt

$$V_1/\ker(\varphi) = \varphi(V_1)$$

und mit Korollar 3.16 folgt die Behauptung. $\square$

Kapitel V.

# 5 Matrizenrechnung

## 1. Lineare Abbildungen und Matrizen

Koordinaten dienen dazu, dass man mit Vektoren rechnen kann. Um eine lineare Abbildung mit Koordinaten zu beschreiben, benutzt man Matrizen. Dies wird der Inhalt dieses Abschnitts sein. Wir arbeiten wie immer mit einem beliebigen Grundkörper $K$.

**1.1 Proposition.** *Sei $b_1, \dots, b_n$ eine Basis des $K$-Vektorraums $V$ und $w_1, \dots, w_n$ beliebige Elemente des $K$-Vektorraums $W$. Dann gibt es genau eine $K$-lineare Abbildung $\varphi\colon V \to W$ mit $\varphi(b_i) = w_i$ für alle $i \in \{1, \dots, n\}$.*

**Beweis:** Zuerst zeigen wir die Existenz von $\varphi$. Sei $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ der Koordinatenvektor von $v \in V$, d.h. $v = \lambda_1 b_1 + \cdots + \lambda_n b_n$. Dann setzen wir

$$\varphi(v) := \lambda_1 w_1 + \cdots + \lambda_n w_n. \tag{V.1}$$

Wegen der Linearität der Koordinatenabbildung (siehe Proposition IV.2.15) folgt sofort die Linearität von $\varphi$. Für $i \in \{1, \dots, n\}$ hat $b_i$ den Koordinatenvektor $(0, \dots, 1, \dots, 0)$ und somit gilt $\varphi(b_i) = w_i$ wie gewünscht. Dies zeigt die Existenz von $\varphi$.

Wir zeigen noch die Eindeutigkeit von $\varphi$. Sei also $\varphi$ irgendeine lineare Abbildung $V \to W$ mit $\varphi(b_i) = w_i$ für alle $i \in \{1, \dots, n\}$. Dann gilt für $v = \lambda_1 b_1 + \cdots + \lambda_n b_n \in V$:

$$\varphi(v) = \lambda_1 \varphi(b_1) + \cdots + \lambda_n \varphi(b_n) = \lambda_1 w_1 + \cdots + \lambda_n w_n.$$

Also wird $\varphi$ durch (V.1) definiert und ist somit eindeutig. □

Proposition 1.1 gilt auch analog für einen unendlich dimensionalen Vektorraum $V$ und wird genauso bewiesen.

**1.2.** Es sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen $K$-Vektorräumen. Wir nehmen an, dass $V$ die Basis $b_1, \dots, b_n$ und $W$ die Basis $c_1, \dots, c_m$ hat. Wir wollen den Koordinatenvektor $\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} \in K^m$ von $\varphi(v)$ mit Hilfe des Koordinatenvektors $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in K^n$ von $v \in V$ ausdrücken. Nach Proposition 1.1 müssen wir dazu nur die Koordinatenvektoren $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m$ der Basisbilder $\varphi(b_j)$ kennen. Aus den Definitionen ergibt sich

$$\varphi(v) = \mu_1 c_1 + \cdots + \mu_m c_m, \tag{V.2}$$

und

$$\varphi(b_j) = a_{1j} c_1 + \cdots + a_{mj} c_m. \tag{V.3}$$

Es folgt

$$\begin{aligned} \varphi(v) &= \lambda_1 \varphi(b_1) + \cdots + \lambda_n \varphi(b_n) = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} a_{ij} c_i \\ &= \sum_{i=1}^{m} \Big( \sum_{j=1}^{n} \lambda_j a_{ij} \Big) c_i. \end{aligned} \tag{V.4}$$

Weil die Koordinaten bzgl. einer Basis eindeutig sind, folgt aus (V.2) und (V.4)

$$\mu_i = \sum_{j=1}^{n} a_{ij}\lambda_j. \tag{V.5}$$

Damit ist unser Ziel erreicht und das motiviert folgende Definitionen:

**1.3 Definition.** Eine $m \times n$ **Matrix** $A$ mit Einträgen $a_{ij}$ in einem Ring $R$ ist von der Form $A = (a_{ij})_{1 \le i \le m; 1 \le j \le n}$. Wir schreiben so eine Matrix in Rechteckform

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{array}{l} \xleftarrow{\text{Zeile } 1} \\ \xleftarrow{\text{Zeile } 2} \\ \\ \xleftarrow{\text{Zeile } m} \end{array}$$

$$\begin{array}{cccc} \uparrow & \uparrow & & \uparrow \\ \text{Spalte } 1 & \text{Spalte } 2 & & \text{Spalte } n \end{array}$$

Zwei $m \times n$ Matrizen sind genau dann gleich, wenn $a_{ij} = a'_{ij}$ für alle $i, j$. Die Menge der $m \times n$ Matrizen mit Einträgen in $R$ wird mit $M(m \times n, R)$ bezeichnet.

**1.4 Definition.** Wir definieren die Multiplikation von $A \in M(m \times n, R)$ mit $\lambda \in R^n$ durch

$$A \cdot \lambda := \Big( \sum_{j=1}^{n} a_{ij}\lambda_j \Big)_{i=1,\dots,m} \in R^m$$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} a_{11}\lambda_1 + \cdots + a_{1n}\lambda_n \\ \vdots \\ a_{m1}\lambda_1 + \cdots + a_{mn}\lambda_n \end{pmatrix}$$

**1.5.** Wenn $\varphi \colon V \to W$ eine lineare Abbildung von $K$-Vektorräumen ist und $b_1, \dots, b_n$ (bzw. $c_1, \dots, c_m$) Basis von $V$ (bzw. $W$) ist, dann definieren wir die **Darstellungsmatrix $\mathbf{A}(\boldsymbol{\varphi})$ zu $\boldsymbol{\varphi}$ bzgl. den Basen $\mathbf{b_1}, \dots, \mathbf{b_n}$ und $\mathbf{c_1}, \dots, \mathbf{c_m}$** als die $m \times n$-Matrix $A(\varphi) = (a_{ij})$ aus 1.2, d.h. $A(\varphi)$ als diejenige Matrix, deren Spaltenvektoren die Koordinatenvektoren von $\varphi(b_1), \dots, \varphi(b_n)$ sind.

Sei $\lambda \in K^n$ (bzw. $\mu \in K^m$) der Koordinatenvektor von $v \in V$ bzw. von $\varphi(v)$ bezüglich der gegebenen Basis, dann gilt

$$\mu = A(\varphi) \cdot \lambda. \tag{V.6}$$

Dies folgt aus (V.5). Beachte, dass $\lambda$ hier und im Folgenden ein Vektor aus $K^n$ ist und kein Skalar!

Durch (V.6) ist die $m \times n$ Matrix $A(\varphi)$ auch eindeutig charakterisiert, weil wir durch Einsetzen von $(0, \dots, 1, \dots, 0)$ die Spaltenvektoren von $A(\varphi)$ erhalten und die den Basisbildern entsprechen.

**1.6 Proposition.** *Es sei $b_1, \dots, b_n$ eine Basis von $V$ und $c_1, \dots, c_m$ eine Basis von $W$. Weiter bezeichneten wir mit $\mathrm{Hom}_K(V, W)$ die Menge aller linearen Abbildungen von $V$ nach $W$. Dann gibt es eine kanonische Bijektion zwischen $\mathrm{Hom}_K(V, W)$ und $M(m \times n, K)$, in dem wir jeder linearen Abbildung $\varphi \colon V \to W$ die zu $\varphi$ gehörige Matrix $A(\varphi)$ bzgl. den Basen $b_1, \dots, b_n$ und $c_1, \dots, c_m$ zuordnet.*

**Beweis:** Wahl einer Matrix $A \in M(m \times n, K) \iff$ Wahl von $n$-Spaltenvektoren $a_1, \dots, a_n \in K^m \iff$ Wahl der Bilder von $b_1, \dots, b_n \iff \varphi \in \mathrm{Hom}_K(V, W)$. Die letzte Äquivalenz folgt aus Proposition 1.1. □

**1.7 Definition.** Für $A = (a_{ij}), B = (b_{ij}) \in M(m \times n, R)$ definieren wir die **Matrixaddition** durch

$$A + B := (a_{ij} + b_{ij})_{1 \le i \le m;\, 1 \le j \le n} \in M(m \times n, R)$$

Wir definieren die **skalare Multiplikation** von $\alpha \in R$ mit $A \in M(m \times n, R)$ durch

$$\alpha \cdot A := (\alpha \cdot a_{ij})_{1 \le i \le m;\, 1 \le j \le n} \in M(m \times n, R)$$

Für $C = (c_{ij}) \in M(m \times n, R)$ und $D = (d_{jk}) \in M(n \times p, R)$ definieren wir die **Matrizenmultiplikation** durch

$$C \cdot D := \Big( \sum_{j=1}^{n} c_{ij} d_{jk} \Big)_{1 \le i \le m; 1 \le k \le p} \in M(m \times p, R)$$

Wir nennen $\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in M(m \times n, R)$ die **Nullmatrix** und $E_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in M(n \times n, R)$ die **Einheitsmatrix** der Länge $n$.

Es gelten die folgenden Rechenregeln für Matrizen mit Einträgen im Ring $R$.

**1.8 Proposition.** *Es sei $A, A' \in M(m \times n, R), B, B' \in M(n \times p, R), C \in M(p \times q, R)$ und $\alpha \in R$. Dann gelten*

*a)* $A + A' = A' + A$;

*b)* $A \cdot E_n = A = E_m \cdot A$;

*c)* $(\alpha \cdot A) \cdot B = \alpha \cdot (A \cdot B)$;
*Weiter gilt* $(\alpha \cdot A) \cdot B = A \cdot (\alpha \cdot B)$, *falls $R$ kommutativ ist.*

*d)* $(A + A') \cdot B = A \cdot B + A' \cdot B$; $\quad A \cdot (B + B') = A \cdot B + A \cdot B'$;

*e)* $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

**Beweis:** Aufgabe 10.1 □

**1.9 Korollar.** *Bzgl. der Matrixaddition und Matrixmultiplikation bildet $M(n \times n, R)$ einen Ring mit der Nullmatrix als neutrales Element der Addition und der Einheitsmatrix $E_n$ als neutrales Element der Multiplikation.*

**Beweis:** Folgt sofort aus Proposition 1.8. □

**1.10.** Für $n \ge 2$ ist $M(n \times n, R)$ nie kommutativ $\big($außer für $R = \{0\}\big)$. Zum Beispiel gilt

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

Es gibt immer nicht-triviale Nullteiler, z.B.

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

Ab jetzt betrachten wir wieder den Körper $K$ statt den Ring $R$.

**1.11 Proposition.** *Mit der Matrixaddition und der skalaren Multiplikation aus Definition 1.7 bildet $M(m \times n, K)$ einen $K$-Vektorraum der Dimension $m \cdot n$.*

**Beweis:** Die Abbildung

$$M(m\times n,K)\to K^{mn},\qquad \begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{m1} & \cdots & a_{mn}\end{pmatrix}\longmapsto \begin{pmatrix} a_{11}\\ \vdots\\ a_{m1}\\ \vdots\\ a_{1n}\\ \vdots\\ a_{mn}\end{pmatrix}$$

ist eine Bijektion, die $+$ und die skalare Multiplikation erhält. Also ist $M(m\times n,K)$ ein zu $K^{mn}$ isomorpher $K$-Vektorraum. Als Urbilder der Standardbasis von $K^{mn}$ erhalten wir die Basis

$$\begin{pmatrix} 1 & 0 & \cdots & 0\\ 0 & 0 & & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 0\end{pmatrix},\begin{pmatrix} 0 & 0 & \cdots & 0\\ 1 & 0 & & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 0\end{pmatrix},\dots,\begin{pmatrix} 0 & 0 & \cdots & 0\\ 0 & 0 & & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 1\end{pmatrix}$$

von $M(m\times n,K)$ und $\dim_K M(m\times n,K)=m\cdot n$. □

Wir haben in Proposition 3.8 gesehen, dass $\mathrm{Hom}_K(V,W):=\{\varphi:V\to W\mid \varphi \text{ linear}\}$ ein $K$-Vektorraum ist bezüglich der Addition

$$\varphi+\psi:V\to W,\quad v\mapsto\varphi(v)+\psi(v)$$

und der skalaren Multiplikation

$$\lambda\varphi:V\to W,\quad v\mapsto\lambda\varphi(v)$$

für $\lambda\in K$ und $\varphi,\psi\in\mathrm{Hom}_K(V,W)$.

**1.12 Proposition.** *Es sei $b_1,\dots,b_n$ eine Basis des $K$-Vektorraums $V$ und $c_1,\dots,c_m$ eine Basis des $K$-Vektorraums $W$. Für die zu $\varphi,\psi\in\mathrm{Hom}_K(V,W)$ gehören Matrizen $A(\varphi),A(\psi)\in M(m\times n,K)$ (vgl. 1.5) und für alle $\alpha\in K$ gilt*

$$A(\varphi+\psi)=A(\varphi)+A(\psi),\quad A(\alpha\cdot\varphi)=\alpha\cdot A(\varphi)$$

**Beweis:** Einfache Übung. □

**1.13 Proposition.** *Unter den Voraussetzungen von Proposition 1.12 betrachten wir einen zusätzlichen Vektorraum $X$ mit Basis $d_1,\dots,d_l$. Für $\varphi\in\mathrm{Hom}_K(V,W)$, $\psi\in\mathrm{Hom}_K(W,X)$ mit den zugehörigen Matrizen $A(\varphi),A(\psi)$ gilt*

$$A(\psi\circ\varphi)=A(\psi)\cdot A(\varphi)$$

**Beweis:** $A(\varphi)=(a_{jk})$, $A(\psi)=(a'_{ij})$. Für das Bild von $b_k$

$$(\psi\circ\varphi)(b_k)\overset{1.6}{=}\psi\Big(\sum_{j=1}^m a_{jk}c_j\Big)=\sum_{j=1}^m a_{jk}\,\psi(c_j)=\sum_{j=1}^m a_{jk}\sum_{i=1}^l a'_{ij}d_i\overset{II.3.2}{=}\sum_{i=1}^l\sum_{j=1}^m(a'_{ij}a_{jk})d_i$$

Also ist der Koordinatenvektor von $(\psi\circ\varphi)(b_k)$ gleich dem $k$-ten Spaltenvektor von $A(\varphi)\cdot A(\psi)$. □

**1.14 Bemerkung.** Man kann die Wahl einer Basis für einen endlich dimensionalen Vektorraum vergleichen mit der Wahl einer Sprache. Damit werden Vektoren aus $V$ in Vektoren aus $K^n$ und linearen Abbildungen in Matrizen übersetzt. Dabei gehen $+$ und $\circ$ von linearen Abbildungen über in die Matrizenaddition und die Matrizenmultiplikation, was eine natürliche Erklärung liefert für deren Definitionen. Die Übersetzungen haben den Vorteil, dass man aus abstrakten Vektoren und linearen Abbildungen Objekte erhält, mit denen man rechnen kann.

Im Folgenden erklären wir die Rückübersetzung im Fall $V=K^n$. Auf diesen Fall kann man sich durch die Wahl von Basen immer beschränken (Proposition 2.15).

**1.15.** Eine Matrix $A \in M(m \times n, K)$ induziert eine Abbildung

$$\varphi_A : K^n \to K^m,\ \lambda \longmapsto A \cdot \lambda$$

Zur Erinnerung

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} a_{11}\lambda_1 + \cdots + a_{1n}\lambda_n \\ \vdots \\ a_{m1}\lambda_1 + \cdots + a_{mn}\lambda_n \end{pmatrix}$$

Wir können $\lambda \in K^n$ auch als $n \times 1$ Matrix auffassen, dann ist $A \cdot \lambda$ dasselbe wie die Matrixmultiplikation. Aus den Regeln der Matrizenrechnung (siehe Proposition 1.8) folgt sofort, dass $\varphi_A$ eine lineare Abbildung ist. Wenn wir für $K^n$ und $K^m$ die Standardbasis wählen, dann sieht man leicht, dass $A$ die zu $\varphi_A$ assoziierten Matrix ist, wie in 1.5 definiert, d.h.

$$A(\varphi_A) = A$$

Mit Proposition 1.8 folgt

$$\varphi_{A+A'} = \varphi_A + \varphi_{A'} \quad \wedge \quad \varphi_{A \cdot B} = \varphi_A \circ \varphi_B.$$

Umgekehrt zeigt man leicht, dass für jede lineare Abbildung $\varphi : K^n \to K^m$ mit Darstellungsmatrix $A := A(\varphi)$ bzgl. den Standardbasen folgendes gilt:

$$\varphi_A = \varphi.$$

Wir hatten mit $\mathrm{Hom}_K(V,W)$ die Menge der linearen Abbildungen zwischen den $K$-Vektorräumen $V$ und $W$ bezeichnet. Bzgl. der punktweise definierten Addition und der skalaren Multiplikation wird $\mathrm{Hom}_K(V,W)$ zu einem $K$-Vektorraum.

Im Spezialfall $W = K$ heißt $V^* := \mathrm{Hom}_K(V,K)$ der Dualraum zu $V$.

**1.16 Korollar.** *Sei $V$ ein $n$-dimensionaler $K$-Vektorraum und $W$ ein $m$-dimensionaler $K$-Vektorraum. Dann ist* $\mathrm{Hom}_K(V,W)$ *ein $n \cdot m$ dimensionaler $K$-Vektorraum. Insbesondere gilt* $\dim(V^*) = \dim(V)$.

**Beweis:** Wir wählen feste Basen $b_1, \ldots, b_n$ von $V$ und $c_1, \ldots, c_m$ von $W$. Dann ist die bijektive Abbildung

$$\mathrm{Hom}_K(V,W) \to M(m \times n, K),\ \varphi \mapsto A(\varphi)$$

aus Proposition 1.6 ein Isomorphismus von $K$-Vektorräumen (siehe Proposition 1.12). Die erste Behauptung folgt nun aus $n \cdot m = \dim M(m \times n, K)$ (siehe Proposition 1.11). Die 2. Behauptung folgt aus dem Spezialfall $W = K$ in der ersten Behauptung. □

**1.17.** Wir erinnern an die Bilinearform

$$V \times V^* \to K,\ (v, f) \mapsto \langle v, f \rangle := f(v)$$

aus Proposition III.3.11. Zu einer Basis $b_1, \ldots, b_n$ von $V$ gibt es genau eine Basis $b_1^*, \ldots, b_n^*$ von $V^*$ mit

$$\langle b_i, b_j^* \rangle = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases} \tag{V.7}$$

Wir nennen $b_1^*, \ldots, b_n^*$ die **duale Basis** zu $b_1, \ldots, b_n$.

Die Existenz und Eindeutigkeit von $b_1^*, \ldots, b_n^* \in V^*$ mit (V.7) folgt aus Proposition 1.1, d.h. lineare Abbildungen sind durch die Basisbilder eindeutig bestimmt und man kann die Basisbilder frei wählen. Wir bezeichnen mit $V^{**}$ den Bidualraum von $V$, d.h. $V^{**} := (V^*)^*$.

**1.18 Theorem.** *Es sei $V$ ein endlich-dimensionaler $K$-Vektorraum. Dann ist die kanonische Abbildung $V \to V^{**}$, $v \mapsto \langle v, \cdot \rangle$ ein Isomorphismus von $K$-Vektorräumen.*

**Beweis:** In Proposition III.3.13 haben wir gesehen, dass diese Abbildung $\Phi$ linear und injektiv ist. Wegen

$$\dim\big(\Phi(V)\big) \overset{IV.3.17}{=} \dim(V) - \dim\ker(\Phi) = \dim(V) \overset{1.16}{=} \dim(V^*) \overset{1.16}{=} \dim(V^{**})$$

ist $\Phi$ surjektiv und damit ein Isomorphismus. $\square$

**1.19.** Für $A = (a_{ij})_{1\le i\le m;\, 1\le j\le n} \in M(m\times n, R)$ mit Einträgen im Ring $R$ definieren wir die **Transponierte** $A^t \in M(n\times m, R)$ durch

$$A := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \Longrightarrow A^t := \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & \ddots & & a_{m2} \\ \vdots & & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nm} \end{pmatrix}$$

Für $A, B \in M(m\times n, R)$ und $C \in m(n\times p, R)$, $\alpha \in R$ gilt

a) $(A+B)^t = A^t + B^t$

b) $(\alpha A)^t = \alpha A^t$

c) Falls $R$ kommutativ, dann $(A\cdot C)^t = C^t \cdot A^t$.

**Beweis:** Aufgabe 10.1 $\square$

**1.20 Lemma.** *Sei $b_1,\dots,b_n$ eine Basis von $V$ und $b_1^*,\dots,b_n^*$ die duale Basis von $V^*$. Für $v\in V$ ist dann $\langle v, b_i^*\rangle = b_i^*(v)$ die $i$-te Koordinate von $v$ bzgl. der Basis $b_1,\dots,b_n$. Für $f\in V^*$ ist $\langle b_k, f\rangle = f(b_k)$ ebenso die $k$-te Koordinate von $f$ bzgl. der Basis $b_1^*,\dots,b_n^*$.*

**Beweis:** Sie $\lambda_i$ die $i$-te Koordinate von $v$ bzgl. $b_1,\dots,b_n$, d.h. $v = \sum_{j=1}^n \lambda_j b_j$

$$\langle v, b_i^*\rangle = b_i^*(v) = b_i^*\Big(\sum_{j=1}^n \lambda_j b_j\Big) = \sum_{j=1}^n \lambda_j b_i^*(b_j) = \sum_{j=1}^n \lambda_j \langle b_j, b_i^*\rangle = \lambda_i.$$

Dies zeigt die erste Behauptung und die zweite folgt analog. $\square$

**1.21 Proposition.** *Sei $\varphi\colon V\to W$ eine lineare Abbildung von $K$-Vektorräumen und seien $b_1,\dots,b_n$ und $c_1,\dots,c_m$ Basen von $V$ bzw. $W$. Auf den Dualräumen $V^*$ und $W^*$ betrachten wir dazu die dualen Basen. Es sei $A(\varphi)$ die zu $\varphi$ assoziierte Matrix und $A(\varphi^*)$ die zu der dualen Abbildung $\varphi^*\colon W^*\to V^*$ assoziierte Matrix. Dann gilt*

$$A(\varphi^*) = A(\varphi)^t.$$

**Beweis:** Wir erinnern daran, dass die Spaltenvektoren von $A(\varphi^*)$ gleich den Koordinatenvektoren von $\varphi^*(c_1^*),\dots,\varphi^*(c_m^*)$ bzgl. der Basis $b_1^*,\dots,b_n^*$ ist. Es gilt

$$\langle b_i, \varphi^*(c_k^*)\rangle = \langle \varphi(b_i), c_k^*\rangle$$

nach Definition der dualen Abbildung. Es sei $A(\varphi) = (a_{ij})$. Dann gilt

$$\varphi(b_i) = \sum_{j=1}^m a_{ji} c_j$$

nach 1.5. Also folgt

$$\langle b_i, \varphi^*(c_k^*)\rangle = \langle \sum_{j=1}^m a_{ji} c_j, c_k^*\rangle = \sum_{j=1}^m a_{ji}\langle c_j, c_k^*\rangle = a_{ki}.$$

Nach Lemma 1.20 ist also die $i$-te Koordinate von $\varphi^*(c_k^*)$ gleich $a_{ki}$. Es folgt, dass der $k$-te Spaltenvektor von $A(\varphi^*)$ gleich dem $k$-ten Zeilenvektor von $A(\varphi)$ ist d.h. $A(\varphi^*) = A(\varphi)^t$. $\square$

## 2. Rang und Invertierbarkeit

Wir haben im letzten Abschnitt gesehen, dass $M(n \times n, K)$ ein Ring ist. Die Matrizen, die bzgl. der Multiplikation invertierbar sind, spielen eine Sonderrolle. Der Rang einer Matrix ist eine Größe, die beschreibt, wie stark ausgeartet eine Matrix ist. Wie immer bezeichnet $K$ eine Körper.

**2.1 Definition.** Eine Matrix $A \in M(n \times n, K)$ heißt **invertierbar** $:\iff \exists A^{-1} \in M(n \times n, K)$ mit $A \cdot A^{-1} = A^{-1} \cdot A = E_n$. Die Menge der invertierbaren $n \times n$-Matrizen wird mit $\mathrm{GL}(n, K)$ bezeichnet und heißt **allgemeine lineare Gruppe**.

**2.2 Proposition.** *$\mathrm{GL}(n, K)$ ist eine Gruppe bzgl. der Matrixmultiplikation.*

**Beweis:** Da $M(n \times n, K)$ ein Ring ist (Proposition 1.11) und da $\mathrm{GL}(n, K) = M(n \times n, K)^*$ (invertierbare Elemente bzg. Multiplikation), folgt die Behauptung aus II.2.7. □

**2.3.** Für $A \in M(m \times n, K)$ definieren wir den **Rang** als die Dimension des von den Spaltenvektoren erzeugten Unterraum in $K^m$. Wir bezeichnen ihn mit $\mathrm{Rang}(A)$. Den **Rang** einer linearen Abbildung $\varphi \colon V \to W$ definiert man als $\mathrm{Rang}(\varphi) := \dim \varphi(V)$.

Wir betrachten im Folgenden die lineare Abbildung $\varphi_A \colon K^n \to K^m$, $\lambda \mapsto A \cdot \lambda$. Da die Spaltenvektoren von $A$ gleich den Bildern der Standardbasis von $K^n$ sind, erzeugen sie das Bild von $\varphi_A$, d.h.

$$\boxed{\mathrm{Rang}(A) = \dim \varphi_A(K^n)}.$$

Wir definieren den **Zeilenrang** von $A$ als die Dimension des von den Zeilenvektoren erzeugten Unterraums von $K^n$. Offenbar ist der Zeilenrang gleich

$$\boxed{\mathrm{Rang}(A^t) = \dim \varphi_{A^t}(K^m)}.$$

**2.4 Beispiel.** $\mathrm{Rang}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$, $\mathrm{Rang}\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} = 1$, $\mathrm{Rang}\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} = 1$, $\mathrm{Rang}\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} = 2$. Aus der folgenden Proposition folgt, dass nur die letzte Matrix invertierbar ist. Später lernen wir die Inversen zu berechnen. Damit findet man leicht $\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix}$.

**2.5 Proposition.** *Für $A \in M(n \times n, K)$ sind folgende Aussagen äquivalent:*

i) *$A$ ist invertierbar;*

ii) *$\exists B \in M(n \times n, K)$ mit $A \cdot B = E_n$;*

iii) *$\exists C \in M(n \times n, K)$ mit $C \cdot A = E_n$;*

iv) *die lineare Abbildung. $\varphi_A \colon K^n \to K^n, \lambda \mapsto A \cdot \lambda$ ist ein Isomorphismus;*

v) *die lineare Abbildung. $\varphi_A \colon K^n \to K^n, \lambda \mapsto A \cdot \lambda$ ist surjektiv;*

vi) *die lineare Abbildung. $\varphi_A \colon K^n \to K^n, \lambda \mapsto A \cdot \lambda$ ist injektiv;*

vii) *$\mathrm{Rang}(A) = n$;*

viii) *der Zeilenrang von $A$ ist $n$;*

ix) *$A^t$ ist invertierbar.*

**Beweis:** Für einen endlich dimensionalen $K$-Vektorraum $V$ und für eine lineare Selbstabbildung $\varphi \colon V \to V$ sind folgende Aussagen äquivalent:

a) $\varphi$ Isomorphismus;

b) $\varphi$ injektiv;

c) $\varphi$ surjektiv.

Um diesen Schritt zu beweisen, benutzen wir die Dimensionsformel für lineare Abbildung aus Korollar IV.3.17

$$\dim(V) = \dim\ker(\varphi) + \dim\varphi(V).$$

Weil $\varphi$ eine Selbstabbildung ist, folgt sofort

$$\dim\ker(\varphi) = 0 \Longleftrightarrow \dim\varphi(V) = \dim V$$

und somit

$$\varphi \text{ injektiv } \Longleftrightarrow \varphi \text{ surjektiv}.$$

Also ist dies äquivalent zu $\varphi$ bijektiv und dies zeigt a) $\Longleftrightarrow$ b) $\Longleftrightarrow$ c). Aus dieser Äquivalenz folgt, dass iv) $\Longleftrightarrow$ v) $\Longleftrightarrow$ vi). Weiter gilt i) $\Longrightarrow$ ii) und i) $\Longrightarrow$ iii). In Übung 11.1 sehen wir ii) $\Longleftrightarrow$ v) und iii) $\Longleftrightarrow$ vi). Folgende Implikationen haben wir schon eingesehen (und einiges mehr):

$$\begin{tikzcd}
 & ii) \arrow[r, Rightarrow, "\text{Übung}"'] & v) \arrow[dr, Rightarrow] & \\
i) \arrow[ur, Rightarrow] \arrow[dr, Rightarrow] & & & iv) \\
 & iii) \arrow[r, Rightarrow, "\text{Übung}"] & vi) \arrow[ur, Rightarrow] &
\end{tikzcd}$$

Wir zeigen als nächstes, dass iv) $\Longrightarrow$ i) gilt, und damit sind dann i) bis vi) äquivalent. Wir nehmen also an, dass $\psi := \varphi_A$ ein Isomorphismus ist. Wir bezeichnen die Inverse von $\psi$ mit $\psi' \colon V \to V$. Wir betrachten die zugehörigen Darstellungsmatrizen $A(\psi), A(\psi')$ von $\psi, \psi'$ bzgl. der Standardbasis von $K^n$. Insbesondere gilt

$$A(\psi) = A(\varphi_A) = A.$$

Es folgt mit Proposition 1.13

$$E_n = A(\mathrm{id}) = A(\psi \circ \psi') = A(\psi) \cdot A(\psi')$$

und analog

$$E_n = A(\psi') \cdot A(\psi)$$

Somit ist $A$ invertierbar, d.h. es gilt i).

In 2.3 haben wir

$$\operatorname{Rang} A = \dim \varphi_A(K^n)$$

gesehen. Damit folgt v) $\Longleftrightarrow$vii). Also sind i) bis vii) äquivalent. Weiter gilt für $C, D \in M(n \times n, K)$

$$C \cdot D = E_n \underset{1.19\ c)}{\Longleftrightarrow} D^t \cdot C^t = E_n$$

Dies zeigt i) $\Longleftrightarrow$ix). Weiter folgt $(A^t)^{-1} = (A^{-1})^t$ falls $A$ invertierbar.

Da der Zeilenrang von $A$ gleich dem (Spalten-)Rang von $A^t$ ist, folgt die Äquivalenz von viii) und xi), wenn man i) $\Longleftrightarrow$vii) für $A^t$ statt $A$ benutzt. Damit sind i) bis xi) äquivalent. $\square$

**2.6 Definition.** Zwei Matrizen $A, B \in M(m \times n, K)$ heißen **äquivalent** $:\Longleftrightarrow \exists S \in \mathrm{GL}(m, K), \exists T \in \mathrm{GL}(n, K)$ mit $B = S \cdot A \cdot T^{-1}$.

Zwei Matrizen $A, B \in M(n \times n, K)$ heißen **ähnlich** $:\Longleftrightarrow \exists S \in \mathrm{GL}(n, K)$ mit $B = S \cdot A \cdot S^{-1}$. Ähnliche Matrizen sind äquivalent, aber nicht umgekehrt.

**2.7 Proposition.** *Die Äquivalenz und die Ähnlichkeit von Matrizen sind Äquivalenzrelationen.*

**Beweis:** Übung 11.2 □

**2.8 Proposition.** *Der Rang von äquivalenten Matrizen ist gleich. Dasselbe gilt für den Zeilenrang.*

**Beweis:** Es seien $A, B$ äquivalente $m \times n$ Matrizen, d.h. $\exists S \in \mathrm{GL}(m,K), T \in \mathrm{GL}(n,K)$ mit

$$B = S \cdot A \cdot T^{-1}.$$

Es seien $\varphi_A, \varphi_B, \varphi_S, \varphi_{T^{-1}}$ die zugehörigen linearen Abbildungen, dann folgt

$$\varphi_B = \varphi_S \circ \varphi_A \circ (\varphi_T)^{-1}.$$

Beachte, dass nach Proposition 2.5 $\varphi_S$ und $\varphi_{T^{-1}}$ Isomorphismen sind.

Es folgt

$$\mathrm{Rang}(B) = \dim \varphi_B(K^n) = \dim \left(\varphi_S \circ \varphi_A \circ \varphi_T^{-1}(K^n)\right)$$

Weil $\varphi_T^{-1}$ ein Isomorphismus ist, muss $\varphi_T^{-1}(K^n) = K^n$ gelten. Da $\varphi_S$ ein Isomorphismus ist und isomorphe Vektorräume dieselbe Dimension haben, folgt

$$\mathrm{Rang}\, B = \dim \varphi_S \circ \varphi_A(K^n) = \dim \varphi_A(K^n) = \mathrm{Rang}\, A$$

Dies zeigt die erste Behauptung. Weiter folgt

$$B^t = (SAT^{-1})^t = (T^{-1})^t \cdot (S \cdot A)^t = (T^{-1})^t \cdot A^t \cdot S^t$$

aus Proposition 1.19 c). Wir haben in Beweis von Proposition 2.5 gesehen, dass

$$(T^{-1})^t = (T^t)^{-1}$$

gilt. Es folgt

$$T^t \cdot B^t \cdot (S^t)^{-1} = A^t$$

und somit sind $B^t$ und $A^t$ äquivalent. Die zweite Behauptung folgt nun aus der ersten Behauptung, wenn man $B^t$ und $A^t$ betrachtet. □

**2.9 Beispiel.** Wir zeigen exemplarisch, wie man die Inverse einer Matrix berechnen kann. Gegeben sei die reelle Matrix $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \\ -1 & 0 & 2 \end{pmatrix}$. Wir wollen entscheiden, ob $A$ invertierbar ist und wenn ja, wollen wir die Inverse berechnen. Nach Proposition 2.5 müssen wir dazu $B \in M(3 \times 3, \mathbb{R})$ bestimmen mit

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

Wenn wir so ein $B$ gefunden haben, ist es auch die Inverse von $A$, wie wir durch Multiplikation der obigen Gleichung mit $A^{-1}$ einsieht. Wir betrachten die Einträge von $B$ als Unbekannte $\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix}$. Dann müssen wir die folgenden drei linearen Gleichungssysteme lösen:

$$A \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad A \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad A \cdot \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Dabei betrachten wir den Gauß-Algorithmus aus I.4. Weil die linken Seiten bei allen drei Gleichungssystemen gleich ist, können wir den Gauß-Algorithmus für alle drei Systeme gleichzeitig durchführen. Es empfiehlt sich dabei nicht nur die Zeilenstufenform, sondern sogar die Einheitsmatrix auf der linken Seite anzustreben.

$$\left(\begin{array}{rrr|rrr} \boxed{1} & 2 & 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 & 1 & 0 \\ -1 & 0 & 2 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \mid\cdot(-3) \\ \\ \\ \end{array}$$

$$\rightsquigarrow \left(\begin{array}{rrr|rrr} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -6 & 1 & -3 & 1 & 0 \\ 0 & 2 & 2 & 1 & 0 & 0 \end{array}\right)$$

Vertauschung von Zeilen ist auch eine Äquivalenzumformung

$$\rightsquigarrow \left(\begin{array}{rrr|rrr} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & \boxed{2} & 2 & 1 & 0 & 1 \\ 0 & -6 & 1 & -3 & 1 & 0 \end{array}\right) \begin{array}{l} \\ \mid\cdot(-1) \quad \mid\cdot 3 \\ \\ \end{array}$$

$$\rightsquigarrow \left(\begin{array}{rrr|rrr} 1 & 0 & -2 & 0 & 0 & -1 \\ 0 & 2 & 2 & 1 & 0 & 1 \\ 0 & 0 & 7 & 0 & 1 & 3 \end{array}\right) \begin{array}{l} \\ \mid\cdot\frac{1}{2} \\ \mid\cdot\frac{1}{7} \end{array}$$

$$\rightsquigarrow \left(\begin{array}{rrr|rrr} 1 & 0 & -2 & 0 & 0 & -1 \\ 0 & 1 & 1 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \boxed{1} & 0 & \frac{1}{7} & \frac{3}{7} \end{array}\right) \begin{array}{l} \\ \\ \mid\cdot 2 \quad \mid\cdot(-1) \end{array}$$

$$\rightsquigarrow \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 0 & \frac{2}{7} & -\frac{1}{7} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{7} & \frac{1}{14} \\ 0 & 0 & 1 & 0 & \frac{1}{7} & \frac{3}{7} \end{array}\right)$$

Offenbar ist $x_1 = 0, x_2 = \frac{1}{2}, x_3 = 0$ die Lösung von $A \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Analog für die anderen Gleichungssysteme. Also ist $\begin{pmatrix} 0 & 2/7 & -1/7 \\ 1/2 & -1/7 & 1/14 \\ 0 & 1/7 & 3/7 \end{pmatrix}$ die Inverse von $A$. Hätte $A$ keine Inverse, dann hätte man die linke Seite nicht auf $E_3$, sondern nur auf $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ oder $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ bringen können.

## 3. Zusammenhang zu linearen Gleichungssystemen

In diesem Abschnitt werden wir sehen, dass die Theorie der linearen Abbildungen und Matrizen Anwendungen für linearen Gleichungssysteme über einem beliebigen Körper $K$ hat.

**3.1.** Wie wir in Beispiel 2.9 gesehen haben, eignet sich die Matrizenschreibweise hervorragend um ein lineares Gleichungssystem kompakt darzustellen als

$$A \cdot x = b,$$

wobei $A \in M(m \times n, K)$, $b \in K^m$ und wir den unbekannten Vektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ in $K^n$ suchen. Wie früher können wir dazu die lineare Abbildung

$$\varphi_A \colon K^n \to K^m, \quad \lambda \mapsto A \cdot \lambda$$

betrachten.

**3.2 Theorem.** *Der Lösungsraum $\mathbb{L}$ des homogenen linearen Gleichungssystems $A \cdot x = 0$ ist ein Unterraum von $K^n$ mit*

$$\dim \mathbb{L} = n - \operatorname{Rang}(A).$$

**Beweis:** Offenbar gilt $\mathbb{L} = \ker(\varphi_A)$. Insbesondere ist damit $\mathbb{L}$ ein Unterraum (siehe Proposition III.2.7). Weiter gilt nach 2.3

$$\operatorname{Rang}(A) = \dim \varphi_A(K^n)$$

Wir benutzen die Dimensionsformel für lineare Abbildungen und Korollar IV.3.17 um

$$n = \dim K^n = \dim \underbrace{\ker(\varphi_A)}_{=\mathbb{L}} + \underbrace{\dim \varphi_A(K^n)}_{=\operatorname{Rang}(A)}$$

zu erhalten. □ □

**3.3 Korollar.** *Für ein homogenes lineare Gleichungssystem $A \cdot x = 0$ mit $A \in M(m \times n, K)$ sind folgende Aussagen äquivalent:*

*a) Es gibt nur die triviale Lösung $x = 0 \in K^n$*

*b)* $\operatorname{Rang}(A) = n$

**Beweis:** a) $\Longleftrightarrow \mathbb{L} = \{0\} \Longleftrightarrow \dim \mathbb{L} = 0 \Longleftrightarrow n = \operatorname{Rang}(A)$ □

**3.4 Proposition.** *Für das lineare Gleichungssystem $A \cdot x = b$, $A \in M(m \times n, K)$, $b \in K^m$, $x \in K^n$ sind folgende Aussagen äquivalent:*

*a) $A \cdot x = b$ hat eine Lösung*

*b) Der Rang der durch die Spalte $b$ erweiterten Matrix $(A, b)$ ist gleich dem Rang von $A$*

*c) $b$ ist im Bild von $\varphi_A : K^n \to K^m$, $\lambda \mapsto A \cdot \lambda$*

**Beweis:** a) $\Longleftrightarrow$ c) ist trivial.

c) $\Longrightarrow$ b) Die Spalten $a_1, \ldots, a_n$ von $A$ erzeugen das Bild von $\varphi_A$. Analog erzeugen die Spalten $a_1, \ldots, a_n, b$ von $(A, b)$ das Bild von $\varphi_{(A,b)}$. Es folgt

$$\operatorname{Bild}(\varphi_{(A,b)}) = \langle a_1, \ldots, a_n, b \rangle \underset{b \in \operatorname{Bild}(\varphi_A)}{=} \langle a_1, \ldots, a_n \rangle = \operatorname{Bild}(\varphi_A).$$

Da die Dimension von $\operatorname{Bild}(\varphi_A)$ gleich dem Rang von $A$ ist (siehe 2.3) und analog für $(A, b)$, folgt

$$\operatorname{Rang}(A, b) = \dim(\operatorname{Bild} \varphi_{(A,b)}) = \operatorname{Rang}(A).$$

Also impliziert c) die Aussage b).

b) $\Longrightarrow$ c) Es gilt

$$\operatorname{Bild}(\varphi_A) = \langle a_1, \ldots, a_n \rangle \subseteq \langle a_1, \ldots, a_n, b \rangle = \operatorname{Bild}(\varphi_{(A,b)}).$$

Nach b) gilt

$$\dim \operatorname{Bild}(\varphi_A) \underset{2.3}{=} \operatorname{Rang}(A) \underset{b)}{=} \operatorname{Rang}(A, b) \underset{2.3}{=} \dim\big(\operatorname{Bild}(\varphi_{(A,b)})\big).$$

Also hat der Unterraum $\operatorname{Bild}(\varphi_A)$ von $\operatorname{Bild}(\varphi_{(A,b)})$ dieselbe Dimension und damit sind die Unterräume gleich $\big($Korollar IV.3.12 b)$\big)$. Da $b \in \operatorname{Bild}(\varphi_{(A,b)})$, folgt auch $b \in \operatorname{Bild}(\varphi_A)$ und damit c). □

**3.5 Korollar.** *Für $A \in M(m \times n, K)$ sind folgende Bedingungen äquivalent.*

*a) Für jedes $b \in K^m$ hat $A \cdot x = b$ eine Lösung.*

*b)* $\operatorname{Rang}(A) = m$.

*c) $\varphi_A$ ist surjektiv.*

**Beweis:** a) $\Longleftrightarrow$ c) folgt aus Proposition 3.4. Nun ist

$$\operatorname{Rang}(A) \underset{2.3}{=} \dim \varphi_A(K^n)$$

und somit

$$\varphi_A \text{ surjektiv } \Longleftrightarrow \varphi_A(K^n) = K^m \underset{IV.3.12}{\Longleftrightarrow} \dim \varphi_A(K^n) = m \Longleftrightarrow \operatorname{Rang}(A) = m.$$

Dies beweist b) $\Longleftrightarrow$ c). $\square$

**3.6 Korollar.** *Folgende Bedingungen sind äquivalent für $M(n \times n, K)$:*

*a) $A$ ist invertierbar.*

*b) Für jedes $b \in K^n$ hat $A \cdot x = b$ eine Lösung.*

*c) Für jedes $b \in K^n$ hat $A \cdot x = b$ genau eine Lösung.*

*Wenn c) erfüllt ist, gilt $x = A^{-1} \cdot b$ für die eindeutige Lösung.*

**Beweis:** a) $\overset{2.5}{\Longleftrightarrow}$ $\varphi_A$ surjektiv $\overset{\text{Kor. } 3.5}{\Longleftrightarrow}$ b).
a) $\Longrightarrow$ c): $A \cdot x = b \overset{\cdot A^{-1}}{\Longleftrightarrow} x = A^{-1} \cdot b$
c) $\Longrightarrow$ b): trivial. $\square$

**3.7 Beispiel.** Wir zeigen mit dem Gauß-Algorithmus, wie man eine Basis des Lösungsraums eines homogenen lineare Gleichungssystems findet. Hier sei $K = \mathbb{R}$.

$$\begin{array}{rrrrrl} & x_2 & -x_3 & +x_4 & +x_5 & = 0 \\ 2x_1 & -x_2 & +x_3 & +x_4 & -x_5 & = 0 \\ 4x_1 & +x_2 & -x_3 & +x_4 & +x_5 & = 0 \\ 2x_1 & +x_2 & -x_3 & +x_4 & +x_5 & = 0 \end{array}$$

In Matrixform:

$$\begin{pmatrix} 0 & 1 & -1 & 1 & 1 \\ 2 & -1 & 1 & 1 & -1 \\ 4 & 1 & -1 & 1 & 1 \\ 2 & 1 & -1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Wir transformieren nur die Matrix, da die rechte Seite bei homogenen linearen Gleichungssysteme sich nicht verändert:

$$\begin{pmatrix} \boxed{2} & -1 & 1 & 1 & -1 \\ 0 & 1 & -1 & 1 & 1 \\ 4 & 1 & -1 & 1 & 1 \\ 2 & 1 & -1 & 1 & 1 \end{pmatrix} \begin{array}{l} \mid \cdot(-2) \quad \mid \cdot(-1) \\ \\ \longleftarrow + \\ \longleftarrow + \end{array} \rightsquigarrow \begin{pmatrix} 2 & -1 & 1 & 1 & -1 \\ 0 & 1 & -1 & 1 & 1 \\ 0 & 3 & -3 & -1 & 3 \\ 0 & 2 & -2 & 0 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 2 & -1 & 1 & 1 & -1 \\ 0 & \boxed{1} & -1 & 1 & 1 \\ 0 & 3 & -3 & -1 & 3 \\ 0 & 2 & -2 & 0 & 2 \end{pmatrix} \begin{array}{l} \\ \mid \cdot(-3) \quad \mid \cdot(-2) \\ \longleftarrow + \\ \longleftarrow + \end{array} \rightsquigarrow \begin{pmatrix} 2 & -1 & 1 & 1 & -1 \\ 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & -2 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 2 & -1 & 1 & 1 & -1 \\ 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & \boxed{-4} & 0 \\ 0 & 0 & 0 & -2 & 0 \end{pmatrix} \begin{array}{l} \\ \\ \mid \cdot\left(-\frac{1}{2}\right) \\ \longleftarrow + \end{array} \rightsquigarrow \begin{pmatrix} 2 & -1 & 1 & 1 & -1 \\ 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Rückwärts auflösen ergibt:

$$-4x_4 = 0 \Longrightarrow x_4 = 0$$
$$x_2 - x_3 + x_4 + x_5 = 0 \Longrightarrow x_2 = x_3 - x_5$$
$$2x_1 - x_2 + x_3 + x_4 - x_5 = 0 \Longrightarrow x_1 = \frac{1}{2}(x_2 - x_3 - x_4 + x_5) = 0$$

Für den Lösungsraum $\mathbb{L}$ ergibt sich mit der Standardbasis $e_1, \ldots, e_5$:

$$\mathbb{L} = \{ \begin{pmatrix} 0 \\ \lambda_3 - \lambda_5 \\ \lambda_3 \\ 0 \\ \lambda_5 \end{pmatrix} \mid \lambda_3, \lambda_5 \in \mathbb{R}\} = \{\lambda_3(e_2 + e_3) + \lambda_5(e_5 - e_2) \mid \lambda_3, \lambda_5 \in \mathbb{R}\}$$

Als Basis des Lösungsraum ergibt sich $e_2 + e_3, e_5 - e_2$, da sie nach obigem ein Erzeugendensystem von $\mathbb{L}$ bilden und offenbar linear unabhängig sind.

*Allgemein*: Freie Parameter gibt es für jede Spalte außer den Pivotspalten. Im Beispiel sind die erste, zweite und vierte Pivotspalten, also freie Parameter $\lambda_3, \lambda_5$. Durch Ausklammern der freien Parameter in der allgemeinen Lösung erhält man eine Basis des allgemeinen Lösungsraums.

## 4. Basiswechsel

Wir haben gesehen, dass man Vektoren durch Koordinaten bezüglich einer Basis beschreiben kann. Wenn man die Basis wechselt, dann transformieren sich auch die Koordinaten. Wie das genau geschieht, sehen wir in diesem Abschnitt. Am Schluss werden wir als Anwendung beweisen, dass der Zeilenrang einer Matrix gleich dem Spaltenrang ist. Dabei wird sehr schön klar, wie wichtig eine geeignete Wahl einer Basis ist. Wie immer ist $K$ ein Körper.

**4.1.** Wir gehen im folgenden von einem endlich dimensionalen $K$-Vektorraum $V$ aus. Wenn wir eine Basis $b_1, \ldots, b_n$ von $V$ wählen, dann ist jeder Vektor $v \in V$ durch seine Koordinaten $\lambda_1, \ldots, \lambda_n$ bestimmt, d.h.

$$v = \lambda_1 b_1 + \ldots + \lambda_n b_n.$$

Wir bezeichnen mit $\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in K^n$ den Koordinatenvektor von $v$ bzgl. der Basis $b_1, \ldots, b_n$.

**4.2 Proposition.** *Sei $c_1, \ldots, c_n$ eine neue Basis von $V$. Dann gibt es genau eine Matrix $T \in M(n \times n, K)$ so, dass für jedes $v \in V$ mit Koordinatenvektor $\lambda \in K^n$ bzgl. $b_1, \ldots, b_n$ und Koordinatenvektor $\mu \in K^n$ bzgl. $c_1, \ldots, c_n$ gilt:*

$$\mu = T \cdot \lambda$$

**Beweis:** Wir stellen den alten Basisvektor $b_j$ bzgl. der neuen Basis $c_1, \ldots, c_n$ dar, d.h.

$$b_j = \sum_{i=1}^{n} t_{ij} c_i$$

mit allen $t_{ij} \in K$. Es gilt

$$v = \sum_{j=1}^{n} \lambda_j b_j = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{n} t_{ij} c_i = \sum_{i=1}^{n} \left( \sum_{j=1}^{n} t_{ij} \lambda_j \right) c_i.$$

Andererseits gilt

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

Wegen der Eindeutigkeit der Koordinaten folgt

$$\mu_i = \sum_{j=1}^{n} t_{ij}\lambda_j.$$

Mit $T = (t_{ij}) \in M(n \times n, K)$ folgt die Existenz in der Proposition.

Wir müssen noch die Eindeutigkeit von $T$ beweisen. Es erfülle also $T \in M(n \times n, K)$ die Anforderungen der Proposition. Sei $e_i$ der $i$-te Vektor aus der Standardbasis von $K^n$. Dann ist $e_i$ der Koordinatenvektor von $b_i$ bzgl. der Basis $b_1, \dots, b_n$. Es gilt dann, dass $T \cdot e_i$ der $i$-te Spaltenvektor von $T$ ist und andererseits der Koordinatenvektor von $b_i$ bzgl. $c_1, \dots c_n$. Somit ist $T$ dieselbe Matrix wie im ersten Teil des Beweises und damit eindeutig. □

**4.3 Bemerkung.** Die Matrix $T$ aus Proposition 4.2 erhält man, in dem man die Koordinatenvektoren der alten Basisvektoren bzgl. der neuen Basis in die Spalten von $T$ schreibt. Sie heißt die **Transformationsmatrix** des Basiswechsels.

**4.4 Proposition.** *Sei $T$ die Transformationsmatrix, die die Koordinaten bzgl. der Basis $b_1, \dots, b_n$ von $V$ in die Koordinaten bzgl. der Basis $c_1, \dots, c_n$ transformiert. Sei $S$ die Transformationsmatrix, die die Koordinaten bzgl. der Basis $c_1, \dots, c_n$ in die Koordinaten bzgl. einer weiteren Basis $d_1, \dots, d_n$ transformiert. Dann ist $S \cdot T$ die Transformationsmatrix, die die Koordinaten von der ursprünglichen Basis $b_1, \dots, b_n$ in die Koordinaten bzgl. $d_1, \dots, d_n$ transformiert.*

**Beweis:** Sei $\lambda \in K^n$ der Koordinatenvektor von $v \in V$ bzgl. $b_1, \dots, b_n$, sei $\mu \in K^n$ der Koordinatenvektor von $v \in V$ bzgl. $c_1, \dots, c_n$ und sei $\rho \in K^n$ der Koordinatenvektor von $v \in V$ bzgl. $d_1, \dots, d_n$. Dann gilt

$$\rho = S \cdot \mu = S \cdot (T \cdot \lambda) = (S \cdot T) \cdot \lambda$$

und nach Proposition 4.2 folgt, dass $S \cdot T$ die Transformationsmatrix des Basiswechsels von $b_1, \dots, b_n$ nach $d_1, \dots, d_n$ ist. □

**4.5 Korollar.** *Die Transformation eines Basiswechsels ist invertierbar.*

**Beweis:** Sei $T$ die Transformationsmatrix des Basiswechsels von $b_1, \dots, b_n$ nach $c_1, \dots, c_n$. Sei $S$ die Transformationsmatrix des Basiswechsels von $c_1, \dots, c_n$ nach $b_1, \dots, b_n$.

Nach Proposition 4.4 ist $S \cdot T$ die Transformationsmatrix, die die Koordinaten bzgl. $b_1, \dots, b_n$ in die Koordinatenform $b_1, \dots, b_n$ transformiert. Da dies dieselbe Basis ist, bleiben die Koordinaten gleich und damit gilt:

$$S \cdot T = E_n$$

Mit Proposition 2.5 ist $T$ invertierbar. □

**4.6 Proposition.** *Sei $\varphi \colon V \to W$ eine lineare Abbildung von $K$-Vektorräumen. Wir betrachten alte Basen $b_1, \dots, b_n$ von $V$ und $c_1, \dots, c_m$ von $W$ und neue Basen $b'_1, \dots, b'_n$ von $V$ und $c'_1, \dots, c'_m$ von $W$. Weiter sei $T$ (bzw. $S$) die Transformationsmatrix des Basiswechsels auf $V$ (bzw. $W$). Weiter sei $A(\varphi)$ (bzw. $A'(\varphi)$) die zu $\varphi$ gehörige Matrix bzgl. den alten (bzw. neuen) Basen. Dann gilt*

$$A'(\varphi) = S \cdot A(\varphi) \cdot T^{-1}.$$

**Beweis:** Es sei $\lambda \in K^n$ der Koordinatenvektor von $v \in V$ bzgl. $b_1, \dots, b_n$. Es sei $\lambda' \in K^n$ der Koordinatenvektor von $v \in V$ bzgl. $b'_1, \dots, b'_n$. Es sei $\mu \in K^m$ der Koordinatenvektor von $\varphi(v) \in W$ bzgl. $c_1, \dots, c_n$. Es sei $\mu' \in K^m$ der Koordinatenvektor von $\varphi(v) \in W$ bzgl. $c'_1, \dots, c'_n$.

Nach Definitionen gilt

$$\mu = A(\varphi) \cdot \lambda \quad \wedge \quad \mu' = A'(\varphi) \cdot \lambda'$$

und

$$\lambda' = T \cdot \lambda \quad \wedge \quad \mu' = S \cdot \mu.$$

Zusammen ergibt sich

$$S \cdot A(\varphi) \cdot \lambda = S \cdot \mu = \mu' = A'(\varphi) \cdot \lambda' = A'(\varphi) \cdot T \cdot \lambda.$$

Offenbar beschreiben sowohl $S \cdot A(\varphi)$ wie auch $A'(\varphi) \cdot T$ die lineare Abbildung $\varphi$ bzgl. den Basen $b_1, \ldots, b_n$ von $V$ und $c'_1, \ldots, c'_m$ von $W$. Wegen der Eindeutigkeit dieser Matrix (siehe 1.5) folgt

$$S \cdot A(\varphi) = A'(\varphi) \cdot T.$$

Nach Korollar 4.5 ist die Basiswechselmatrix $T$ invertierbar und es folgt:

$$A'(\varphi) = S \cdot A(\varphi) \cdot T^{-1}.$$ □

**4.7 Bemerkung.** Hat man eine Basis $b_1, \ldots, b_n$ von $V$ und $T \in \mathrm{GL}(n,K)$ gegeben, dann gibt es genau eine Basis $b'_1, \ldots, b'_n$ von $V$ so, dass $T$ die Transformationsmatrix des Basiswechsels ist (siehe Aufgabe 12.2).

**4.8 Korollar.** *Seien $V, W$ $K$-Vektorräume mit $n := \dim(V) < \infty$ und $m := \dim(W) < \infty$.*

*a) Zwei Matrizen $A, A' \in M(m \times n, K)$ sind genau dann äquivalent, wenn es eine lineare Abbildung $\varphi\colon V \to W$ gibt so, dass $A$ die Matrix zu $\varphi$ ist bzgl. einer Wahl von Basen von $V$ und $W$ und $A'$ die Matrix zu $\varphi$ ist bzgl. einer weiteren Wahl von Basen von $V$ und $W$.*

*b) Zwei Matrizen $B, B' \in M(n \times n, K)$ sind genau dann ähnlich, wenn $B$ die Matrix zu $\varphi$ ist bzgl. einer Wahl **einer** Basis von $V$ und $B'$ die Matrix zu $\varphi$ ist bzgl. einer weiteren Wahl **einer** Basis von $V$.*

**Beweis:** Wir zeigen zuerst „$\Longrightarrow$“ in a). Es seien $A, A'$ äquivalent, d.h. $\exists S \in \mathrm{GL}(m,K)$ und $T \in \mathrm{GL}(n,K)$ mit

$$A' = S \cdot A \cdot T^{-1}.$$

Wir wählen eine beliebige Basis $b_1, \ldots, b_n$ von $V$ und eine beliebige Basis $c_1, \ldots, c_m$ von $W$. Nach Proposition 1.6 gibt es eine lineare Abbildung $\varphi\colon V \to W$ so, dass $A$ die Matrix zu $\varphi$ ist bzgl. diesen Basen. Nach Bemerkung 4.7 gibt es neue Basen $b_1, \ldots, b'_n$ von $V$ und $c'_1, \ldots, c'_m$ von $W$ so, dass $S$ und $T$ die Transformationsmatrizen der Basiswechsel sind. Mit Proposition 4.6 folgt, dass $A'$ die Matrix zu $\varphi$ ist bzgl. den neuen Basen. Dies zeigt „$\Longrightarrow$“ in a). Umkehrung in a) folgt direkt aus Proposition 4.6. b) folgt analog wie a), wenn man $V = W$ und $b_i = c_i$ benutzt. □

**4.9 Bemerkung.** Zusammen mit Theorem 4.10 wird folgen, dass $A, A' \in M'(m \times n, K)$ genau dann äquivalent sind, wenn $\mathrm{Rang}(A) = \mathrm{Rang}(A')$ gilt. Viel schwieriger ist ein Kriterium zu erhalten für ähnliche Matrizen. Dies wird uns in den kommenden Kapiteln beschäftigen.

**4.10 Theorem.** *Es sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen endlich dimensionalen Vektorräumen. Dann gibt es Basen $b_1, \ldots, b_n$ von $V$ und $c_1, \ldots, c_m$ von $W$ so, dass die zu $\varphi$ gehörige Matrix die Form*

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \in M(m \times n, K)$$

*hat, wobei $r = \dim \varphi(V)$ ist.*

**Beweis:** Es sei $c_1, \ldots, c_r$ eine Basis von $\varphi(V)$. Mit dem Basisergänzungssatz (Theorem IV.2.5) können wir das zu einer Basis $c_1, \ldots, c_m$ von $W$ ergänzen. Nach Korollar IV.2.10 gibt es einen Unterraum $U$ von $V$ mit

$$V = U \oplus \ker(\varphi).$$

Offenbar gilt $\varphi(V) = \varphi(U)$. Damit existieren $b_1, \ldots, b_r \in U$ mit $\varphi(b_i) = c_i$ für $i = 1, \ldots, r$. Wir zeigen nun, dass $b_1, \ldots, b_r$ eine Basis von $U$ ist. Die lineare Unabhängigkeit folgt sofort aus derjenigen von $c_1, \ldots, c_r$. Weiter sei $u \in U$. Dann gilt $\varphi(u) = \sum_{i=1}^{r} \lambda_i c_i$ mit den Koordinaten $\lambda_i \in K$ von $\varphi(u)$. Aus der

Linearität von $\varphi$ folgt, dass $u - \sum_{i=1}^{r} \lambda_i b_i \in \ker(\varphi)$ gilt. Wegen $U \cap \ker(\varphi) = \{0\}$ muss $u = \sum_{i=1}^{r} \lambda_i b_i$ gelten und damit ist $b_1, \ldots, b_r$ auch ein Erzeugendensystem von $U$.

Wir wählen eine Basis $b_{r+1}, \ldots, b_n$ von $\ker(\varphi)$, dann ist $b_1, \ldots, b_n$ eine Basis von $V$ (siehe Beweis von Proposition IV.3.14). Es gilt $\varphi(b_j) = 0$ für $r < j \leq n$ und damit hat die zu $\varphi$ gehörige Matrix bzgl. den Basen $b_1, \ldots, b_n$ und $c_1, \ldots, c_m$ die gewünschte Form. □

**4.11 Korollar.** *Jede $m \times n$ Matrix $A$ ist äquivalent zu genau einer Matrix der Form*

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \in M(m \times n, K).$$

*Es gilt $r = \mathrm{Rang}(A)$.*

**Beweis:** Korollar 4.8 und Theorem 4.10. Die Eindeutigkeit folgt aus Proposition 2.8. □

**4.12 Theorem.** *Für $m \times n$ Matrizen ist der Rang gleich dem Zeilenrang.*

1\. Beweis: Sei $A \in M(m \times n, K)$. Wir wählen $K$-Vektorräume $V, W$ mit Basen $b_1, \ldots, b_n$ bzw. $c_1, \ldots, c_m$. Nach Proposition 1.6 gibt es eine lineare Abbildung $\varphi \colon V \to W$ so, dass $A$ die zu $\varphi$ gehörige Matrix ist. Man kann z.B. $V = K^n$ und $W = K^m$ und $\varphi = \varphi_A$ nehmen. Nach Theorem 4.10 gibt es neue Basen $b'_1, \ldots, b'_n$ von $V$ und $c'_1, \ldots, c'_m$ von $W$ so, dass die neue zu $\varphi$ gehörige Matrix die Form

$$B = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

hat mit

$$\mathrm{Rang}(B) = r = \dim \varphi(V) = \mathrm{Rang}(A).$$

Andererseits gilt für den Zeilenrang $R$ von $A$

$$R = \mathrm{Rang}(A^t) \underset{1.21}{=} \dim \varphi^*(W^*) \underset{1.21}{=} \mathrm{Rang}(B^t).$$

Weil offenbar die ersten $r$ Zeilen (bzw. Spalten) von $B$ linear unabhängig sind, muss

$$\mathrm{Rang}(B^t) = \mathrm{Rang}(B) = r$$

gelten. Es folgt

$$R = \mathrm{Rang}(B^t) = r$$

wie gewünscht. □

2\. Beweis: Nach Korollar 4.11 gibt es $S \in \mathrm{GL}(m, K)$ und $T \in \mathrm{GL}(n, K)$ mit

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = SAT^{-1}.$$

Transponieren liefert mit den Rechenregeln in 1.19

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}^t = (SAT^{-1})^t = (T^{-1})^t A^t S^t.$$

Wir haben im Beweis von Proposition 2.5 gesehen, dass $(T^{-1})^t = (T^t)^{-1}$ gilt und somit ist die Matrix $A^t$ äquivalent zu $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}^t$. Nach Proposition 2.7 müssen sie deshalb denselben Rang haben. Nun gilt

$$\mathrm{Rang}(A) = r = \mathrm{Rang}(\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}^t) = \mathrm{Rang}(A^t).$$

Weil $\mathrm{Rang}(A^t)$ der Zeilenrang von $A$ ist, folgt die Behauptung. □

# 6 Kapitel VI. Determinanten

## 1. Permutationen

In II.1.10 haben wir gesehen, dass die Menge $S(X)$ der bijektiven Selbstabbildungen einer Menge $X$ eine Gruppe bzgl. der Komposition von Abbildungen bildet. Im Spezialfall $X := \{1,\dots,n\}$ erhalten wir die Gruppe der Permutationen von $1,\dots,n$ die uns in diesem Abschnitt beschäftigen wird. Die Permutationen werden bei der Definition und Berechnung von Determinanten eine wichtige Rolle spielen.

**1.1 Definition.** Eine **Permutation** der Zahlen $1,\dots,n$ ist eine bijektive Abbildung $\pi\colon \{1,\dots,n\} \to \{1,\dots,n\}$. Wir schreiben $\pi$ in der Form

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \pi(1) & \pi(2) & \pi(3) & \cdots & \pi(n) \end{pmatrix}.$$

Zum Beispiel bedeutet $\left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{smallmatrix}\right)$, dass $\pi(1) = 2$, $\pi(2) = 3$, $\pi(3) = 1$ und $\pi(4) = 4$. Die Menge aller Permutationen von $1,\dots,n$ wird mit $S_n$ bezeichnet und heißt **symmetrische Gruppe zum Index n**.

**1.2 Proposition.** *Zusammen mit der Komposition von Selbstabbildungen ist $S_n$ eine Gruppe mit $n!$ Elementen.*

**Beweis:** Wir hatten in Proposition II.1.10 schon gesehen, dass $S_n = S(\{1,\dots,n\})$ eine Gruppe ist. Wenn wir eine Permutationen festlegen, dann haben wir für $\pi(1)$ insgesamt $n$ Möglichkeiten. Wenn $\pi(1)$ festgelegt ist, dann bleiben $n-1$ Möglichkeiten für $\pi(2)$ usw. Insgesamt ergeben sich

$$n \cdot (n-1) \cdot (n-2) \cdots 1 = n!$$

Möglichkeiten. □

**1.3 Beispiel.** Wir sehen an folgender Rechnung, dass $S_n$ nicht abelsch sein muss für $n \geq 3$:

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

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

**1.4 Definition.** Eine **Transposition** $\tau$ ist eine Permutation, die zwei verschiedene Zahlen vertauscht und die anderen Zahlen fest lässt, d.h. $\exists i \neq j$ mit $\tau(i) = j$, $\tau(j) = i$ und $\tau(k) = k \, \forall k \in \{1,\dots,n\} \setminus \{i,j\}$.

Wir schreiben $\tau = (i,j)$ für so eine Transposition. Es gilt immer $\tau^2 = \mathrm{id}$.

**1.5 Theorem.** *Für $n \geq 2$ ist jedes $\pi \in S_n$ ein Produkt von Transpositionen der Form $(i, i+1)$.*

**Beweis:** Mit Induktion nach $n$. Der Induktionsanfang ist trivial. Induktionsschritt: Wir nehmen an, dass $n \geq 3$ ist und dass die Behauptung stimmt für $n-1$. Es sei $k := \pi(n)$. Dann gilt:

$$\underbrace{(n,n-1) \circ \cdots \circ (k+1,k) \circ \pi}_{=\pi'}(n) = n$$

Weil $\pi'(n) = n$ gilt, induziert $\pi'$ eine Permutation von $1,\dots,n-1$. Mit vollständiger Induktion folgt:

$$\pi' = \tau_1 \circ \dots \circ \tau_r$$

für Transpositionen benachbarter Elemente in $\{1,\dots,n-1\}$. Diese Gleichung macht Sinn als Permutation von $1,\dots,n$ (die $n$ fest lassen) und es folgt

$$\begin{aligned}\pi &= (k+1,k)^{-1} \circ \dots \circ (n,n-1)^{-1} \circ \tau_1 \circ \dots \circ \tau_r \\ &= (k,k+1) \quad \circ \dots \circ (n-1,n) \quad \circ \tau_1 \circ \dots \circ \tau_r.\end{aligned}$$

Dies zeigt den Induktionsschritt. Mit vollständiger Induktion folgt die Behauptung. □

**1.6 Definition.** Das **Signum** von $\pi \in S_n$ ist

$$\operatorname{sign}(\pi) := \prod_{i<j} \frac{\pi(j)-\pi(i)}{j-i},$$

wobei die Indexmenge des Produkts gleich $\{(i,j) \in \{1,\dots,n\}^2 \mid i<j\}$ ist.

**1.7 Beispiel.**

$$\operatorname{sign}\begin{pmatrix}1 & 2 & 3 & 4\\ 2 & 3 & 4 & 1\end{pmatrix} = \frac{(3-2)(4-2)(1-2)(4-3)(1-3)(1-4)}{(2-1)(3-1)(4-1)(3-2)(4-2)(4-3)} = (-1)^{1+1+1} = -1$$

**1.8.** Wir bemerken ganz allgemein, dass im Zähler und im Nenner von $\operatorname{sign}(\pi)$ bis auf das Vorzeichen dieselben Faktoren vorkommen, weil $\pi$ eine Permutation ist. Wenn im Zähler ein Faktor $i-j$ mit $i<j$ vorkommt, dann sprechen wir von einem Fehlstand. Also gilt:

$$\operatorname{sign}(\pi) = (-1)^N,$$

wobei $N$ die Anzahl der Fehlstände ist.

**1.9 Proposition.** *Die Abbildung* $\operatorname{sign}\colon S_n \to \{-1,1\}$ *ist ein Gruppenhomomorphismus, d.h.*

$$\operatorname{sign}(\sigma\circ\pi) = \operatorname{sign}(\sigma)\cdot\operatorname{sign}(\pi).$$

**Beweis:**

$$\begin{aligned}\operatorname{sign}(\sigma\circ\pi) &= \prod_{i<j} \frac{\sigma\big(\pi(j)\big)-\sigma\big(\pi(i)\big)}{j-i} \\ &= \prod_{i<j} \frac{\sigma\big(\pi(j)\big)-\sigma\big(\pi(i)\big)}{\pi(j)-\pi(i)} \cdot \underbrace{\prod_{i<j}\frac{\pi(j)-\pi(i)}{j-i}}_{=\operatorname{sign}(\pi)}\end{aligned}$$

Wir behaupten, dass

$$\prod_{i<j} \frac{\sigma\big(\pi(j)\big)-\sigma\big(\pi(i)\big)}{\pi(j)-\pi(i)} = \underbrace{\prod_{k<l}\frac{\sigma(l)-\sigma(k)}{l-k}}_{=\operatorname{sign}(\sigma)} \tag{VI.1}$$

gilt und damit folgt dann sofort die Proposition. Um (VI.1) zu beweisen, bemerken wir zuerst, dass es zu jedem $k<l$ genau ein $i<j$ gibt mit $\{k,l\} = \{\pi(i),\pi(j)\}$. Dies folgt, weil $\pi$ eine Permutation ist.

**1. Fall** $\pi^{-1}(k) < \pi^{-1}(l)$. Dann gilt $i = \pi^{-1}(k)$ und $j = \pi^{-1}(l)$.

$$\Longrightarrow \frac{\sigma\big(\pi(j)\big)-\sigma\big(\pi(i)\big)}{\pi(j)-\pi(i)} = \frac{\sigma(l)-\sigma(k)}{l-k}$$

**2. Fall** $\pi^{-1}(k) > \pi^{-1}(l)$. Dann gilt $i = \pi^{-1}(l)$ und $j = \pi^{-1}(k)$. Es folgt

$$\frac{\sigma(\pi(j)) - \sigma(\pi(i))}{\pi(j) - \pi(i)} = \frac{\sigma(k) - \sigma(l)}{k - l} = \frac{\sigma(l) - \sigma(k)}{l - k}$$

Also sind auf beiden Seiten in (VI.1) dieselben Faktoren und deshalb gilt (VI.1). □

**1.10 Korollar.** *Ist $\pi \in S_n$ ein Produkt von $N$ Transpositionen. Dann gilt* $\operatorname{sign}(\pi) = (-1)^N$.

**Beweis:** Es sei $\pi = \tau_1 \circ \cdots \circ \tau_N$ das Produkt von $N$ Transpositionen $\tau_i$. Dann gilt mit Proposition 1.9:

$$\operatorname{sign}(\pi) = \prod_{i=1}^{N} \operatorname{sign}(\tau_i)$$

Die Signatur einer Transposition $\tau = (i, j)$ ist $-1$, da sie genau einen Fehlstand hat und zwar in $(i, j)$

$$\Longrightarrow \operatorname{sign}(\pi) = (-1)^N.$$ □

**1.11.** Weil sich nach Theorem 1.5 jedes $\pi \in S_n$ als Produkt von Transpositionen schreiben lässt, kann man Korollar 1.10 benutzen um $\operatorname{sign}(\pi)$ zu berechnen. Eine Permutation $\pi \in S_n$ heißt **gerade** (bzw. **ungerade**), falls $\operatorname{sign}(\pi) = 1$ (bzw. $(\operatorname{sign}(\pi) = -1)$ ist. Nach Korollar 1.10 ist dies äquivalent dazu, dass $\pi$ ein Produkt einer geraden (bzw. ungeraden) Anzahl Transpositionen ist.

**1.12 Beispiel.**

$$\begin{aligned}\pi := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} &= (1,4) \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} = (1,4) \circ (1,3) \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \\ &= (1,4) \circ (1,3) \circ (1,2) \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} = (1,4) \circ (1,3) \circ (1,2)\end{aligned}$$

und damit ergibt sich wieder $\operatorname{sign}(\pi) = (-1)^3 = -1$, d.h. $\pi$ ist ungerade.

## 2. Determinantenformen

Die Determinante ist eine Zahl, die einer quadratischen Matrix zugeordnet wird. Sie ist genau dann Null, wenn die Matrix nicht invertierbar ist. Man kann sie benutzt, um dann die Inverse zu berechnen. Geometrisch misst sie das orientierte Volumen des Parallelotops, das durch die Spalten der Matrix aufgespannt wird. Dies alles wird uns in den kommenden Abschnitten beschäftigen. Jetzt führen wir nur die Determinanten ein. Dabei ist $K$ ein beliebiger Körper und $V$ ein endlich-dimensionaler Vektorraum. Weiter sei $n := \dim(V)$.

**2.1 Definition.** Eine **k-fache Multilinearform** auf $V$ ist eine Abbildung

$$F \colon V^k \to K, \quad (v_1, \ldots, v_k) \mapsto F(v_1, \ldots, v_k)$$

mit der Eigenschaft, dass $F$ linear ist in jedem der $k$ Argumente, d.h.

$$F(v_1, \ldots, v_{j-1}, \lambda' v_j' + \lambda'' v_j'', v_{j+1}, \ldots, v_k) = \lambda' F(v_1, \ldots, v_j', \ldots, v_k) + \lambda'' F(v_1, \ldots, v_j'', \ldots, v_k)$$

für alle $\lambda', \lambda'' \in K$, $v_1, \ldots, v_{j-1}, v_j', v_j'', v_{j+1}, \ldots, v_k \in V$ und $j = 1, \ldots, k$.

**2.2 Beispiel.** Für $k = 1$ erhalten wir: 1-fache Multilinearform = lineares Funktional, denn $F \colon V \to K$ muss linear sein. Wir erhalten also die Elemente des Dualraums $V^*$. Für beliebiges $k \geq 1$ haben wir folgendes Beispiel einer $k$-fachen Multilinearform auf $V = K$:

$$F \colon V^k = K^k \to K, \quad \lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix} \mapsto \lambda_1 \cdots \lambda_k$$

Aufgrund der Distributivität von $K$ ist $F$ linear in jedem $\lambda_j$ und damit eine $k$-fache Multilinearform. Beachte, dass $F$ nicht linear ist, falls $k \geq 2$.

**2.3 Definition.** Eine $k$-fache Multilinearform $F$ auf $V$ heißt **alternierend** $:\Longleftrightarrow F(v_1,\ldots,v_k) = 0$ falls $v_i = v_j$ für ein $i \neq j$. Der Name alternierend erklärt sich mit folgenden Resultat:

**2.4 Proposition.** *Sei $\pi \in S_k$ und $F$ eine alternierende $k$-fache Multilinearform. Dann gilt*

$$F(v_{\pi(1)},\ldots,v_{\pi(k)}) = \operatorname{sign}(\pi)\cdot F(v_1,\ldots,v_k).$$

**Beweis:** Nach Theorem 1.5 ist $\pi$ ein Produkt von Transpositionen. Wir argumentieren mit Induktion nach der Anzahl Faktoren $N$. Der Fall $\pi = \mathrm{id}$ ist trivial. Wir machen den Induktionsanfang bei $N = 1$, d.h. $\pi$ ist eine Transposition $(i,j)$. Der Einfachheit halber nehmen wir $(i,j) = (1,2)$ an, die anderen Fälle gehen analog. Es gilt

$$\begin{aligned}
0 &\underset{\text{alternierend}}{=} F(v_1+v_2, v_1+v_2, v_3,\ldots,v_k) \\
&\underset{\text{multilinear}}{=} F(v_1,v_1,v_3,\ldots,v_k) + F(v_1,v_2,v_3,\ldots,v_k) + F(v_2,v_1,v_3,\ldots,v_k) + F(v_2,v_2,v_3,\ldots,v_k) \\
&\underset{\text{alternierend}}{=} F(v_1,v_2,\ldots,v_k) + F(v_2,v_1,v_3,\ldots,v_k).
\end{aligned}$$

Damit folgt

$$F(v_2,v_1,v_3,\ldots,v_k) = -F(v_1,v_2,\ldots,v_k).$$

Weil $\operatorname{sign}((1,2)) = -1$, zeigt dies die Behauptung im Fall $N = 1$. Induktionsschritt: Es sei $N \geq 2$ und wir nehmen an, dass die Behauptung richtig ist für $N-1$. Also gilt $\pi = \pi' \circ \tau$ mit $\pi' \in S_k$ Produkt von $N-1$ Transpositionen und $\tau$ Transposition. Nach Induktionsannahme im Fall $N-1$ gilt

$$F(v_{\pi(1)},\ldots,v_{\pi(k)}) = F(v_{\pi'\circ\tau(1)},\ldots,v_{\pi'\circ\tau(k)}) = \operatorname{sign}(\pi')F(v_{\tau(1)},\ldots,v_{\tau(k)}).$$

Es folgt aus dem Induktionsanfang $N = 1$

$$\begin{aligned}
F(v_{\pi(1)},\ldots,v_{\pi(k)}) &= \operatorname{sign}(\pi')\operatorname{sign}(\tau)F(v_1,\ldots,v_k) \\
&\underset{1.9}{=} \operatorname{sign}(\pi'\circ\tau)F(v_1,\ldots,v_k) \\
&= \operatorname{sign}(\pi)F(v_1,\ldots,v_k).
\end{aligned}$$

Dies beweist den Induktionsschritt. Mit vollständiger Induktion folgt die Behauptung. □

**2.5 Proposition.** *Es sei $F$ eine alternierende $k$-fache Multilinearform auf $V$. Falls $v_1,\ldots,v_k$ linear abhängig sind, dann gilt*

$$F(v_1,\ldots,v_k) = 0.$$

**Beweis:** Nach Proposition IV.1.8 lässt sich einer der Vektoren als Linearkombination der anderen Vektoren schreiben. O.B.d.A.

$$v_1 = \sum_{i=2}^{k} \lambda_i v_i.$$

Aufgrund der Multilinearität folgt:

$$F(v_1,v_2,\ldots,v_k) = \sum_{i=2}^{k} \lambda_i F(v_i,v_2,\ldots,v_k) = \sum_{i=2}^{k} \lambda_i \cdot 0 = 0.$$

Dies zeigt die Behauptung. □

**2.6.** Wir haben in Beispiel III.1.3 gesehen, dass für eine beliebige Menge $X$ die Menge der Funktionen $f\colon X \to K$ unter punktweiser Addition und skalarer Multiplikation zu einem $K$-Vektorraum wird. Wir benutzen dies für $X := V^k$ und haben somit eine Rezept um $k$-fache Multilinearformen $F_i$ zu addieren und mit einem Skalar $\lambda$ zu multiplizieren:

$$\begin{aligned}(F_1+F_2)(v_1,\dots,v_k) &:= F_1(v_1,\dots,v_k)+F_2(v_1,\dots,v_k)\\ (\lambda\cdot F)(v_1,\dots,v_k) &:= \lambda\cdot F(v_1,\dots,v_k)\end{aligned}$$

Es ist trivial, dass $F_1+F_2$ und $\lambda F$ wieder $k$-fache Multilinearformen auf $V$ sind. Wenn $F_1,F_2,F$ alternierend sind, dann bleiben $F_1+F_2,\lambda\cdot F$ offensichtlich alternierend. Wir fassen diese Resultate zusammen:

**2.7 Proposition.** *Mit den in 2.6 definierten Operationen bildet die Menge der $k$-fachen Multilinearformen auf $V$ ein $K$-Vektorraum. Die Menge der alternierenden $k$-fachen Multilinearformen auf $V$ ist darin ein Unterraum.*

**Beweis:** Wir haben in 2.6 gesehen, dass diese beiden Mengen Unterräume des Vektorraums aller Funktionen $f\colon X = V^k \to K$ sind . □

**2.8.** Eine **Determinantenform** $\Delta$ ist eine alternierende $n$-fache Multilinearform auf $V$, d.h.

$$\Delta\colon V^n \to K, (v_1,\dots,v_n) \mapsto \Delta(v_1,\dots,v_n) \text{ mit}$$

- $\Delta(v_1,\dots,v_n)$ ist linear in jedem $v_i$ (aber nicht unbedingt linear als Abbildung $V^n \to K$).
- $\Delta(v_1,\dots,v_n) = 0$ falls $v_i = v_j$ für ein $i \neq j$.

Wir werden zeigen, dass die Determinantenformen einen 1-dimensionalen $K$-Vektorraum bilden. Zuerst führen wir die Determinante einer $n\times n$-Matrix mit Hilfe einer Formel ein. Später geben wir dann auch eine intrinsische Charakterisierung.

**2.9 Definition.** Für $A = (a_{ij}) \in M(n\times n, K)$ definieren wir die **Determinante** als

$$\det(A) := \sum_{\pi\in S_n} \operatorname{sign}(\pi) a_{\pi(1)1}\cdots a_{\pi(n)n}.$$

**2.10 Lemma.** *Es sei $\Delta$ eine Determinantenform auf $V$ und $b_1,\dots,b_n$ eine Basis von $V$. Für Vektoren $v_1,\dots,v_n \in V$ mit Koordinatenvektoren $a_1,\dots,a_n \in K^n$ bzgl. dieser Basis gilt dann*

$$\Delta(v_1,\dots,v_n) = \det(a_1,\dots,a_n)\Delta(b_1,\dots,b_n),$$

*wobei $(a_1,\dots,a_n)$ die $n\times n$-Matrix mit dem Spaltenvektor $a_1,\dots,a_n$ ist.*

Beachte, dass wir hier die Koordinaten der Vektoren folgendermassen bezeichnen, damit es mit den Einträgen der Matrix passt:

$$a_j = \begin{pmatrix} a_{1j}\\ \vdots \\ a_{nj}\end{pmatrix}.$$

**Beweis:** Aufgrund der Multilinearität von $\Delta$ folgt

$$\begin{aligned}\Delta(v_1,\dots,v_n) &= \Delta\Big(\sum_{i_1=1}^n a_{i_11}b_{i_1},\dots,\sum_{i_n=1}^n a_{i_nn}b_{i_n}\Big) = \sum_{i_1=1}^n a_{i_11}\cdot\Delta\Big(b_{i_1},\sum_{i_2=1}^n a_{i_22}b_{i_2},\dots,\sum_{i_n=1}^n a_{i_nn}b_{i_n}\Big)\\ &= \sum_{i_1,\dots,i_n\in\{1,\dots,n\}} a_{i_11}\cdots a_{i_nn}\cdot\Delta(b_{i_1},\dots,b_{i_n}).\end{aligned}$$

Weil $\Delta$ alternierend ist, sind alle Summanden gleich 0, bei denen $i_r = i_s$ ist für ein $r \neq s$. Also spielen nur diejenigen $i_1, \ldots, i_n$ eine Rolle, bei denen $k \mapsto i_k$ eine Permutation von $\{1, \ldots, n\}$ ist. Also gilt

$$\begin{aligned}\Delta(v_1, \ldots, v_n) &= \sum_{\pi \in S_n} a_{\pi(1)1} \cdots a_{\pi(n)n} \cdot \Delta(b_{\pi(1)}, \ldots, b_{\pi(n)}) \\ &= \sum_{\pi \in S_n} \operatorname{sign}(\pi) a_{\pi(1)1} \cdots a_{\pi(n)n} \cdot \Delta(b_1, \ldots, b_n) = \det(a_1, \ldots, a_n)\Delta(b_1, \ldots, b_n),\end{aligned}$$

wobei der letzte Schritt aus Proposition 2.4 folgt. □ □

**2.11 Theorem.** *Der Raum der Determinantenformen auf $V$ ist ein 1-dimensionaler $K$-Vektorraum.*

**Beweis:** Wir haben in Proposition 2.7 gesehen, dass die alternierenden $n$-fachen Multilinearformen auf $V$ einen $K$-Vektorraum bilden. Wir wählen eine Basis $b_1, \ldots, b_n$ von $V$. Für $v_1, \ldots, v_n \in V$ mit Koordinatenvektoren $a_1, \ldots, a_n \in K^n$ bzgl. dieser Basis definieren wir

$$\Delta(v_1, \ldots, v_n) := \det(a_1, \ldots, a_n).$$

Wir zeigen zuerst, dass $\Delta$ eine Determinantenform ist. Aufgrund der Distributivität im Körper $K$ ist

$$(a_1, \ldots, a_n) \mapsto a_{\pi(1)1} \cdots a_{\pi(n)n}$$

eine $n$-fache Multilinearform auf $K^n$ für jedes $\pi \in S_n$ (analog zu Bsp 2.2). Nach Proposition 2.7 ist

$$(a_1, \ldots, a_n) \mapsto \det(a_1, \ldots, a_n)$$

ebenfalls eine $n$-fache Multilinearform auf $K^n$. Nach Proposition IV.2.15 ist die Koordinatenabbildung $V \to K^n$ bzgl. der Basis $b_1, \ldots, b_n$ ein Isomorphismus und damit muss auch $\Delta$ eine $n$-fache Multilinearform auf $V$ sein. Wir wollen nur zeigen, dass $\Delta$ alternierend ist. Sie also $a_i = a_j \in K^n$ für ein $i \neq j$. Wir benutzen die Transposition $\tau = (i, j) \in S_n$. Wir bezeichnen die Menge der geraden Permutationen mit $A_n$. Nach Aufgabe 13.2 ist $\pi \mapsto \pi \circ \tau$ eine Bijektion von $A_n$ auf die Menge der ungeraden Permutationen.

$$\begin{aligned}\Longrightarrow \det(a_1, \ldots, a_n) &= \sum_{\pi \in S_n} \operatorname{sign}(\pi) a_{\pi(1)1} \cdots a_{\pi(n)n} \\ &= \sum_{\pi \in A_n} a_{\pi(1)1} \cdots a_{\pi(n)n} - \sum_{\pi \in S_n,\ \operatorname{sign}(\pi) = -1} a_{\pi(1)1} \cdots a_{\pi(n)n} \\ &= \sum_{\pi \in A_n} a_{\pi(1)1} \cdots a_{\pi(n)n} - \sum_{\pi \in A_n} a_{\pi\circ\tau(1)1} \cdots a_{\pi\circ\tau(n)n}\end{aligned}$$

Wegen $\tau = (i, j)$ und $a_i = a_j \in K^n$ gilt

$$a_{\pi\circ\tau(i),i} = a_{\pi(j),i} = a_{\pi(j),j}$$

$$a_{\pi\circ\tau(j),j} = a_{\pi(i),j} = a_{\pi(i),i}$$

und

$$a_{\pi\circ\tau(k),k} = a_{\pi(k),k}$$

für $k \neq i, j$. Also folgt,

$$\det(a_1, \ldots, a_n) = \sum_{\pi \in A_n} a_{\pi(1)1} \cdots a_{\pi(n)n} - \sum_{\pi \in A_n} a_{\pi(1)1} \cdots a_{\pi(n)n} = 0$$

Damit ist det eine alternierenden $n$-fache Multilinearform auf $K^n$, d.h. eine Determinantenform. Mit Hilfe des Koordinatenisomorphismus $V \to K^n$ folgt, dass $\Delta$ eine Determinantenform auf $V$ ist. Wir haben bis jetzt gesehen, dass

$$\Delta\colon V^n \to K, (v_1, \ldots, v_n) \mapsto \Delta(v_1, \ldots, v_n) := \sum_{\pi \in S_n} \operatorname{sign}(\pi) a_{\pi(1)1} \cdots a_{\pi(n)n}$$

eine Determinantenform auf $V$ ist. Wir wollen als nächstes zeigen, dass $\Delta$ eine **nicht-triviale** Determinantenform ist, d.h. $\Delta$ ist nicht identisch Null ist. Das folgt aus:

$$\Delta(b_1,\ldots,b_n)=\sum_{\pi\in S_n}\operatorname{sign}(\pi)\delta_{\pi(1)1}\cdots\delta_{\pi(n)n}=1$$

wobei $\delta_{ij}$ das **Kroneckersymbol** ist, d.h.

$$\delta_{ij}:=\begin{cases}1, & \text{falls } i=j\\ 0, & \text{falls } i\neq j\end{cases}$$

Hier benutzen wir, dass der Koordinatenvektor von $b_j$ gleich dem Standardbasisvektor $e_j=(\delta_{ij})_{i=1,\ldots,n}\in K^n$ ist.

Wir wollen zeigen, dass der Vektorraum der Determinantenformen 1-dimensional ist. Es genügt also zu zeigen, dass eine weitere Determinantenform $\Delta'$ auf $V$ ein Vielfaches von unserem $\Delta$ ist. Aus Lemma 2.10 folgt

$$\Delta'(v_1,\ldots,v_n)=\det(a_1,\ldots,a_n)\Delta'(b_1,\ldots,b_n)=\Delta'(b_1,\ldots,b_n)\Delta(v_1,\ldots,v_n)$$

für alle $v_1,\ldots,v_n\in V$ mit Koordinatenvektoren $a_1,\ldots,a_n\in K^n$ bzgl. der Basis $b_1,\ldots,b_n$. Also folgt

$$\Delta'=\Delta'(b_1,\ldots,b_n)\cdot\Delta$$

und somit ist der Vektorraum der Determinantenformen 1-dimensional. □

**2.12 Korollar.** *Sei $\Delta$ eine nicht-triviale Determinantenform auf $V$, $n:=\dim(V)<\infty$ und $v_1,\ldots,v_n\in V$. Dann sind die folgenden Aussagen äquivalent:*

*a)* $v_1,\ldots,v_n$ *Basis von* $V$

*b)* $\Delta(v_1,\ldots,v_n)\neq 0$

**Beweis:** Falls $v_1,\ldots,v_n$ linear abhängig sind, dann folgt $\Delta(v_1,\ldots,v_n)=0$ aus Proposition 2.5. Damit folgt a) aus b). Falls $v_1,\ldots,v_n$ linear unabhängig sind, dann bilden sie eine Basis von $V$. Wenn nun $\Delta$ die im Beweis von Theorem 2.11 betrachtete Determinantenform ist bzgl. der Basis $b_1:=v_1,\ldots,b_n:=v_n$, dann gilt

$$\Delta(v_1,\ldots,v_n)=1.$$

Nach Theorem 2.11 ist jede andere nicht-triviale Determinantenform $\Delta'$ ein nicht-triviales Vielfaches von $\Delta$. Also gilt

$$\Delta'(v_1,\ldots,v_n)\neq 0$$

in jedem Fall. Dies zeigt, dass b) aus a) folgt. □

**2.13 Theorem.** *Die Determinante* det *ist die Funktion $\Delta\colon M(n\times n,K)\to K$, die eindeutig charakterisiert wird durch folgende Eigenschaft:*

*a)* *$\Delta$ ist linear in den Spalten, d.h.*

$$\Delta(a_1,\ldots,a_{j-1},\lambda'a_j'+\lambda''a_j'',a_{j+1},\ldots,a_n)=\lambda'\Delta(a_1,\ldots,a_j',\ldots,a_n)+\lambda''\Delta(a_1,\ldots,a_j'',\ldots,a_n)$$

*für alle $\lambda',\lambda''\in K$, $a_1,\ldots,a_{j-1},a_j',a_j'',a_{j+1},\ldots,a_n\in K^n$ und $j=1,\ldots,n$.*

*b)* *$\Delta$ ist alternierend in den Spalten, d.h. sind zwei Spalten von $A$ gleich, dann gilt $\Delta(A)=0$.*

*c)* $\Delta(E_n)=1$

**Beweis:** Im Beweis von Theorem 2.11 haben wir gesehen, dass $\Delta(A) := \det(A)$ die Eigenschaften (a)-(c) erfüllt. Dies zeigt die Existenz von $\Delta$. Wenn $\Delta'$ eine weitere Funktion ist mit (a)-(c), dann muss $\Delta'$ auch eine Determinantenform auf $K^n$ induzieren via Spalten wegen (a) und (b). Wegen Theorem 2.11 gilt $\Delta' = \lambda \det$ für ein $\lambda \in K$. Mit (c) folgt

$$1 = \Delta'(E_n) = \lambda \det(E_n) = \lambda \cdot 1 = \lambda$$

und somit gilt $\Delta' = \det$ wie gewünscht. □

**2.14 Beispiel.** Sei $A = \left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right)$. Dann gilt

$$\det(A) = a_{11}a_{22} - a_{21}a_{12}$$

Analog zeigt man, dass die Determinante einer $3 \times 3$-Matrix $A = \left(\begin{smallmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{smallmatrix}\right)$ mit der **Regel von Sarrus**

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

berechnen lässt (siehe Aufgabe 13.3).

Es sei $\varphi\colon V \to V$ eine lineare Selbstabbildung von $V$. Eine lineare Selbstabbildung heißt **Endomorphismus**. Die Menge der Endomorphismen von $V$ bezeichnen wir mit

$$\mathrm{End}_K(V) := \mathrm{Hom}_K(V,V).$$

Wir erinnern daran, dass $\mathrm{End}_K(V)$ ein Ring ist bezüglich der Verknüpfung von Abbildungen. Dies folgt sofort aus der entsprechenden Aussage für quadratische Matrizen (siehe Korollar V.1.9).

**2.15 Proposition.** *Für $\varphi \in \mathrm{End}_K(V) \Longrightarrow \exists! d_\varphi \in K$ mit folgender Eigenschaft: Für alle Determinantenformen $\Delta$ gilt:*

$$\Delta(\varphi(v_1),\dots,\varphi(v_n)) = d_\varphi \cdot \Delta(v_1,\dots,v_n) \tag{VI.2}$$

*für alle $v_1,\dots,v_n \in V$.*

**Beweis:** Weil $\varphi$ eine lineare Selbstabbildung ist, muss

$$\Delta'\colon V^n \to K, (v_1,\dots,v_n) \mapsto \Delta(\varphi(v_1),\dots,\varphi(v_n))$$

eine Determinantenform sein. Nach Theorem 2.11 gilt

$$\Delta' = d_\varphi \cdot \Delta$$

für genau ein $d_\varphi \in K$, wenn $\Delta$ eine nicht-triviale Determinantenform ist. Also gilt (VI.2) für dieses ausgewählte $\Delta$. Weil jede andere Determinantenform auf $V$ ein Vielfaches von $\Delta$ ist, gilt (VI.2) auch für alle anderen Determinantenformen. □

**2.16 Definition.** Wir definieren $d_\varphi$ als die **Determinante von $\boldsymbol{\varphi}$** und bezeichnen sie mit $\det(\varphi)$.

**2.17 Proposition.** *Wenn $A(\varphi) \in M(n \times n, K)$ die Matrix von $\varphi$ ist bzgl. der Basis $b_1,\dots,b_n$ von $V$, dann gilt* $\det(\varphi) = \det\bigl(A(\varphi)\bigr)$.

**Beweis:** Es seien $a_1,\dots,a_n$ die Koordinatenvektoren von $\varphi(b_1),\dots,\varphi(b_n)$ bzgl. der Basis $b_1,\dots,b_n$. Wir nehmen dieselbe nicht-triviale Determinantenform $\Delta$, die wir im Beweis von Theorem 2.11 konstruiert haben. Man muss dabei die Koordinatenvektoren als Spalten einer $n \times n$-Matrix schreiben und dann die Determinante dieser Matrix nehmen. $\Longrightarrow \Delta(\varphi(b_1),\dots,\varphi(b_n)) = \det(a_1,\dots,a_n) = \det A(\varphi)$ Weiter gilt

$$\Delta(b_1,\dots,b_n) = 1$$

Mit Proposition 2.15 folgt, dass $\det(\varphi) := d_\varphi = \det A(\varphi)$. □

**2.18 Proposition.** *Für $\varphi \in \operatorname{End}_K(V)$ sind folgende Aussagen äquivalent:*

*i)* $\varphi$ *invertierbar im Endomorphismenring* $\operatorname{End}_K(V)$*;*

*ii)* $\exists \psi \in \operatorname{End}_K(V)$*; mit* $\varphi \circ \psi = \mathrm{id}_V$*;*

*iii)* $\exists \psi' \in \operatorname{End}_K(V)$ *mit* $\psi' \circ \phi = \mathrm{id}_V$*;*

*iv)* $\varphi$ *bijektiv;*

*v)* $\varphi$ *surjektiv;*

*vi)* $\varphi$ *injektiv;*

*vii)* $\operatorname{Rang}(\varphi) = \dim(V)$*;*

*viii)* $\varphi^*$ *invertierbar;*

*ix)* $\det(\varphi) \neq 0$.

**Beweis:** Durch die Wahl *einer* festen Basis von $V$ erhält man äquivalente Aussagen für $n \times n$ Matrizen. Deshalb folgt die Äquivalenz von i)-viii) aus Proposition V.2.5.

„v) $\Longrightarrow$ ix)“: Es sei $b_1, \dots, b_n$ eine Basis von $V$. Dann ist $\varphi(b_1), \dots, \varphi(b_n)$ Erzeugendensystem von $\varphi(V)$. Weil $\varphi$ als surjektiv in v) vorausgesetzt wird, muss $\varphi(b_1), \dots, \varphi(b_n)$ ein Erzeugendensystem von $V$ sein. Weil $n = \dim(V)$, muss $\varphi(b_1), \dots, \varphi(b_n)$ eine Basis von $V$ sein (Korollar IV.2.6). Sei $\Delta$ eine nicht-triviale Determinantenform. Aus Korollar 2.12 folgt dann, dass $\Delta(\varphi(b_1), \dots, \varphi(b_n)) \neq 0$. Wegen

$$\Delta\big(\varphi(b_1), \dots, \varphi(b_n)\big) \overset{2.15}{=} \det(\varphi)\Delta(b_1, \dots, b_n)$$

folgt $\det(\varphi) \neq 0$ und damit ix).

„ix) $\Longrightarrow$ v)“: Wieder sei $b_1, \dots, b_n$ eine Basis von $V$. Dann folgt aus

$$\Delta\big(\varphi(b_1), \dots, \varphi(b_n)\big) \overset{2.15}{=} \det(\varphi)\Delta(b_1, \dots, b_n)$$

und der Annahme $\det(\varphi) \neq 0$ aus ix), dass $\Delta\big(\varphi(v_1), \dots, \varphi(n)\big) \neq 0$, denn $\Delta(b_1, \dots, b_n) \neq 0$ nach Korollar 2.12. Noch einmal Korollar 2.12 zeigt, dass $\varphi(b_1), \dots, \varphi(b_n)$ eine Basis von $V$ ist. Also gilt $\varphi(V) = V$. Dies zeigt v). □

**2.19 Proposition.** *Für $\varphi, \psi \in \operatorname{End}_K(V)$ gilt* $\det(\psi \circ \varphi) = \det(\psi)\det(\varphi)$.

**Beweis:** Für $v_1, \dots, v_n \in V$ und eine Determinantenform $\Delta$ gelten

$$\Delta\big(\psi \circ \varphi(v_1), \dots, \psi \circ \varphi(v_n)\big) = \det(\psi \circ \varphi)\Delta(v_1, \dots, v_n)$$

und

$$\begin{aligned}\Delta\Big(\psi\big(\varphi(v_1)\big), \dots, \psi\big(\varphi(v_n)\big)\Big) &= \det(\psi)\Delta\big(\varphi(v_1), \dots, \varphi(v_n)\big)\\ &= \det(\psi)\det(\varphi)\Delta(v_1, \dots, v_n)\end{aligned}$$

mit Proposition 2.15. Es folgt

$$\det(\psi \circ \varphi) = \det(\psi)\det(\varphi)$$

wie gewünscht. □

**2.20 Bemerkung.** Es folgt sofort

$$\det(\psi \circ \varphi) = \det(\varphi \circ \psi).$$

Analoge Aussagen haben wir dank Proposition 2.13 und der Übersetzung von linearen Abbildungen in Matrizen (siehe Bemerkung V.1.19) für $n \times n$ Matrizen $A, B$:

$$\det(A \cdot B) = \det(A) \cdot \det(B) = \det(B \cdot A)$$

## 3. Eigenschaften von Determinanten

Es sei $K$ ein Körper. In diesem Abschnitt werden wir die wesentlichen Eigenschaften der Determinanten von $n \times n$ Matrizen mit Einträgen in $K$ erarbeiten. Mit Hilfe von Proposition 2.17 erhält man entsprechende Aussagen für Endomorphismen eines endlich dimensionalen $K$-Vektorraum.

**3.1 Bemerkung.** In Aufgabe 13.4 zeigen wir für $A \in M(n \times n, K)$, dass

$$\det(A) = \det(A^t)$$

gilt. Als Anwendung sehen wir, dass alle Eigenschaften der Determinante bzgl. den Spaltenvektoren analog für die Zeilenvektoren gelten. Übrigens folgt

$$\det(\varphi) = \det(\varphi^*)$$

aus der obigen Identität und aus $A(\varphi^*) = A(\varphi)^t$, wenn man die duale Basis benutzt.

**3.2 Definition.** Eine Matrix $A \in M(n \times n, K)$ heißt **obere Dreiecksmatrix**, falls $a_{ij} = 0$ für $i > j$, d.h.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

**3.3 Proposition.** *Wenn $A$ eine obere Dreiecksmatrix ist, dann gilt* $\det(A) = a_{11}a_{22}\cdots a_{nn}$.

**Beweis:** Wir erinnern an die Definition

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sign}(\pi) \cdot a_{\pi(1)1} \cdot a_{\pi(2)2} \cdots a_{\pi(n)n}.$$

Wenn $\pi \neq \mathrm{id}$, dann $\exists j \in \{1, \ldots, n\}$ mit $\pi(j) > j$. Dann ist der entsprechende Summand in $\det(A)$ gleich 0, weil $A$ eine obere Dreiecksmatrix ist. Also bleibt nur $\pi = \mathrm{id}$ und damit folgt die Behauptung. $\square$

Als nächstes betrachten wir die Determinante unter den Zeilenoperationen des Gauß-Algorithmus. Das wird nützlich sein, für die Berechnung der Determinante.

**3.4 Proposition.** *Für eine Matrix $A = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in M(n \times n, K)$ mit Zeilen $z_1, \ldots, z_n$ gilt:*

*a) Wenn wir $A'$ durch Multiplikation der $j$-ten Zeile mit $\lambda \in K$ aus $A$ erhalten für ein* einziges $j \in \{1, \ldots, n\}$, *dann gilt*

$$\det(A') = \lambda \det(A).$$

*b) Wenn wir $A'$ durch Vertauschung zweier Zeilen erhalten, dann gilt*

$$\det(A') = -\det(A).$$

*c) Wann wir $A'$ dadurch aus $A$ erhalten, in dem wir zur $i$-ten Zeile noch ein Vielfaches einer* anderen *Zeile dazu addieren, dann gilt*

$$\det(A') = \det(A).$$

*Alles gilt sinngemäß für Spalten- statt Zeilenoperationen. Beachte, dass die Determinante im Allgemeinen nicht linear ist, d.h.*

$$\det(A + B) \neq \det(A) + \det(B).$$

*Wegen a) gilt*

$$\det(\lambda A) = \lambda^n \det(A).$$

**Beweis:** a) und b) ist klar, weil die Determinante nach Theorem 2.13 und Bemerkung 3.1 linear und alternierend in den Zeilen ist. Sei $A \in M(n \times n, K)$ mit den Zeilenvektoren $z_1, \dots, z_n \in K^n$. Dann gilt

$$\det(A') = \det \begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ z_i + \lambda z_j \\ z_{i+1} \\ \vdots \\ z_n \end{pmatrix} = \det(A) + \lambda \det \begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ z_j \\ z_{i+1} \\ \vdots \\ z_n \end{pmatrix}$$

wegen der Linearität der Determinanten in der $i$-ten Zeile. Wegen $i \neq j$ hat die Matrix des zweiten Summanden zweimal den Zeilenvektor $z_j$. Nach b) muss sie deshalb Determinante 0 haben und damit folgt c). $\square$

**3.5 Beispiel.** Wir benutzen Proposition 3.4 um eine Determinante mit Hilfe des Gauss-Algorithmus zu berechnen

$$\det \begin{pmatrix} 0 & 1 & 1 & 1 \\ 2 & -1 & 1 & -1 \\ 4 & 1 & 1 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix} = -\det \begin{pmatrix} \boxed{2} & -1 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 4 & 1 & 1 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix} \begin{matrix} \cdot(-2) \;\; \cdot(-1) \\ \\ + \\ + \end{matrix}$$

$$= -\det \begin{pmatrix} 2 & -1 & 1 & -1 \\ 0 & \boxed{1} & 1 & 1 \\ 0 & 3 & -1 & 3 \\ 0 & 2 & 0 & 1 \end{pmatrix} \begin{matrix} \\ \cdot(-3) \;\; \cdot(-2) \\ + \\ + \end{matrix} = -\det \begin{pmatrix} 2 & -1 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & -2 & -1 \end{pmatrix}$$

$$\underset{3.4\ a)}{=} 4 \det \begin{pmatrix} 2 & -1 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & -2 & -1 \end{pmatrix} \begin{matrix} \\ \\ \cdot 2 \\ + \end{matrix} = 4 \det \begin{pmatrix} 2 & -1 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$= 4 \cdot 2 \cdot 1 \cdot 1 \cdot (-1) = -8$$

**3.6 Proposition.** *Sei $A \in M(n \times n, K)$ von der Form $A = \left(\begin{smallmatrix} A_1 & B \\ 0 & A_2 \end{smallmatrix}\right)$ mit $A_i \in M(n_i \times n_i, K)$ und $n_1 + n_2 = n, B \in M(n_1 \times n_2, K)$, dann gilt*

$$\det(A) = \det(A_1)\det(A_2).$$

**Beweis:** Mit Zeilenoperationen wie aus dem Gauß-Algorithmus kann man $(A_1 B)$ auf Zeilenstufenform bringen und erhält

$$\left(\begin{array}{cccc|c} a'_{11} & \cdots & \cdots & a'_{1n_1} & \\ 0 & \ddots & & \vdots & B' \\ \vdots & \ddots & \ddots & \vdots & \\ 0 & \cdots & 0 & a'_{n_1 n_1} & \end{array}\right)$$

mit $B' \in M(n_1 \times n_2, K)$. Dabei muss man nur Zeilenoperationen aus Proposition 3.4 b) und c) machen und mit Proposition 3.3 folgt

$$\det(A_1) = (-1)^{N_1} a'_{11} a'_{22} \cdots a'_{n_1 n_1},$$

wobei $N_1$ die Anzahl Zeilenvertauschungen ist. Analog kann man $A_2$ auf die Form

$$\begin{pmatrix} a'_{n_1+1,n_1+1} & \cdots & \cdots & a'_{n_1+1,n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a'_{n,n} \end{pmatrix}$$

und somit gilt

$$\det(A_2) = (-1)^{N_2} a_{n_1+1,n_1+1} \cdots a'_{n,n},$$

wobei $N_2$ wiede die Anzahl Zeilenvertauschungen ist. Wenn man dieselben Operationen auf die ganze Matrix $A$ anwendet, erhält man

$$\begin{pmatrix} a'_{11} & \cdots & \cdots & a'_{1,n_1} & & & & \\ 0 & \ddots & & \vdots & & B' & & \\ \vdots & \ddots & \ddots & \vdots & & & & \\ 0 & \cdots & 0 & a'_{n_1,n_1} & & & & \\ & & & & a'_{n_1+1,n_1+1} & \cdots & \cdots & a'_{n_1+1,n} \\ & & & & 0 & \ddots & & \vdots \\ & \mathbf{0} & & & \vdots & \ddots & \ddots & \vdots \\ & & & & 0 & \cdots & 0 & a'_{n,n} \end{pmatrix},$$

wobei $N_1 + N_2$ Zeilenvertauschungen vorkommen. Also gilt

$$\det(A) = (-1)^{N_1+N_2} a'_{11} \cdots a'_{nn}$$

Setzen wir alles zusammen, folgt

$$\det(A) = (-1)^{N_1} a'_{11} \cdots a'_{n_1 n_1} \cdot (-1)^{N_2} a'_{n_1+1,n_1+1} \cdots a'_{n,n} = \det(A_1)\det(A_2)$$

und damit die Behauptung. □

**3.7 Beispiel.** Gegeben seien $\lambda_1, \ldots, \lambda_n \in K$. Wir wollen die Gleichung

$$\det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & & \lambda_n^{n-1} \end{pmatrix} = \prod_{1 \le i < j \le n} (\lambda_j - \lambda_i)$$

zeigen. Diese berühmte Determinante heißt **Vandermondsche Determinante** und wird oft gebraucht.

Wir beweisen dies per Induktion über $n$. Der Induktionsanfang $n = 1$ ist trivial, da das Produkt über die leere Indexmenge per Konvention als 1 erklärt wird.

Induktionsschritt $n - 1 \mapsto n$: Nach Proposition 3.4 lassen folgende Zeilenoperationen die Vandermondsche Determinante invariant:

$$n\text{-te Zeile } -\lambda_1 (n-1) \text{ Zeile}, \ (n-1) \text{ Zeile } -\lambda_1 \cdot (n-2)\text{-te Zeile}, \ldots, 2 \text{ te Zeile } -\lambda_1 \text{ 1. Zeile}.$$

Also gilt:

$$
\begin{aligned}
\det\begin{pmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_n \\ \vdots & & \vdots \\ \lambda_1^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix} &= \det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & \lambda_2-\lambda_1 & & \lambda_n-\lambda_1 \\ \vdots & \vdots & & \vdots \\ 0 & \lambda_2^{n-2}(\lambda_2-\lambda_1) & \cdots & \lambda_n^{n-2}(\lambda_n-\lambda_1) \end{pmatrix} \\
&= 1\cdot\det\begin{pmatrix} \lambda_2-\lambda_1 & \cdots & \lambda_n-\lambda_1 \\ \vdots & & \vdots \\ \lambda_2^{n-2}(\lambda_2-\lambda_1) & \cdots & \lambda_n^{n-2}(\lambda_n-\lambda_1) \end{pmatrix} \\
&= (\lambda_2-\lambda_1)\cdots(\lambda_n-\lambda_1)\cdot\det\begin{pmatrix} 1 & \cdots & 1 \\ \lambda_2 & \cdots & \lambda_n \\ \vdots & & \vdots \\ \lambda_2^{n-2} & \cdots & \lambda_n^{n-2} \end{pmatrix} \\
&\overset{\text{IV}}{=} (\lambda_2-\lambda_1)\cdots(\lambda_n-\lambda_1)\prod_{2\le i<j\le n}(\lambda_j-\lambda_i) = \prod_{1\le i<j\le n}(\lambda_j-\lambda_i).
\end{aligned}
$$

Dies zeigt den Induktionsschritt und per Induktion folgt die Behauptung. □

## 4. Adjungierte Matrix

In diesem Abschnitt werden wir sehen, wie die Determinante einer $n\times n$-Matrix mit den Determinanten ihrer $(n-1)\times(n-1)$-Teilmatrizen zusammenhängt. Das führt auf die adjungierte Matrix, die für die Lösung inhomogener linearer Gleichungssysteme und für die Inverse einer Matrix wichtig ist.

Wie immer bezeichne $K$ einen Körper. Weiter sei $A \in M(n\times n, K)$.

**4.1 Definition.** Für $1 \le i, j \le n$ definieren wir $A_{ij} \in M(n\times n, K)$ durch

$$
A_{ij} := \begin{pmatrix}
a_{11} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\
0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn}
\end{pmatrix}
$$

und $A'_{ij} \in M((n-1)\times(n-1), K)$ durch Streichen der $i$-ten Zeile und der $j$-ten Spalte von $A$:

$$
A'_{ij} := \begin{pmatrix}
a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\
\vdots & & \vdots & \vdots & & \vdots \\
a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\
a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\
\vdots & & \vdots & \vdots & & \vdots \\
a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn}
\end{pmatrix}
$$

**4.2 Lemma.** *Wenn $a_1,\dots,a_n$ die Spalten der Matrix $A$ bezeichnen, dann gilt*

$$
\det(A_{ij}) = (-1)^{i+j}\det(A'_{ij}) = \det(a_1,\dots,a_{j-1},e_i,a_{j+1},\dots,a_n).
$$

**Beweis:** Durch $i-1$ Zeilenvertauschungen und $j-1$ Spaltenvertauschungen erhalten wir

$$\det(A_{ij}) = (-1)^{i+j} \det \left( \begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A'_{ij} & \\ 0 & & & \end{array} \right)$$

nach Proposition 3.4 b). Mit Hilfe von Proposition 3.6 folgt die erste Behauptung. Für $k \neq j$ subtrahiert man von der $k$-ten Spalte der Matrix $(a_1, \ldots, a_{j-1}, e_i, a_{j+1}, \ldots, a_n)$ des $a_{ik}$-fache der $j$-ten Spalte $(= e_i)$ und erhält die Matrix $A_{ij}$. Nach Proposition 3.4 c) ändert sich dabei die Determinante nicht, d.h.

$$\det(a_1, \ldots, a_{j-1}, e_i, a_{j+1}, \ldots, a_n) = \det(A_{ij}).$$ □

**4.3 Definition.** Wir nennen $\det(A_{ij})$ den **Cofaktor** zu $(i, j)$ und

$$A^{ad} := \big(\det(A_{ji})\big)_{1 \leq i \leq n, 1 \leq j \leq n} \in M(n \times n, K)$$

heißt **die zu A adjungierte Matrix**. Beachte, dass sie *transponiert* ist zur Matrix der Cofaktoren, da $i \leftrightarrow j$.

**4.4 Theorem.** *Es gilt* $A^{ad} \cdot A = A \cdot A^{ad} = \det(A) \cdot E_n$

**Beweis:** Wir berechnen den $(i, k)$-ten Eintrag von $A^{ad} \cdot A$:

$$\begin{aligned}
(A^{ad} \cdot A)_{i,k} &= \sum_{j=1}^{n} \det(A_{ji}) \cdot a_{jk} = \sum_{j=1}^{n} a_{jk} \det(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n) \\
&= \det(a_1, \ldots, a_{i-1}, \sum_{j=1}^{n} a_{jk} \cdot e_j, a_{i+1}, \ldots, a_n) \\
&= \det(a_1, \ldots, a_{i-1}, a_k, a_{i+1}, \ldots, a_n)
\end{aligned}$$

Wenn $i = k$, dann ist das gleich $\det(A)$. Falls $i \neq k$ ist dies nach 2.13 gleich 0. Dies zeigt $A^{ad} \cdot A = \det(A) \cdot E_n$. Benutzt man dies für $A^t$ statt $A$ und die triviale Gleichung $(A^t)^{ad} = (A^{ad})^t$ erhalten wir:

$$(A^{ad})^t \cdot A^t = \det(A^t) \cdot E_n.$$

Mit V.1.19 folgt:

$$(A \cdot A^{ad})^t = \det(A^t) \cdot E_n = \det(A) \cdot E_n.$$

Erneutes Transponieren liefert $A \cdot A^{ad} = \det(A) \cdot E_n$. Dies zeigt die Behauptungen. □

**4.5 Korollar.** *Man hat folgende Entwicklung der Determinante nach der $j$-ten Spalte:*

$$\det(A) = \sum_{i=1}^{n} a_{ij} \cdot \det(A_{ij}) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \det(A'_{ij}).$$

*Damit lässt sich die Determinante der $n \times n$-Matrix $A$ aus den Determinanten der $(n-1) \times (n-1)$ Matrizen $A'_{ij}$ berechnen.*

**Beweis:** Die erste Behauptung ist die Auswertung von Theorem 4.4 im Eintrag $(j, j)$ und die zweite Behauptung ergibt sich aus Lemma 4.2. □

**4.6 Korollar.** *Falls $A \in \mathrm{GL}(n, K)$, dann gilt $A^{-1} = \det(A)^{-1} \cdot A^{ad}$.*

**Beweis:** Da $A$ invertierbar ist, gilt $\det(A) \neq 0$. Dies folgt aus dem Matrixanalogon von Proposition 2.18. Also folgt Korollar 4.6 direkt aus Theorem 4.4. □

**4.7 Beispiel.** Sei $A = \left(\begin{smallmatrix} 1 & 1 \\ 2 & 3 \end{smallmatrix}\right)$. Dann ist $\det(A'_{11}) = 3$, $\det(A'_{12}) = 2$, $\det(A'_{21}) = 1$, $\det(A'_{22}) = 1$ und damit mit den Vorfaktoren $(-1)^{i+j}$ und Transponieren: $A^{ad} = \left(\begin{smallmatrix} 3 & -1 \\ -2 & 1 \end{smallmatrix}\right)$. Nach Korollar 4.6 gilt $A^{-1} = \det(A)^{-1} \cdot \left(\begin{smallmatrix} 3 & -1 \\ -2 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 3 & -1 \\ -2 & 1 \end{smallmatrix}\right)$ da offenbar $\det(A) = 1$. Dies beendet Beispiel V.2.4

**4.8 Korollar (Cramersche Regel).** *Sei $A \in \mathrm{GL}(n,K)$ und $b \in K^n$. Dann hat das lineare Gleichungssystem $A \cdot x = b$ genau eine Lösung $\lambda \in K^n$. Es gilt*

$$\lambda_i = (\det(A))^{-1} \cdot \det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n),$$

*wobei $a_j$ den $j$-ten Spaltenvektor von $A$ beschreibt für $j = 1, \ldots, n$.*

**Beweis:** In Korollar V.3.6 haben wir gesehen, dass $\lambda = A^{-1}b$ die einzige Lösung ist. Nach Korollar 4.6 ist der $(i, j)$-te Eintrag von $A^{-1}$ gleich

$$\det(A)^{-1} \cdot \det(A_{ji}) \underset{\text{Lemma 4.2}}{=} \det(A)^{-1} \cdot \det(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n)$$

Es folgt:

$$\begin{aligned} \lambda_i = (A^{-1}b)_i &= \sum_{j=1}^{n} \det(A)^{-1} \cdot \det(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n) \cdot b_j \\ &= \det(A)^{-1} \cdot \det(a_1, \ldots, a_{i-1}, \underbrace{\sum_{j=1}^{n} b_j e_j}, a_{i+1}, \ldots, a_n) \\ &= \det(A)^{-1} \cdot \det(a_1, \ldots, a_{i-1}, \quad b \quad, a_{i+1}, \ldots, a_n). \end{aligned}$$

□

Wir geben noch 3 interessante Resultate ohne Beweis an. Für $1 \leq r \leq n$ betrachten wir folgende Verallgemeinerungen der Entwicklungssatzes nach der $j$-ten Zeile (Korollar 4.5).

**4.9 Theorem.** *Es ist*

$$\det(A) = \sum_{\substack{H \subseteq \{1,\ldots,n\}, \\ |H|=r}} \rho_H \cdot \det(a_{ij})_{\substack{1 \leq i \leq r, \\ j \in H}} \cdot \det(a_{kl})_{\substack{r+1 \leq k \leq n, \\ l \in H^c}}$$

*wobei $H^c$ das Komplement von $H$ in $\{1, \ldots, n\}$ beschreibt und $\rho_H := (-1)^\nu$, mit $\nu := |\{(h, h') \in H \times H^c \mid h > h'\}|$*

**Beweis:** Siehe [Bo] Abschnitt 4.5, insbesondere Theorem 4.5.7. □

Es gilt auch folgende Verallgemeinerung des Determinanten-Multiplikationssatzes (Proposition 2.19).

**4.10 Theorem.** *Es sei $m \leq n$ und $A, B \in M(m \times n, K)$. Dann gilt*

$$\det(A \cdot B^t) = \sum_{\substack{H \subseteq \{1,\ldots,n\}, \\ |H|=m}} \det(A_H) \cdot \det(B_H),$$

*wobei* $\det(A_H)$ *der* ***m-reihe Minor*** *von $A$ ist gegeben durch*

$$\det(A_H) := \det(a_{ij})_{\substack{1 \leq i \leq m, \\ j \in H}}$$

**Beweis:** [KoMi, §34] oder [Fi, 3.3.7]. □

**4.11 Bemerkung.** Im Spezialfall $A = B$ nennt man $\det(A \cdot A^t)$ die **Gramsche-Determinante**. Falls $K = \mathbb{R}$, dann ist das Volumen des Parallelotops aufgespannt durch die $m$ Zeilenvektoren von $A$ gleich $\det(A \cdot A^t)^{1/2}$ (siehe [Bo, Korollar 7.2.10]). Wenn $m = n$, dann ist das Volumen gleich $\big(\det(A) \cdot \det(A^t)\big)^{1/2} = |\det(A)|$.

7 Kapitel VII.

# Eigenwerte

## 1. Ergänzungen zu Polynomen und Determinanten

Wir werden im diesem Abschnitt die algebraische Grundlagen über Polynome erweitern im Hinblick auf die Anwendungen in diesem Kapitel. Wir hatten in Abschnitt II.3 den Ring der Polynome $K[x]$ studiert. Als wichtigste Eigenschaft hatten wir die Polynomdivision und das Abspalten von Nullstellen. Jetzt wollen wir zuerst das Studium der Nullstellen fortführen, ihre Multiplizität und den Fundamentalsatz der Algebra ohne Beweis formulieren. Dann werden wir Polynome in mehreren Variablen mit Koeffizienten in einem kommutativen Ring $R$ einführen. Wir werden sehen, dass sie einen kommutativen Ring bilden. Allerdings steht uns nun die Polynomdivision nicht mehr zur Verfügung. Wir benutzen diesen Abstecher in die Algebra um Determinanten von quadratischen Matrizen mit Einträgen in $R$ zu definieren. Um zu zeigen, dass sie dieselben Eigenschaften haben wie im Fall der Körper $K$, betrachten wir die Koeffizienten als Variablen und erhalten universelle Formeln, die auch gelten, wenn man Ringelemente einsetzt.

In diesem Abschnitt ist $K$ ein Körper und $R$ ein *kommutativer* Ring.

**1.1 Proposition.** *Sei $p(x) \in K[x] \setminus \{0\}$ und $\alpha \in K$. Dann existiert genau ein $q(x) \in K[x]$ mit $q(\alpha) \neq 0$ und ein $m \in \mathbb{N}$ so, dass gilt:*

$$p(x) = (x-\alpha)^m \cdot q(x)$$

**Beweis:** *Existenz:* Wenn $p(\alpha) \neq 0$, setzen wir $q := p$. Wenn $p(\alpha) = 0$, dann spalten wir die Nullstelle $\alpha$ ab (Korollar II.3.12) und erhalten $p(x) = (x-\alpha)p_1(x)$. Dann wiederholen wir das Verfahren für $p_1$ statt $p$, usw. Das Verfahren bricht spätestens nach $d := \deg(p)$ Schritten ab, weil dann $p_d \in K \setminus \{0\}$.

*Eindeutigkeit:* Sei $p(x) = (x-\alpha)^n \cdot r(x)$, mit $r(\alpha) \neq 0$ eine weitere Darstellung. OBdA sei $m \geq n$. Wegen der Eindeutigkeit der Polynomdivision (Theorem II.3.11) ergibt sich $(x-\alpha)^{m-n} \cdot q(x) = r(x)$. Wegen $r(\alpha) \neq 0$ folgt $n = m$ und somit $q(x) = r(x)$. □

**1.2 Definition.** Wir nennen die Zahl $m$ aus Proposition 1.1 die **Multiplizität** oder die **Vielfachheit** der Nullstelle $\alpha$ von $p(x)$.

**1.3 Bemerkung.** Die Multiplizität $m$ darf auch 0 sein. Das bedeutet, dass $\alpha$ keine Nullstelle von $p(x)$ ist. Wenn $m \geq 2$ ist, dann sprechen wir von einer **mehrfachen** Nullstelle.

**1.4 Korollar.** *Sei $p(x) \in K[x] \setminus \{0\}$ mit $\deg(p) = d$. Dann gibt es höchstens $d$ verschiedene Nullstellen in $K$, die wir mit $\alpha_1, \ldots, \alpha_r$ bezeichnen. Wenn $m_1, \ldots, m_r$ ihre Multiplizitäten sind, dann gilt*

$$p(x) = (x-\alpha_1)^{m_1} \cdots (x-\alpha_r)^{m_r} \cdot q(x)$$

*für ein eindeutig bestimmtes Polynom $q(x) \in K[x]$ ohne Nullstelle in $K$.*

**Beweis:** Klar durch sukzessives Anwenden von Proposition 1.1. Die Eindeutigkeit von $q(x)$ ergibt sich aus der Eindeutigkeit der Polynomdivision (Theorem II.3.11). □

**1.5.** Es kann passieren, dass $p(x)$ überhaupt keine Nullstelle in $K$ besitzt, z.B. hat $p(x) = x^2 + 1$ keine Nullstelle in $\mathbb{R}$. Denn $\alpha^2 + 1 \geq 1$ für alle reelle Zahlen $\alpha$. Optimal ist der Fall wenn in Korollar 1.4 $\deg(q) = 0$ gilt. Dafür benutzen wir folgende:

**1.6 Definition.** Wir sagen, dass $p(x) \in K[x] \setminus \{0\}$ **vollständig in Linearfaktoren aus K[x] zerfällt**, wenn

$$p(x) = a_d(x-\alpha_1)^{m_1} \cdots (x-\alpha_r)^{m_r}$$

gilt für $a_d \in K$, $m_1 \geq 1, \ldots, m_r \geq 1$ und verschiedene $\alpha_1, \ldots, \alpha_r \in K$.

**1.7 Bemerkung.** In obiger Definition ist klar, dass $\alpha_1, \ldots, \alpha_r$ alle Nullstellen von $p(x)$ in $K$ und $m_1, \ldots, m_r$ ihre Multiplizitäten sind.

Durch Ausmultiplizieren sehen wir, dass $a_d$ der höchste Koeffizient von $p(x)$ ist, d.h. $p(x) = \sum_{i=0}^{d} a_i x^i$ mit $a_d \neq 0$ der konstante Faktor aus Definition 1.6.

**1.8 Theorem (Fundamentalsatz der Algebra).** *Jedes komplexe Polynom vom Grad $\geq 1$ hat mindestens 1 komplexe Nullstelle.*

**Beweis:** Mit Hilfe des Satzes von *Liouville* wird in der Funktionentheorie ein sehr eleganter Beweis gegeben (Analysis III oder [Ahlfors]). Für einen Beweis mit Analysis I siehe [Gu, Satz 5.4.10]. □

**1.9 Korollar.** *Jedes Polynom $p(x) \in \mathbb{C}[x] \setminus \{0\}$ zerfällt vollständig in Linearfaktoren über $\mathbb{C}[x]$.*

**Beweis:** Nach Korollar 1.4 gilt $p(x) = (x-\alpha_1)^{n_1} \cdots (x-\alpha_r)^{n_r} q(x)$ mit $q(x)$ ohne Nullstelle in $\mathbb{C}$. Nach dem Fundamentalsatz der Algebra muss $q(x)$ ein konstantes Polynom sein. Das zeigt die Behauptung. □

**1.10 Beispiel.** Ein komplexes Polynom zu faktorisieren ist im Allgemeinen schwierig, da es für $\deg(p) \geq 5$ keine Formel gibt. Meistens sind die Aufgaben aber so gemacht, dass man Nullstellen raten kann, wie im Beispiel:

$$p(x) = x^4 + x^3 - 3x^2 - 4x - 4.$$

Wenn es bei einem Polynom mit ganzzahligen Koeffizienten ganzzahlige Nullstellen gibt, dann folgt leicht mit Polynomdivision, dass diese Nullstellen Teiler von $a_0$ sind. Wir probieren zuerst $\pm 1$, aber dies sind keine Nullstellen. Dann sehen wir $p(2) = 0$. Abspalten von $\alpha_1 = 2$ liefert

$$x^4 + x^3 - 3x^2 - 4x - 4 = (x-2)(x^3 + 3x^2 + 3x + 2).$$

Nun ist 2 keine Nullstelle von $x^3 + 3x^2 + 3x + 2$, aber $\alpha_2 = -2$ ist eine Nullstelle. Abspalten von $-2$ liefert:

$$x^3 + 3x^2 + 3x + 2 = (x+2)(x^2 + x + 1).$$

Mit der Lösungsformel für quadratische Gleichungen folgt

$$\alpha_{3/4} = \frac{-1 \pm \sqrt{1-4}}{2} = \frac{-1 \pm \sqrt{-3}}{2} = -\frac{1}{2} \pm \frac{\sqrt{3}}{2} \cdot \mathrm{i}.$$

In $\mathbb{Q}$ oder $\mathbb{R}$ sieht die Faktorisierung von $p(x)$ aus Korollar 1.4 folgendermaßen aus

$$p(x) = (x-2)(x+2)(x^2 + x + 1).$$

Falls $K = \mathbb{C}$, zerfällt $p(x)$ vollständig in Linearfaktoren aus $\mathbb{C}[x]$.

$$p(x) = (x-2)(x+2)(x + \frac{1}{2} - \frac{\sqrt{3}}{2}\mathrm{i})(x + \frac{1}{2} + \frac{\sqrt{3}}{2}\mathrm{i}).$$

**1.11 Definition.** Es sei $F$ ein Körper. Dann heißt $E \subseteq F$ **Teilkörper** von $F$, falls $E$ mit der Addition und Multiplikation von $F$ selbst ein Körper ist. $F$ heißt dann **Körpererweiterung** von $E$.

Zum Beispiel ist $\mathbb{R}$ eine Körpererweiterung von $\mathbb{Q}$ und $\mathbb{C}$ ist eine Körpererweiterung von $\mathbb{R}$ und somit auch von $\mathbb{Q}$ (siehe Satz II.2.14). Mit Hilfe von Korollar 1.9 kann man folgendes Korollar beweisen (siehe Aufgabe 15.1).

**1.12 Korollar.** *Sei $p(x) \in \mathbb{R}[x] \setminus \{0\}$ und seien $\alpha_1, \dots, \alpha_r$ die verschiedenen reellen Nullstellen von $p$ mit den Multiplizitäten $m_1, \dots, m_r$. Dann gibt es paarweise verschiedene Polynome $q_1, \dots, q_s \in \mathbb{R}[x]$ vom Grad 2 ohne reelle Nullstelle und $n_1, \dots, n_s \in \mathbb{N} \setminus \{0\}$ mit*

$$p(x) = (x-\alpha_1)^{m_1} \cdots (x-\alpha_r)^{m_r} q_1(x)^{n_1} \cdots q_s(x)^{n_s}.$$

**1.13 Korollar.** *Jedes reelle Polynom von ungeradem Grad hat mindestens eine reelle Nullstelle.*

**Beweis:** Mit Korollar 1.12 folgt:

$$\deg(p) = m_1 + \cdots + m_r + 2n_1 + \cdots + 2n_s$$

Links steht eine ungerade Zahl, also gibt es mindestens eine reelle Nullstelle. Einen weiteren Beweis findet man in der Analysis mit Hilfe des Zwischenwertsatzes. □

Nicht jeder Körper ist Teilkörper von $\mathbb{C}$ z.B. ist $\mathbb{Z}/p\mathbb{Z}$ ($p$ Primzahl) sicher kein Teilkörper von $\mathbb{C}$. Denn es gilt $\underbrace{1 + \cdots + 1}_{p\text{-mal}} = 0$ in $\mathbb{Z}/p\mathbb{Z}$. Dann man trotzdem eine Faktorisierung in Nullstellen hat, zeigt folgendes:

**1.14 Theorem.** *Sei $p(x) \in K[x] \setminus \{0\}$. Dann gibt es eine Körpererweiterung $L$ von $K$ so, dass $p(x)$ vollständig in Linearfaktoren aus $L[x]$ zerfällt.*

**Beweis:** Wird in der Algebra bewiesen, siehe [Gu, Satz 5.4.5]. □

**1.15.** Beachte, dass $L$ von der Wahl des Polynoms $p(x)$ abhängen kann. Es ist aber richtig, dass es eine Körpererweiterung $F$ von $K$ gibt, so dass jedes Polynom $p(x) \in F[x] \setminus \{0\}$ vollständig in Linearfaktoren aus $F[x]$ zerfällt, analog zu $\mathbb{C}$. So ein Körper $F$ heißt **algebraisch abgeschlossen**. Die Existenz von $F$ ist viel schwieriger als die von $L$ zu zeigen und $F$ wird im Allgemeinen auch viel größer sein als $L$ (siehe Algebra I).

**1.16 Definition.** Wir definieren die **Determinante** für eine $n \times n$-Matrix $A = (a_{ij})$ mit Einträgen aus einem kommutativen Ring $R$ als

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sign}(\pi) a_{\pi(1)1} \cdots a_{\pi(n)n}$$

analog zur Definition VI.2.3 im Körperfall. Für diese Determinanten im Fall eines kommutativen Rings gelten dieselben Eigenschaften wie in §VI.3 und §VI.4. Man muss nur aufpassen bei der Division und deshalb ist Bedingung $\det(A) \neq 0$ jeweils zu ersetzen durch $\det(A) \in R^*$.

Für die Begründungen werden wir einen Abstecher in die Algebra machen. Wer sich nicht dafür interessiert, kann direkt zu Abschnitt 2 weitergehen.

**1.17 Definition.** Ein **Ringhomomorphismus** ist eine Abbildung $\varphi \colon R_1 \to R_2$ zwischen Ringen so, dass für alle $a, b \in R_1$ gilt:

$$\varphi(a+b) = \varphi(a) + \varphi(b), \quad \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b) \qquad \wedge \qquad \varphi(1) = 1.$$

**1.18.** Wir können auch Polynome in der Variablen $x$ mit Koeffizienten im kommutativen Ring $R$ betrachten. Analog zu II.3 bilden sie einen kommutativen Ring.

Wir nehmen nun an, dass $R$ ein **Teilring** eines beliebigen Ringes $S$ ist, d.h. $R \subseteq S$ und die Addition, Multiplikation und das Einselement stimmen auf $R$ überein. Für jedes $\alpha \in S$ haben wir einen **Einsetzungshomomorphismus**

$$\Phi_\alpha \colon R[x] \to S, \ p(x) \mapsto p(\alpha).$$

Wie in Proposition II.3.8 beweist man, dass $\Phi_\alpha$ ein Ringhomomorphismus ist.

**1.19.** Wir definieren nun den Ring $R[x_1,\ldots,x_n]$ der Polynome in den Variablen $x_1,\ldots,x_n$ rekursiv. Der Fall $R[x_1]$ ist in 1.18 behandelt worden. Wir nehmen nun rekursiv an, dass $R_{n-1} := R[x_1,\ldots,x_{n-1}]$ schon definiert wurde und sehen

$$R[x_1,\ldots,x_n] := R_{n-1}[x_n].$$

Nach 1.18 ist dies ein kommutativer Ring. Neben der rekursiven Darstellung

$$p(x_1,\ldots,x_n) = \sum_{j=0}^{m} p_j(x_1,\ldots,x_{n-1})x_n^j$$

mit $p_j \in R[x_1,\ldots,x_{n-1}]$, können wir auch durch ausmultiplizieren die explizite Darstellung

$$p(x_1,\ldots,x_n) = \sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n} a_{j_1\cdots j_n}x_1^{j_1}\cdots x_n^{j_n}$$

mit $a_{j_1\cdots j_n} \in R$. Wenn $R$ ein Teilring eines *kommutativen* Ringes $S$ ist, dann haben wir für $\alpha_1,\ldots,\alpha_n \in S$ ein Einsetzungshomomorphismus

$$\Phi_{\alpha_1,\ldots,\alpha_n} : R[x_1,\ldots,x_n] \to S, \quad p(x_1,\ldots,x_n) \mapsto p(\alpha_1,\ldots,\alpha_n) = \sum a_{j_1\cdots j_n}\alpha_1^{j_1}\cdots\alpha_n^{j_n}$$

Da $\Phi_{\alpha_1,\ldots,\alpha_n}$ die Verknüpfung von Einsetzungshomomorphismen

$$\begin{aligned} R[x_1,\ldots,x_n] = R[x_1,\ldots,x_{n-1}][x_n] &\xrightarrow{\Phi_{\alpha_n}} S[x_1,\ldots,x_{n-1}] = S[x_1,\ldots,x_{n-2}][x_{n-1}] \\ &\xrightarrow{\Phi_{\alpha_{n-1}}} \cdots \xrightarrow{\Phi_{\alpha_2}} S[x_1] \xrightarrow{\Phi_{\alpha_1}} S \end{aligned}$$

ist, muss $\Phi_{\alpha_1,\ldots,\alpha_n}$ auch ein Ringhomomorphismus sein.

**1.20.** Wir haben einen natürlichen Ringhomomorphismus $\rho : \mathbb{Z} \to R$, definiert durch

$$\rho(n) := \underbrace{1+\cdots+1}_{n\text{-mal}} \quad \wedge \quad \rho(-n) := -\rho(n)$$

für $n \in N$. Die Ringaxiome zeigen tatsächlich, dass dies ein Ringhomomorphismus ist. Wir können $\rho$ erweitern zu einem Ringhomomorphismus

$$\rho_n : \mathbb{Z}[x_1,\ldots,x_n] \to R[x_1,\ldots,x_n], \; \sum a_{j_1,\ldots,j_n}x_1^{j_1}\cdots x_n^{j_n} \mapsto \sum \rho(a_{j_1,\ldots,j_n})x_1^{j_1}\cdots x_n^{j_n}$$

**1.21 Lemma.** *Seien $S_1,\ldots,S_n$ unendliche Teilmengen des Körpers $F$ und es seien $p,q \in F[x_1,\ldots,x_n]$ mit $p(\alpha_1,\ldots,\alpha_n) = q(\alpha_1,\ldots,\alpha_n)$ für alle $(\alpha_1,\ldots,\alpha_n) \in S_1\times\cdots\times S_n$. Dann gilt*

$$p(x_1,\ldots,x_n) = q(x_1,\ldots,x_n).$$

**Beweis:** Mit Induktion über $n$. Der Fall $n = 1$ folgt sofort aus Korollar 1.4.

„$n-1 \mapsto n$“: Sei jetzt $n \geq 2$. Es gilt

$$p(x_1,\ldots,x_n) = \sum_{j=0}^{m} p_j(x_1,\ldots,x_{n-1})x_n^j \quad \wedge \quad q(x_1,\ldots,x_n) = \sum_{j=0}^{m} q_j(x_1,\ldots,x_n)x_n^j$$

für geeignete $p_j,q_j \in K[x_1,\ldots,x_{n-1}]$, die auch Null sein dürfen. Wir wählen $(\alpha_1,\ldots,\alpha_{n-1}) \in S_1\times\cdots\times S_{n-1}$, dann gilt

$$p(\alpha_1,\ldots,\alpha_{n-1},x_n) = q(\alpha_1,\ldots,\alpha_{n-1},x_n)$$

nach dem Fall $n = 1$, da die Gleichheit nach einsetzen aller $\alpha_n \in S_n$ gilt. Also folgt

$$p_j(\alpha_1,\ldots,\alpha_{n-1}) = q_j(\alpha_1,\ldots,\alpha_{n-1})$$

für alle $(\alpha_1,\ldots,\alpha_{n-1}) \in S_1\times\cdots\times S_{n-1}$. Nach dem Induktionsannahme gilt

$$p_j(x_1,\ldots,x_{n-1}) = q_j(x_1,\ldots,x_{n-1})$$

für $j = 0,\ldots,m$ und damit $p(x_1,\ldots,x_n) = q(x_1,\ldots,x_n)$. □

**1.22 Proposition.** *Sei $\varphi\colon R \to S$ ein Homomorphismus kommutativer Ringen, $A = (a_{ij}) \in M(n \times n, R)$ und $\varphi(A) := (\varphi(a_{ij}))$. Dann gilt* $\det(\varphi(A)) = \varphi(\det(A))$.

**Beweis:** Klar aufgrund der Definition der Determinante und der Homomorphismuseigenschaft. □

**1.23 Proposition.** *Für $A, B \in M(n \times n, R)$ gilt* $\det(AB) = \det(A)\det(B)$.

**Beweis:** Wir betrachten die Einträge der Matrizen als Unbekannte, d.h. wir betrachten

$$X := (x_{ij})_{1\le i,j\le n}, Y := (y_{ij})_{1\le i,j\le n} \in M\big(n \times n, , \mathbb{Z}[x_{ij}, y_{ij}]_{1\le i,j\le n}\big).$$

Wir zeigen zuerst

$$\det(X \cdot Y) = \det(X) \cdot \det(Y). \tag{VII.1}$$

Dies ist eine Gleichung von zwei Polynomen in $2n^2$ Variablen und Koeffizienten von $\mathbb{Z}$. Da $\mathbb{Z}$ ein Teilring von $\mathbb{Q}$ ist, können wir (VII.1) auch in $\mathbb{Q}[x_{ij}, y_{ij}]_{1\le i,j\le n}$ beweisen. Nach Proposition VI.2.19 gilt (VII.1), wenn wir rationale Zahlen in die Einträge der Matrix einsetzen. Also folgt (VII.1) aus Lemma 1.21. Nun benutzen wir den Ringhomomorphismus $\varphi$, definiert als Verknüpfung der Homomorphismen

$$\mathbb{Z}[x_{ij}, y_{ij}]_{1\le i,j\le n} \xrightarrow{\rho_{2n^2}} R[x_{ij}, y_{ij}]_{1\le i,j\le n} \xrightarrow{\Phi_{A,B}} R$$

aus 1.19 und 1.20. Weil wir die Matrizen $A, B$ für $X$ und $Y$ einsetzen, erhalten wir $\varphi(X) = A$ und $\varphi(Y) = B$. Mit Proposition 1.22 folgt

$$\begin{aligned}\det(A \cdot B) &= \det\big(\varphi(X) \cdot \varphi(Y)\big) = \varphi\big(\det(X) \cdot \det(Y)\big) \underset{\text{(VII.1)}}{=} \varphi\big(\det(XY)\big)\\ &= \det\big(\varphi(XY)\big) = \det\big(\varphi(X)\varphi(Y)\big) = \det(A \cdot B)\end{aligned}$$

wie gewünscht. □

**1.24 Bemerkung.** Mit derselben Technik kann man zeigen, dass alle Eigenschaften der Determinante aus VI.3 und VI.4 für beliebige kommutative Ring $R$ gelten (außer Bemerkung VI.4.11, die nur für $\mathbb{R}$ gilt). Man ersetzt dabei immer die Einträge der Matrizen durch Unbekannte und zeigt dann die entsprechende polynomiale Gleichungen in mehreren Variablen mit Hilfe von Lemma 1.21 aus dem entsprechenden Resultat für $K = \mathbb{Q}$.

**1.25 Korollar.** *Für $A \in M(n \times n, R)$ gilt: $A \in \mathrm{GL}(n \times n, R) \iff \det(A) \in R^*$.*

**Beweis:** „$\Longrightarrow$“ $\exists B \in M(n \times n, R)$ mit $AB = E_n \overset{1.23}{\Longrightarrow} \det(A)\det(B) = \det(AB) = 1$, d.h. $\det(A) \in R^*$
„$\Longleftarrow$“ Aus der Verallgemeinerung von Theorem VI.4.4. □

## 2. Eigenwerte und Eigenvektoren

Ein Eigenvektor eines Endomorphismus $\varphi$ wird durch $\varphi$ gestreckt (oder gestaucht). Er hat damit eine ausgezeichnete geometrische oder physikalische Bedeutung. Besonders einfach sind solche $\varphi$, die eine Basis aus Eigenvektoren haben. Man nennt sie diagonalisierbar. In diesem Abschnitt kümmern wir uns um die Grundlagen, weiterreichende Resultate betrachten in den folgenden Abschnitten. Wie immer ist $K$ ein Körper, $V$ ein $K$-Vektorraum und $\varphi \in \mathrm{End}_K(V)$.

**2.1 Definition.** Wir nennen $\lambda \in K$ einen **Eigenwert** von $\varphi$, wenn es ein $v \in V \setminus \{0\}$ gibt mit $\varphi(v) = \lambda \cdot v$. So ein Vektor $v$ heißt **Eigenvektor** zum Eigenwert $\lambda$.

**2.2 Beispiel.** Falls $\dim(V) = 1$, dann ist jedes $v \in V \setminus \{0\}$ eine Basis von $V$. Für $\varphi \in \mathrm{End}_K(V)$ gibt es somit ein $\lambda \in K$ mit $\varphi(v) = \lambda \cdot v$. Offenbar gilt dann $\varphi = \lambda \cdot \mathrm{id}_V$, weil jeder Vektor aus $V$ ein Vielfaches von $v$ ist. Fazit: Jeder Vektor aus $V \setminus \{0\}$ ist Eigenvektor zu diesem Eigenwert $\lambda$.

**2.3 Beispiel.** Es sei $\varphi = \varphi_A \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^2)$ für $A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$ mit $\alpha \in \mathbb{R}$. Aus der Geometrie wissen wir, dass $\varphi_A$ die Drehung im Ursprung $\binom{0}{0} \in \mathbb{R}^2$ um den Winkel $\alpha$ ist. Falls $\alpha \not\equiv 0, \pi$ mod $2\pi$, dann gibt es keine Eigenwerte und keine Eigenvektoren. Falls $\alpha \equiv 0$ mod $2\pi$, dann ist $\varphi = \mathrm{id}$ und damit ist $\lambda = 1$ der einzige Eigenwert und jedes $v \in V \setminus \{0\}$ ist Eigenvektor zum Eigenwert 1. Falls $\alpha \equiv \pi$ mod $2\pi$, dann ist $\varphi = -\mathrm{id}$ die Punktspiegelung am Ursprung und $\lambda = -1$ der einzige Eigenwert. Wieder ist jedes $v \in V \setminus \{0\}$ ein Eigenvektor zum Eigenwert $-1$.

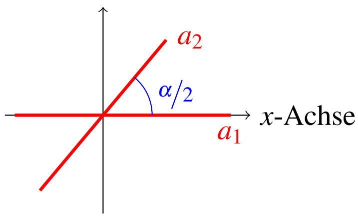
**2.4 Beispiel.** Es sei $\varphi = \varphi_A \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^2)$ für $A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}$ mit $\alpha \in \mathbb{R}$. Es gilt:

$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Also gilt:

$$\varphi = \varphi_{\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}} \circ \varphi_{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} = (\text{Drehung um } \alpha) \circ (\text{Spiegelung an der } x\text{-Achse})$$

Aus der Geometrie weiß man, dass die Drehung um $\alpha$ gleich dem Produkt von 2 Geradenspiegelungen $\sigma_{a_1}$, $\sigma_{a_2}$ an Achsen $a_1, a_2$ durch 0 ist. Dabei ist die Lage der Achsen unerheblich, sie müssen nur den Winkel $\alpha/2$ in $\binom{0}{0}$ einschließen. Wir legen $a_1$ auf die $x$-Achse und erhalten:




$$\varphi = \sigma_{a_2} \circ \sigma_{a_1} \circ \sigma_{x\text{-Achse}}$$

Weil eine Spiegelung idempotent ist, d.h. $\sigma^2 = \mathrm{id}$, muss $\varphi$ die Spiegelung an der Achse $a_2$ sein. Damit sind die Eigenwerte von $\varphi$ gleich 1 und $-1$. Die Eigenvektoren zum Eigenwert 1 (bzw. $-1$) sind alle von 0 verschiedenen Vektoren auf $a_2$ (bzw. senkrecht zu $a_2$).

**2.5 Proposition.** *Wir betrachten Eigenvektoren $v_1, \ldots, v_m$ von $\varphi$ bezüglich paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_m$. Dann sind $v_1, \ldots, v_m$ linear unabhängig. Insbesondere hat $\varphi$ höchstens $n := \dim(V)$ verschieden Eigenwerte.*

**Beweis:** Mit Induktion nach $m$. *Induktionsanfang:* $m = 1$: Klar, weil jeder Eigenvektor verschieden von 0 ist. *Induktionsschritt:* „$m - 1 \mapsto m$“ Es sei $m \geq 2$: Wir nehmen an, dass

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0. \tag{VII.2}$$

für $\alpha_1, \ldots, \alpha_m \in K$. Wenden wir $\varphi$ auf (VII.2), bzw. multiplizieren wir (VII.2) mit $\lambda_m$, dann erhalten wir

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \cdots + \alpha_m \lambda_m v_m = 0 \tag{VII.3}$$

$$\lambda_m \alpha_1 v_1 + \lambda_m \alpha_2 v_2 + \cdots + \lambda_m \alpha_m v_m = 0 \tag{VII.4}$$

Subtrahieren wir (VII.4) von (VII.3), so erhalten wir

$$\alpha_1(\lambda_1 - \lambda_m)v_1 + \alpha_2(\lambda_2 - \lambda_m)v_2 + \cdots + \alpha_{m-1}(\lambda_{m-1} - \lambda_m)v_{m-1} = 0$$

Weil die Eigenwerte paarweise verschieden sind, gilt $\lambda_1 - \lambda_m \neq 0$ für alle $i \in \{1, \ldots, m-1\}$. Nach Induktionsannahme müssen $v_1, \ldots, v_{m-1}$ linear unabhängig sein und somit gilt $\alpha_1 = \cdots = \alpha_{m-1} = 0$. Aus (VII.2) folgt $\alpha_m v_m = 0$. Weil $v_m \neq 0$, muss auch $\alpha_m = 0$ sein. Also sind $v_1, \ldots, v_m$ linear unabhängig. Mit vollständiger Induktion folgt die Behauptung. $\square$

**2.6 Definition.** (a) Für $\lambda \in K$ sei $V_\lambda := \ker(\varphi - \lambda \cdot \mathrm{id}_V)$ der **Eigenraum** zu $\lambda$.

(b) Wir nennen $\varphi \in \mathrm{End}_K(V)$ **diagonalisierbar**, wenn $V$ eine Basis aus Eigenvektoren von $\varphi$ hat.

(c) Die Matrix $A \in M(n \times n, K)$ heißt **diagonalisierbar**, wenn sie ähnlich ist zu einer **Diagonalmatrix** $D$, d.h. zu einer $n \times n$ Matrix der Form

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

mit allen $\lambda_i \in K$.

Zu diesen Definitionen bemerken wir folgendes:

In (a) ist $\lambda$ genau dann ein Eigenwert, wenn $V_\lambda \neq \{0\}$ gilt. Dann ist $V_\lambda \setminus \{0\}$ die Menge der Eigenvektoren zum Eigenwert $\lambda$. Weiter ist aufgrund der Definition von $V_\lambda$ als Kern einer linearen Abbildung klar, dass $V_\lambda$ ein Unterraum von $V$ ist.

Zu (b) und (c) werden wir in Theorem 2.9 sehen, dass $A$ genau dann diagonalisierbar ist, wenn $\varphi_A$ diagonalisierbar ist. Wir haben in V.1 gesehen, dass $n \times n$ Matrizen die Übersetzung von Endomorphismen bzgl. *einer* gewählten Basis von $V$ sind. Dabei sei betont, dass wir bei einem Endomorphismus nur *eine* Basis wählen, weil der Definitionsbereich und der Wertebereich gleich sind.

**2.7 Korollar.** *Sei $\varphi \in \mathrm{End}_K(V)$ und $n := \dim(V) < \infty$. Falls $\varphi$ $n$ verschiedene Eigenwerte hat, dann ist $\varphi$ diagonalisierbar.*

**Beweis:** Es seien $\lambda_1, \dots, \lambda_n$ $n$ verschiedene Eigenwerte von $\varphi$. Nach Definition gibt es dazu Eigenvektoren $v_1, \dots, v_n$, d.h. $\varphi(v_i) = \lambda_i v_i$ für $i = 1, \dots, n$. Nach Proposition 2.5 sind $v_1, \dots, v_n$ linear unabhängig. Weil $n = \dim(V)$, muss $v_1, \dots, v_n$ eine Basis von $V$ sein (Proposition IV.3.11). □

**2.8 Lemma.** *Es sei $\varphi \in \mathrm{End}_K(V)$ und $W$ der von allen Eigenvektoren von $\varphi$ erzeugte Unterraum von $V$. Dann gilt*

$$W = \bigoplus_\lambda V_\lambda,$$

*wobei $\lambda$ über alle Eigenwerte von $\varphi$ läuft und $V_\lambda$ der Eigenraum zu $\lambda$ ist.*

**Beweis:** Wir nehmen zuerst an, dass $V$ endlich dimensional ist. Das ist auch der Fall, der für unsere Anwendungen relevant ist. Nach Definition ist $W$ erzeugt von $\bigcup_{\lambda\,\mathrm{EW}} V_\lambda \setminus \{0\}$: Also enthält $W$ den Unterraum $\sum_{\lambda\,\mathrm{EW}} V_\lambda \supseteq \bigcup_\lambda V_\lambda \setminus \{0\}$. Weil $W$ aber der kleinste solche Unterraum ist, gilt

$$W = \sum_\lambda V_\lambda.$$

Wir müssen zeigen, dass die Summe von Unterräumen auf der rechten Seite eine direkte Summe ist. Sei $v = \sum_\lambda v_\lambda$, wobei $\lambda$ über alle Eigenwerte von $\varphi$ läuft und $v_\lambda \in V_\lambda$ ist. Um die Behauptung zu zeigen, müssen wir beweisen, dass diese Darstellung eindeutig ist. Sei also $v = \sum_\lambda v'_\lambda$ eine weitere solche Darstellung. Dann gilt

$$0 = \sum_\lambda (v'_\lambda - v_\lambda). \tag{VII.5}$$

weil die Eigenräume als Kern von linearen Abbildungen Unterräume von $V$ sind, gilt $v'_\lambda - v_\lambda \in V_\lambda$. Nach Proposition 2.5 sind Eigenvektoren linear unabhängig. Da jedes $v'_\lambda - v_\lambda \neq 0$ ein Eigenvektor ist, würde (VII.5) Proposition 2.5 widersprechen. Also folgt $v_\lambda = v'_\lambda$ für alle Eigenwerte $\lambda$ und damit die Behauptung. Falls $V$ unendlich dimensional ist, geht der Beweis analog. Man muss nur aufpassen, dass bei Summen $\sum_\lambda v_\lambda$ immer nur endlich viele Vektoren $v_\lambda \neq 0$ zu setzen. □

**2.9 Theorem.** *Für den Endomorphismus $\varphi$ des endlich dimensionalen $K$-Vektorraums $V$ sind folgenden Aussagen äquivalent:*

*a)* $\varphi$ *ist diagonalisierbar;*

*b)* $V$ *ist die Summe der Eigenräume* $V_\lambda$ *zu den Eigenwerten* $\lambda$*;*

*c)* $V$ *ist die direkte Summe der Eigenräume* $V_\lambda$ *zu den Eigenwerten* $\lambda$*;*

*d)* $\dim(V) = \sum_\lambda \dim(V_\lambda)$*, wobei* $\lambda$ *über alle Eigenwerte läuft;*

*e)* $\forall$ *Basen von* $V$ *ist die Matrix* $A(\varphi)$ *zu* $\varphi$ *diagonalisierbar;*

*f)* $\exists$ *Basis von* $V$ *so, dass die Matrix* $A(\varphi)$ *zu* $\varphi$ *diagonalisierbar ist;*

*g)* $\exists$ *Basis von* $V$ *so, dass die Matrix* $A(\varphi)$ *zu* $\varphi$ *eine Diagonalmatrix ist.*

**Beweis:** „a) $\Longrightarrow$ b)" Es sei $\varphi$ diagonalisierbar, d.h. $\exists$ Basis $b_1, \ldots, b_n$ aus Eigenvektoren zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$. Für $v \in V$ gilt:

$$v = \sum_{j=1}^{n} \alpha_j b_j$$

mit $\alpha_j \in K$. Wir betrachten einen Eigenwert $\lambda$ aus $\lambda_1, \ldots, \lambda_n$ und setzen

$$v_\lambda := \sum_{j=1,\ldots,n;\, \lambda_j=\lambda} \alpha_j b_j$$

Weil $V_\lambda$ ein Unterraum ist, gilt $v_\lambda \in V_\lambda$. Weil $v$ Summe von solchen $v_\lambda$ ist, folgt sofort b).

„b) $\Longleftrightarrow$ c) $\Longleftrightarrow$ d)" Nach Lemma 2.8 gilt für $W := \sum_\lambda V_\lambda$ folgendes:

$$W = \bigoplus_\lambda V_\lambda, \tag{VII.6}$$

wobei $\lambda$ wie immer über alle Eigenwerte von $\varphi$ läuft. Damit sieht man schon die Äquivalenz von b) und c). Weiter folgt aus (VII.6), dass

$$\dim(W) = \sum_\lambda \dim(V_\lambda)$$

nach IV.3.15. Also gilt c) $\Longleftrightarrow W = V \overset{IV.3.12}{\Longleftrightarrow} \dim(W) = \dim(V) \Longleftrightarrow \sum_\lambda \dim(V_\lambda) = \dim(V) \Longleftrightarrow$ d).

„c) $\Longrightarrow$ a)" Für jeden Eigenwert $\lambda$ wählen wir eine Basis im Eigenraum $V_\lambda$. Setzen wir diese Basen zusammen, erhalten wir wegen $V = \bigoplus_\lambda V_\lambda$ eine Basis von $V$. Da dies eine Basis aus Eigenvektoren ist, muss $\varphi$ diagonalisierbar sein nach Definition. Damit folgt a).

„a) $\Longrightarrow$ g)" Es sei $\varphi$ diagonalisierbar, d.h. $\exists$ Basis $b_1, \ldots, b_n$ aus Eigenvektoren zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$. Zu dieser Basis gilt offenbar

$$A(\varphi) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Dies zeigt g).

„g) $\Longrightarrow$ a)" Sei also $A(\varphi) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ bzgl. der Basis $b_1, \ldots, b_n$. Dann ist $b_j$ ein Eigenvektor zum Eigenwert $\lambda_j$ und damit ist $\varphi$ diagonalisierbar.

„e) $\Longrightarrow$ f)" trivial.

„f) $\Longrightarrow$ g)“ Es sei $b_1, \dots, b_n$ eine Basis so, dass $A(\varphi)$ diagonalisierbar ist, d.h. $A(\varphi)$ ist ähnlich zu einer Diagonalmatrix $D$. Also $\exists S \in \mathrm{GL}(n,K)$ mit

$$D = S \cdot A \cdot S^{-1}.$$

Nun gibt es einen Basiswechsel zu der Basis $b'_1, \dots, b'_n$ so, dass $S$ die Transformationsmatrix ist (Aufgabe 12.2). Dann ist $D$ aber die Matrix von $\varphi$ bzgl. der neuen Basis $b'_1, \dots, b'_n$ (Proposition V.4.6). Dies zeigt g).

„g) $\Longrightarrow$ e)“ Sei $b_1, \dots, b_n$ eine Basis von $V$ so, dass die zu $\varphi$ gehörige Matrix $A(\varphi)$ eine Diagonalmatrix ist. Weiter sei $b'_1, \dots, b'_n$ eine beliebige Basis und $A'(\varphi)$ die zu $\varphi$ gehörige Matrix bezüglich $b'_1, \dots, b'_n$. Nach Proposition V.4.6 gilt

$$A'(\varphi) = S \cdot A(\varphi) \cdot S^{-1},$$

wobei $S$ die Transformationsmatrix des Basiswechsels von $b_1, \dots, b_n$ nach $b'_1, \dots, b'_n$ ist. Also ist $A'(\varphi)$ ähnlich zur Diagonalmatrix $A(\varphi)$. Es folgt, dass $A'(\varphi)$ diagonalisierbar ist und damit e). $\square$

**2.10 Beispiel.** Es sei $\varphi\colon \mathbb{R}^3 \to \mathbb{R}^3$ gegeben durch die Matrix $A = \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix}$, d.h. $\varphi\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = A \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. Es gilt für alle $\lambda \in \mathbb{R}$: $\lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \varphi\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \Longleftrightarrow (A - \lambda E_3) \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$. Die rechte Seite ist ein homogenes lineare Gleichungssystem mit Parameter $\lambda$ und den Unbekannten $x_1, x_2, x_3$. Mit dem Gauß-Algorithmus folgt.

$$\begin{array}{ccc|c} -2-\lambda & 1 & & 0 \\ 1 & -2-\lambda & 1 & 0 \\ & 1 & -2-\lambda & 0 \end{array} \rightsquigarrow \begin{array}{ccc|cl} \boxed{1} & -2-\lambda & 1 & 0 & \cdot(2+\lambda) \\ -2-\lambda & 1 & 0 & 0 & + \\ 0 & 1 & -2-\lambda & 0 & \end{array}$$

$$\begin{array}{ccc|c} 1 & -2-\lambda & 1 & 0 \\ \rightsquigarrow 0 & 1-(2+\lambda)^2 & 2+\lambda & 0 \\ 0 & 1 & -2-\lambda & 0 \end{array} \rightsquigarrow \begin{array}{ccc|cl} 1 & -2-\lambda & 1 & 0 & \\ 0 & \boxed{1} & -2-\lambda & 0 & \cdot(\lambda^2+4\lambda+3) \\ 0 & -\lambda^2-4\lambda-3 & 2+\lambda & 0 & + \end{array}$$

$$\begin{array}{ccc|c} 1 & -2-\lambda & 1 & 0 \\ \rightsquigarrow 0 & 1 & -2-\lambda & 0 \\ 0 & 0 & p(\lambda) & 0 \end{array}$$

wobei $p(\lambda) = (\lambda^2+4\lambda+3)(-2-\lambda)+2+\lambda = -(2+\lambda)(\lambda^2+4\lambda+2)$.

Falls $p(\lambda) \neq 0$, dann gibt es nur die triviale Lösung $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, d.h. $\lambda$ ist *kein* Eigenwert. Falls $p(\lambda) = 0$, dann hat die Zeilenstufenform den Rang 2 und damit gibt es nicht-triviale Lösungen, also ist $\lambda$ dann ein Eigenwert. Konkret finden wir die Nullstellen von $p(\lambda)$ und damit die Eigenwerte $-2, -2 \pm \sqrt{2}$ von $A$. Wir berechnen die zugehörigen Eigenräume. Beachte, dass der Eigenraum $V_\lambda$ gerade der Lösungsraum des obigen Gleichungssystems ist.

*1. Fall:* $\lambda = -2$: Gesucht ist der Lösungsraum von

$$\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}.$$

Also folgt

$$V_{-2} = \left\{ \begin{pmatrix} -x_3 \\ 0 \\ x_3 \end{pmatrix} \middle| x_3 \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\rangle.$$

*2. Fall:* $\lambda = -2 \pm \sqrt{2}$: Gesucht ist der Lösungsraum von

$$\begin{array}{ccc|c} 1 & \mp\sqrt{2} & 1 & 0 \\ 0 & 1 & \mp\sqrt{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}.$$

Also folgt

$$V_{-2\pm\sqrt{2}} = \left\{ \begin{pmatrix} x_3 \\ \pm\sqrt{2}x_3 \\ x_3 \end{pmatrix} \middle| x_3 \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} 1 \\ \pm\sqrt{2} \\ 1 \end{pmatrix} \right\rangle.$$

**2.11 Theorem.** *Seien $\varphi, \varphi' \in \mathrm{End}_K(V)$ diagonalisierbar. Dann haben $\varphi$ und $\varphi'$ genau dann eine gemeinsame Basis aus Eigenvektoren, wenn $\varphi \circ \varphi' = \varphi' \circ \varphi$.*

**Beweis:** „$\Longrightarrow$" Sei $b_1, \ldots, b_n$ eine Basis von $V$ so, dass jedes $b_i$ ein Eigenvektor von $\varphi$ zum Eigenwert $\lambda_i$ und auch ein Eigenvektor von $\varphi'$ zum Eigenwert $\lambda_i'$ ist. Für $v \in V$ gilt

$$v = \alpha_1 b_1 + \cdots + \alpha_n b_n$$

für geeignete $\alpha_1, \ldots, \alpha_n \in K$.

$$\begin{aligned} \varphi \circ \varphi'(v) &= \varphi \circ \varphi'(\alpha_1 b_1 + \cdots + \alpha_n b_n) = \alpha_1 \varphi\big(\varphi'(b_1)\big) + \cdots + \alpha_n \varphi\big(\varphi'(b_n)\big) \\ &= \alpha_1 \varphi(\lambda_1' b_1) + \cdots + \alpha_n \varphi(\lambda_n' b_n) = \alpha_1 \lambda_1' \varphi(b_1) + \cdots + \alpha_n \lambda_n' \varphi(b_n) \\ &= \alpha_1 \lambda_1' \lambda_1 b_1 + \cdots + \alpha_n \lambda_n' \lambda_n b_n. \end{aligned}$$

Analog gilt

$$\varphi' \circ \varphi(v) = \alpha_1 \lambda_1 \lambda_1' b_1 + \cdots + \alpha_n \lambda_n \lambda_n' b_n.$$

Weil die Multiplikation von $K$ kommutativ ist, folgt $\varphi \circ \varphi' = \varphi' \circ \varphi$ wie gewünscht.

„$\Longleftarrow$" Wir nehmen an, dass $\varphi \circ \varphi' = \varphi' \circ \varphi$ gilt. Weil $\varphi$ diagonalisierbar ist, gilt:

$$V = \bigoplus_{\lambda} V_\lambda, \tag{VII.7}$$

wobei $\lambda$ über alle Eigenwerte von $\varphi$ läuft und $V_\lambda$ der zugehörige Eigenraum ist. Für $v \in V_\lambda$ gilt:

$$\varphi(\varphi'(v)) = \varphi'(\varphi(v)) = \varphi'(\lambda v) = \lambda \varphi'(v).$$

Damit folgt $\varphi'(v) \in V_\lambda$. Also folgt $\varphi'(V_\lambda) \subseteq V_\lambda$ für jeden Eigenwert $\lambda$ von $\varphi$. Analog hat man für jeden Eigenwert $\lambda'$ von $\varphi'$ mit zugehörigem Eigenraum $V'_{\lambda'}$

$$V = \bigoplus_{\lambda'} V'_{\lambda'} \tag{VII.8}$$

und zeigt, dass $\varphi(V'_{\lambda'}) \subseteq V'_{\lambda'}$. Es genügt zu zeigen, dass $V_\lambda$ eine Basis aus Eigenvektoren von $\varphi'$ hat. Wenn man dann diese Basis zusammensetzt, erhält man eine Basis von $V$ wegen (VII.7). Weiter besteht dann diese Basis nach Konstruktion aus gemeinsamen Eigenvektoren wie gewünscht. Sei $w \in V_\lambda$. Wegen (VII.8) gilt $w = \sum_{\lambda'} v'_{\lambda'}$ für geeignete $v'_{\lambda'} \in V'_{\lambda'}$. Es gilt

$$\sum_{\lambda'} \lambda v'_{\lambda'} = \lambda w = \varphi(w) = \sum_{\lambda'} \varphi(v'_{\lambda'}),$$

wobei $\lambda'$ über alle Eigenwerte von $\varphi'$ läuft. Da $\varphi(V_{\lambda'}) \subseteq V_{\lambda'}$ ist, muss $\lambda v_{\lambda'} = \varphi(v_{\lambda'})$ gelten aufgrund der Eindeutigkeit der Summendarstellung in (VII.8). Es folgt $v'_{\lambda'} \in V_\lambda$. Also ist jedes $w \in V_\lambda$ Summe von Eigenvektoren von $\varphi'$. Nach Theorem 2.9 ist $\varphi'|_{V_\lambda}$ diagonalisierbar wie gewünscht. □

## 3. Charakteristisches Polynom

In diesem Abschnitt zeigen wir, wie man die Eigenwerte mit Hilfe der Determinante berechnen kann. Dabei benutzt man die Determinante, um das charakteristische Polynom zu definieren und die Eigenwerte sind dann die Nullstellen dieses Polynoms. Als Anwendung geben wir ein Kriterium das entscheidet, ob eine Matrix diagonalisierbar ist.

In diesem Abschnitt sei $K$ ein Körper.

**3.1 Definition.** Für $A = (a_{ij}) \in M(n \times n, K)$ definieren wir das **charakteristische Polynom**

$$\chi_A(x) := \sum_{\pi \in S_n} \operatorname{sign}(\pi) \prod_{j=1}^{n} (x\delta_{\pi(j),j} - a_{\pi(j),j}) \in K[x].$$

Dabei ist $\delta_{ij} := \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$ das Kroneckersymbol.

**3.2.** In VII.1 haben wir gesehen, dass wir die Determinante auf $n \times n$ Matrizen mit Einträgen in einem kommutativen Ring verallgemeinern können. Das wollen wir hier anwenden, um eine einfache Interpretation des charakteristischen Polynoms zu geben. Als kommutativen Ring benutzen wir den Polynomring $K[x]$ und erhalten

$$\chi_A(x) = \det(x \cdot E_n - A) \in K[x].$$

**3.3 Definition.** Die **Spur** von $A \in M(n \times n, K)$ ist definiert als die Summe der Diagonaleinträge, d.h.

$$\operatorname{Tr}(A) := a_{11} + a_{22} + \cdots + a_{nn}.$$

**3.4 Proposition.** *Die Spur ist eine lineare Abbildung* $\operatorname{Tr}\colon M(n \times n, K) \to K$. *Weiter gilt*

$$\operatorname{Tr}(A \cdot B) = \operatorname{Tr}(B \cdot A)$$

*für* $A, B \in M(n \times n, K)$.

**Beweis:** Die Spur ist Summe von Koordinatenabbildungen bezüglich der Standardbasis von $M(n \times n, K)$. Also ist die Spur ein lineares Funktional als Summe von linearen Funktionalen. Wir setzen $C := A \cdot B$ mit Einträgen $c_{ij}$. Dann gilt

$$\operatorname{Tr}(A \cdot B) = \operatorname{Tr}(C) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ji} = \sum_{(i,j) \in \{1,\ldots,n\}^2} a_{ij} b_{ji}.$$

Analog gilt

$$\operatorname{Tr}(B \cdot A) = \sum_{(i,j) \in \{1,\ldots,n\}^2} b_{ij} a_{ji}.$$

Weil $a_{ij} b_{ji} = b_{ji} a_{ij}$ gilt, folgt $\operatorname{Tr}(B \cdot A) = \operatorname{Tr}(A \cdot B)$. □

**3.5 Proposition.** *Für* $A \in M(n \times n, K)$ *gilt*

$$\chi_A(x) = x^n - \operatorname{Tr}(A) x^{n-1} + \cdots + (-1)^n \det(A),$$

*wobei wir über die Koeffizienten von* $x^{n-2}, \ldots, x$ *keine Aussagen machen.*

**Beweis:** Wir betrachten die Definition von $\chi_A(x)$ aus 3.1. Wenn $\pi \in S_n$ verschieden von id ist, dann gibt es mindestens zwei $j \in \{1, \ldots, n\}$. mit $\pi(j) \neq j$. Wenn wir den entsprechenden Summanden zu $\pi$ in Definition 3.1 betrachten, dann hat er Grad $< n-2$. Wenn $\pi = \mathrm{id}$, dann ist der entsprechende Summand in Definition 3.1 gleich

$$\begin{aligned} \prod_{j=1}^{n} (x - a_{jj}) &= x^n - \sum_{j=1}^{n} a_{jj} x^{n-1} + \text{ Terme vom Grad } \leq n-2 \\ &= x^n - \operatorname{Tr}(A) x^{n-1} + \text{ Terme vom Grad } \leq n-2. \end{aligned}$$

Also gilt:

$$\chi_A(x) = x^n - \operatorname{Tr}(A) x^{n-1} + \alpha_{n-2} x^{n-2} + \cdots + \alpha_0$$

Aus

$$\alpha_0 = \chi_A(0) = \sum_{\pi \in S_n} \operatorname{sign}(\pi) \prod_{j=1}^{n} (-a_{\pi(j),j}) = (-1)^n \det(A)$$

folgt die Behauptung. □

**3.6 Proposition.** *Seien $A, B \in M(n \times n, K)$ ähnlich, dann gilt $\chi_A = \chi_B$.*

**Beweis:** Weil $A$ und $B$ ähnlich sind, $\exists S \in \mathrm{GL}(n, K)$ mit $B = S \cdot A \cdot S^{-1}$

$$\Longrightarrow \chi_B(x) = \det(x \cdot E_n - B) = \det(x \cdot E_n - S \cdot A \cdot S^{-1}). \tag{VII.9}$$

Mit Proposition V.1.8 folgt

$$S \cdot (xE_n) \cdot S^{-1} = S \cdot (x \cdot S^{-1}) = x \cdot S \cdot S^{-1} = x \cdot E_n.$$

Also erhalten wir

$$xE_n - SAS^{-1} = S(xS^{-1}) - SAS^{-1} = S(xE_n - A)S^{-1} \tag{VII.10}$$

Wir setzen (VII.10) in (VII.9) ein und benutzen

$$\det(S^{-1}) \cdot \det(S) = \det(S^{-1} \cdot S) = \det(E_n) = 1,$$

dann folgt

$$\chi_B(x) = \det\big(S \cdot (xE_n - A) \cdot S^{-1}\big) = \det(S)\det(xE_n - A) \cdot \det(S)^{-1} = \chi_A(x).$$

Also folgt die Behauptung. □

**3.7 Korollar.** *Ähnliche Matrizen haben dieselbe Spur.*

**Beweis:** *1.Beweis:*

$$\mathrm{Tr}(B) = \mathrm{Tr}(SAS^{-1}) \overset{3.4}{=} \mathrm{Tr}(AS^{-1}S) = \mathrm{Tr}(A).$$

*2.Beweis:* Nach Proposition 3.6 gilt: $\chi_A = \chi_B$. Mit Proposition 3.5 folgt $\mathrm{Tr}(A) = \mathrm{Tr}(B)$. □

**3.8.** Es sei $V$ ein endlich dimensionaler $K$-Vektorraum und $\varphi \in \mathrm{End}_K(V)$. Wir wählen eine Basis $b_1, \ldots, b_n$ von $V$, dann wird $\varphi$ durch eine $n \times n$-Matrix $A(\varphi)$ dargestellt (siehe V.1.5). Wir definieren das **charakteristische Polynom** von $\varphi$ als

$$\chi_\varphi(x) := \chi_{A(\varphi)}(x) = \det\big(x \cdot E_n - A(\varphi)\big) \in K[x].$$

Da wir bei Basiswechsel ähnliche Matrizen erhalten für $\varphi$ (Proposition V.4.6), ist das charakteristische Polynom $\chi_\varphi(x)$ unabhängig von der Wahl der Basis definiert. Die Bedeutung des charakteristischen Polynoms ergibt sich aus:

**3.9 Proposition.** *Eine Zahl $\lambda \in K$ ist genau dann Nullstelle von $\chi_\varphi$, wenn $\lambda$ ein Eigenwert von $\varphi$ ist.*

**Beweis:** Sei $\lambda \in K$. Dann gilt:

$$\begin{aligned}\chi_\varphi(\lambda) = 0 \Longleftrightarrow \det\big(\lambda \cdot E_n - A(\varphi)\big) = 0 &\overset{\text{VI.2.17}}{\Longleftrightarrow} \det(\lambda \cdot \mathrm{id} - \varphi) = 0 \\ \overset{\text{VI.2.18}}{\Longleftrightarrow} \lambda \cdot \mathrm{id} - \varphi \text{ nicht injektiv } &\Longleftrightarrow \ker(\lambda \cdot \mathrm{id} - \varphi) \neq \{0\} \\ \overset{2.6}{\Longleftrightarrow} \lambda \text{ Eigenwert von } \varphi &\end{aligned}$$

□

**3.10.** Es sei weiter $V$ ein $n$-dimensionaler $K$-Vektorraum und $\varphi \in \mathrm{End}_K(V)$. Für $\lambda \in K$ definieren wir die **geometrische Vielfachheit** von $\lambda$ als $\dim(V_\lambda)$, d.h. als die Dimension des Eigenraums $V_\lambda = \ker(\lambda \cdot \mathrm{id} - \varphi)$. Die **algebraische Vielfachheit** von $\lambda \in K$ ist die Multiplizität von $\lambda$ als Nullstelle von $\chi_\varphi(x)$. Wenn $\lambda$ keine Nullstelle ist (und äquivalent kein Eigenwert von $\varphi$ nach Proposition 3.9), dann ist sowohl die geometrische wie auch die algebraische Vielfachheit Null.

**3.11 Proposition.** *Sei $\lambda$ ein Eigenwert von $\varphi$. Dann gilt:* $1 \leq$ *geometrische Vielfachheit* $\leq$ *algebraische Vielfachheit.*

**Beweis:** Nach Definition ist die geometrische Vielfachheit eines Eigenwertes mindestens 1. Wir wählen eine Basis $b_1, \dots, b_r$ des Eigenraums $V_\lambda$ und ergänzen sie zu einer Basis $b_1, \dots, b_n$ von $V$ (Theorem IV.2.5). Weil $b_1, \dots, b_r$ Eigenvektoren zum Eigenwert $\lambda$ sind, folgt

$$A(\varphi) = \begin{pmatrix} \lambda E_r & B \\ 0 & C \end{pmatrix},$$

wobei $B \in M\big(r \times (n-r), K\big)$, $C \in M\big((n-r) \times (n-r), K\big)$. Damit gilt

$$\chi_\varphi(x) = \det\big(x \cdot \mathrm{id} - A(\varphi)\big) = \det \begin{pmatrix} (x-\lambda)E_r & -B \\ 0 & xE_{n-r} - C \end{pmatrix} \overset{VI.3.6}{=} \det\big((x-\lambda)E_r\big) \cdot \det(x \cdot E_{n-r} - C)$$
$$= (x-\lambda)^r \cdot \det(xE_{n-r} - C).$$

Also ist die algebraische Vielfachheit von $\lambda$ mindestens $r =$ geometrische Vielfachheit. □

**3.12 Theorem.** *Ein Endomorphismus $\varphi$ des endlich dimensionalen $K$-Vektorraums $V$ ist genau dann diagonalisierbar, wenn das charakteristische Polynom $\chi_\varphi(x)$ vollständig in Linearfaktoren aus $K[x]$ zerfällt und die geometrische Vielfachheit jedes Eigenwertes gleich der algebraischen Vielfachheit ist.*

**Beweis:** „$\Longrightarrow$“ Es sei $\varphi$ diagonalisierbar, d.h. $\exists$ Basis $b_1, \dots, b_n$ von $V$ aus Eigenvektoren zu den Eigenwerten $\lambda_1, \dots, \lambda_n$. Bezüglich dieser Basis haben wir

$$A(\varphi) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

und damit zerfällt das charakteristische Polynom in die Linearfaktoren

$$\chi_\varphi(x) = \det(xE_n - A(\varphi)) = (x-\lambda_1)\cdots(x-\lambda_n).$$

Die algebraische Vielfachheit $\mu_\lambda$ des Eigenwerts $\lambda$ ist gleich $|\{j | \lambda_j = \lambda\}|$. Also besteht die Menge $\{b_j | \lambda_j = \lambda\}$ aus $\mu_\lambda$ linear unabhängigen Vektoren in $V_\lambda$. Somit gilt: geometrische Vielfachheit von $\lambda = \dim(V_\lambda) \geq \mu_\lambda$. Mit Proposition 3.11 folgt „=“ und damit „$\Longrightarrow$“.

„$\Longleftarrow$“ Wir nehmen an, dass das charakteristische Polynom $\chi_\varphi(x)$ vollständig in Linearfaktoren aus $K[x]$ zerfällt. Damit gilt

$$\chi_\varphi(x) = \prod_{\lambda \in Z} (x-\lambda)^{\mu_\lambda}$$

für eine endliche Teilmenge $Z$ aus $K$ und $\mu_\lambda \in \mathbb{N} \setminus \{0\}$. Nach Proposition 3.9 ist $Z$ die Menge der Eigenwerte von $\varphi$ und $\mu_\lambda$ ist nach Definition die algebraische Vielfachheit des Eigenwertes $\lambda$. Wir nehmen weiter an, dass $\mu_\lambda$ auch die geometrische Vielfachheit von $\lambda$ ist. Dann folgt

$$\dim(V) = n = \deg(\chi_\varphi(x)) = \sum_{\lambda \in Z} \mu_\lambda = \sum_{\lambda \in Z} \dim(V_\lambda).$$

Aus Theorem 2.9 folgt, dass $\varphi$ diagonalisierbar ist. □

**3.13 Beispiel.** In Beispiel 2.10 haben wir für $A = \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix}$

$$\det(A - \lambda E_3) = -(2+\lambda)(\lambda^2 + 4\lambda + 2)$$

berechnet. Wenn man $\lambda$ durch die Variable $x$ ersetzt, erhält man

$$\chi_A(x) = (2+x)(x^2 + 4x + 2).$$

Das Beispiel bestätigt, dass die Eigenwerte die Nullstellen von $\chi_A(x)$ sind. Weil es drei verschiedene reelle Nullstellen gibt, muss $A$ diagonalisierbar sein. Weiter gilt

$$\det(A) = -\chi_A(0) = -4, \quad \mathrm{Tr}(A) = -6.$$

## 4. Trigonalisierung

In diesem Abschnitt betrachten wir $\varphi \in \mathrm{End}_K(V)$ für einen endlich dimensionalen Vektorraum $V$ über dem Körper $K$. Wir geben ein Kriterium an, wann $\varphi$ sich durch eine obere Dreiecksmatrix darstellen lässt bzgl. einer geeigneten Basis. Als Anwendung beweisen wir den Satz von Cayley-Hamilton.

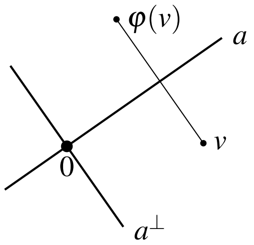
**4.1 Definition.** Ein Unterraum $W$ von $V$ heißt **$\varphi$-invariant**, wenn $\varphi(W) \subseteq W$.

**4.2 Bemerkung.** $\{0\}$ und $V$ sind die trivialen $\varphi$-invarianten Unterräume von $V$. Jeder Eigenvektor $v$ von $\varphi$ erzeugt einen 1-dimensionalen $\varphi$-invarianten Unterraum $\langle v\rangle = Kv$, denn für $w \in \langle v\rangle$ gilt $w = \alpha v, \alpha \in K$, und damit

$$\varphi(w) = \varphi(\alpha v) = \alpha\varphi(v) = \alpha\lambda v \in \langle v\rangle.$$

Mit demselben Argument erkennt man, dass der Eigenraum $V_\lambda$ $\varphi$-invariant ist.

**4.3 Beispiel.** Wenn $\varphi$ die Geradenspiegelung in $\mathbb{R}^2$ an der Achse $a$ durch den Ursprung $\binom{0}{0}$ ist, dann sind $a$ und $a^\perp$ die einzigen nicht-trivialen $\varphi$-invarianten Unterräume von $\mathbb{R}^2$.




Wenn $\varphi$ eine Drehung ist um den Winkel $\alpha \not\equiv 0 \pmod{\pi}$ mit Zentrum $\binom{0}{0}$, dann gibt es keine nicht-trivialen $\varphi$-invarianten Unterräume.

**4.4 Proposition.** *Die Summe und Durchschnitte von $\varphi$-invarianten Unterräumen sind wieder $\varphi$-invariant.*

**Beweis:** Seien $(W_i)_{i\in I}$ $\varphi$-invariante Unterräume. Für $v \in \bigcap_{i\in I} W_i$ gilt $\varphi(v) \in W_i$ für alle $i \in I$. Also gilt $\varphi(v) \in \bigcap_{i\in I} W_i$ und damit ist $\bigcap_{i\in I} W_i$ $\varphi$-invariant. Sei $v \in \sum_{i\in I} W_i$, d.h. $v = \sum_{i\in I_0} v_i$ für ein endliches $I_0 \subset I$ und $v_i \in W_i$.

$$\Longrightarrow \varphi(v) = \varphi\Big(\sum_{i\in I_0} v_i\Big) = \sum_{i\in I_0} \underbrace{\varphi(v_i)}_{\in W_i}$$

Es folgt $\varphi(v) \in \sum_{i\in I} W_i$ und damit ist $\sum_{i\in I} W_i$ $\varphi$-invariant. □

**4.5 Proposition.** *Wenn $W$ ein $\varphi$-invarianter Unterraum von $V$ ist, dann haben wir Endomorphismen*

$$\varphi_1 \colon W \to W,\ w \mapsto \varphi(w)$$

*und*

$$\varphi_2 \colon V/W \to V/W,\ [v] \mapsto \big[\varphi(v)\big].$$

*Weiter gilt $\chi_\varphi(x) = \chi_{\varphi_1}(x)\chi_{\varphi_2}(x)$.*

**Beweis:** Wir wählen eine Basis $b_1,\ldots,b_r$ von $W$ und ergänzen sie zu einer Basis $b_1,\ldots,b_n$ von $V$ (Theorem IV.2.5). Die zu $\varphi$ gehörige Matrix bzgl. $b_1,\ldots,b_n$ hat die Form $A(\varphi) = \begin{pmatrix} A_1 & B \\ 0 & A_2\end{pmatrix}$ mit $A_1 \in M(r\times r,K)$, $B \in M\big(r\times(n-r),K\big)$ und $A_2 \in M\big((n-r)\times(n-r),K\big)$, denn wegen der $\varphi$-Invarianz von $W$ sind $\varphi(b_1),\ldots,\varphi(b_r)$ Linearkombinationen von $b_1,\ldots,b_r$. Mit Proposition VI.3.6 folgt

$$\begin{aligned}\chi_\varphi(x) = \det(x\cdot E_n - A) &= \det\left(\begin{array}{c|c} x\cdot E_r - A_1 & -B \\ \hline 0 & x\cdot E_{n-r} - A_2\end{array}\right)\\ &= \det(x\cdot E_r - A_1)\det(x\cdot E_{n-r} - A_2).\end{aligned} \tag{VII.11}$$

Wegen $\varphi(W) \subseteq W$ ist $\varphi_1$ wohldefiniert und wird durch die Matrix $A_1$ dargestellt. Wegen $\varphi(W) \subseteq W$ ist auch $\varphi_2$ wohldefiniert: Sei $[v] = [v']$, dann gilt $v - v' \in W$ und damit $\varphi(v) - \varphi(v') \in W$, d.h. $\big[\varphi(v)\big] = \big[\varphi(v')\big]$. Da $\varphi$ linear ist, ist $\varphi_2$ linear. Offenbar ist $[b_1] = \cdots = [b_r] = 0$ und $[b_{r+1}], \cdots, [b_n]$ ein Erzeugendensystem von $V/W$, also, eine Basis wegen $\dim V/W = n - r$ (IV.2.10/IV.3.16). Dazu wird $\varphi_2$ durch $A_2$ dargestellt. Aus (VII.11) folgt die Behauptung. □

**4.6 Definition.** Der Endomorphismus $\varphi$ heißt **trigonalisierbar** genau dann, wenn es eine Basis von $V$ gibt bzgl. der $\varphi$ durch eine obere Dreiecksmatrix darstellbar ist.

**4.7 Theorem.** *Der Endomorphismus $\varphi$ ist genau dann trigonalisierbar, wenn das charakteristische Polynom $\chi_\varphi$ vollständig in Linearfaktoren aus $K[x]$ zerfällt.*

**Beweis:** „$\Longrightarrow$“ Es gibt eine Basis mit $A(\varphi) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ obere Dreiecksmatrix wie oben. Dann gilt

$$\chi_\varphi(x) = \det(xE_n - A) \overset{\text{VI.3.3}}{=} (x-\lambda_1)\cdots(x-\lambda_n).$$

Also zerfällt $\chi_\varphi(x)$ vollständig in Linearfaktoren aus $K[x]$.

„$\Longleftarrow$“ Wir nehmen an, dass $\chi_\varphi(x)$ vollständig in Linearfaktoren aus $K[x]$ zerfällt. Also

$$\chi_\varphi(x) = (x-\lambda_1)\cdots(x-\lambda_n)$$

mit $\lambda_1, \ldots, \lambda_n \in K$, wobei $n := \dim(V)$. Wir beweisen mit Induktion nach $n$, dass es eine Basis von $V$ gibt so, dass $A(\varphi) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$. Induktionsanfang $n = 1$: Dann gilt $\varphi = \lambda_1 \cdot \mathrm{id}$, wobei $\chi_\varphi(x) = x - \lambda_1$. Induktionsschritt „$n-1 \mapsto n$“: Sei jetzt $n \geq 2$. Weil $\lambda_1$ eine Nullstelle von $\chi_\varphi(x)$ ist, muss $\lambda_1$ ein Eigenwert von $\varphi$ sein (Proposition 3.9). Sei $b_1$ ein Eigenvektor zum Eigenwert $\lambda_1$. Da $W := \langle b_1 \rangle$ ein $\varphi$-invarianter Unterraum von $V$ ist (siehe Bemerkung 4.2), induziert $\varphi$ Endomorphismen:

$$\varphi_1 \colon W \to W,\ w \mapsto \varphi(w) \quad \wedge \quad \varphi_2 \colon V/W \to V/W,\ [v] \mapsto [\varphi(v)]$$

nach Proposition 4.5. Weiter gilt

$$\chi_\varphi(x) = \chi_{\varphi_1}(x) \cdot \chi_{\varphi_2}(x).$$

Es gilt nach Induktionsanfang $\chi_{\varphi_1}(x) = (x - \lambda_1)$. Mit Polynomdivision folgt

$$\chi_{\varphi_2}(x) = (x-\lambda_2)\cdots(x-\lambda_n)$$

(siehe Theorem II.3.11). Aus der Induktionsannahme im Fall $n-1$ folgt, dass es eine Basis $[b_2], \ldots, [b_n]$ von $V/W$ gibt so, dass $\varphi$ durch die obere Dreiecksmatrix $\begin{pmatrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ dargestellt wird. Es ist leicht zu sehen, dass $b_1, \ldots, b_n$ eine Basis von $V$ bildet. Es sei $A(\varphi) = (a_{ij})$ die Matrix von $\varphi$ bezüglich der Basis $b_1, \ldots, b_n$.

$$\Longrightarrow \varphi(b_j) = \sum_{i=1}^n a_{ij} b_i$$

Aus

$$\varphi_2\big([b_j]\big) = \big[\varphi(b_j)\big] = \Big[\sum_{i=1}^n a_{ij} b_i\Big] = \sum_{i=1}^n a_{ij}[b_i] \overset{b_1 \in W}{=} \sum_{i=2}^n a_{ij}[b_i]$$

folgt, dass $A(\varphi_2) = (a_{ij})_{2 \leq i \leq n;\, 2 \leq j \leq n}$. Weil $A(\varphi) = \begin{pmatrix} \lambda_1 & B \\ 0 & A(\varphi_2) \end{pmatrix}$ gilt und $A(\varphi)$ eine Dreiecksmatrix ist, muss auch $A(\varphi)$ eine obere Dreiecksmatrix sein und dies beweist den Induktionsschritt. Mit vollständiger Induktion folgt die Behauptung. □

**4.8 Korollar.** *Jeder Endomorphismus eines endlich dimensionalen* ***komplexen*** *Vektorraums ist trigonalisierbar.*

**Beweis:** Nach dem Fundamentalsatz der Algebra zerfällt jedes komplexe Polynom vom Grad $\geq 1$ vollständig in komplexe Linearfaktoren (siehe Korollar 1.9). Mit Theorem 4.7 folgt die Behauptung. $\square$

**4.9 Definition.** Ein Endomorphismus $\varphi$ heißt **nilpotent**, wenn $\varphi^k = 0$ für ein $k \geq 1$.

**4.10 Korollar.** *Für einen Endomorphismus $\varphi$ des $n$-dimensionalen $K$-Vektorraums $V$ sind folgende Aussagen äquivalent.*

*a) $\varphi$ ist nilpotent;*

*b) $\chi_\varphi(x) = x^n$;*

*c) $\exists$ Basis von $V$ so dass $\varphi$ durch die Matrix $A(\varphi) = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$ dargestellt wird;*

*d) $\varphi^k = 0$ für ein $k \in \{1, \dots, n\}$.*

**Beweis:** „a)$\Longrightarrow$b)" Wir nehmen an, dass $\varphi$ nilpotent ist. Wir beweisen b) mit verbesserter vollständiger Induktion nach $n$. Induktionsanfang: $n = 1$: $\Longrightarrow \varphi = \lambda \cdot \mathrm{id}$. Weil $\varphi^k = 0$, folgt $\lambda = 0$. Also folgt $\varphi = 0$ und damit $\chi_\varphi(x) = x$. Induktionsschritt: Es sei $n \geq 2$ und wir nehmen an, dass b) stimmt für alle $n' < n$. Wäre $\varphi$ injektiv, dann wäre auch $\varphi^k$ injektiv für alle $k \geq 1$. Da aber $\varphi^k = 0$ für ein $k \geq 1$, wäre das ein Widerspruch. Also ist $\varphi$ nicht injektiv. Betrachte $W := \ker(\varphi)$. Wegen $\varphi(W) = \{0\}$, muss $W$ $\varphi$-invariant sein. Aus Proposition 4.5 haben wir Endomorphismen

$$\varphi_1 \colon W \to W,\ w \mapsto \varphi(w) \quad \wedge \quad \varphi_2 \colon V/W \to V/W,\ [v] \mapsto \big[\varphi(v)\big]$$

mit $\chi_\varphi(x) = \chi_{\varphi_1}(x)\chi_{\varphi_2}(x)$. Weil $W = \ker(\varphi)$, folgt $\varphi_1 = 0$ und $\chi_{\varphi_1}(x) = x^{\dim(W)}$. Also gilt:

$$\chi_\varphi(x)) = x^{\dim(W)} \cdot \chi_{\varphi_2}(x)$$

Es gilt $\dim(V/W) = \dim(V) - \dim(W) < \dim(V) = n$ und $\varphi_2$ ist nilpotent (da $\varphi$ nilpotent). Nach Induktionsannahme gilt

$$\chi_{\varphi_2}(x) = x^{\dim(V/W)}$$

und daraus folgt

$$\chi_\varphi(x) = x^{\dim(V)}$$

wie gewünscht.

„b)$\Longrightarrow$ c)" Nach dem Trigonalisierungssatz (Theorem 4.7) gibt es eine Basis von $V$ so, dass $A(\varphi) = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$. Also gilt

$$\chi_\varphi(x) = (x - a_{11})(x - a_{22}) \cdots (x - a_{nn}).$$

Andererseits haben wir $\chi_\varphi(x) = x^n$ in b) angenommen und somit gilt $a_{11} = a_{22} = \cdots = a_{nn} = 0$ wie in c) gewünscht.

„c)$\Longrightarrow$ d)" Wir betrachten die Basis $b_1, \dots, b_n$ mit $A(\varphi) = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$ wie in c). Wir setzen $V_i := \langle b_1, \dots, b_i \rangle$. Wegen der obigen Gestalt von $A(\varphi)$ muss $\varphi(b_i)$ eine Linearkombination von $b_1, \dots, b_{i-1}$ sein. Also folgt: $\varphi(V_i) \subseteq V_{i-1}$. Es folgt

$$\varphi^n(V) = \varphi^{n-1}(\varphi(V_n)) \subseteq \varphi^{n-1}(V_{n-1}) \subseteq \varphi^{n-2}(V_{n-2}) \subseteq \ldots \subset \varphi(V_1) \overset{\varphi(b_1)=0}{=} \{0\}$$

Also gilt $\varphi^n = 0$ und damit d).

„d)$\Longrightarrow$a)" trivial. $\square$

**4.11 Theorem (Satz von Cayley-Hamilton).** *Für den Endomorphismus $\varphi$ des endlich dimensionalen $K$-Vektorraums $V$ gilt:*

$$\chi_\varphi(\varphi) = 0 \in \mathrm{End}_K(V)$$

**Beweis:** Indem wir eine Basis von $V$ wählen, können wir die Aussage in eine entsprechende Aussage für $n \times n$-Matrizen $A$ übersetzen (siehe V.1). Es genügt also

$$\chi_A(A) = 0 \in M(n \times n, K) \tag{VII.12}$$

zu zeigen. Nach Theorem 1.14 gibt es eine Körpererweiterung $L$ von $K$ so, dass $\chi_A(x)$ vollständig in Linearfaktoren aus $L[x]$ zerfällt. Nach Theorem 4.7 gibt es eine Basis $b_1, \ldots, b_n$ von $L^n$ so, dass $\psi = \varphi_A$ durch eine obere Dreiecksmatrix $D = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ dargestellt wird. Dann gilt

$$\chi_A(x) = \chi_\psi(x) = (x - \lambda_1) \cdots (x - \lambda_n). \tag{VII.13}$$

Also ist (VII.12) äquivalent zu

$$\chi_\psi(\psi) = 0 \in \mathrm{End}_L(L^n) \tag{VII.14}$$

Um dies zu beweisen, setzen wir $\psi$ ein in (VII.13) und erhalten nach 1.18

$$\chi_\psi(\psi) = \chi_A(\psi) = (\psi - \lambda_1 \,\mathrm{id}) \circ (\psi - a_{22} \,\mathrm{id}) \circ \cdots \circ (\psi - \lambda_n \,\mathrm{id}). \tag{VII.15}$$

Wir betrachten die Unterräume $W_i := \langle b_1, \ldots, b_i \rangle$ von $V$ für $i = 1, \ldots, n$. Weil $\psi$ durch eine obere Dreiecksmatrix bzgl. $b_1, \ldots, b_n$ dargestellt wird, folgt

$$(\psi - \lambda_i \cdot \mathrm{id})(b_j) \in W_j$$

für alle $i, j \in \{1, \ldots, n\}$ und $(\psi - \lambda_i \cdot \mathrm{id})(b_i) \in W_{i-1}$. Also gilt

$$(\psi - \lambda_i \cdot \mathrm{id})(W_i) \subseteq W_{i-1}. \tag{VII.16}$$

Wir wenden (VII.16) sukzessive in (VII.15) an und erhalten

$$\begin{aligned}
\chi_\psi(\psi)(V) = \chi_\psi(\psi)(W_n) &= (\psi - \lambda_1 \,\mathrm{id}) \circ \cdots \circ (\psi - \lambda_n \,\mathrm{id})(W_n) \\
&\subseteq (\psi - \lambda_1 \,\mathrm{id}) \circ \cdots \circ (\psi - \lambda_{n-1} \,\mathrm{id})(W_{n-1}) \\
&\underset{(\cdots)}{\subseteq} (\psi - \lambda_1 \,\mathrm{id})(W_1) \subseteq W_0 = \{0\}.
\end{aligned}$$

Also gilt (VII.14) wie gewünscht. $\square$

# Teil II.

# Lineare Algebra II

Kapitel VIII.

# 8 Euklidische und unitäre Vektorräume

## 1. Bilinearformen

In diesem Abschnitt studieren wir die Bilinearformen auf einem endlich dimensionalen Vektorraum $V$. Wir definieren zuerst die Matrix einer gegebenen Bilinearform bzgl. einer Basis. Später werden wir sehen, dass die Bilinearformen in den Anwendungen häufig vorkommen und das dabei immer diejenigen Isomorphismen interessant sind, die die Bilinearformen erhalten.

**1.1 Definition.** Eine **Bilinearform** $b$ auf $V$ ist eine Abbildung $V \times V \to K$, die bilinear ist, d.h.

$$b(\alpha_1 v_1 + \alpha_2 v_2, w) = \alpha_1 b(v_1, w) + \alpha_2 b(v_2, w)$$

$$b(v, \alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 b(v, w_1) + \alpha_2 b(v, w_2)$$

für alle $\alpha_1, \alpha_2 \in K$ und $v_1, v_2, v, w_1, w_2, w \in V$.

**1.2 Beispiel.** Sei $b\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$,

$$\left(\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}, \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}\right) \mapsto \lambda_1\mu_2 - \lambda_2\mu_1$$

Dies ist eine Bilinearform auf $V = \mathbb{R}^2$, da $b$ linear ist in $\binom{\lambda_1}{\lambda_2}$ bei festem $\binom{\mu_1}{\mu_2}$ und umgekehrt.

**1.3 Beispiel.** Es sei $b\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(\lambda, \mu) \mapsto \mathrm{e}^{\lambda+\mu}$. Dies ist keine Bilinearform auf $V = \mathbb{R}$, da $b(0, \mu) \neq 0$ im Widerspruch zur Linearität im ersten Argument.

**1.4 Definition.** Wir nennen die $n \times n$-Matrix $B := \big(b(v_i, v_j)\big)_{1 \le i,j \le n}$ die Matrix zu der Bilinearform $b$ bzgl. der Basis $v_1, \ldots, v_n$ von $V$.

**1.5 Proposition.** *Sei $b$ eine Bilinearform auf $V$ und $B$ die Matrix von $b$ bzgl. der Basis $v_1, \ldots, v_n$. Für $v, w \in V$ mit Koordinatenvektoren $\lambda, \mu \in K^n$ aufgefasst als $n \times 1$-Matrizen gilt*

$$b(v, w) = \lambda^t \cdot B \cdot \mu \in M(1 \times 1, K) = K$$

**Beweis:**

$$b(v,w) = b\Big(\sum_{i=1}^n \lambda_i v_i, \sum_{j=1}^n \mu_j v_j\Big) \overset{\text{Linearität im 1. Argument}}{=} \sum_{i=1}^n \lambda_i b\Big(v_i, \sum_{j=1}^n \mu_j v_j\Big)$$

$$\overset{\text{Linearität im 2. Argument}}{=} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j \underbrace{b(v_i, v_j)}_{b_{ij}}.$$

Hier seien $b_{ij}$ die Einträge der $n \times n$ Matrix $B$.

$$b(v,w) = \sum_{i=1}^n \lambda_i \underbrace{\left(\sum_{j=1}^n b_{ij}\mu_j\right)}_{(B\cdot\mu)_i} = \lambda^t \cdot (B \cdot \mu) \qquad \square$$

**1.6 Bemerkung.** Umgekehrt sei $B \in M(n \times n, K)$ gegeben. Dann ist $b(v,w) := \lambda^t \cdot B \cdot \mu$ eine Bilinearform von $V$, wobei $\lambda, \mu \in K^n$ die Koordinaten von $v$ und $w$ bzgl. einer festen Basis $(v_i)$ sind. Dies sieht man, wenn man das erste Argument $\lambda$ (bzgl. das zweite Argument $\mu$) fixiert analog zu Beispiel 1.2. Die zugehörige Matrix dieser Bilinearform ist wieder $B$, denn

$$b(v_i, v_j) = e_i^t \cdot B \cdot e_j = b_{ij}.$$

Explizit gilt

$$b(v,w) = \sum_{i,j=1}^{n} b_{ij} \lambda_i \mu_j$$

wie wir im Beweis von Proposition 1.5 gesehen haben.

**1.7.** Wenn wir die Menge der Bilinearformen auf $V$ als Teilmenge des Vektorraums der Funktionen von $V \times V$ nach $K$ mit der punktweise definierten Addition und skalaren Multiplikation betrachten (siehe Beispiel III.1.3), dann erhalten wir einen Unterraum. Also bildet die Menge der Bilinearformen auf $V$ einen kanonischen $K$-Vektorraum.

**1.8 Proposition.** *Wir fixieren eine Basis $v_1, \ldots, v_n$ auf $V$ und ordnen jeder Bilinearform $b$ auf $V$ die zugehörige Matrix $B$ zu. Diese ergibt einen Isomorphismus vom $K$-Vektorraum der Bilinearformen auf $V$ auf der Vektorraum $M(n \times n, K)$ (siehe Proposition V.1.11).*

**Beweis:** Wir skizzieren diesen leichten Beweis nur, in dem wir bemerken, dass die Umkehrabbildung aus Beispiel 1.6 genommen werden kann. □

**1.9 Definition.** Wir betrachten eine Bilinearform $b$ auf $V$. Wir sagen, dass $v \in V$ **orthogonal** zu $w \in W$ ist (bzgl. b), wenn $b(v,w) = 0$. Dann schreiben wir $v \perp w$. Für $S \subseteq V$ definieren wir

$$S^\perp := \{w \in V \mid b(v,w) = 0 \; \forall v \in S\}$$

als Menge aller Vektoren, die **rechtsorthogonal** auf $S$ bzgl. $b$ sind. Analog sei

$${}^\perp S := \{v \in V \mid b(v,w) = 0 \; \forall w \in S\}$$

als Menge der **linksorthogonalen** Vektoren auf $S$.

**1.10 Beispiel.** $b\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\left(\binom{\lambda_1}{\lambda_2}, \binom{\mu_1}{\mu_2}\right) \mapsto \lambda_1\mu_1 + \lambda_2\mu_1$. Diese Abbildung ist bilinear mit Matrix $B = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ bzgl. der Standardbasen von $\mathbb{R}^2$. Sei $S = \{\binom{0}{1}\}$, dann gilt

$$b\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}\right) = \mu_1, \qquad b\left(\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 0.$$

Also gilt

$$S^\perp = \left\{ \begin{pmatrix} 0 \\ \mu_2 \end{pmatrix} \in \mathbb{R}^2 \mid \mu_2 \in \mathbb{R} \right\}, \qquad {}^\perp S = \mathbb{R}^2.$$

**1.11 Proposition.** *$S^\perp$ und ${}^\perp S$ sind Unterräume von $V$.*

**Beweis:** Dies folgt aus dem zweiten (bzw. ersten) Axiom der Bilinearformen, z.B. seien $w_1, w_2 \in S^\perp$, dann gilt für alle $v \in S$

$$b(v, w_1 + w_2) = b(v, w_1) + b(v, w_2) \overset{w_1, w_2 \in S^\perp}{=} 0 + 0 = 0$$

und damit folgt $w_1 + w_2 \in S^\perp$. Die anderen Unterraum-Axiome folgen analog. □

**1.12 Definition.** Eine Bilinearform $b$ auf $V$ heißt **nicht ausgeartet** genau dann, wenn $V^\perp = \{0\}$ und $^\perp V = \{0\}$.

**1.13 Proposition.** *Für eine Bilinearform $b$ auf $V$ mit zugehöriger Matrix $B$ bzgl. der Basis $v_1, \dots, v_n$ von $V$ sind folgende Bedingungen äquivalent:*

*a) $b$ ist nicht ausgeartet;*

*b) $V^\perp = \{0\}$;*

*c) $^\perp V = \{0\}$;*

*d) $B$ ist invertierbar.*

**Beweis:** Weil $v_1, \dots, v_n$ eine Basis von $V$ und $b$ bilinear ist, gilt

$$V^\perp = \{w \in V \mid b(v,w) = 0\ \forall v \in V\} = \{w \in V \mid b(v_i, w) = 0\ \forall i = 1, \dots, n\}.$$

Also ist $w = \sum_{j=1}^n \lambda_j v_j$ in $V^\perp$ genau dann, wenn der Koordinatenvektor $\lambda \in K^n$ von $w$ im Lösungsraum des homogenen linearen Gleichungssystems

$$B \cdot \lambda = 0$$

ist, denn es gilt

$$b(v_i, w) = \sum_{j=1}^n \lambda_j b(v_i, v_j) = (B \cdot \lambda)_i.$$

Die Bedingung b) ist also äquivalent dazu, dass homogene lineare Gleichungssystem nur die Nulllösung hat und nach Proposition V.2.5 ist dies äquivalent zu $B$ invertierbar. Dies zeigt b) $\iff$ d). Analog folgt c) $\iff$ d). Weil a) $\iff$ b) $\wedge$ c), folgt also d) $\iff$ a). □

**1.14 Proposition.** *Es sei $b$ eine nicht ausgeartete Bilinearform auf $V$. Dann gilt für jeden Unterraum $U$ von $V$:*

*a)* $\dim(U) + \dim(U^\perp) = \dim(V), \quad \dim(U) + \dim(^\perp U) = \dim(V)$

*b)* $U = ^\perp(U^\perp) = (^\perp U)^\perp$

*c) Die Abbildung $\Phi \colon V \to U^*,\ v \mapsto b(\cdot, v)$ ist linear und surjektiv.*

**Beweis:** Wegen der Linearität von $b$ im ersten Argument ist $\Phi$ wohldefiniert. Wegen der Linearität von $b$ im zweiten Argument ist $\Phi$ linear. Es gilt $\ker \Phi = U^\perp$ nach Definition von $U^\perp$. Wir wenden das zuerst für den Spezialfall $U = V$ an. Weil $b$ nicht ausgeartet ist, ist $\ker \Phi = \{0\}$, d.h. $\Phi$ ist ein Isomorphismus. Also hat jede Linearform auf $V$ die Form $b(\cdot, v)$.

Für beliebiges $U$ zeigen wir, dass $\Phi$ surjektiv ist: Sei $f \in U^*$. Nach Korollar IV.2.11 hat $f$ eine Erweiterung zu einer Linearform $g \in V^*$. Aus dem Spezialfall folgt, dass es ein $v \in V$ gibt mit $g = b(\cdot, v)$. Somit gilt $f = g|_U = b(\cdot, v)|_U = \Phi(v)$. Also ist $\Phi$ surjektiv und damit gilt c).

Nach der Dimensionsformel in Korollar IV.3.17 folgt

$$\dim V = \dim \ker \Phi + \dim \Phi(V) = \dim U^\perp + \dim U^*.$$

Wegen $\dim(U^*) = \dim(U)$ (Korollar V.1.16) folgt a) (die zweite Relation folgt analog).

Beachte, dass $U \subseteq {}^\perp(U^\perp)$ für jede Bilinearform direkt aus der Definition folgt. Wegen

$$\dim {}^\perp(U^\perp) \overset{\text{a)}}{=} \dim(V) - \dim(U^\perp) \overset{\text{a)}}{=} \dim(U)$$

folgt b) mit Korollar IV.3.12. □

## 2. Symmetrische Bilinearformen

In dieser Vorlesung werden symmetrische Bilinearformen eine Rolle spielen, d.h. man darf die Argumente $v$ und $w$ vertauschen. In diesem Fall gilt $S^\perp = {}^\perp S$ und man spricht vom **Orthogonalkomplement** von $S$. Eng verbunden mit symmetrischen Bilinearformen sind quadratische Formen, die man durch Einsetzen von $v = w$ erhält. Wir werden sehen, dass es immer eine Basis gibt, so dass die Matrix zur symmetrischen Bilinearform diagonal ist. Im reellen Fall kann man sogar verlangen, dass die Diagonalmatrix aus $+1, -1$ und $0$ bestehen. Dies ist der **Satz von Sylvester**.

In diesem Abschnitt ist $V$ immer ein endlich-dimensionaler $K$-Vektorraum der Dimension $n$.

**2.1 Definition.** Eine Bilinearform $b$ auf $V$ heißt **symmetrisch** $:\Longleftrightarrow b(v,w) = b(w,v)\ \forall v,w \in V$.

**2.2 Beispiel.** Im Abschnitt 1 haben wir gesehen, dass

$$b_1\left(\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}, \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}\right) = \alpha_1\beta_1 + \alpha_2\beta_2,$$

$$b_2\left(\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}, \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}\right) = \alpha_1\beta_2 - \alpha_2\beta_1,$$

$$b_3\left(\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}, \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}\right) = \alpha_1\beta_2 + \alpha_2\beta_1,$$

Bilinearformen auf $V = \mathbb{R}^2$ sind. Die Formen $b_1$ und $b_3$ sind symmetrisch, aber $b_2$ nicht!

**2.3 Proposition.** *Sei $b$ eine Bilinearform auf $V$ mit Matrix $B$ bzgl. der Basis $v_1, \ldots, v_n$ von $V$. Dann ist $b$ genau dann symmetrisch, wenn $B = B^t$ gilt.*

**Beweis:** Wir erinnern an die Definition: $B = (b(v_i, v_j))_{1 \le i,j \le n}$.

„$\Longrightarrow$“ Es sei $b$ eine symmetrische Bilinearform.

$$\Longrightarrow b(v_i, v_j) = b(v_j, v_i)$$

Somit gilt $B = B^t$.

„$\Longleftarrow$“ Es gelte nun umgekehrt $B = B^t$, d.h.

$$b(v_i, v_j) = b(v_j, v_i)\ \forall i,j = 1, \ldots, n.$$

Es seien $v, w \in V$. Weil $v_1, \ldots, v_n$ eine Basis von $V$ ist, lassen sich $v$ und $w$ als Linearkombination von $v_1, \ldots, v_n$ schreiben. Aufgrund der Bilinearität folgt $b(v,w) = b(w,v)$. □

**2.4 Bemerkung.** Wir nennen deshalb $A \in M(n, \times n, K)$ **symmetrisch**, wenn $A = A^t$ gilt.

**2.5 Definition.** Eine **quadratische Form** auf $V$ ist eine Funktion $q\colon V \to K$ mit folgenden Eigenschaften:

a) $q(\alpha \cdot v) = \alpha^2 \cdot q(v)\ \forall \alpha \in K, v \in V$;

b) $b_q\colon V \times V \to K,\ (v,w) \mapsto b_q(v,w). = q(v+w) - q(v) - q(w)$ ist eine Bilinearform auf $V$.

Wir nennen $b_q$ die zu $q$ **assoziierte Bilinearform**. Offensichtlich ist sie symmetrisch.

**2.6 Proposition.** *Sei $v_1, \ldots, v_n$ eine Basis von $V$. Wir betrachten eine Funktion $q\colon V \to K$ und die entsprechende Funktion $f\colon K^n \to K$, $\lambda \mapsto q(\lambda_1 v_1 + \cdots + \lambda_n v_n)$, bzgl. den Koordinatenvektoren. Dann ist $q$ genau dann eine quadratische Form, wenn $f$ ein homogenes Polynom vom Grad 2 ist, d.h.*

$$f(\lambda) = \sum_{1 \le i \le j \le n} a_{ij} \lambda_i \lambda_j$$

*für geeignete $a_{ij} \in K$.*

**Beweis:** „$\Longrightarrow$“ Es sei $q$ eine quadratische Form auf $V$. Für $v, w \in V$ gilt

$$q(v+w) = q(v) + q(w) + b_q(v,w).$$

Mit Induktion nach $m$ gilt

$$q(w_1 + \cdots + w_m) = \sum_{i=1}^{m} q(w_i) + \sum_{1 \le i < j \le m} b_q(w_i, w_j).$$

Für

$$v = \sum_{i=1}^{n} \lambda_i v_i$$

folgt

$$\begin{aligned} q(v) = q\Big(\sum_{i=1}^{n} \lambda_i v_i\Big) &= \sum_{i=1}^{n} q(\lambda_i v_i) + \sum_{1 \le i < j \le n} b_q(\lambda_i v_i, \lambda_j v_j) \\ &= \sum_{i=1}^{n} \lambda_i^2 q(v_i) + \sum_{1 \le i < j \le n} \lambda_i \lambda_j b_q(v_i, v_j). \end{aligned}$$

Wegen $f(\lambda) = q(v)$ folgt die Behauptung, wenn wir $a_{ij} := b_q(v_i, v_j)$ setzen für $i < j$ und $a_{ii} := q(v_i)$.
„$\Longleftarrow$“ Sei also $f(\lambda)$ homogenes Polynom vom Grad 2 wie in der Behauptung. Dann ist $q(\alpha v) = f(\alpha\lambda) = \alpha^2 f(\lambda) = \alpha^2 q(v)$ klar.

$$\begin{aligned} q(v+w) = f(\lambda+\mu) &= \sum_{i \le j} a_{ij}(\lambda_i + \mu_i)(\lambda_j + \mu_j) \\ &= \underbrace{\sum_{i \le j} a_{ij}\lambda_i\lambda_j}_{q(v)} + \underbrace{\sum_{i \le j} a_{ij}\mu_i\mu_j}_{q(w)} + \sum_{i \le j} a_{ij}(\lambda_i\mu_j + \mu_i\lambda_j). \end{aligned}$$

Also gilt

$$b_q(v,w) = \sum_{1 \le i \le j \le n} a_{ij}(\lambda_i\mu_j + \lambda_j\mu_i)$$

und dies ist bilinear nach Beispiel 1.6. □

**2.7 Proposition.** *Es sei $b$ eine Bilinearform auf $V$. Dann ist $q(v) := b(v,v)$ eine quadratische Form von $V$. Weiter gilt $b_q(v,w) = b(v,w) + b(w,v)$.*

**Beweis:** Wir testen die Axiome für eine quadratische Form:

a)
$$q(\alpha v) = b(\alpha v, \alpha v) = \alpha b(v, \alpha v) = \alpha^2 b(v,v) = \alpha^2 q(v);$$

b)
$$\begin{aligned} b_q(v,w) &= q(v+w) - q(v) - q(w) = b(v+w, v+w) - b(v,v) - b(w,w) \\ &= b(v,v) + b(v,w) + b(w,v) + b(w,w) - b(v,v) - b(w,w) \\ &= b(v,w) + b(w,v). \end{aligned}$$

Also gilt $b_q(v,w) = b(v,w) + b(w,v)$ und dies ist offenbar bilinear. □

**2.8 Bemerkung.** Wir definieren $2 := 1 + 1 \in K$. Falls $2 \neq 0 \in K$ und $b$ eine symmetrische Bilinearform auf $V$ ist, dann können wir $b$ aus der quadratischen Form $q(v) := b(v,v)$ zurückgewinnen durch die Formel $b = \frac{1}{2} b_q$, d.h.

$$b(v,w) = \frac{1}{2}\big(q(v+w) - q(v) - q(w)\big) = \frac{1}{2}\big(b(v+w, v+w) - b(v,v) - b(w,w)\big).$$

**2.9.** Die bekannte Formel
$$\lambda \cdot \mu = \frac{1}{2}\big((\lambda+\mu)^2 - \lambda^2 - \mu^2\big)$$
für $\lambda, \mu \in K$ ist ein Spezialfall der obigen Identität für $b\colon K \times K \to K,\ (\lambda,\mu) \mapsto \lambda \cdot \mu$.

Das bekannteste Beispiel für einen Körper mit $2 = 0$ ist $\mathbb{Z}/2\mathbb{Z}$. In diesem Fall gibt es quadratische Formen $q(v)$ so, dass für keine symmetrische Bilinearform $b$ gilt $q(v) = b(v,v)\ \forall v \in V$.

**2.10 Theorem.** *Wir nehmen $2 \neq 0 \in K$ an. Weiter sei $b$ eine symmetrische Bilinearform auf $V$. Dann gibt es eine Basis $v_1, \dots, v_n$ von $V$ so, dass die zugehörige Matrix $B$ der Bilinearform $b$ diagonal ist.*

**Beweis:** Wir argumentieren mit Induktion nach $n = \dim(V)$. Im Fall $n = 1$ können wir irgendein $v \in V \setminus \{0\}$ als Basis wählen, weil jede $1 \times 1$ Matrix diagonal ist.

„*Induktionsschritt $n-1 \mapsto n$*“: Falls $b$ identisch Null ist, können wir irgendeine Basis $v_1, \dots, v_n$ wählen, weil $B$ dann immer die Nullmatrix ist und die ist diagonal. Also dürfen wir annehmen, dass $b$ nicht identisch Null ist. Dann gibt es ein $v_1 \in V$ mit $b(v_1, v_1) \neq 0$. Dies folgt sofort aus der Form in Bemerkung 2.8, die $b(v,w)$ in Funktion von $q$ ausdrückt. Wir zeigen zuerst für $U = \langle v_1 \rangle$, dass $V = U \oplus U^\perp$ gilt. Sei $u \in U \cap U^\perp$. Dann gilt $u = \alpha v_1$ für ein $\alpha \in K$ wegen $u \in U$. Aus $u \in U^\perp$ folgt
$$0 = b(u,u) = b(\alpha v_1, \alpha v_1) = \alpha^2 b(v_1, v_1).$$
Wegen $b(v_1, v_1) \neq 0$ folgt $\alpha^2 = 0$ und damit $\alpha = 0$. Also gilt $U \cap U^\perp = \{0\}$. Für $v \in V$ sei
$$u := b(v, v_1) b(v_1, v_1)^{-1} v_1, \quad u' := v - u.$$
Es gilt $u \in U$. Weiter gilt
$$b(u', v_1) = b(v - u, v_1) = b(v_1, v_1) - b(u, v_1) = b(v, v_1) - b(v, v_1) = 0.$$
Wegen der Linearität im zweiten Argument gilt auch $b(u', \alpha v_1) = 0$ und damit $u' \in U^\perp$. Wegen $v = u + u'$ folgt $V = U + U^\perp$. Insgesamt gilt $V = U \oplus U^\perp$. Aus $V = U \oplus U^\perp$ folgt $\dim(U^\perp) = \dim(V) - \dim(U) = n - 1$. Nach Induktionsannahme gibt es eine Basis $v_2, \dots, v_n$ von $U^\perp$ so, dass $b(v_i, v_j) = 0$ für $i \neq j$ aus $\{2, \dots, n\}$. Wegen $v_1 \in U$ und $v_2, \dots, v_n \in U^\perp$ gilt dies auch für $i$ oder $j$ gleich 1. Also ist $B = (b(v_i, v_j))_{1 \leq i,j \leq n}$ eine Diagonalmatrix. □

**2.11 Bemerkung.** Das Theorem besagt, dass es eine **Orthogonalbasis** von $V$ bzgl. $b$ gibt, d.h. eine Basis $v_1, \dots, v_n$ von $V$ mit $b(v_i, v_j) = 0$ für $i \neq j$. Das Verfahren im Beweis von Theorem 2.10 wird später die Grundlage des Gram-Schmidtschen Orthogonalisierungsverfahren sein.

**2.12 Korollar.** *Wenn $K = \mathbb{C}$ und $b$ eine symmetrische Bilinearform auf dem endlich dimensionalen komplexen Vektorraum $V$ ist, dann gibt es eine Basis $v_1, \dots, v_n$ von $V$ so, dass die Matrix $B$ zu $b$ die Form $B = \left(\begin{smallmatrix} E_r & 0 \\ 0 & 0 \end{smallmatrix}\right)$ hat mit der Einheitsmatrix $E_r \in M(r \times r, \mathbb{C})$.*

**Beweis:** Wir wählen die Basis $v_1, \dots, v_n$ aus Theorem 2.10. Damit hat die Matrix von $b$ Diagonalform. Wir dürfen dabei annehmen, dass $\beta_i := b(v_i, v_i) \neq 0$ für $i = 1, \dots, r$ und $b(v_i, v_i) = 0$ für $i = r+1, \dots, n$ (durch Vertauschen). In $\mathbb{C}$ kann man aus jeder Zahl die Wurzel ziehen. Also gibt es $\alpha_i \in \mathbb{C}$ mit $\alpha_i^2 = \beta_i$ für $i = 1, \dots, r$. Durch Ersetzen von $v_1, \dots, v_n$ durch $\alpha^{-1} v_1, \dots, \alpha_r^{-1} v_r$ erhalten wir die gewünschte Basis. □

**2.13 Korollar (Trägheitssatz von Sylvester).** *Sei $K = \mathbb{R}$ und $b$ eine symmetrische Bilinearform auf $V$. Dann gibt es eine Basis $v_1, \dots, v_n$ von $V$ so, dass die zu $b$ gehörige Matrix $B$ die Form $B = \begin{pmatrix} E_s & 0 & 0 \\ 0 & -E_t & 0 \\ 0 & 0 & 0 \end{pmatrix}$ hat, wobei $E_s \in M(s \times s, \mathbb{R})$, $E_t \in M(t \times t, \mathbb{R})$ die Einheitsmatrizen sind. Die Zahlen $s, t \in \mathbb{N}$ sind unabhängig von der Wahl der Basis.*

**Beweis:** Wir nehmen die Basis $v_1, \dots, v_n$ aus Theorem 2.10. Durch Umordnen dieser Basis dürfen wir annehmen, dass $b(v_i, v_i) > 0$ mit $1 \le i \le s$, $b(v_i, v_i) < 0$ für $s+1 \le i \le s+t$ und $b(v_i, v_i) = 0$ für $s+t+1 \le i \le n$. Für $i \in \{1, \dots, s+t\}$ setzen wir $v_i' := |b(v_i, v_i)|^{-1/2} v_i$. Aufgrund der Bilinearität von $b$ gilt dann

$$b(v_i', v_i') = \begin{cases} 1 & \text{für } i = 1, \dots, s \\ -1 & \text{für } i = s+1, \dots, s+t. \end{cases}$$

Also ist $v_1', \dots, v_{s+t}', v_{s+t+1}, \dots, v_n$ die gesuchte Basis. Die Eindeutigkeit wird später in Theorem 7.7 bewiesen. $\square$

**2.14 Bemerkung.** Als Folgerung aus Bemerkung 2.8 und Korollar 2.12 sehen wir, dass jede komplexe Form bzgl. einer geeigneten Basis die Form

$$q(v) = \lambda_1^2 + \cdots + \lambda_r^2$$

hat für den Koordinatenvektor $\lambda \in \mathbb{C}^n$ des Vektors $v$. Analog folgt aus Korollar 2.13, dass jede reelle quadratische Form bzgl. geeigneten Koordinaten die Form

$$q(v) = \lambda_1^2 + \cdots + \lambda_s^2 - \lambda_{s+1}^2 - \cdots - \lambda_{s+t}^2$$

hat für den Koordinatenvektor $\lambda \in \mathbb{R}^n$.

## 3. Skalarprodukte

Ein Skalarprodukt auf einem reellen Vektorraum ist eine spezielle symmetrische Bilinearform, mit der man die Länge von Vektoren und Winkel zwischen Vektoren berechnen kann. Damit kann man eine Metrik und somit Konvergenz auf diesen Vektorräumen definieren und erhält so eine Verbindung zwischen linearer Algebra und Analysis. Wir werden ein analoges Konzept für komplexe Vektorräume einführen und benötigen dazu den Begriff der **Sesquilinearform**, der die Bilinearform ersetzt. Um beide Fälle parallel zu behandeln, bezeichnen wir mit $\mathbb{K}$ entweder $\mathbb{R}$ oder $\mathbb{C}$. Zuerst einmal betrachten wir einen reellen Vektorraum $V$.

**3.1 Definition.** Eine reelle symmetrische Bilinearform $V \times V \to \mathbb{R}$, $(v, w) \mapsto \langle v, w \rangle$ heißt **positiv semi-definit**, falls $\langle v, v \rangle \ge 0 \ \forall v \in V$. Wenn sogar $\langle v, v \rangle > 0$ gilt für alle $v \in V \setminus \{0\}$, dann heißt $\langle\ ,\ \rangle$ **positiv definit**. Ein **Skalarprodukt** ist eine reelle positiv definite symmetrische Bilinearform auf $V$. Ein **euklidischer Vektorraum** ist ein endlich dimensionaler reeller Vektorraum mit einem gegebenen Skalarprodukt $\langle\ ,\ \rangle$.

**3.2 Bemerkung.** Wir haben im Trägheitssatz von Sylvester (Korollar 2.13) gesehen, dass jede symmetrische reelle Bilinearform $\langle\ ,\ \rangle$ auf einem $n$-dimensionalen $V$ bzgl. geeigneten Koordinaten von der Form

$$\begin{aligned} \left\langle \sum_{i=1}^n \lambda_i v_i, \sum_{j=1}^n \mu_j v_j \right\rangle &= \sum_{i,j=1}^n \lambda_i \mu_j \langle v_i, v_j \rangle \\ &= \lambda_1 \mu_1 + \cdots + \lambda_s \mu_s - \lambda_{s+1} \mu_{s+1} - \cdots - \lambda_{s+t} \mu_{s+t} \end{aligned}$$

ist. Insbesondere gilt

$$\left\langle \sum_{i=1}^n \lambda_i v_i, \sum_{i=1}^n \lambda_i v_i \right\rangle = \lambda_1^2 + \cdots + \lambda_s^2 - \lambda_{s+1}^2 - \cdots - \lambda_{s+t}^2.$$

Also ist $\langle\ ,\ \rangle$ genau dann positiv semi-definit, wenn $t = 0$ gilt. Weiter ist $\langle\ ,\ \rangle$ genau dann ein Skalarprodukt, wenn $t = 0$ und $s = n$.

**3.3 Beispiel.** Insbesondere ist $\langle\ ,\ \rangle\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $(\lambda,\mu) \mapsto \sum_{i=1}^n \lambda_i \mu_i$ ein Skalarprodukt auf $\mathbb{R}^n$. Wir sprechen vom **Standardskalarprodukt auf $\mathbb{R}^n$**. Für $n = 2,3$ entspricht es dem Skalarprodukt aus der analytischen Geometrie in der Schule.

**3.4 Bemerkung.** Für reelle Skalarprodukte übernehmen wir die Begriffe wie Orthogonalität, orthogonales Komplement $S^\perp$ und Matrix bzgl. einer Basis aus den vorangegangenen Abschnitten. Insbesondere ist $q(v) := \langle v,v\rangle$ die quadratische Form aus Proposition 2.7.

**3.5.** Als nächstes betrachten wir eine komplexen endlich dimensionalen Vektorraum $V$ der Dimension $n$. Unser Ziel ist es, Skalarprodukte auch für komplexe Vektorräume zu definieren. Zuerst erklären wir, wieso symmetrische komplexe Bilinearformen $b\colon V \times V \to \mathbb{C}$ nicht funktionieren. In Korollar 2.12 haben wir gesehen, dass

$$b\Big(\sum_{i=1}^n \lambda_i v_i, \sum_{j=1}^n \mu_j v_j\Big) = \sum_{i=1}^r \lambda_i \mu_i$$

gilt bezüglich geeigneten Koordinaten. Insbesondere gilt

$$b\Big(\sum_{j=1}^n \lambda_j v_j, \sum_{j=1}^n \lambda_j v_j\Big) = \sum_{j=1}^r \lambda_j^2.$$

Wegen $\mathrm{i}^2 = -1 \in \mathbb{C}$ gibt es immer Vektoren, für die das negativ wird außer wenn $r = 0$ ist. Also ist es uninteressant im Hinblick auf Längenmessung, komplexe symmetrische positiv definite Bilinearformen zu betrachten. Folgende Definition löst dieses Problem:

**3.6 Definition.** Eine **Sesquilinearform** auf dem komplexen Vektorraum $V$ ist eine Funktion $h\colon V \times V \to \mathbb{C}$ mit folgenden Axiomen

a) $h(\alpha_1 v_1 + \alpha_2 v_2, w) = \alpha_1 h(v_1,w) + \alpha_2 h(v_1,w)$ („linear im 1. Argument“)

b) $h(v, \alpha_1 w_1 + \alpha_2 w_2) = \overline{\alpha_1} h(v,w_1) + \overline{\alpha_2} h(v,w_2)$ („konjugiert linear im 2. Argument“)

für alle $\alpha_1, \alpha_2 \in \mathbb{C}$ und alle $v_1, v_2, w_1, w_2, v, w \in V$. Eine Sesquilinearform $h$ heißt **Hermitesche Form**, falls $h(w,v) = \overline{h(v,w)}$ gilt für alle $v,w \in V$. Insbesondere gilt $h(v,v) \in \mathbb{R}$ für eine Hermitesche Form. Eine Hermitesche Form $h$ heißt **positiv semi-definit**, falls $h(v,v) \geq 0\ \forall v \in V$. Falls sogar $h(v,v) > 0$ gilt $\forall v \in V \setminus \{0\}$, dann heißt $h$ **positiv definit** und wir sprechen von einem **Skalarprodukt**, das wir üblicherweise wieder mit $\langle\ ,\ \rangle$ bezeichnen. Ein **unitärer Vektorraum** ist ein endlich dimensionaler komplexer Vektorraum mit gegebenem Skalarprodukt.

**3.7.** Für eine Sesquilinearform $h$ auf einem endlich dimensionalen komplexen Vektorraum $V$ der Dimension $n$ übertragen wir die alten Begriffe wie Matrix $H = \big(h(v_i,v_j)\big)_{1\leq i,j\leq n}$ bzgl. einer Basis $v_1,\ldots,v_n$ von $V$, orthogonal und nicht ausgeartet. Beachte, dass $S^\perp = {}^\perp S$ für Hermitesche Formen. Weiter definieren wir für $A \in M(m\times n,\mathbb{C})$ die Matrix $\overline{A}$, in dem wir alle Einträge der Matrix komplex konjugieren, z.B. $A = \begin{pmatrix} 1 & \mathrm{i} \\ 1+\mathrm{i} & 2\end{pmatrix} \Longrightarrow \overline{A} = \begin{pmatrix} 1 & -\mathrm{i} \\ 1-\mathrm{i} & 2\end{pmatrix}$ Für diese Konjugation zeigt man leicht folgende Rechenregeln:

$$\overline{A_1 + A_2} = \overline{A_1} + \overline{A_2}, \quad \overline{\alpha A} = \overline{\alpha}\,\overline{A},$$
$$\overline{A\cdot B} = \overline{A}\cdot\overline{B}, \quad \overline{A^t} = \overline{A}^t$$

für alle $\alpha \in \mathbb{C}, A, A_1, A_2 \in M(m\times n,\mathbb{C}), B \in M(n\times p,\mathbb{C})$. Dies folgt aus den Rechenregeln $\overline{z+w} = \overline{z} + \overline{w}$ und $\overline{z\cdot w} = \overline{z}\cdot\overline{w}$ für komplexe Zahlen $z,w \in \mathbb{C}$.

**3.8 Proposition.** *Sei $h$ eine Sesquilinearform mit Matrix $H = (h_{ij})$ bzgl. der Basis $v_1,\ldots,v_n$ von $V$. Dann gilt*

$$h\Big(\sum_{i=1}^n \lambda_i v_i, \sum_{j=1}^n \mu_j v_j\Big) = \lambda^t \cdot H \cdot \overline{\mu} = \sum_{i,j=1}^n h_{ij} \lambda_i \overline{\mu_j}.$$

**Beweis:** Siehe Aufgabe 3.3. □

**3.9 Bemerkung.** Umgekehrt definiert jede komplexe $n\times n$-Matrix $H=(h_{ij})$ durch obige Formel eine Sesquilinearform $h$ mit zugehöriger Matrix $H$ bzgl. der gegebenen Basis $v_1,\dots,v_n$. Analog zu Proposition 3.8 ist klar, dass $h$ genau dann eine Hermitesche Form ist, wenn $H=\overline{H}^t$ gilt. Deshalb nennen wir eine Matrix $A\in M(n\times n,\mathbb{C})$ mit $A=\overline{A}^t$ eine **Hermitesche Matrix**.

**3.10 Beispiel.** Nach 3.8 und 3.9 ist $\langle\;,\;\rangle:\mathbb{C}^n\times\mathbb{C}^n\to\mathbb{C},\ (\lambda,\mu)\mapsto\sum_{j=1}^n\lambda_j\overline{\mu_j}$ eine Hermitesche Form auf $\mathbb{C}^n$ mit zugehöriger Matrix $H=E_n$ bzgl. der Standardbasis. Wegen

$$\langle\lambda,\lambda\rangle=\sum_{j=1}^n\lambda_i\overline{\lambda_i}=\sum_{j=1}^n|\lambda_j|^2,$$

ist $\langle\;,\;\rangle$ ein Skalarprodukt. Wir nennen es das **Standardskalarprodukt** auf $\mathbb{C}^n$. Wir werden später sehen, dass jedes Skalarprodukt bzgl. geeigneter Koordinaten auf diese Form gebracht werden kann.

**3.11.** Wir haben gesehen, dass wir Skalarprodukte im komplexen Fall analog zum reellen Fall behandeln können, wenn wir Sesquilinearformen (bzw. Hermitesche Formen) benutzen statt Bilinearformen (bzw. symmetrische Bilinearformen). Wir werden ab jetzt diese beiden Fälle parallel behandeln und deshalb $\mathbb{K}=\mathbb{R}$ oder $\mathbb{C}$ setzen. Unter einer Sesquilinearform (bzw. einer Hermiteschen Form) versteht man im Fall $\mathbb{K}=\mathbb{R}$ eine Bilinearform (bzw. symmetrische Bilinearform), was im Einklang mit der Definition ist, weil dann $\alpha=\overline{\alpha}\ \forall\alpha\in\mathbb{K}=\mathbb{R}$ gilt.

**3.12 Theorem (Schwarz'sche Ungleichung).** *Sei $\langle\;,\;\rangle$ eine positiv semi-definite Hermitesche Form auf dem komplexen Vektorraum $V$. Dann gilt*

$$|\langle v,w\rangle|^2\le\langle v,v\rangle\cdot\langle w,w\rangle$$

*für alle $v,w\in V$. Falls $\langle\;,\;\rangle$ ein Skalarprodukt auf $V$ ist, dann gilt: „$=$“ $\Longleftrightarrow$ $v,w$ linear abhängig.*

**Beweis:** Für $\alpha,\beta\in\mathbb{K}$ gilt

$$0\underset{\text{pos. semi-def}}{\le}\langle\alpha v+\beta w,\alpha v+\beta w\rangle\underset{\text{sesquilinear}}{=}\alpha\overline{\alpha}\langle v,v\rangle\alpha\overline{\beta}\langle v,w\rangle+\beta\overline{\alpha}\langle w,v\rangle+\beta\overline{\beta}\langle w,w\rangle.$$

Für eine komplexe Zahl $z$ gilt

$$z+\overline{z}=\mathrm{Re}(z)+\mathrm{i}\cdot\mathrm{Im}(z)+\mathrm{Re}(z)-\mathrm{i}\cdot\mathrm{Im}(z)=2\,\mathrm{Re}(z).$$

Also folgt zusammen mit $\overline{z_1\cdot z_2}=\overline{z_1}\cdot\overline{z_2}$ für $z_1,z_2\in\mathbb{C}$, dass

$$\alpha\overline{\beta}\langle v,w\rangle+\beta\overline{\alpha}\langle w,v\rangle=\alpha\overline{\beta}\langle v,w\rangle+\beta\overline{\alpha}\overline{\langle v,w\rangle}=\alpha\overline{\beta}\langle v,w\rangle+\overline{\alpha\overline{\beta}\langle v,w\rangle}=2\,\mathrm{Re}\big(\alpha\overline{\beta}\langle v,w\rangle\big).$$

Zusammen mit $\alpha\overline{\alpha}=|\alpha|^2$ folgt

$$0\le|\alpha|^2\langle v,v\rangle+2\,\mathrm{Re}\big(\alpha\overline{\beta}\langle v,w\rangle\big)+|\beta|^2\langle w,w\rangle. \tag{VIII.1}$$

*1.Fall:* $\langle v,v\rangle=\langle w,w\rangle=0$ Dann setzt man $\alpha:=-1$ und $\beta:=\langle v,w\rangle$ in (VIII.1) ein und erhält

$$0\le\langle v,v\rangle+2\,\mathrm{Re}\big((-1)\cdot\overline{\langle v,w\rangle}\cdot\langle v,w\rangle\big)+|\langle v,w\rangle|^2\cdot\langle w,w\rangle=2\,\mathrm{Re}\big(-|\langle v,w\rangle|^2\big)=-2\,\mathrm{Re}\big(|\langle v,w\rangle|^2\big).$$

Also gilt $0\le-2|\langle v,w\rangle|^2$ und somit folgt $|\langle v,w\rangle|=0$ wie gewünscht.

*2.Fall:* $\langle v,v\rangle\ne0$ oder $\langle w,w\rangle\ne0$. Wegen $\langle w,v\rangle=\overline{\langle v,w\rangle}$ und damit $|\langle w,v\rangle|=|\langle v,w\rangle|$ dürfen wir annehmen, dass $\langle v,v\rangle\ne0$. Aus positiv semi-definit folgt dann $\langle v,v\rangle>0$. In diesem Fall setzen wir $\alpha:=-\overline{\langle v,w\rangle}$ und $\beta:=\langle v,v\rangle$. Dann folgt aus (VIII.1):

$$0\le|\langle v,w\rangle|^2\langle v,v\rangle+2\,\mathrm{Re}\big(-\langle v,v\rangle|\langle v,w\rangle|^2\big)+\langle v,v\rangle^2\langle w,w\rangle=-|\langle v,w\rangle|^2\langle v,v\rangle+\langle v,v\rangle^2\langle w,w\rangle.$$

Weil $\langle v,v\rangle$ eine reelle Zahl $> 0$ ist, dürfen wir mit $\langle v,v\rangle$ kürzen und erhalten dann die Schwarz'sche Ungleichung.

Wenn $v$ und $w$ linear abhängig sind, dann ist der eine Vektor ein Vielfaches des anderen Vektors. OBdA $w = \alpha \cdot v$ für ein $\alpha \in \mathbb{K}$.

$$\begin{aligned}|\langle v,w\rangle|^2 &= |\langle v,\alpha v\rangle|^2 = |\overline{\alpha}\langle v,v\rangle|^2 = |\overline{\alpha}|^2\langle v,v\rangle|^2 = |\alpha|^2\langle v,v\rangle^2 = \alpha\cdot\overline{\alpha}\langle v,v\rangle\langle v,v\rangle = \langle v,v\rangle\overline{\alpha}\langle \alpha v,v\rangle \\ &= \langle v,v\rangle\langle \alpha v,\alpha v\rangle = \langle v,v\rangle\langle w,w\rangle.\end{aligned}$$

Dies zeigt „=“ in der Schwarz'schen Ungleichung im Fall linear abhängiger Vektoren. Umgekehrt gelte

$$|\langle v,w\rangle|^2 = \langle v,v\rangle\langle w,w\rangle. \tag{VIII.2}$$

Wir setzen wieder $\alpha := -\overline{\langle v,w\rangle}$ und $\beta := \langle v,v\rangle$, wie im 2.Fall. Dann folgt

$$\begin{aligned}\langle \alpha v+\beta w,\alpha v+\beta w\rangle &= |\alpha|^2\langle v,v\rangle + \alpha\overline{\beta}\langle v,w\rangle + \beta\overline{\alpha}\langle w,v\rangle + |\beta|^2\langle w,w\rangle \\ &= |\langle v,w\rangle|^2\langle v,v\rangle - \overline{\langle v,w\rangle}\langle v,v\rangle\langle v,w\rangle - \langle v,v\rangle\langle v,w\rangle\langle w,v\rangle + \langle v,v\rangle^2\langle w,w\rangle \\ &\underset{\text{hermitesch}}{=} |\langle v,w\rangle|^2\langle v,v\rangle + \langle v,v\rangle^2\langle w,w\rangle - 2\langle w,v\rangle\langle v,v\rangle\langle v,w\rangle \\ &= -|\langle v,w\rangle|^2\langle v,v\rangle + \langle v,v\rangle^2\langle w,w\rangle.\end{aligned}$$

Wir setzen (VIII.2) ein und es folgt

$$\langle \alpha v+\beta w,\alpha v+\beta w\rangle = -\langle v,v\rangle\langle w,w\rangle\langle v,v\rangle + \langle v,v\rangle^2\langle w,w\rangle^2 = 0.$$

Weil $\langle\ ,\ \rangle$ positiv semi-definit ist, folgt $0 = \alpha v + \beta w$ und damit sind $v$ und $w$ linear abhängig. □

Wie in der Einleitung angedeutet, ist eine der wichtigsten Eigenschaften des Skalarprodukts, dass man damit Längen messen kann. Dazu gibt es auf einem $\mathbb{K}$-Vektorraum $V$ folgenden Begriff:

**3.13 Definition.** Eine Funktion $\|\ \|\colon V \to \mathbb{R}$ heißt **Seminorm**, wenn folgende Axiome erfüllt sind

(i) $\|\alpha v\| = |\alpha|\|v\|$,

(ii) $\|v+w\| \leq \|v\| + \|w\|$ (**Dreiecksungleichung**)

für alle $\alpha \in \mathbb{K}$ und $v,w \in V$. Falls zusätzlich das Axiom

(iii) $\|v\| = 0 \Longrightarrow v = 0$

gilt für alle $v \in V$, dann sprechen wir von einer **Norm**.

**3.14 Bemerkung.**

- Es gilt $\|0\| = 0$, was aus (i) mit $\alpha = 0$ folgt.
- Es gilt $\|v\| = \|-v\|$, was wieder aus (i) mit $\alpha = -1$ folgt.
- Jede Seminorm ist nicht negativ.

  **Beweis:** $0 = \|0\| = \|v+(-v)\| \leq \|v\| + \|-v\| = 2\|v\|$. □

Also sind Normen für Längenmessung geeignet.

**3.15 Korollar.** *Für eine positiv semi-definite Hermitesche Form $\langle\ ,\ \rangle$ setzen wir $\|v\| := \sqrt{\langle v,v\rangle}$. Damit erhalten wir eine Seminorm auf $V$.*

**Beweis:** Für $\alpha \in \mathbb{K}$ und $v \in V$ erhalten wir

$$\|\alpha v\| = \langle \alpha v, \alpha v\rangle^{1/2} = \big(\alpha\overline{\alpha}\langle v,v\rangle\big)^{1/2} = \big(|\alpha|^2\langle v,v\rangle\big)^{1/2} = |\alpha|\|v\|.$$

Also gilt Axiom (i) einer Seminorm. Weiter folgt

$$\begin{aligned}\|v+w\|^2 &\underset{\text{sesquilinear}}{=} \langle v,v\rangle + \langle v,w\rangle + \langle w,v\rangle + \langle w,w\rangle \underset{\text{hermitesch}}{=} \langle v,v\rangle + \langle v,w\rangle + \overline{\langle v,w\rangle} + \langle w,w\rangle \\ &= \langle v,v\rangle + 2\operatorname{Re}\langle v,w\rangle + \langle w,w\rangle \leq \langle v,v\rangle + 2|\langle v,w\rangle| + \langle w,w\rangle \leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \\ &= \big(\|v\| + \|w\|\big)^2.\end{aligned}$$

Weil alles nicht negative Zahlen vorkommen, darf man die Wurzel ziehen und es folgt Axiom (ii) für $\|\ \|$. $\square$

**3.16 Bemerkung.** Falls $\langle\ ,\ \rangle$ ein Skalarprodukt ist, dann definiert $\|v\| := \sqrt{\langle v,v\rangle}$ eine Norm (wegen positiv definit folgt Axiom (iii)). Wir sprechen von der **Norm des Skalarprodukts**. Es gilt die **Parallelogrammgleichung**

$$\|v+w\|^2 + \|v-w\|^2 = 2\big(\|v\|^2 + \|w\|^2\big).$$

**Beweis:**

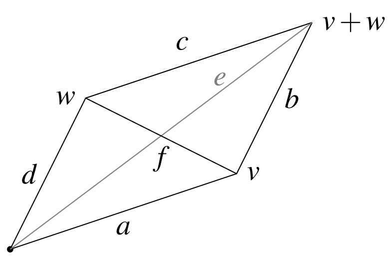
$$\begin{aligned}\|v+w\|^2 + \|v-w\|^2 &= \langle v+w, v+w\rangle + \langle v-w, v-w\rangle \\ &= \langle v,v\rangle + \langle v,w\rangle + \langle w,v\rangle + \langle w,w\rangle + \langle v,v\rangle - \langle v,w\rangle - \langle w,v\rangle + \langle w,w\rangle \\ &= 2\|v\|^2 + 2\|w\|^2.\end{aligned}$$ $\square$

Umgekehrt kann man zeigen, dass jede Norm, die die Parallelogrammgleichung erfüllt, die Norm eines Skalarprodukts ist (siehe [JN]).

**3.17 Beispiel.** Sei $\mathbb{K} = \mathbb{R}$ und $V = \mathbb{R}^2$. Dann ist die Norm zum Standard-Skalarprodukt gegeben durch $\|\binom{\lambda_1}{\lambda_2}\| = \sqrt{\lambda_1^2 + \lambda_2^2}$. Die Parallelogrammgleichung besagt, dass

$$a^2 + b^2 + c^2 + d^2 = e^2 + f^2$$

im Parallelogramm:




,

da $v - w$ der zweite Diagonalvektor ist.

**3.18 Beispiel.** Sei wieder $\mathbb{K} = \mathbb{R}$ und $V = \mathbb{R}^2$. Durch

$$\left\|\begin{pmatrix}\lambda_1\\\lambda_2\end{pmatrix}\right\| := \max\{|\lambda_1|, |\lambda_2|\}$$

wird ebenfalls eine Norm auf $V$ definiert:

(i)
$$\left\|\alpha\begin{pmatrix}\lambda_1\\\lambda_2\end{pmatrix}\right\| = \max\{|\alpha\lambda_1|, |\alpha\lambda_2|\} = |\alpha|\max\{|\lambda_1|, |\lambda_2|\} = |\alpha|\left\|\begin{pmatrix}\lambda_1\\\lambda_2\end{pmatrix}\right\|;$$

(ii)

$$\begin{aligned}\left\| \begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix} + \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} \right\| &= \max\{|\lambda_1+\mu_1|, |\lambda_2+\mu_2|\} \\ &\leq \max\{|\lambda_1|,|\lambda_2|\} + \max\{|\mu_1|,|\mu_2|\} = \left\|\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}\right\| + \left\|\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}\right\|;\end{aligned}$$

(iii) $\|\binom{\lambda_1}{\lambda_2}\| = 0 \Longrightarrow |\lambda_i| = 0$ für $i = 1,2$. Also ist $\lambda_i = 0$ für $i = 1,2$.

Sei $v = \binom{1}{0}$ und $w = \binom{0}{1}$, dann gilt

$$\|v+w\|^2 + \|v-w\|^2 = 1+1 = 2,$$

aber

$$2\big(\|v\| + \|w\|^2\big) = 2\cdot(1+1) = 4.$$

Somit gilt die Parallelogrammgleichung für diese Norm nicht und deshalb ist sie nicht die Norm eines Skalarproduktes!

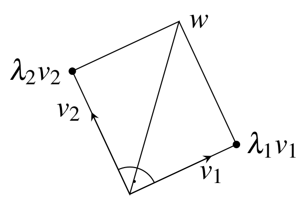
**3.19 Bemerkung.** Sei $\|\ \|$ die Norm des Skalarprodukts $\langle\ ,\ \rangle$ auf dem $\mathbb{K}$-Vektorraum $V$. Dann besagt die Schwarz'sche Ungleichung (Theorem 3.12)

$$\langle v,w\rangle \leq \|v\|\cdot\|w\| \qquad \forall v,w \in V.$$

Wir nehmen an, dass $v \neq 0 \neq w$. Dann können wir den unorientierten Zwischenwinkel $\varphi$ von $v$ und $w$ definieren durch

$$\cos(\varphi) = \frac{\langle v,w\rangle}{\|v\|\cdot\|w\|} \in [-1,1].$$

Er ist bestimmt bis auf das Vorzeichen in $\mathbb{R}/2\pi\mathbb{Z}$ (deshalb unorientiert). Das stimmt mit unserer geometrischen Anschauung überein. Falls $v,w$ linear unabhängig sind, dann ist der von $v$ und $w$ aufgespannte Unterraum $U$ zweidimensional. In geeigneten Koordinaten ist die Einschränkung des Skalarprodukts (von $V$) auf $U$ gleich dem Standardskalarprodukt auf $\mathbb{R}^2$ (siehe Bemerkung 3.2). Explizit können wir als den ersten Basisvektor $v_1 = \frac{v}{\|v\|}$ nehmen und $v_2 \in U$ einen normierten Vektor aus $U$, der orthogonal zu $v_1$ ist.




$v = \|v\|\cdot v_1$, $w = \lambda_1 v_1 + \lambda_2 v_2$

$$\langle v,w\rangle = \|v\|\lambda_1 \Longrightarrow \frac{\langle v,w\rangle}{\|v\|\cdot\|w\|} = \frac{\lambda_1}{\|w\|}$$

und dies ist auch der Kosinus des geometrischen Zwischenwinkels.

## 4. Orthogonalität

Beim Studium der Bilinearformen sind wir schon auf die Bedeutung von Orthogonalität gestoßen. Wir werden die Resultate für den Fall eines Skalarprodukts $\langle\ ,\ \rangle$ auf einem $\mathbb{K}$-Vektorraum $V$ der Dimension $n < \infty$ vertiefen, wobei $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$. Eine spezielle Bedeutung hat die orthogonale Projektion.

**4.1 Proposition.** *Es seien $e_1, \dots, e_r \in V \setminus \{0\}$ paarweise orthogonal zueinander. Dann sind $e_1, \dots, e_r$ $\mathbb{K}$-linear unabhängig.*

**Beweis:** Es sei $\alpha_1 e_1 + \ldots + \alpha_r e_r = 0$ für $\alpha_i \in \mathbb{K}$. Für $j \in \{1, \dots, r\}$ gilt dann

$$0 = \langle \alpha_1 e_1 + \cdots + \alpha_r e_r, e_j \rangle = \alpha_j \langle e_j, e_j \rangle.$$

Aus positiv definit folgt $\alpha_j = 0$. □

**4.2 Definition.** Eine Basis von $V$ heißt **Orthogonalbasis**, falls die Basisvektoren paarweise zueinander orthogonal sind. Falls die Basisvektoren zusätzlich alle die Norm 1 haben, dann sprechen wir von einer **Orthonormalbasis** (ONB).

**4.3 Korollar.** *In $V$ bilden $n$ paarweise, orthogonale Vektoren $e_1, \dots, e_n \in V \setminus \{0\}$ immer eine Orthogonalbasis.*

**Beweis:** Nach Proposition 4.1 sind sie linear unabhängig und $n = \dim(V)$ bilden sie damit eine Basis (Proposition IV.3.11). □

**4.4 Proposition.** *Es sei $e_1, \dots, e_n$ eine Orthonormalbasis von $V$. Dann sind die Koordinaten von $v \in V$ bzgl. $e_1, \dots, e_n$ gegeben durch*

$$v = \sum_{j=1}^{n} \langle v, e_j \rangle e_j.$$

*Wir nennen das die **Fourierentwicklung** von $v$ bzgl. $e_1, \dots, e_n$.*

**Beweis:** Sei $v = \sum_{j=1}^{n} \lambda_j e_j$ Dann gilt $\langle v, e_i \rangle = \sum_{j=1}^{n} \lambda_j \langle e_j, e_i \rangle = \lambda_i \; \forall i$. □

**4.5.** Wir zeigen nun, wie man aus einer gegebenen Basis $v_1, \dots, v_n$ von $V$ eine Orthonormalbasis $e_1, \dots, e_n$ erhält. Dies wird durch das **Gram-Schmidtsche-Orthogonalisierungsverfahren** gemacht, das an den Beweis von Theorem 3.12 anlehnt. Wir beschreiben den Algorithmus:

1.*Schritt:* $e_1 := \|v_1\|^{-1} v_1$. Offenbar ist das eine Orthonormalbasis von $\langle v_1 \rangle$. Wir nehmen nun an, dass wir eine Orthonormalbasis $e_1, \dots, e_r$ von $\langle v_1, \dots, v_r \rangle$ schon konstruiert haben im $r$.Schritt.

$r+1$.*Schritt:*

$$e'_{r+1} := v_{r+1} - \sum_{j=1}^{r} \langle v_{r+1}, e_j \rangle e_j, \quad e_{r+1} := \|e'_{r+1}\|^{-1} e'_{r+1}.$$

Offenbar ist $e'_{r+1}$ ein Element des $r+1$-dimensionalen Untervektorraums $\langle v_1, \dots, v_{r+1} \rangle$, das verschieden vom Nullvektor ist. Weiter gilt für $i \in \{1, \dots, r\}$:

$$\langle e'_{r+1}, e_i \rangle = \langle v_{r+1}, e_i \rangle - \sum_{j=1}^{r} \langle v_{r+1}, e_j \rangle \langle e_j, e_i \rangle \overset{\text{ONB}}{=} \langle v_{r+1}, e_i \rangle - \langle v_{r+1}, e_i \rangle = 0.$$

Nach Korollar 4.3 ist $e_1, \dots, e_r, e'_{r+1}$ eine Orthogonalbasis von $\langle v_1, \dots, v_{r+1} \rangle$ und damit $e_1, \dots, e_r, e_{r+1}$ eine Orthonormalbasis von $\langle v_1, \dots, v_{r+1} \rangle$. Nach dem $n$-ten Schritt erhalten wir eine Orthonormalbasis von $V$.

**4.6 Theorem.** *Sei $U$ ein Unterraum von $V$. Dann existiert eine Orthonormalbasis von $U$. Weiter können wir jede Orthonormalbasis von $U$ zu einer Orthonormalbasis von $V$ ergänzen.*

**Beweis:** Wähle eine Basis $v_1, \dots, v_m$ von $U$ und ergänze sie zu einer Basis $v_1, \dots, v_n$ von $V$ (Theorem IV.2.5). Mit dem Gram-Schmidt-Orthogonalisierungsverfahren folgt die Behauptung. □

**4.7 Korollar.** *Sei $U$ ein Unterraum von $V$. Dann gilt:*

*a)* $V = U \oplus U^{\perp}$;

*b)* $\dim(U^\perp) = \dim(V) - \dim(U)$*;*

*c)* $(U^\perp)^\perp = U$*.*

**Beweis:** Nach Theorem 4.6 gibt es eine Orthonormalbasis von $U$, die wir zu einer Orthonormalbasis $e_1, \dots, e_n$ von $V$ ergänzen können. Dann gilt:

$$\langle e_{m+1}, \dots, e_n \rangle \subseteq U^\perp.$$

Die Fourier Entwicklung zeigt andererseits, dass jedes $v \in U^\perp$ in $\langle e_{m+1}, \dots, e_n \rangle$ liegt. Also gilt

$$U^\perp = \langle e_{m+1}, \dots, e_n \rangle$$

und dies zeigt a) bis c). □

**4.8 Korollar.** *Es sei $v \in V$ und $U$ ein Unterraum von $V$. Dann existiert genau ein Element $(u, w) \in U \times U^\perp$ mit $v = u + w$.*

**Beweis:** Folgt direkt aus Korollar 4.7a). □

Wir nennen $u$ aus Korollar 4.8 die **Orthogonalprojektion** von $v$ auf $U$ und bezeichnen sie mit $P_U(v)$. Für eine beliebige Orthonormalbasis $e_1, \dots, e_m$ von $U$ folgt aus dem Beweis von Korollar 4.7, dass

$$P_U(v) = \sum_{j=1}^{m} \langle v, e_j \rangle e_j. \tag{VIII.3}$$

**4.9 Proposition.** *Es gilt: $P_U \in \mathrm{End}_K(V)$ mit $P_U(V) = U$ und $P_U^2 = P_U$. Solche Abbildungen nennt man Projektionen auf $U$.*

**Beweis:** Die Linearität folgt aus der Linearität von $\langle \cdot, e_j \rangle$ und der Formel (VIII.3). Für $i \in \{1, \dots, m\}$ gilt wieder mit (VIII.3) $P_U(e_i) = e_i$ und damit $P_U(V) = U$. Weil nun jedes $u \in U$ eine Linearkombination von $e_1, \dots, e_m$ ist, gilt auch $P_U^2 = P_U$. □

**4.10 Theorem.** *Die Orthogonalprojektion $P_U(v)$ ist gleich dem Element $u_0 \in U$ mit minimalem Abstand zu $v$, das heißt*

$$\|v - u_0\| \leq \|v - u\| \quad \textit{für alle } u \in U.$$

**Beweis:** Für $u \in U$ gilt:

$$\begin{aligned} \|v-u\|^2 = \|v - P_U(v) + P_U(v) - u\|^2 &\underset{v - P_U(v) \in U^\perp}{=} \|v - P_U(v)\|^2 + \|P_U(v) - u\|^2 \\ &\geq \|v - P_U(V)\|^2. \end{aligned}$$

Weiter gilt „=“ $\iff P_U(v) = u$. □

## 5. Adjungierte Abbildung

Wir betrachten in diesem Abschnitt einen $\mathbb{K}$-Vektorraum $V$ der Dimension $n < \infty$ mit einem gegebenen Skalarprodukt $\langle \ , \ \rangle$, wobei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Für $\varphi \in \mathrm{End}_\mathbb{K}(V)$ werden wir die adjungierte Abbildung $\varphi^* \in \mathrm{End}_\mathbb{K}(V)$ definieren, die verwandt ist mit der dualen Abbildung aus III.3. Mit Hilfe des Skalarproduktes kann man $V$ mit seinem Dual $V^*$ mit dem folgendem **Lemma von Riesz** verbinden.

**5.1 Lemma (Riesz).** *Für $f \in V^* \Longrightarrow \exists! w \in V$ mit $f(\,\cdot\,) = \langle \,\cdot\,, w \rangle$.*

**Beweis:** Nach Theorem 4.6 gibt es eine orthonormale Basis $e_1,\dots,e_n$ von $V$. Für beliebiges $w = \sum_{j=1}^{n} \lambda_j e_j \in V$ mit Koordinaten $\lambda_1,\dots,\lambda_n \in \mathbb{K}$ gilt

$$\langle e_i, w\rangle = \left\langle e_i, \sum_{j=1}^{n} \lambda_j e_j \right\rangle = \sum_{j=1}^{n} \overline{\lambda_j} \langle e_i, e_j\rangle = \overline{\lambda_i}.$$

Also gelten folgende Äquivalenzen:

$$\begin{aligned} f(\cdot) = \langle \cdot, w\rangle \overset{e_1,\dots,e_n}{\underset{\text{Basis}}{\iff}} & f(e_i) = \langle e_i, w\rangle \quad \forall i = 1,\dots,n \\ \iff & f(e_i) = \overline{\lambda_i} \quad \forall i = 1,\dots,n \\ \iff & \lambda_i = \overline{f(e_i)} \quad \forall i = 1,\dots,n. \end{aligned}$$

Dies zeigt die Existenz und die Eindeutigkeit von $w$. □

**5.2 Proposition.** *Sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$. Dann $\exists! \varphi^* \in \mathrm{End}_{\mathbb{K}}(V)$ mit*

$$\langle \varphi(v), w\rangle = \langle v, \varphi^*(w)\rangle$$

*für alle $v, w \in V$.*

**Beweis:** Es sei $w \in V$. Dann sind $\varphi(\cdot)$ und $\langle \cdot, w\rangle$ lineare Abbildungen und somit gilt $\langle \varphi(\,\cdot\,), w\rangle \in V^*$. Nach dem Lemma von Riesz $\exists! \varphi^*(w) \in V$ mit

$$\langle \varphi(v), w\rangle = \langle v, \varphi^*(w)\rangle$$

für alle $v \in V$. Es bleibt zu zeigen, dass die Abbildung $\varphi^* : V \to V,\ w \mapsto \varphi^*(w)$, linear ist. Für $\alpha_1, \alpha_2 \in \mathbb{K}$ und $w_1, w_2 \in V$ gilt

$$\begin{aligned} \langle v, \varphi^*(\alpha_1 w_1 + \alpha_2 w_2)\rangle & \underset{\text{Def. von } \varphi^*}{=} \langle \varphi(v), \alpha_1 w_1 + \alpha_2 w_2\rangle \\ & \underset{\text{konjugiert linear im 2. Arg.}}{=} \overline{\alpha_1}\langle \varphi(v), w_1\rangle + \overline{\alpha_2}\langle \varphi(v), w_2\rangle \\ & \underset{\text{Def.von } \varphi^*}{=} \overline{\alpha_1}\langle v, \varphi^*(w_1)\rangle + \overline{\alpha_2}\langle v, \varphi^*(w_2)\rangle \\ & \underset{\text{konjugiert linear im 2. Arg.}}{=} \langle v, \alpha_1 \varphi^*(w_1) + \alpha_2 \varphi^*(w_2)\rangle \end{aligned}$$

für alle $v \in V$. Nun sind also $\langle\,\cdot\,, \varphi^*(\alpha_1 w_1 + \alpha_2 w_2)\rangle$ und $\langle\,\cdot\,, \alpha_1 \varphi^*(w_1) + \alpha_2 \varphi^*(w_2)\rangle$ zwei Linearformen auf $V$, die gleich sind. Wegen der Eindeutigkeit im Lemma von Riesz folgt

$$\varphi^*(\alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 \varphi^*(w_1) + \alpha_2 \varphi^*(w_2)$$

und damit die Linearität von $\varphi^*$. □

**5.3 Definition.** $\varphi^*$ heißt **die zu $\varphi$ adjungierte Abbildung**.

**5.4 Bemerkung.** Das Lemma von Riesz ergibt einen konjugiert linearen Isomorphismus $V \to V^*$, $w \mapsto \langle\,\cdot\,, w\rangle$, wobei bei konjugiert linear $\varphi(\alpha v) = \overline{\alpha}\varphi(v)$ statt $\varphi(\alpha v) = \alpha\varphi(v)$ gilt. Damit kann man $V$ mit seinem Dualraum $V^*$ „identifizieren“ und dabei entspricht die duale Abbildung aus Definition III.3.16 der adjungierten Abbildung von hier. Das erklärt auch dieselbe Notation. Für Einzelheiten siehe [Bo, §7.4].

**5.5 Proposition.** *Sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ mit zugehöriger Matrix $A$ bzgl. der orthonormalen Basis $e_1,\dots,e_n$. Dann hat $\varphi^*$ die Matrix $\overline{A}^t$ bzgl. $e_1,\dots,e_n$.*

**Beweis:** Für $v = \sum_{i=1}^n \lambda_i e_i$ und $w = \sum_{j=1}^n \mu_j e_j$ folgt aus Proposition 3.8.

$$\langle \varphi(v), w \rangle = (A \cdot \lambda)^t \cdot E_n \cdot \overline{\mu}$$

denn die Matrix der Sesquilinearform $\langle\ ,\ \rangle$ ist $E_n$, weil $e_1, \ldots, e_n$ eine orthonormale Basis ist. Nach den Rechenregeln aus 3.9 gilt

$$\langle \varphi(v), w \rangle = \lambda^t \cdot A^t \cdot \overline{\mu} = \lambda^t \cdot \overline{\overline{A}^t \cdot \mu}.$$

Andererseits gilt

$$\langle \varphi(v), w \rangle = \langle v, \varphi^*(w) \rangle$$

und das charakterisiert $\varphi^*(w)$. Somit gilt

$$\langle v, \varphi^*(w) \rangle = \lambda^t \cdot \overline{\overline{A}^t \cdot \mu}$$

für alle $v \in V$. Es sei $\varphi^*(w) = \sum_{k=1}^n \rho_k e_k$ die Koordinatendarstellung von $\varphi^*(w)$, dann gilt

$$\langle v, \varphi^*(w) \rangle = \lambda^t \cdot \overline{\rho}.$$

Also folgt

$$\lambda^t \cdot \overline{\overline{A}^t \cdot \mu} = \lambda^t \cdot \overline{\rho}$$

für alle $v \in V$. Wieder mit dem Lemma von Riesz (für das Standardskalarprodukt) folgt

$$\overline{\overline{A}^t \cdot \mu} = \overline{\rho}$$

und somit $\rho = \overline{A}^t \mu$, d.h. $\overline{A}^t \mu$ ist der Koordinatenvektoren von $\varphi^*(w)$ bzgl. $e_1, \ldots, e_n$. Also ist $\overline{A}^t$ die Matrix von $\varphi^*$ bzgl. $e_1, \ldots, e_n$. □

**5.6 Definition.** Ein Endomorphismus $\varphi$ von $V$ heißt **selbstadjungiert** $\Longleftrightarrow \varphi = \varphi^*$.

**5.7 Korollar.** *Es sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ mit Matrix $A$ bzgl. der orthonormalen Basis $e_1, \ldots, e_n$. Dann ist $\varphi$ genau dann selbstadjungiert, wenn $A = \overline{A}^t$ gilt.*

Dies folgt sofort aus Proposition 5.5. Im Fall $\mathbb{K} = \mathbb{R}$ (bzw. $\mathbb{K} = \mathbb{C}$) bedeutet das Korollar, dass die Matrix $A$ symmetrisch (bzw. hermitesch) ist. Wenn man eine nicht orthonormierte Basis wählt, dann muss das nicht mehr stimmen (siehe Übung 5.4).

**5.8 Proposition.** *Es sei $U$ ein Unterraum von $V$ und $P \in \mathrm{End}_{\mathbb{K}}(V)$ mit $P(V) = U$. Dann ist $P$ genau dann die Orthogonalprojektion auf $U$, wenn $P^2 = P$ und $P = P^*$ gilt.*

**Beweis:** „$\Longrightarrow$" Sei $P_U$ die Orthogonalprojektion auf $U$. Dann gilt

$$\langle P_U(v), w \rangle = \big\langle P_U(v), P_U(w) + \big(w - P_U(w)\big)\big\rangle = \langle P_U(v), P_U(w) \rangle + \langle P_U(v), w - P_U(w) \rangle.$$

Jetzt benutzen wir $P_U(v) \in U$ und $w - P_U(w) \in U^\perp$.

$$\Longrightarrow \langle P_U(v), w \rangle = \langle P_U(v), P_U(w) \rangle$$

für alle $v, w \in V$. Analog erhalten wir

$$\langle v, P_U(w) \rangle = \langle P_U(v), P_U(w) \rangle.$$

Es folgt

$$\langle P_U(v), w \rangle = \langle v, P_U(w) \rangle$$

für alle $v, w \in V$. Nach Definition gilt also $P_U^*(w) = P_U(w)$, d.h. $P_U^* = P_U$. Nach Proposition 4.9 folgt auch $P_U^2 = P_U$.

„$\Longleftarrow$" Es gelte $P^2 = P$ und $P = P^*$. Zu zeigen $P = P_U$. Dazu genügt es zu zeigen, dass für ein beliebiges $v \in V$ gilt $u := P(v) \in U$ und $w := v - u \in U^\perp$ (siehe Korollar 4.7). Nach Voraussetzung gilt $u \in P(V) = U$ und somit ist die erste Bedingung erfüllt. Für $u' \in U$ folgt

$$\begin{aligned}
\langle u', w \rangle &= \langle u', v - u \rangle = \langle u', v - P(v) \rangle \\
&= \langle u', v \rangle - \langle u', P(v) \rangle \\
&\underset{P=P^*}{=} \langle u', v \rangle - \langle P(u'), v \rangle.
\end{aligned}$$

Wegen $u' \in U = P(V)$ gibt es ein $v' \in V$ mit $u' = P(v')$.

$$\Longrightarrow P(u') = P(P(v')) = P^2(v') = P(v') = u'$$

$$\Longrightarrow \langle u', w \rangle = \langle u', v \rangle - \langle P(u'), v \rangle = \langle u', u \rangle - \langle u', v \rangle = 0$$

Das zeigt $w \in U^\perp$ und damit ist auch die zweite Bedingung erfüllt. Es folgt $P = P_U$. $\square$

**5.9 Definition.** Ein Endomorphismus $\varphi$ von $V$ heißt **normal** $\Longleftrightarrow$ $\varphi^* \circ \varphi = \varphi \circ \varphi^*$ gilt.

**5.10 Bemerkung.** Jede selbstadjungierte Abbildung ist normal.

**5.11 Proposition.** *Es sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$. Dann ist $\varphi$ genau dann normal, wenn*

$$\langle \varphi(v), \varphi(w) \rangle = \langle \varphi^*(v), \varphi^*(w) \rangle$$

*für alle $v, w \in V$.*

**Beweis:** Für $v, w \in V$ gilt

$$\langle \varphi(v), \varphi(w) \rangle = \langle v, \varphi^* \circ \varphi(w) \rangle$$

und

$$\langle \varphi^*(v), \varphi^*(w) \rangle = \langle v, \varphi^{**} \circ \varphi^*(w) \rangle \underset{\varphi^{**}=\varphi \text{ nach Aufgabe 5.1}}{=} \langle v, \varphi \circ \varphi^*(w) \rangle.$$

Nach dem Lemma von Riesz gilt

$$\begin{aligned}
\langle \varphi(v), \varphi(w) \rangle = \langle \varphi^*(v), \varphi^*(w) \rangle \quad \forall v, w \in V &\Longleftrightarrow \langle v, \varphi^* \circ \varphi(w) \rangle = \langle v, \varphi \circ \varphi^*(w) \rangle \quad \forall v, w \in V \\
&\overset{\text{Riesz}}{\Longleftrightarrow} \varphi^* \circ \varphi(w) = \varphi \circ \varphi^*(w) \quad \forall w \in V.
\end{aligned}$$ $\square$

**5.12 Korollar.** *Sei $\varphi$ ein normaler Endomorphismus von $V$. Dann gelten*

*a)* $\ker(\varphi) = \ker(\varphi^*)$.

*b) $v$ ist genau dann Eigenvektor von $\varphi$ mit Eigenwert $\lambda$, wenn $v$ Eigenvektor von $\varphi^*$ ist mit Eigenwert $\overline{\lambda}$.*

*c) Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal aufeinander.*

**Beweis:** a) Für $v \in V$ gilt

$$\|\varphi(v)\| = \|\varphi^*(v)\|$$

und somit folgt a).

b) Es sei $\psi := \lambda \cdot \mathrm{id}_V - \varphi$. Es gilt $\psi \in \mathrm{End}_{\mathbb{K}}(V)$. Nach Übung 5.1 gilt $\psi^* = \overline{\lambda} \cdot \mathrm{id}_V - \varphi^*$. Wegen

$$\psi^* \circ \psi = (\overline{\lambda} \cdot \mathrm{id}_V - \varphi^*) \circ (\lambda \cdot \mathrm{id}_V - \varphi) = \overline{\lambda} \cdot \lambda \cdot \mathrm{id}_V - \overline{\lambda}\varphi - \lambda\varphi^* + \varphi^* \circ \varphi$$

und

$$\psi \circ \psi^* = \lambda\overline{\lambda} \cdot \mathrm{id}_V - \overline{\lambda}\varphi - \lambda\varphi^* + \varphi \circ \varphi^*$$

ist auch $\psi$ normal. Mit a) folgt

$$\ker(\lambda \cdot \mathrm{id}_V - \varphi) = \ker(\overline{\lambda} \cdot \mathrm{id}_V - \varphi^*)$$

und damit b).

c) Falls $\varphi(v) = \lambda v$, $\varphi(w) = \mu w$ und $\lambda \neq \mu$

$$\Longrightarrow \lambda \langle v,w\rangle = \langle \varphi(v), w\rangle = \langle v, \varphi^*(w)\rangle \overset{\text{b)}}{=} \mu \langle v,w\rangle \Longrightarrow \langle v,w\rangle = 0.$$ □

**5.13 Korollar.** *Selbstadjungierte Endomorphismen haben nur reelle Eigenwerte.*

**Beweis:** Im Fall $\mathbb{K} = \mathbb{R}$ sind sowieso alle Eigenwerte reell. Also OBdA $\mathbb{K} = \mathbb{C}$. Sei $\lambda$ ein Eigenwert von der selbstadjungierten Abbildung $\varphi\colon V \to V$ mit Eigenvektor $v$. Weil selbstadjungierte Abbildung normal sind, können wir Korollar 5.12 anwenden und schließen, dass $v$ ein Eigenvektor von $\varphi^*$ zum Eigenwert $\overline{\lambda}$. Wegen $\varphi = \varphi^*$ ist also $v$ ein Eigenvektor von $\varphi$ zum Eigenvektor $\overline{\lambda}$. Also gilt

$$\overline{\lambda} v = \varphi^*(v) = \varphi(v) = \lambda v$$

und damit $(\overline{\lambda} - \lambda)v = 0$. Wegen $v \neq 0 \Longrightarrow \lambda - \overline{\lambda} = 0$, d.h. $\lambda \in \mathbb{R}$. □

Der wichtigste Satz in diesem Abschnitt ist der **Spektralsatz für normale Abbildungen**.

**5.14 Theorem.** *Sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$. Dann sind folgende Bedingungen äquivalent.*

*(i) $\varphi$ ist normal und das charakteristische Polynom zerfällt vollständig in Linearfaktoren.*

*(ii) Es existiert eine Orthonormalbasis von $V$ aus Eigenvektoren bzgl. $\varphi$.*

**Beweis:** „$\Longrightarrow$“ Mit Induktion nach $n = \dim(V)$. Der Fall $n = 0$ ist trivial. Wem dabei unwohl ist, darf den Induktionsanfang bei $n = 1$ machen. Dieser Fall ist ebenfalls trivial, da jeder Endomorphismus eines 1-dimensionalen Vektorraums ein Vielfaches von $\mathrm{id}_V$ ist. *Induktionsschritt* „$n-1 \mapsto n$“. Es sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ normal und $\chi_\varphi(x)$ zerfalle vollständig in Linearfaktoren aus $\mathbb{K}[x]$. Zu zeigen ist (ii) für $\varphi$. Weil $\chi_\varphi(x)$ in Linearfaktoren zerfällt, gibt es eine Nullstelle $\lambda_1$ von $\chi_\varphi(x)$. Nach Proposition VII.3.9 sind die Nullstellen von $\chi_\varphi(x)$ gerade die Eigenwerte von $\varphi$ und damit ist $\lambda_1$ ein Eigenwert von $\varphi$. Sei $e_1$ ein Eigenvektor zum Eigenwert $\lambda_1$. Nach Korollar 4.7 gilt

$$V = \mathbb{K}e_1 \oplus (\mathbb{K}e_1)^\perp.$$

Für $w \in W := (\mathbb{K}e_1)^\perp$ gilt

$$\langle \varphi(w), e_1\rangle = \langle w, \varphi^*(e_1)\rangle \overset{5.12\text{ b)}}{=} \langle w, \overline{\lambda_1} e_1\rangle = \lambda_1 \langle w, e_1\rangle \overset{w \in (\mathbb{K}e_1)^\perp}{=} 0.$$

Es folgt $\varphi(w) \in (\mathbb{K}e_1)^\perp = W$. Also induziert $\varphi$ einen Endomorphismus

$$\psi\colon W \to W,\ w \mapsto \varphi(w).$$

Nach Proposition VII.4.5 ist $\chi_\psi(x)$ ein Teiler von $\chi_\varphi(x)$ und zerfällt somit vollständig in Linearfaktoren in $\mathbb{K}[x]$. Wenn wir das Skalarprodukt $\langle\ ,\ \rangle$ von $V$ auf $W$ einschränken, muss $\varphi^*(w) = \psi^*(w)$ gelten für alle $w \in W$, denn

$$\langle \psi(w_1), w_2\rangle = \langle \varphi(w_1), w_2\rangle = \langle w_1, \varphi^*(w_2)\rangle \qquad w_1, w_2 \in W.$$

Aber damit folgt sofort, $\psi \circ \psi^* = \psi^* \circ \psi$ durch Einschränkung der Identität $\varphi^* \circ \varphi = \varphi^* \circ \varphi$. Also ist auch $\psi$ normal. Wegen $\dim(W) = n-1$ (siehe Korollar 4.7), können wir die Induktionsannahme auf $\psi$ anwenden. Somit existiert eine Orthonormalbasis $e_2, \dots, e_n$ von $W$ aus Eigenvektoren bzgl. $\psi$. Weil $\psi = \varphi|_W$, müssen $e_2, \dots, e_n$ auch Eigenvektoren von $\varphi$ sein. Wenn wir $e_1$ noch normieren, dann ergibt $e_1, \dots, e_n$ eine orthogonale Familie in $V$ aus normierten Eigenvektoren. Nach Korollar 4.3 muss das eine Basis von $V$ sein. Also ist $e_1, \dots, e_n$ eine Orthonormalbasis von $V$ aus Eigenvektoren bzgl. $\varphi$. Mit vollständiger Induktion gilt (i) $\Longrightarrow$ (ii).

„$\Longleftarrow$“ Sei $e_1, \dots, e_n$ eine Orthonormalbasis aus Eigenvektoren bzgl. $\varphi$ . Also gilt $\varphi(w_j) = \lambda_j e_j$ für den Eigenwert $\lambda_j$ von $\varphi$. Dann ist die zu $\varphi$ gehörige Matrix $A$ bzgl. der Basis $e_1, \dots, e_n$ gleich

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

mit $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, die nicht paarweise verschieden sein müssen. Aus Proposition 5.5 folgt, dass die zu $\varphi^*$ gehörige Matrix gleich

$$\overline{A}^t = \begin{pmatrix} \overline{\lambda}_1 & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda}_n \end{pmatrix}$$

ist. Da Diagonalmatrizen trivialerweise miteinander kommutieren, gilt $A\overline{A}^t = \overline{A}^t A$. Das sind aber die Matrizen von $\varphi \circ \varphi^*$ bzw. $\varphi^* \circ \varphi$ bzgl. $e_1, \dots, e_n$ (siehe Proposition V.1.13) und damit ist $\varphi$ normal. Weiter gilt $\chi_\varphi(x) = (x - \lambda_1) \cdots (x - \lambda_n)$. Also folgt (i) aus (ii). $\square$

**5.15 Korollar.** *Es sei $\varphi \in \mathrm{End}_{\mathbb{C}}(V)$ eines endlich dimensionalen* ***komplexen*** *Vektorraums. Dann sind folgende Aussagen äquivalent:*

*a)* *$\varphi$ normal.*

*b)* *$\exists$ Orthonormalbasis von $V$ aus Eigenvektoren zu $\varphi$.*

**Beweis:** Nach dem Fundamentalsatz der Algebra (Theorem VII.1.8) zerfällt jedes nicht-konstante komplexe Polynom vollständig in Linearfaktoren. Insbesondere gilt das für das charakteristische Polynom $\chi_\varphi(x)$. Also folgt a) $\Longleftrightarrow$ b) nach dem Spektralsatz (Theorem 5.14). $\square$

**5.16 Proposition.** *Sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ für einen endlich-dimensionalen $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Dann gilt:*

*a)* $\ker(\varphi^*) = \varphi(V)^\perp$*;*

*b)* $\ker(\varphi) = \varphi^*(V)^\perp$.

**Beweis:** Weil $\langle\ ,\ \rangle$ positiv definit ist, gilt:

$$\begin{aligned} w \in \ker(\varphi^*) &\Longleftrightarrow \varphi^*(w) = 0 \Longleftrightarrow \langle v, \varphi^*(w) \rangle = 0 \ \forall v \in V \\ &\Longleftrightarrow \langle \varphi(v), w \rangle = 0 \ \forall v \in V \Longleftrightarrow w \in \varphi(V)^\perp. \end{aligned}$$

Das zeigt a). Teil b) folgt aus a), wenn man $\varphi^*$ statt $\varphi$ benutzt und dann $\varphi^{**} = \varphi$ benützt (Übung 5.1). $\square$

## 6. Isometrien

Eine Isometrie in $\mathbb{R}^2$ ist entweder eine Drehung mit Zentrum 0 oder eine Geradenspiegelung an einer Geraden durch 0. In diesem Abschnitt verallgemeinern wir diese Betrachtung auf endlich dimensionale $\mathbb{K}$-Vektorräume mit Skalarprodukt. Wie immer ist $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Es seien $(V_1, \langle\ ,\ \rangle_1)$ und $(V_2, \langle\ ,\ \rangle_2)$ endlich dimensionale $\mathbb{K}$-Vektorräume mit Skalarprodukt.

**6.1 Definition.** Eine **Isometrie** von $(V_1, \langle\ ,\ \rangle_1)$ nach $(V_2, \langle\ ,\ \rangle_2)$ ist eine $\mathbb{K}$-lineare Abbildung $\varphi\colon V_1 \to V_2$ mit
$$\langle \varphi(v), \varphi(w)\rangle_2 = \langle v, w\rangle_1$$
für alle $v, w \in V_1$.

**6.2 Proposition.** *Folgende Bedingungen sind äquivalent für eine lineare Abbildung $\varphi\colon V_1 \to V_2$:*

*a) $\varphi$ ist eine Isometrie;*

*b) $\|\varphi(v)\|_2 = \|v\|_1\ \forall v \in V$*

*wobei $\|\ \|_i$ die Norm zum Skalarprodukt $\langle\ ,\ \rangle_i$ ist ($i = 1, 2$).*

**Beweis:** a) $\Longrightarrow$ b)
$$\|\varphi(v)\|_2 = \big(\langle \varphi(v), \varphi(v)\rangle_2\big)^{1/2} \underset{\varphi \text{ Isometrie}}{=} \big(\langle v, v\rangle_1\big)^{1/2} = \|v\|_1.$$

b) $\Longrightarrow$ a) Analog zu oben folgt
$$\langle \varphi(v), \varphi(v)\rangle_2 = \langle v, v\rangle_1$$
für alle $v \in V$ (aus b)). Im Fall $\mathbb{K} = \mathbb{R}$ gilt
$$\langle v, w\rangle_1 = \frac{1}{2}\big(\langle v+w, v+w\rangle_1 - \langle v, v\rangle_1 - \langle w, w\rangle_1\big) \tag{VIII.4}$$
und analog für $\langle\ ,\ \rangle_2$ erhält man
$$\langle \varphi(v), \varphi(w)\rangle_2 = \frac{1}{2}\big(\langle \varphi(v+w), \varphi(v+w)\rangle_2 - \langle \varphi(v), \varphi(v)\rangle_2 - \langle \varphi(w), \varphi(w)\rangle_2\big).$$
Zusammen mit obigen Identitäten ergibt sich
$$\langle \varphi(v), \varphi(w)\rangle_2 = \frac{1}{2}\big(\langle v+w, v+w\rangle_1 - \langle v, v\rangle_1 - \langle w, w\rangle_1\big) = \langle v, w\rangle_1.$$
Im Fall $\mathbb{K} = \mathbb{C}$ können wir statt (VIII.4) die Polarisationsidentität
$$\langle v, w\rangle_1 = \frac{1}{4}\big(\langle v+w, v+w\rangle_1 - \langle v-w, v-w\rangle_1 + i\langle v+iw, v+iw\rangle_1 - i\langle v-iw, v-iw\rangle_1\big)$$
aus Aufgabe 4.4 benutzen und dann analog schließen, dass $\varphi$ eine Isometrie ist. ☐

**6.3 Proposition.** *Sei $V$ ein endlich dimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$ und $n =$* $\dim(V)$*. Weiter sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ mit Matrix $A$ bzgl. der Orthonormalbasis $e_1, \ldots, e_n$. Folgende Bedingungen sind dann äquivalent:*

*a) $\varphi$ ist eine Isometrie;*

*b) $\|\varphi(v)\| = \|v\| \quad \forall v \in V$, wobei $\|\ \| =$ Norm zu $\langle\ ,\rangle$;*

*c) $\varphi(e_1), \ldots, \varphi(e_n)$ ist Orthonormalbasis von $V$;*

*d) $\varphi$ ist eine Isomorphismus und $\varphi^{-1} = \varphi^*$;*

*e) $A \in \mathrm{GL}(n, \mathbb{K})$ und $A^{-1} = \overline{A}^t$;*

*f) Die Spalten von $A$ sind orthonormiert bzgl. dem Standardskalarprodukt auf $\mathbb{K}^n$;*

*g) Die Zeilen von $A$ sind orthonormiert bzgl. dem Standardskalarprodukt auf $\mathbb{K}^n$.*

**Beweis:** a) $\Longleftrightarrow$ b) Folgt aus Proposition 6.2 im Spezialfall $V = V_1 = V_2$.

a) $\Longrightarrow$ c) $\langle \varphi(e_i), \varphi(e_j)\rangle = \langle e_i, e_j\rangle = \delta_{ij}$. Also ist $\varphi(e_1), \dots, \varphi(e_n)$ orthonormierte Familie der Länge $n = \dim(V)$. Nach Korollar 4.3 ist damit $\varphi(e_1), \dots, \varphi(e_n)$ eine Basis von $V$.

c) $\Longrightarrow$ f) Sei $a_j$ der $j$-te Spaltenvektor von $A$. Weiter bezeichnen wir mit $\langle\ ,\ \rangle_{\text{std}}$ des Standard-Skalarprodukt auf $\mathbb{K}^n$. Dann gilt

$$\delta_{ij} \underset{\text{c)}}{=} \langle \varphi(e_i), \varphi(e_j)\rangle \underset{\text{Def. von } A}{=} \left\langle \sum_{g=1}^n a_{gi} e_g, \sum_{h=1}^n a_{hj} e_h \right\rangle = \sum_{g,h=1}^n a_{gi} \overline{a}_{hj} \langle e_g, e_h\rangle \underset{\text{ONB}}{=} \sum_{h=1}^n a_{hi} \overline{a_{hj}} = \langle a_i, a_j\rangle_{\text{std}}$$

f) $\Longrightarrow$ e) Es gilt $\overline{A}^t \cdot A = E_r$ und mit Proposition V.2.5 folgt $A^{-1} = \overline{A}^t$.

e) $\Longrightarrow$ d) Folgt aus Proposition 5.5.

d) $\Longrightarrow$ a) $\langle \varphi(v), \varphi(w)\rangle = \langle v, \varphi^* \circ \varphi(w)\rangle = \langle v, w\rangle$.

e) $\Longleftrightarrow$ g) Beachte, dass g) äquivalent ist zu $A \cdot \overline{A}^t = E_n$ und damit folgt die Behauptung. $\square$

**6.4 Bemerkung.** Im Fall $\mathbb{K} = \mathbb{R}$ (bzw. $\mathbb{K} = \mathbb{C}$) nennen wir eine Isometrie $\varphi\colon V \to V$ eine **orthogonale** (bzw. **unitäre**) Abbildung. Weiter heißt $A \in \mathrm{GL}(n, \mathbb{R})$ mit $A^{-1} = A^t$ eine **orthogonale Matrix**. Analog heißt $A \in \mathrm{GL}(n, \mathbb{C})$ mit $A^{-1} = \overline{A}^t$ eine **unitäre Matrix**.

Wir erinnern daran, dass invertierbare lineare Selbstabbildungen **Automorphismen** heißen. Es sind also die invertierbaren Elemente aus dem Endomorphismenring $\big(\mathrm{End}(V), \circ\big)$ und deshalb bilden die Automorphismen von $V$ eine Gruppe bzgl. der Komposition von Selbstabbildungen. Sie heißt die **Automorphismengruppe** von $V$ und wir mit $\mathrm{Aut}_{\mathbb{K}}(V)$ bezeichnet. Wenn man eine feste Basis von $V$ wählt, dann gehört zu jedem $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ eine Matrix und dabei entspricht die $\mathrm{Aut}_{\mathbb{K}}(V)$ der Gruppe $\mathrm{GL}(n, \mathbb{K})$. Wir werden zeigen, dass orthogonale (bzw. unitäre) Abbildungen eine Untergruppe von $\mathrm{Aut}_{\mathbb{K}}(V)$ bilden und entsprechend bilden orthogonale (bzw. unitäre) Matrizen eine Untergruppe von $\mathrm{GL}(n, \mathbb{K})$.

**6.5 Proposition.** *Die Isometrien eines $\mathbb{K}$-Vektorraums $V$ mit Skalarprodukt bilden eine Gruppe bzgl. der Komposition und Selbstabbildung.*

**Beweis:** Wir müssen zeigen, dass Isometrien eine Untergruppe von $\mathrm{Aut}_{\mathbb{K}}(V)$ bilden. Aus Proposition 6.3 folgt, dass jede Isometrie auch wirklich in $\mathrm{Aut}_{\mathbb{K}}(V)$ enthalten ist. Offensichtlich ist $\mathrm{id}_V$ eine Isometrie und damit ist das Neutralelement von $\mathrm{Aut}_{\mathbb{K}}(V)$ eine Isometrie. Wir zeigen nun, dass die Verknüpfung zweier Isometrien $\varphi_1, \varphi_2$ wieder eine Isometrie ist. Es gilt

$$\big\langle \varphi_1 \circ \varphi_2(v), \varphi_1 \circ \varphi_2(w)\big\rangle = \big\langle \varphi_2(v), \varphi_2(w)\big\rangle = \langle v, w\rangle$$

für alle $v, w \in V$. Also ist auch $\varphi_1 \circ \varphi_2$ eine Isometrie. Wir müssen noch zeigen, dass die Inverse $\varphi^{-1}$ einer Isometrie wieder eine Isometrie ist. Es gilt

$$\big\langle \varphi^{-1}(v), \varphi^{-1}(w)\big\rangle \underset{\varphi \text{ Isometrie}}{=} \big\langle \varphi(\varphi^{-1}(v)), \varphi(\varphi^{-1}(w))\big\rangle = \langle v, w\rangle$$

für alle $v, w \in V$. Also bilden die Isometrien von $V$ eine Untergruppe von $\mathrm{Aut}_{\mathbb{K}}(V)$ und damit auch eine Gruppe wie gewünscht. $\square$

**6.6 Theorem.** *Jeder endlich dimensionale $\mathbb{K}$-Vektorraum $V$ mit Skalarprodukt $\langle\ ,\ \rangle$ ist isometrisch isomorph zu $\mathbb{K}^n$ mit Standard-Skalarprodukt $\langle\ ,\ \rangle_{\text{std}}$ und $n := \dim(V)$.*

**Beweis:** Nach Theorem 4.6 existiert Orthonormalbasis $e_1, \dots, e_n$ von $V$. Wir wählen als Isomorphismus $\varphi\colon V \to \mathbb{K}^n$ die Koordinatenabbildung bzgl. $e_1, \dots, e_n$. Wir müssen zeigen, dass $\varphi$ eine Isometrie ist. Seien also $v, w \in V$ mit Koordinatendarstellung $v = \sum_{i=1}^n \lambda_i e_i$, $w = \sum_{j=1}^n \mu_j e_j$. Dann gilt

$\varphi(v) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ und $\varphi(w) = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$ nach Definition. Wegen

$$\begin{aligned}\langle \varphi(v), \varphi(w) \rangle_{\text{std}} &= \sum_{i=1}^{n} \lambda_i \overline{\mu}_i = \sum_{i,j=1}^{n} \lambda_i \overline{\mu_j} \delta_{ij} \overset{\text{ONB}}{=} \sum_{i,j=1}^{n} \lambda_i \overline{\mu_j} \langle e_i, e_j \rangle \\ &= \Big\langle \sum_{i=1}^{n} \lambda_i e_i, \sum_{j=1}^{n} \mu_j e_j \Big\rangle = \langle v, w \rangle\end{aligned}$$

ist $\varphi$ eine Isometrie wie gewünscht. □

**6.7 Proposition.** *Sei $G$ die Basiswechselmatrix von einer Orthonormalbasis $e_1, \ldots, e_n$ zu einer weiteren Basis $v_1, \ldots, v_n$ von $V$. Dann ist $G$ genau dann orthogonal (bzw. unitär), wenn $v_1, \ldots, v_n$ eine Orthonormalbasis ist.*

**Beweis:** Sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ gegeben durch die Vorschrift, dass $\varphi(e_i) = v_i$ für $i = 1, \ldots, n$. Dann ist die Matrix $A$ von $\varphi$ bzgl. der Basis $e_1, \ldots, e_n$ gleich der Basiswechselmatrix von $v_1, \ldots, v_n$ zur Basis $e_1, \ldots, e_n$ (vgl. Bemerkung V.4.3). Also gilt $A = G^{-1}$ (Korollar V.4.5). Nach Proposition 6.3 gilt $A$ orthogonal (bzw. unitär) $\underset{\text{Def}}{\Longleftrightarrow} \overline{A}^t = A^{-1} \underset{\text{Proposition 6.3}}{\Longleftrightarrow} v_1 = \varphi(e_1), \ldots, v_n = \varphi(e_n)$ Orthonormalbasis von $V$. Nach Bemerkung 6.4 bilden die orthogonalen (bzw. unitären) Matrizen eine Gruppe, d.h. $A$ ist orthogonal (bzw. unitär) $\Longleftrightarrow G = A^{-1}$ orthogonal (bzw. unitär). Damit folgt die Behauptung. □

Der **Spektralsatz für unitäre Abbildungen** lautet:

**6.8 Theorem.** *Sei $(V, \langle\ ,\ \rangle)$ ein unitärer Vektorraum und $\varphi \in \mathrm{End}_{\mathbb{C}}(V)$. Dann ist $\varphi$ genau dann eine unitäre Abbildung, wenn es eine Orthonormalbasis $e_1, \ldots, e_n$ von $V$ gibt mit $\varphi(e_j) = \lambda_j e_j$ und $|\lambda_j| = 1$ $\forall j = 1, \ldots, n$. Insbesondere ist $\varphi$ dann diagonalisierbar und alle Eigenwerte von $\varphi$ haben Betrag* 1.

**Beweis:** „$\Longrightarrow$“ Sei $\varphi$ unitär. Dann ist $\varphi$ normal wegen

$$\varphi \circ \varphi^* = \varphi \circ \varphi^{-1} = \mathrm{id}_V = \varphi^{-1} \circ \varphi = \varphi^* \circ \varphi$$

Nach dem Spektralsatz für komplexe normale Abbildungen (Korollar 5.15) gibt es eine Orthonormalbasis $e_1, \ldots, e_n$ von $V$ mit $\varphi(e_j) = \lambda_j e_j$ für $j = 1, \ldots, n$. Da $\varphi$ eine Isometrie ist, gilt

$$1 \underset{ONB}{=} \|e_j\| = \|\varphi(e_j)\| = \|\lambda_j e_j\| = |\lambda_j| \|e_j\| = |\lambda_j|.$$

„$\Longleftarrow$“ Es sei $e_1, \ldots, e_n$ eine Orthonormalbasis von $V$ mit $\varphi(e_j) = \lambda_j e_j$ und $|\lambda_j| = 1$ für $j = 1, \ldots, n$. Bzgl. der Orthonormalbasis $e_1, \ldots, e_n$ hat $\varphi$ die Darstellungsmatrix

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Weil $A \cdot \overline{A}^t = E_n = \overline{A}^t \cdot A$ gilt, muss $A$ eine unitäre Matrix sein. Also ist $\varphi$ eine unitäre Abbildung nach Proposition 6.3. Dies zeigt "$\Longleftarrow$". Weil alle Eigenwerte von $\varphi$ in der Diagonalen von $A$ auftauchen, ist jeder Eigenwert von der Form $\lambda_j$, also folgt auch die letzte Behauptung. □

**6.9 Lemma.** *Jedes normierte reelle Polynom $p(x) \in \mathbb{R}[x] \setminus \mathbb{R}$ lässt sich in $\mathbb{R}[x]$ faktorisieren durch $p(x) = (x - \lambda_1) \cdots (x - \lambda_p) \cdot q_1(x) \cdots q_s(x)$ mit $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ und $q_j(x) = (x - \mu_j)(x - \overline{\mu_j})$ für $\mu_j \in \mathbb{C} \setminus \mathbb{R}$ und $j = 1, \ldots, s$. Die Faktorisierung von $p(x)$ ist eindeutig bis auf Vertauschung.*

**Beweis:** Die Existenz folgt aus Aufgabe 1.1. Die Eindeutigkeit folgt aus der Eindeutigkeit der komplexen Nullstellen mit ihren Vielfachheiten (Korollar VII.4.1). □

**6.10 Beispiel.** Wir studieren orthogonale Abbildungen zuerst im 2-dimensionalen Fall. Nach Theorem 6.6 dürfen wir $V = \mathbb{R}^2$ mit dem Standardskalarprodukt annehmen. Eine orthogonale Abbildung $\varphi \in \mathrm{End}(\mathbb{R}^2)$ ist damit durch eine eine orthogonale $2 \times 2$ Matrix $A$ gegeben.

*1. Fall:* Das charakteristische Polynom $\chi_A(x)$ zerfällt in Linearfaktoren in $\mathbb{R}[x]$. Also gilt $\chi_A(x) = (x-\lambda_1)(x-\lambda_2)$ mit $\lambda_1, \lambda_2 \in \mathbb{R}$. Eine orthogonale Abbildung ist normal (analog zum Beweis von Theorem 6.8). Also können wir den Spektralsatz 5.14 für unser $\varphi$ anwenden (weil $\chi_\varphi$ zerfällt). Also gibt es eine Orthonormalbasis $v_1, v_2$ von $\mathbb{R}^2$ mit $\varphi(v_j) = \lambda_j v_j$ für $j = 1,2$. Weil $\varphi$ eine Isometrie ist, gilt $|\lambda_j| = 1$ analog zum Beweis von Theorem 6.8. Also folgt $\lambda_j \in \{-1, 1\}$. Bis auf Vertauschung der Basisvektoren haben wir folgende orthogonale Abbildung im 1.Fall:

| $\lambda_1$ | $\lambda_2$ | Matrix bzgl. $v_1, v_2$ | $\varphi$ |
|---|---|---|---|
| $1$ | $1$ | $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ | Identität |
| $1$ | $-1$ | $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ | Geradenspiegelung an $\mathbb{R}v_1$ |
| $-1$ | $-1$ | $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ | Punktspiegelung an 0 |

*2. Fall:* $\chi_A(x)$ zerfällt nicht in Linearfaktoren aus $\mathbb{R}[x]$. Wir wenden dann einen wichtigen Trick an, den man **Komplexifizierung** nennt. Wir betrachten die reelle orthogonale Matrix $A$ als komplexe $2 \times 2$ Matrix. Wegen $\overline{A}^t = A^t = A^{-1}$ ist $A$ eine unitäre $2 \times 2$ Matrix. Nach dem Spektralsatz (Theorem 6.8) gibt es eine Orthonormalbasis $w_1, w_2 \in \mathbb{C}^2$ mit $Aw_j = \lambda_j w_j$ für $j = 1,2$ und $|\lambda_j| = 1$. Es gilt

$$\chi_A(x) = (x-\lambda_1)(x-\lambda_2)$$

und nach Annahme im 2.Fall $\lambda_1 \notin \mathbb{R}$ (OBdA). Weiter haben wir

$$A \cdot \overline{w_1} \underset{A \text{ reell}}{=} \overline{A} \cdot \overline{w_1} = \overline{A \cdot w_1} = \overline{\lambda_1 w_1} = \overline{\lambda_1}\,\overline{w_1}.$$

Es folgt $\lambda_2 = \overline{\lambda_1} \notin \mathbb{R}^2$ und $\overline{w_1}$ ist eine Eigenvektor zu $\lambda_2$ und damit ein Vielfaches von $w_2$. Wir ersetzen $w_2$ durch $\overline{w_1}$, d.h. wir nehmen $w_2 = \overline{w_1}$ an. Dann bleibt $w_1, w_2$ eine Orthonormalbasis aus Eigenvektoren. Wir setzen

$$v_1 := \frac{1}{\sqrt{2}}(w_1 + w_2), \quad v_2 := \frac{1}{\sqrt{2}\mathrm{i}}(w_1 - w_2).$$

Wir definieren jetzt für $u \in \mathbb{C}^2$ folgendes $\mathrm{Re}(u) := \binom{\mathrm{Re}(u_1)}{\mathrm{Re}(u_2)}$ und analog $\mathrm{Im}(u) := \binom{\mathrm{Im}(u_1)}{\mathrm{Im}(u_2)}$. Wegen $w_2 = \overline{w_1}$ gilt $v_1, v_2 \in \mathbb{R}^2$ $\big($da $w_1 + w_2 = 2\,\mathrm{Re}(w_1)$ und $\frac{1}{\mathrm{i}}(w_1 - w_2) = 2\,\mathrm{Im}(w_1)\big)$. Wir zeigen zuerst, dass $v_1, v_2$ eine Orthonormalbasis von $\mathbb{R}^2$ ist:

$$\langle v_1, v_1 \rangle = \frac{1}{2}\big(\langle w_1, w_1 \rangle + \langle w_2, w_2 \rangle\big) = 1, \qquad \langle v_2, v_2 \rangle = \frac{1}{2} \cdot \frac{1}{\mathrm{i}} \cdot \frac{1}{-\mathrm{i}}\big(\langle w_1, w_1 \rangle + \langle w_2, w_2 \rangle\big) = 1$$
$$\langle v_1, v_2 \rangle = -\frac{1}{2\mathrm{i}}\big(\langle w_1, w_1 \rangle - \langle w_2, w_2 \rangle\big) = 0.$$

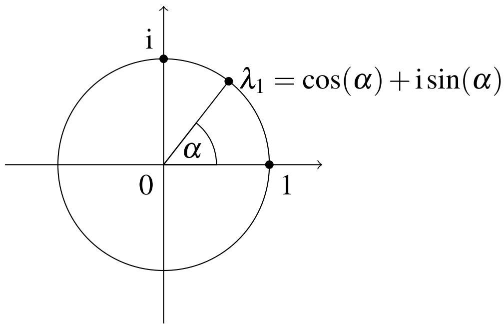
Also ist $v_1, v_2$ eine orthonormierte Basis von $\mathbb{R}^2$ nach Korollar 4.3. Durch Vertauschen dürfen wir annehmen, dass die Basis orientiert ist, dh. $\det(v_1 v_2) > 0$. Wir bestimmen jetzt die Matrix $B$ von $\varphi$ bzgl. dieser Basis $v_1, v_2$.

$$\begin{aligned}
\varphi(v_1) &= \frac{1}{\sqrt{2}}(\lambda_1 w_1 + \overline{\lambda_1}\,\overline{w_1}) \\
&= \frac{1}{\sqrt{2}}\big(\lambda_1 \mathrm{Re}(w_1) + \mathrm{i}\lambda_1 \mathrm{Im}(w_1) + \overline{\lambda_1}\mathrm{Re}(w_1) - \mathrm{i}\overline{\lambda_1}\mathrm{Im}(w_1)\big) \\
&= \frac{1}{\sqrt{2}}\big((\lambda_1 + \overline{\lambda_1})\mathrm{Re}(w_1) + \mathrm{i}(\lambda_1 - \overline{\lambda_1})\mathrm{Im}(w_1)\big) \\
&= \mathrm{Re}(\lambda_1)v_1 - \mathrm{Im}(\lambda_1)v_2.
\end{aligned}$$

Analog findet man

$$\varphi(v_2) = \mathrm{Im}(\lambda_1)v_1 + \mathrm{Re}(\lambda_1)v_2.$$

Nun benutzen wir die **Polardarstellung** $\lambda_1 = \cos(\alpha) + \mathrm{i}\sin(\alpha)$ von $\lambda_1$:




Somit gilt

$$B = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

und es handelt sich also bei $\varphi$ um die Drehung um den Winkel $-\alpha$ mit Zentrum 0. Dabei ist $\alpha \not\equiv 0 \pmod{\pi}$, da $\lambda_1 \notin \mathbb{R}$.

**6.11 Theorem.** *Sei $\varphi \in \mathrm{End}_{\mathbb{R}}(V)$ für den euklidischen Vektorraum $(V, \langle\ ,\ \rangle)$. Dann ist $\varphi$ genau dann eine orthogonale Abbildung, wenn es eine Orthonormalbasis $v_1, \dots, v_n$ von $V$ gibt so, dass die Darstellungsmatrix von $\varphi$ die Form*

$$\begin{pmatrix}
1 &&&&&&&&&& \\
& \ddots &&&&&&&& \mathbf{0} & \\
&& 1 &&&&&&&& \\
&&& -1 &&&&&&& \\
&&&& \ddots &&&&&& \\
&&&&& -1 &&&&& \\
&&&&&& \cos(\alpha_1) & \sin(\alpha_1) &&& \\
&&&&&& -\sin(\alpha_1) & \cos(\alpha_1) &&& \\
&&&&&&&& \ddots && \\
& \mathbf{0} &&&&&&&& \cos(\alpha_s) & \sin(\alpha_s) \\
&&&&&&&&& -\sin(\alpha_s) & \cos(\alpha_s)
\end{pmatrix}$$

*hat mit $0 < \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_s < \pi$. Diese Darstellung ist eindeutig.*

**Beweis:** Nach Theorem 6.6 dürfen wir OBdA annehmen, dass $V = \mathbb{R}^n$ und $\langle\ ,\ \rangle =$ Standardskalarprodukt. Wir argumentieren mit Induktion nach $n = \dim(V)$. Der Fall $n = 1$ ist trivial, da $\varphi$ dann gleich $\mathrm{id}_V$ oder $\varphi = -\mathrm{id}_V$ ist.

„$n-1 \mapsto n$“: Falls $\varphi$ einen reellen Eigenvektor $\lambda_1$ mit Eigenvektor $v_1$ hat, dann gilt $|\lambda_1| = 1$, weil $\varphi$ eine Isometrie ist. Also gilt $\lambda_1 = -1$ oder $\lambda_1 = 1$. Wie im Beweis von Theorem 5.14 folgt

$$V = \mathbb{R}v_1 \oplus (\mathbb{R}v_1)^{\perp}$$

und weiter sind das $\varphi$-invariante Unterräume. Deshalb schränkt sich $\varphi$ zu einem Endomorphismus $\psi\colon W \to W$ für $W := (\mathbb{R}v_1)^{\perp}$ ein. Schränken wir das Skalarprodukt auf $W$ ein, dann ist $\psi$ eine orthogonale Abbildung. Wegen $\dim(W) = \dim(V) - \dim(\mathbb{R}v_1) = n-1$ dürfen wir die Induktionsannahme benutzen. Also gibt es eine Orthonormalbasis von $v_2, \dots, v_n$ von $W$ so, dass $\psi$ eine Kästchenform $B$ hat. Dann hat $\varphi$ bzgl. der Orthonormalbasis $v_1, \dots, v_n$ die Kästchenform $A = \left(\begin{smallmatrix} \pm 1 & 0 \\ 0 & B \end{smallmatrix}\right)$ wie gewünscht.

Falls $\varphi$ keinen reellen Eigenwert hat, dann komplexifizieren wir $\varphi$ wie in Beispiel 6.10. Sei $A$ die reelle $n \times n$ Matrix, die $\varphi$ bzgl. der Standardbasis $e_1, \ldots, e_n$ darstellt. Dann sieht die Komplexifizierung von $\varphi$ folgendermaßen aus:

$$\varphi_{\mathbb{C}} \colon \mathbb{C}^n \to \mathbb{C}^n,\ \lambda \mapsto A \cdot \lambda.$$

Da $\overline{A} = A$ ist, folgt $A^{-1} = \overline{A}^t$ (weil $A$ orthogonal ist). Also ist $\varphi_{\mathbb{C}}$ eine unitäre Abbildung. Nach Theorem 6.8 hat $\varphi_{\mathbb{C}}$ einen Eigenvektor $w_1$ zu einem Eigenwert $\lambda_1$ mit $|\lambda_1| = 1$. Es gilt $Aw_1 = \lambda_1 w_1$.

$$\Longrightarrow A\overline{w_1} = \overline{A} \cdot \overline{w_1} = \overline{A \cdot w_1} = \overline{\lambda_1 w_1} = \overline{\lambda_1} \cdot \overline{w_1}$$

Also ist $\overline{\lambda_1}$ Eigenwert von $\varphi_{\mathbb{C}}$ mit Eigenvektor $\overline{w_1}$. Weil $\varphi$ nach Annahme keine reellen Eigenwerte hat, muss $\lambda_1 \in \mathbb{C} \setminus \mathbb{R}$ gelten und damit folgt $\lambda_1 \neq \overline{\lambda_1}$. Jede unitäre Abbildung ist normal (siehe Beweis von Theorem 6.8). Für solche Abbildungen wissen wir aus Korollar 5.12 c), dass Eigenvektoren zu verschiedenen Eigenwerten orthogonal sind. Also folgt $w_1 \perp \overline{w_1}$. Wir normieren $w_1$ auf 1 und damit hat auch $w_2 := \overline{w_1}$ die Norm 1. Wie im Beispiel 6.10 definieren wir

$$v_1 := \frac{1}{\sqrt{2}}(w_1 + w_2),\ v_2 := \frac{1}{\sqrt{2} \cdot \mathrm{i}}(w_1 - w_2).$$

Wir erinnern daran, dass $\mathrm{Re}(w_1) := \begin{pmatrix} \mathrm{Re}(w_{11}) \\ \vdots \\ \mathrm{Re}(w_{1n}) \end{pmatrix} \in \mathbb{R}^n$ und $\mathrm{Im}(w_1) := \begin{pmatrix} \mathrm{Im}(w_{11}) \\ \vdots \\ \mathrm{Im}(w_{1n}) \end{pmatrix} \in \mathbb{R}^n$. Wieder gilt

$$v_1 = \sqrt{2}\,\mathrm{Re}(w_1) \in \mathbb{R}^n, \quad v_2 = \sqrt{2}\,\mathrm{Im}(w_1) \in \mathbb{R}^n.$$

Wie im Beispiel 6.10 sieht man ein, dass $v_1, v_2$ orthonormale Vektoren sind. Wir definieren $U := \mathbb{R}v_1 \oplus \mathbb{R}v_2$. Offensichtlich ist $U$ $\varphi$-invariant. Nach Beispiel 6.10 wird $U \xrightarrow{\varphi} U$ bzgl. der Orthonormalbasis $v_1, v_2$ durch die Matrix

$$B = \begin{pmatrix} \cos(\alpha_1) & \sin(\alpha_1) \\ -\sin(\alpha_1) & \cos(\alpha_1) \end{pmatrix}$$

dargestellt mit $\alpha_1 \in ]-\pi, \pi[ \setminus \{0\}$ bestimmt durch

$$\lambda_1 = \cos(\alpha_1) + i \sin(\alpha_1).$$

Falls $\alpha \notin ]0, \pi[$, dann vertauschen wir $v_1$ und $v_2$ und erhalten die Matrix

$$\begin{pmatrix} \cos(\alpha_1) & -\sin(\alpha_1) \\ \sin(\alpha_1) & \cos(\alpha_1) \end{pmatrix} = \begin{pmatrix} \cos(\alpha_1') & \sin(\alpha_1') \\ -\sin(\alpha_1') & \cos(\alpha_1') \end{pmatrix}$$

für $\alpha_1' := -\alpha_1 \in ]0, \pi[$. Wir betrachten wieder die Zerlegung

$$V = U \oplus U^\perp.$$

Weil $\varphi$ orthogonal ist, muss $\varphi(U^\perp) \subseteq U^\perp$. Damit induziert $\varphi$ wieder eine orthogonale Abbildung $\psi \colon U^\perp \to U^\perp$. Wegen $\dim(U^\perp) = \dim(V) - \dim(U) = n - 2$ dürfen wir verbesserte vollständige Induktion anwenden. Also gibt es eine Orthonormalbasis $v_3, \ldots, v_n$ von $U^\perp$ so, dass $\psi$ durch die Kästchenform $C$ dargestellt wird. Dann wird $\varphi$ bzgl. der Orthonormalbasis $v_1, v_2, v_3, \ldots, v_n$ durch die Matrix

$$\begin{pmatrix} \cos(\alpha_1) & -\sin(\alpha_1) & 0 \\ \sin(\alpha_1) & \cos(\alpha_1) & \\ & 0 & C \end{pmatrix}$$

dargestellt. Durch Umordnen der $2 \times 2$ Kästchen können wir die Bedingung $0 < \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_s < \pi$ erreichen. Also haben wir eine Kästchenform wie gewünscht.

Wir haben gezeigt, dass jede orthogonale Abbildung eine Kästchenform hat. Nun nehmen wir umgekehrt an, dass $\varphi$ Kästchenform $A$ bzgl. der Orthonormalbasis $v_1, \ldots, v_n$ hat. Da die Zeilen von $A$

orthonormiert sind, muss $\varphi$ eine orthogonale Abbildung (siehe Proposition 6.3). Es bleibt die Eindeutigkeit der Kästchenform zu zeigen. Dazu betrachten wir das charakteristische Polynom.

$$\begin{aligned}\chi_\varphi(x) = \det(x\cdot E_n - A) &= (x-1)^a\cdot(x+1)^b \prod_{j=1}^{s} \det\begin{pmatrix} x-\cos(\alpha_j) & -\sin(\alpha_j) \\ \sin(\alpha_j) & x-\cos(\alpha_j)\end{pmatrix} \\ &= (x-1)^a\cdot(x+1)^b \prod_{j=1}^{s}\big(x^2-2\cos(\alpha_j)x+1\big).\end{aligned}$$

Nach Lemma 6.9 sind diese Faktoren eindeutig, denn die Nullstellen $\cos(\alpha_j)\pm \mathrm{i}\sin(\alpha_j)$ der quadratischen Polynome sind nicht reell. Damit sind $a, b, s = n-a-b$ durch $\varphi$ eindeutig bestimmt. Weil der Kosinus eine streng monoton fallende Funktion auf $]0,\pi[$ ist, muss auch $\alpha_j$ eindeutig durch den quadratischen Faktor $x^2-2\cos(\alpha_j)x+1$ bestimmt sein. Also ist die Kästchenform durch $\varphi$ eindeutig bestimmt. □

**6.12 Beispiel.** Wir benutzen Theorem 6.11, um die orthogonalen Abbildungen eines 3-dimensionalen euklidischen Vektorraums eindeutig zu klassifizieren. Wieder OBdA $V=\mathbb{R}^3, \langle\ ,\ \rangle =$ Standardskalarprodukt. Wenn wir $\alpha\in[0,\pi]$ in der Kästchenform zulassen, erhalten wir folgende Fälle:

| Kästchenform | komplexe Eigenwerte | Interpretation |
|---|---|---|
| $\begin{pmatrix}1 & 0 & 0\\ 0 & \cos(\alpha) & \sin(\alpha)\\ 0 & -\sin(\alpha) & \cos(\alpha)\end{pmatrix}$ | $1, \cos(\alpha)+\mathrm{i}\sin(\alpha)$, $\cos(\alpha)-\mathrm{i}\sin(\alpha)$ | Drehung mit Achse entlang dem EV $v_1$ |
| $\begin{pmatrix}-1 & 0 & 0\\ 0 & \cos(\alpha) & \sin(\alpha)\\ 0 & -\sin(\alpha) & \cos(\alpha)\end{pmatrix}$ | $-1, \cos(\alpha)+\mathrm{i}\sin(\alpha)$, $\cos(\alpha)-\mathrm{i}\sin(\alpha)$ | Drehung mit Achse entlang dem EV $v_1$ verknüpft mit Spiegelung an der Ebene $\{v_1\}^\perp$ |

Beim zweiten Typ spricht man von einer **Drehspiegelung**. Beachte, dass dabei die Reihenfolge der Verknüpfung keine Rolle spielt wegen

$$\begin{aligned}\begin{pmatrix}-1 & 0 & 0\\ 0 & \cos(\alpha) & \sin(\alpha)\\ 0 & -\sin(\alpha) & \cos(\alpha)\end{pmatrix} &= \begin{pmatrix}-1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\cdot\begin{pmatrix}1 & 0 & 0\\ 0 & \cos(\alpha) & \sin(\alpha)\\ 0 & -\sin(\alpha) & \cos(\alpha)\end{pmatrix} \\ &= \begin{pmatrix}1 & 0 & 0\\ 0 & \cos(\alpha) & \sin(\alpha)\\ 0 & -\sin(\alpha) & \cos(\alpha)\end{pmatrix}\cdot\begin{pmatrix}-1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.\end{aligned}$$

Dabei rechnet man immer im Koordinatensystem bzgl. der Orthonormalbasis $v_1, v_2, v_3$ aus Theorem 6.11. Wenn man den Drehwinkel bestimmen will, sollte man eventuell $v_2, v_3$ vertauschen, damit $v_1, v_2, v_3$ eine positiv orientierte Basis bilden $\big($d.h. $\det(v_1,v_2,v_3)>0\big)$. Dann ist $-\alpha$ der Drehwinkel um die Achse nach Beispiel 6.10.

## 7. Selbstadjungierte Abbildungen

Selbstadjungierte Abbildungen spielen in der Theorie der Endomorphismen die Rolle der reellen Zahlen. Wir werden in diesem Abschnitt die für Anwendungen wichtige Hauptachsentransformation in diesem Zusammenhang herleiten.

Es sei dabei immer $V$ ein endlich dimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$.

**7.1 Theorem.** *Sei $\varphi\in \mathrm{End}_{\mathbb{K}}(V)$. Dann sind folgende Aussagen äquivalent.*

*a) $\varphi$ ist selbstadjungiert;*

*b) $\exists$ Orthonormalbasis $v_1,\dots,v_n$ von $V$ mit $\lambda_j v_j = \varphi(v_j)$ und $\lambda_j\in\mathbb{R}$ für alle $j=1,\dots,n$.*

**Beweis:** „$\Longrightarrow$" Wir zeigen zuerst, dass das charakteristische Polynom $\chi_\varphi(x)$ vollständig in Linearfaktoren zerfällt. Falls $\mathbb{K}=\mathbb{C}$, gilt dies nach dem Fundamentalsatz der Algebra (Korollar VII.1.9). Falls $\mathbb{K}=\mathbb{R}$ benutzen wir dazu, dass $\varphi=\varphi^*$. Durch die Wahl einer Orthonormalbasis $e_1,\dots,e_n$ führen wir das Problem der Faktorisierung auf den Fall $V=\mathbb{R}^n$ mit dem Standardskalarprodukt zurück (Theorem 6.6). Sei $A$ die Darstellungsmatrix von $\varphi$, d.h. $\varphi(\lambda)=A\cdot\lambda$. Dann gilt $A=A^t$ wegen $\varphi=\varphi^*$ (Korollar 5.7). Wenn wir $A$ als komplexe $n\times n$ Matrix auffassen, dann gilt

$$\chi_\varphi(x)=\chi_A(x)=(x-\lambda_1)\cdots(x-\lambda_n)$$

nach dem Fundamentalsatz der Algebra mit $\lambda_i\in\mathbb{C}$. Nach Korollar 5.13 gilt sogar $\lambda_i\in\mathbb{R}$. Also zerfällt $\chi_\varphi(x)$ vollständig in Linearfaktoren aus $\mathbb{R}[x]$. Weil $\varphi$ eine selbstadjungierte Abbildung ist, muss sie normal sein. Nach dem Spektralsatz 5.14 gibt es eine Orthonormalbasis $v_1,\dots,v_n$ aus Eigenvektoren. Nach Korollar 5.13 sind alle Eigenwerte reell. Dies zeigt b).

„$\Longleftarrow$" Die Darstellungsmatrix $A$ von $\varphi$ bzgl. $v_1,\dots,v_n$ hat die Form $A=\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}$ und somit gilt $A=\overline{A}^t$. Nach Korollar VIII.5.7 ist $\varphi$ selbstadjungiert, weil $v_1,\dots,v_n$ eine Orthonormalbasis ist. $\square$

**7.2 Korollar.** *Sei $h$ eine hermitesche Form auf $V$. Dann existiert eine Orthonormalbasis $v_1,\dots,v_n$ von $V$ bzgl. dem gegebenen Skalarprodukt $\langle\ ,\ \rangle$ so, dass $h(v_i,v_j)=0\ \forall i\neq j$, d.h. $v_1,\dots,v_n$ ist eine simultane orthogonale Basis bzgl. $\langle\ ,\ \rangle$ und $h$.*

**Beweis:** Aus den Übungen 5.2 und 5.3 gibt es eine selbstadjungiertes $\varphi\in\mathrm{End}_{\mathbb{K}}(V)$ mit

$$h(v,w)=\langle\varphi(v),w\rangle$$

für alle $v,w\in V$. Nach Theorem 7.1 gibt es eine Orthonormalbasis $v_1,\dots,v_n$ von $V$ mit $\varphi(v_i)=\lambda_i v_i$ für den Eigenwert $\lambda_i$ und $i=1,\dots,n$. Für $i\neq j$ gilt

$$h(v_i,v_j)=\langle\varphi(v_i),v_j\rangle=\langle\lambda_i v_i,v_j\rangle=\lambda_i\langle v_i,v_j\rangle=\lambda_i\delta_{ij}=0$$

wie gewünscht. $\square$

**Zweiter Beweis:** Nach Gram-Schmidt existiert eine Orthonormalbasis $e_1,\dots,e_n$ von $V$ bzgl. $\langle\ ,\ \rangle$. Sei $H:=\big(h(e_i,e_j)\big)_{1\le i,j\le n}$ die Matrix von $h$ bzgl. $e_1,\dots,e_n$. Nach 3.9 gilt $H=\overline{H}^t$ und somit ist die lineare Abbildung $\varphi\colon V\to V$ mit Darstellungsmatrix $\overline{H}$ bzgl. der Orthonormalbasis $e_1,\dots,e_n$ selbstadjungiert (Korollar 5.7). Wir rechnen jetzt in den Koordinaten bzgl. dieser Basis und damit wird $\langle\ ,\ \rangle$ zum Standardskalarprodukt auf $V=\mathbb{K}^n$. Nach Theorem 7.1 hat $\varphi$ eine Orthonormalbasis $v_1,\dots,v_n$ aus Eigenvektoren, d.h. $\overline{H}v_j=\lambda_j v_j$ mit $\lambda_j\in\mathbb{R}$.

$$\Longrightarrow h(v_i,v_j)=v_i^t\cdot H\cdot\overline{v_j}=v_i^t\cdot\overline{\overline{H}\cdot v_j}=v_i^t\cdot\overline{\lambda_j v_j}=\langle v_i,\lambda_j v_j\rangle=\overline{\lambda_j}\langle v_i,v_j\rangle=\lambda_j\delta_{ij}.$$

Dies zeigt die Behauptung. $\square$

**7.3 Korollar.** *Sei $H$ eine symmetrische (oder hermitesche) $n\times n$ Matrix. Dann existiert eine orthogonale (bzw. unitäre) $n\times n$ Matrix $U$ so, dass $U\cdot H\cdot U^{-1}$ eine reelle Diagonalmatrix ist.*

**Beweis:** Wir betrachten den Endomorphismus

$$\varphi\colon\mathbb{K}^n\to\mathbb{K}^n,\ \lambda\mapsto H\cdot\lambda$$

zu $H$. Wegen $H=\overline{H}^t$ ist $\varphi$ selbstadjungiert bzgl. dem Standardskalarprodukt auf $\mathbb{K}^n$. Nach Theorem 7.1 existiert eine Orthonormalbasis $v_1,\dots,v_n$ aus Eigenvektoren zu reellen Eigenwerten $\lambda_j$, d.h. $H\cdot v_j=\lambda_j v_j$ mit $\lambda_j\in\mathbb{R}$. Nach Proposition 6.7 ist die Basiswechselmatrix von der Standardbasis $e_1,\dots,e_n$ in die neue Basis $v_1,\dots,v_n$ eine orthogonale (bzw. unitäre) Matrix $U$. Nach Proposition V.4.6 gilt

$$\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}=U\cdot H\cdot U^{-1}.$$ $\square$

**7.4 Beispiel.** Wir betrachten die symmetrischen reellen Bilinearformen

$$b_1\left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\right)=x_1y_1-x_2y_2$$

und

$$b_2\left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\right)=x_1y_1-x_1y_2-x_2y_1+2x_2y_2$$

auf $V=\mathbb{R}^2$. Wir suchen eine Basis $v_1,v_2$ von $\mathbb{R}^2$, die bzgl. beiden Bilinearformen orthogonal ist. Nun gibt der Beweis von Korollar 7.2 eine Verfahren, wie man eine solche Basis finden kann. Dabei brauchen wir ein Skalarprodukt, sonst funktioniert das Verfahren nicht. Dafür kommt $b_1$ nicht in Frage, weil $b_1$ wegen

$$b_1\left(\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right)=-1$$

nicht positiv definit ist. Aber es gilt

$$b_2\left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}x_1\\x_2\end{pmatrix}\right)=x_1^2-2x_1x_2+2x_2^2=(x_1-x_2)^2+x_2^2$$

und damit ist $b_2$ positiv definit. Also definiert $b_2$ ein Skalarprodukt. Als nächstes brauchen wir eine Orthonormalbasis $w_1,w_2$ bzgl. $b_2$ damit wir eine symmetrische (bzw. hermitesche) Matrix für das $\varphi$ erhalten. Wir benutzen Gram-Schmidt und wählen $w_1:=\binom{1}{0}$. Das ist schon normiert bzgl. $b_2$. Weiter haben wir

$$w_2':=\begin{pmatrix}0\\1\end{pmatrix}-b_2\left(\begin{pmatrix}0\\1\end{pmatrix},w_1\right)w_1=\begin{pmatrix}0\\1\end{pmatrix}-(-1)\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}.$$

Normiert ergibt sich

$$w_2:=b_2(w_2',w_2')^{-1/2}w_2'=w_2'.$$

Also ist $w_1,w_2$ eine Orthonormalbasis von $\mathbb{R}^2$ bzgl. $b_2$. Jetzt bestimmen wir die $2\times 2$ Matrix $H:=\big(b_1(w_i,w_j)\big)$ der symmetrischen Bilinearform $b_1$. Es gilt

$$\begin{matrix}b_1(w_1,w_1)=1 & b_1(w_1,w_2)=1\\ b_1(w_2,w_1)=1 & b_1(w_2,w_2)=0\end{matrix}\quad\Longrightarrow\quad H=\begin{pmatrix}1&1\\1&0\end{pmatrix}.$$

Wir benutzen, dass die symmetrische Matrix $H$ die Matrix zum Endomorphismus $\varphi$ aus dem Beweis von Korollar 7.2 ist (braucht eine kleine Rechnung, die in Übung 5.2/5.3 gemacht wird). Wir müssen jetzt die Eigenvektoren zu $H$ berechnen und und erhalten nach Theorem 7.1 eine simultan orthogonale Basis. Wir suchen die Eigenvektoren $v=\binom{x_1}{x_2}$ zum Eigenwert $\lambda$.

$$H\cdot v=\lambda\cdot v\quad\Longleftrightarrow\quad\begin{aligned}(1-\lambda)x_1+\quad x_2&=0\\1\cdot x_1-\lambda x_2&=0\end{aligned}$$

Mit dem Gauß-Algorithmus folgt

$$\left(\begin{array}{cc|c}1-\lambda & 1 & 0\\ \boxed{1} & -\lambda & 0\end{array}\right)\rightsquigarrow\left(\begin{array}{cc|c}0 & (1-\lambda)\lambda+1 & 0\\ 1 & -\lambda & 0\end{array}\right).$$

Das hat genau dann eine nicht-triviale Lösung wenn $-\lambda^2+\lambda+1=0$ gilt. Also sind $\lambda_{1/2}=\frac{1\pm\sqrt{5}}{2}$ die beiden Eigenwerte und zu $\lambda_i$ gehört der Eigenvektor $\binom{\lambda_i}{1}$. Nun muss man aufpassen, dass diese Eigenvektoren in den Koordinaten bzgl. der Basis $w_1,w_2$ berechnet werden und nicht bzgl. der Standardbasis. Also ergibt der Beweis von Korollar 7.2, dass

$$v_1=\lambda_1w_1+w_2=\lambda_1\cdot\begin{pmatrix}1\\0\end{pmatrix}+\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}\lambda_1+1\\1\end{pmatrix}=\begin{pmatrix}\frac{3+\sqrt{5}}{2}\\1\end{pmatrix}$$

und

$$v_2 = \lambda_2 w_1 + w_2 = \lambda_2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda_2 + 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{3-\sqrt{5}}{2} \\ 1 \end{pmatrix}.$$

**7.5 Definition.** Eine Matrix $H \in M(n \times n, \mathbb{K})$ heißt **positiv definit**, wenn sie die Matrix eines Skalarprodukts auf einem endlich-dimensionalen $\mathbb{K}$-Vektorraum bzgl. einer beliebigen Basis ist.

Nach 3.8 ist $H$ genau dann eine positiv definite Matrix, wenn $h(\lambda, \mu) := \lambda^t \cdot H \cdot \overline{\mu}$ ein Skalarprodukt auf $\mathbb{K}^n$ definiert. Insbesondere gilt $H = \overline{H}^t$ für positiv definite Matrizen.

**7.6 Korollar.** *Eine Matrix $H \in M(n \times n, \mathbb{K})$ ist genau dann positiv definit, wenn $H = \overline{H}^t$ und $H$ nur positive Eigenwerte hat (d.h. $\lambda > 0$).*

**Beweis:** Es gelte $H = \overline{H}^t$. Dann ist $h(\lambda, \mu) := \lambda^t \cdot H \cdot \overline{\mu}$ eine hermitesche Form auf $\mathbb{K}^n$. Nach dem Beweis von Korollar 7.2 liefert eine Orthonormalbasis aus Eigenvektoren von $\overline{H}$ eine simultane orthogonale Basis $v_1, \ldots, v_n$ bzgl. dem Standardskalarprodukt und bzgl. $h$ (vgl. auch Beispiel 7.4). Also gilt $h(v_i, v_j) = \lambda_j \delta_{ij}$. Seien $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ die Koordinaten von $v \in \mathbb{K}^n$ bzgl. der Basis $v_1, \ldots, v_n$. Dann gilt

$$h(v,v) = h(\sum_{i=1}^n \alpha_i v_i, \sum_{j=1}^n \alpha_j v_j) = \sum_{i,j=1}^n \lambda_i \alpha_i \overline{\alpha_j} \delta_{ij} = \sum_{i=1}^n \lambda_i |\alpha_i|^2.$$

Also ist $h$ genau dann ein Skalarprodukt, wenn alle $\lambda_i > 0$ sind. Dies zeigt die Behauptung. □

**7.7 Theorem (Trägheitssatz von Sylvester).** *Sei $h$ eine symmetrische Bilinearform (bzw. eine hermitesche Form) auf einem reellen (bzw. komplexen) Vektorraum der Dimension $n < \infty$. Dann gibt es eine Basis dieses Vektorraums so, dass die Matrix von $h$ die Form*

$$\begin{pmatrix} E_r & & \\ & -E_s & \\ & & 0 \end{pmatrix} \in M(n \times n, \mathbb{K})$$

*hat, wobei $E_r, E_s$ die Einheitsmatrizen der Länge $r$ und $s$ bezeichnen. Die Zahlen $r$ und $s$ sind unabhängig von der Wahl der Basis.*

**Beweis:** Wir können jeden $\mathbb{K}$-Vektorraum $V$ mit einem Skalarprodukt $\langle \, , \, \rangle$ versehen. Durch Wahl einer Basis können wir einfach das Standardskalarprodukt von $\mathbb{K}^n$ übertragen. Nach Korollar 7.2 gibt es also eine Orthonormalbasis $v_1, \ldots, v_n$ von $V$ bzgl. $\langle \, , \, \rangle$, die auch orthogonal bzgl. $h$ ist. Weil $h$ eine hermitesche Form ist, gilt $h(v,v) = \overline{h(v,v)}$ und damit $h(v,v) \in \mathbb{R}$ für jedes $v \in V$. Durch Umordnen könne wir annehmen, dass

$$h(v_j, v_j) \begin{cases} > 0 & \text{für } j = 1, \ldots, r \\ < 0 & \text{für } j = r+1, \ldots, r+s \\ = 0 & \text{für } j = r+s+1, \ldots, n \end{cases}$$

Ersetzen wir für $j = 1, \ldots, r+s$ den Vektor $v_j$ durch $|h(v_j, v_j)|^{-1/2} v_j$, dann erhalten wir die gesuchte Basis. Nun zeigen wir die Eindeutigkeit von $r$ und $s$. Sei dazu $v'_1, \ldots, v'_n$ eine weitere Basis mit

$$\left(h(v'_i, v'_j)\right) = \begin{pmatrix} E_{r'} & & \\ & -E_{s'} & \\ & & 0 \end{pmatrix}.$$

Wir definieren folgende Unterräume

$$\begin{aligned} V_+ &= \langle v_1, \ldots, v_r \rangle, & V'_+ &:= \langle v'_1, \ldots, v'_{r'} \rangle, \\ V_- &= \langle v_{r+1}, \ldots, v_{r+s} \rangle, & V'_- &:= \langle v'_{r'+1}, \ldots, v'_{r'+s'} \rangle, \\ V_0 &= \langle v_{r+s+1}, \ldots, v_n \rangle, & V'_0 &:= \langle v'_{r'+s'+1}, \ldots, v'_n \rangle. \end{aligned}$$

Dann gilt

$$V_+ \oplus V_- \oplus V_0 = V = V'_+ \oplus V'_- \oplus V'_0.$$

Die besondere Gestalt der Matrix zu $h$ zeigt sofort, dass

$$V_0 = \{v \in V \mid h(v,w) = 0 \; \forall w \in V\} = V'_0.$$

Da $h$ auf $V_+$ positiv definit und auf $V'_- \oplus V'_0$ negativ semi-definit ist, folgt

$$V_+ \cap (V'_- \oplus V'_0) = \{0\}.$$

Also gilt, dass $V_+ + (V'_- \oplus V'_0)$ eine direkte Summe ist und somit

$$\underbrace{\dim(V_+)}_{r} + \underbrace{\dim(V'_-)}_{s'} + \dim(V'_0) \leq n.$$

Wegen $r' + s' + \dim(V'_0) = n$, folgt also $r \leq r'$. Aus Symmetriegründen folgt $r' \geq r$ und damit $r = r'$. Mit

$$s = n - r - \dim(V_0) = n - r' - \dim(V'_0) = s'$$

folgt die Behauptung. □

Kapitel IX.

# 9 Anwendungen in der Geometrie

## 1. Affine Räume

Vom geometrischen Standpunkt ist es unnatürlich einen Ursprung und Koordinaten festzulegen. Das führt auf das Konzept der affinen Räume, die man sich als Vektorräume ohne festgelegten Nullpunkt vorstellen sollte.

In diesem Abschnitt ist $K$ ein beliebiger Körper.

**1.1 Definition.** Ein **affiner Raum** $A$ über $K$ ist eine Menge $A$ zusammen mit einem $K$-Vektorraum $V_A$ und einer Vektorfunktion

$$A \times A \to V_A, \ (p,q) \mapsto \overrightarrow{pq}$$

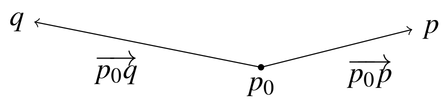
mit folgenden Axiomen

(1) Zu $p \in A$ und $v \in V_A$ existiert genau ein $q \in A$ mit $\overrightarrow{pq} = v$;

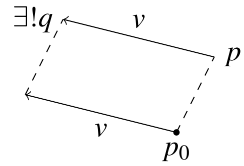
(2) $\overrightarrow{pq} + \overrightarrow{qr} = \overrightarrow{pr}$ für alle $p,q,r \in A$.

Wir betrachten auch die leere Menge als affinen Raum, dazu gehört aber kein Vektorraum. Wir nennen die Elemente von $A$ **Punkte**. Beim Vektor $\overrightarrow{pq}$ heißt $p$ der **Anfangspunkt** und $q$ der **Endpunkt**.

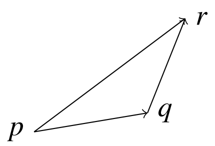
**1.2 Beispiel.** Wir betrachten die euklidische Ebene $E$. Dann ist $E$ ein affiner Raum mit $V_E$ gleich dem Vektorraum der Ortsvektoren mit vorgegebenem Anfangspunkt $p_0$.



Axiom (1). Gegeben sei $p \in E$ und $v \in V_E$. $\Longrightarrow v = \overrightarrow{p_0 q_0}$.



Axiom (2).



Bekanntlich gilt $\overrightarrow{pq} + \overrightarrow{qr} = \overrightarrow{pr}$.

**1.3 Definition.** Für einen affinen Raum $A$ definieren wir $\dim(A) = \dim(V_A)$, falls $A \neq \emptyset$, und $\dim(\emptyset) := -1$.

**1.4 Proposition.** *Für einen affinen Raum $A$ und $p,q \in A$ gilt:*

*a)* $\overrightarrow{pp} = 0$;

*b)* $\overrightarrow{pq} = 0 \Longleftrightarrow p = q$;

*c)* $\overrightarrow{qp} = -\overrightarrow{pq}$.

**Beweis:** a) $\overrightarrow{pp} + \overrightarrow{pp} = \overrightarrow{pp}$ nach Axiom (2) für $p = q = r \Longrightarrow \overrightarrow{pp} = 0$ und damit a).

b) Nach a) gilt $\overrightarrow{pq} = 0$ für $p = q$. Nach Axiom (1) ist $q$ aber eindeutig bestimmt und damit folgt b).

c) $\overrightarrow{pq} + \overrightarrow{qp} \underset{(2)}{=} \overrightarrow{pp} \underset{a)}{=} 0$. Damit folgt $\overrightarrow{qp} = -\overrightarrow{pq}$. □

**1.5 Definition.** Eine Teilmenge $B$ von $A$ heißt **affiner Unterraum** von $A$, falls für alle $p \in B$ die Menge $\{\overrightarrow{pq} \mid q \in B\}$ ein Unterraum von $V_A$ ist.

**1.6 Bemerkung.** Damit $B$ ein affiner Unterraum von $A$ ist, genügt es zu zeigen, dass $V_B := \{\overrightarrow{pq} \mid q \in B\}$ ein Unterraum von $V_A$ ist für *ein* $p \in B$. Weiter ist dann der Unterraum $V_B$ unabhängig von der Wahl von $p$.

**Beweis:** Wir nehmen an, dass es ein $p \in B$ gibt so, dass $V_B = \{\overrightarrow{pq} \mid q \in B\}$ ein Unterraum. Es genügt zu zeigen, dass $V_B$ unabhängig von der Wahl von $p$ ist. Sei $p' \in B$ ein weiterer Punkt und $q$ ein beliebiger Punkt in $A$. Es gilt

$$\overrightarrow{p'q} \underset{(2)}{=} \overrightarrow{p'p} + \overrightarrow{pq} \underset{1.4\,c)}{=} -\overrightarrow{pp'} + \overrightarrow{pq}.$$

Weil $\overrightarrow{pp'} \in V_B$ ist und weil $V_B$ ein Unterraum von $V_B$ ist, gilt

$$\overrightarrow{p'q} \in V_B \Longleftrightarrow \overrightarrow{pq} \in V_B.$$

Also gilt $V_B = \{\overrightarrow{p'q} \mid q \in B\}$. □

**1.7 Beispiel.** Die affinen Unterräume der euklidischen Ebene $E$ bestehen aus $\emptyset$, Punkte, Geraden, $E$.

**1.8 Proposition.** *Der Durchschnitt von einer Familie von affinen Unterräumen von A ist wieder ein affiner Unterraum von A.*

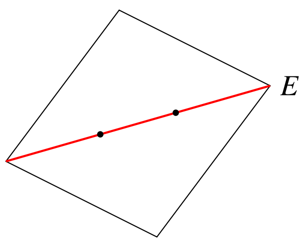
**Beweis:** Es sei $(B_i)_{i\in I}$ die Familie der affinen Unterräume aus der Behauptung und $B := \bigcap_{i\in I} B_i$ . Für $p \in B$ gilt

$$\{\overrightarrow{pq} \mid q \in B\} = \bigcap_{i\in I} \{\overrightarrow{pq} \mid q \in B_i\}.$$

Weil $B_i$ ein affiner Unterraum ist, muss $\{\overrightarrow{pq} \mid q \in B_i\}$ ein Unterraum von $V_A$ sein. Der Durchschnitt von Unterräumen eines Vektorraums ist bekanntlich wieder ein Unterraum. Damit folgt, dass $B$ ein affiner Unterraum ist. □

**1.9.** Nach Proposition 1.8 existiert für jede Teilmenge $Y \subseteq A$ ein kleinster affiner Unterraum, der $Y$ enthält. Dazu nehmen wir, wie im Fall der Vektorräume, einfach den Durchschnitt aller affiner Unterräume, die $Y$ enthalten. Der kleinste solche affine Unterraum wird mit $\langle Y\rangle_{\text{aff}}$ bezeichnet und heißt der **von Y erzeugte affine Unterraum**.

**1.10 Beispiel.** Der von zwei verschiedenen Punkten erzeugte affine Unterraum in der euklidischen Ebene $E$ ist die Verbindungsgerade.



**1.11 Beispiel.** Jeder $K$-Vektorraum $V$ ist ein affiner Raum mit Vektorfunktion

$$V \times V \to V, \quad (p,q) \mapsto \overrightarrow{pq} := q - p.$$

Wir prüfen die Axiome des affinen Raumes.

(1) Sei $p \in V$ und $v \in V$. Dann gilt

$$\overrightarrow{pq} = v \Longleftrightarrow q - p = v \Longleftrightarrow q = p + v.$$

Damit existiert so ein $q$ und es ist eindeutig.

(2) $\overrightarrow{pq} + \overrightarrow{qr} = (q-p) + (r-q) = r - p = \overrightarrow{qr}$.

**1.12 Beispiel.** Sei $A \in M(m \times n, K)$ und $b \in K^m$. Dann ist der Lösungsraum $\mathbb{L}$ des linearen Gleichungssystems $A \cdot x = b$ ein affiner Unterraum von $K^n$, wobei $K^n$ versehen wird mit der Standardstruktur als affiner Raum aus Beispiel 1.11.

**Beweis:** Wir müssen zeigen, dass für $p \in \mathbb{L}$ die Menge $\{\overrightarrow{pq} \mid q \in \mathbb{L}\}$ ein Unterraum von $K^n$ ist. Nach Definition ist $\overrightarrow{pq} = q - p$. Also gilt

$$q \in \mathbb{L} \Longleftrightarrow A \cdot q = b \overset{A \cdot p = b}{\Longleftrightarrow} A(q-p) = 0.$$

Also gilt, dass $\{\overrightarrow{pq} \mid q \in \mathbb{L}\} = \{q - p \mid q \in \mathbb{L}\}$ der Lösungsraum des homogenen linearen Gleichungssystems $A \cdot x = 0$ ist. Wir wissen aber, dass der Lösungsraum eines homogenen linearen Gleichungssystems ein Unterraum von $K^n$. □

**1.13 Definition.** Ein **Koordinatensystem** für einen endlich-dimensionalen affinen Raum $A$ der Dimension $n$ über $K$ besteht aus $o, p_1, \dots, p_n \in A$ so, dass $\overrightarrow{op_1}, \dots, \overrightarrow{op_n}$ eine Basis von $V_A$ ist. Dann heißt $o$ der **Ursprung** und $p_1, \dots, p_n$ sind die **Einheitspunkte**.

**1.14.** Wenn wir so ein Koordinatensystem haben, dann hat jedes $p \in A$ eine eindeutige Darstellung $\overrightarrow{op} = \sum_{j=1}^{n} \alpha_j \overrightarrow{op_j}$ mit geeigneten $\alpha_j \in K$. Wir nennen $\alpha_1, \dots, \alpha_n$ die **Koordinaten** von $p$.

**1.15 Proposition.** *Sei $o', p'_1, \dots, p'_n$ ein weiteres Koordinatensystem für $A$. Dann gibt es genau eine $n \times n$ Matrix $T$ und ein $\lambda \in K^n$ so, dass der Koordinatenvektor $\alpha \in K^n$ von $p \in A$ bzgl. dem alten Koordinatensystem $o, p_1, \dots, p_n$ sich folgendermaßen in den Koordinatenvektor $\alpha' \in K^n$ von $p$ bzgl. dem neuen Koordinatensystem umrechnen lässt:*

$$\alpha' = T \cdot \alpha + \lambda$$

**Beweis:** Übung 8.2. □

## 2. Affine Abbildungen

Eine affine Abbildung zwischen affinen Räumen muss man sich als lineare Abbildung + Translation vorstellen, wobei man dabei jeweils einen Ursprung fixiert hat, um eine Vektorraumstruktur zu erhalten. Wir werden dies intrinsisch studieren.

In diesem Abschnitt sind $A, B$ affine Räume über dem Körper $K$.

**2.1 Definition.** Eine Abbildung $\varphi\colon A \to B$ heißt **affin**, wenn es eine lineare Abbildung $\hat{\varphi}\colon V_A \to V_B$ gibt mit

$$\hat{\varphi}(\overrightarrow{pq}) = \overrightarrow{\varphi(p)\varphi(q)}$$

für alle $p, q \in A$.

**2.2 Proposition.** *Für jede lineare Abbildung $\psi: V_A \to V_B$ und $(p_0, q_0) \in A \times B$ gibt es genau eine affine Abbildung $\varphi: A \to B$ mit $\hat{\varphi} = \psi$ und $\varphi(p_0) = q_0$.*

**Beweis:** Für $p \in A$ gibt es genau ein $q \in B$ mit $\overrightarrow{q_0 q} = \psi(\overrightarrow{p_0 p})$ nach Axiom (1) für den affinen Raum $B$. Wir setzen $\varphi(p) := q$. Es gilt $\psi(\overrightarrow{p_0 p_0}) = \psi(0) = 0$. Da $q = q_0$ das einzige $q \in B$ ist mit $\overrightarrow{q_0 q} = 0$, folgt $\varphi(p_0) = q_0$. Es bleibt zu zeigen, dass

$$\psi(\overrightarrow{p_1 p_2}) = \overrightarrow{\varphi(p_1)\varphi(p_2)}$$

für alle $p_1, p_2 \in A$.

$$\begin{aligned}\overrightarrow{\varphi(p_1)\varphi(p_2)} &= \overrightarrow{\varphi(p_1)\varphi(p_0)} + \overrightarrow{\varphi(p_0)\varphi(p_2)} = -\psi(\overrightarrow{p_0 p_1}) + \psi(\overrightarrow{p_0 p_2}) \\ &= \psi(-\overrightarrow{p_0 p_1} + \overrightarrow{p_0 p_2}) = \psi(\overrightarrow{p_1 p_0} + \overrightarrow{p_0 p_2}) = \psi(\overrightarrow{p_1 p_2}).\end{aligned}$$

Dies zeigt die Existenz einer affinen Abbildung $\psi$ mit den geforderten Eigenschaften und die Eindeutigkeit folgt aus der Konstruktion. □

**2.3 Proposition.** *Für eine affine Abbildung $\varphi: A \to B$ gelten folgende Eigenschaften:*

*a) $\varphi$ ist injektiv (bzw. surjektiv) genau dann, wenn $\hat{\varphi}$ injektiv (bzw. surjektiv) ist;*

*b) $C$ affiner Unterraum von $A \Longrightarrow \varphi(C)$ affiner Unterraum von $B$;*

*c) $D$ affiner Unterraum von $B \Longrightarrow \varphi^{-1}(D)$ affiner Unterraum von $A$.*

**Beweis:** Siehe Übung 8.3. □

**2.4 Proposition.** *Es sei $\varphi: A \to B$ eine Abbildung zwischen affinen Räumen. Weiter seien $o, p_1, \dots, p_n$ bzw. $o', p'_1, \dots, p'_m$ Koordinatensysteme von $A$ und $B$. Dann ist $\varphi$ genau dann eine affine Abbildung, wenn es eine Matrix $C \in M(m \times n, K)$ und ein $c \in K^m$ gibt so, dass für alle $p \in A$ der Koordinatenvektor von $\varphi(p)$ gegeben ist durch $C \cdot \alpha + c$, wobei $\alpha$ der Koordinatenvektor von $p$ ist.*

**Beweis:** „$\Longrightarrow$“ Sei also $\varphi$ eine affine Abbildung. Nach Definition gehört dazu eine lineare Abbildung $\hat{\varphi}: V_A \to V_B$. Sei $C$ die Matrix von $\hat{\varphi}$ bzgl. den Basen $\overrightarrow{op_1}, \dots, \overrightarrow{op_n}$ von $V_A$ und $\overrightarrow{op'_1}, \dots, \overrightarrow{op'_m}$ von $V_B$. Nun ist der Koordinatenvektor von $\varphi(p)$ bzgl. dem affinen Koordinatensystem $o', p'_1, \dots, p'_m$ gleich dem Koordinatenvektor von

$$\overrightarrow{o'\varphi(p)} = \overrightarrow{o'\varphi(o)} + \overrightarrow{\varphi(o)\varphi(p)} = \overrightarrow{o'\varphi(o)} + \hat{\varphi}(\overrightarrow{op})$$

in $V_B$ und somit gleich $c + C \cdot \alpha$, wobei $c \in K^m$ definiert ist als der Koordinatenvektor von $\varphi(o)$ bzgl. dem affinen Koordinatensystem $o', p'_1, \dots, p'_m$. Dies zeigt diese Richtung.

„$\Longleftarrow$“ Sei die Abbildung $\varphi: A \to B$ bzgl. den Koordinaten gegeben als $K^n \ni \alpha \mapsto C \cdot \alpha + c \in K^m$. Es sei $\psi: V_A \to V_B$ die lineare Abbildung, die bzgl. den Basen $\overrightarrow{op_1}, \dots, \overrightarrow{op_n}$ von $V_A$ und $\overrightarrow{o'p'_1}, \dots, \overrightarrow{o'p'_m}$ von $V_B$ gegeben ist durch die Matrix $C$. Für $p \in A$ gilt dann, dass

$$\overrightarrow{\varphi(o)\varphi(p)} = \overrightarrow{\varphi(o)o'} + \overrightarrow{o'\varphi(p)} = -\overrightarrow{o'\varphi(o)} + \overrightarrow{o'\varphi(p)}$$

durch den Koordinatenvektor

$$-(C \cdot 0 + c) + (C \cdot \alpha + c) = C \cdot \alpha$$

repräsentiert wird, wobei $\alpha$ der Koordinatenvektor von $p$ ist. Also folgt, dass $\psi(\overrightarrow{op}) = \overrightarrow{\varphi(o)\varphi(p)}$. Mit Axiom (2) für affine Räume folgt

$$\psi(\overrightarrow{pq}) = \psi(-\overrightarrow{op} + \overrightarrow{oq}) = -\psi(\overrightarrow{op}) + \psi(\overrightarrow{oq}) = -\overrightarrow{\varphi(o)\varphi(p)} + \overrightarrow{\varphi(o)\varphi(q)} = \overrightarrow{\varphi(p)\varphi(q)}$$

für alle $p, q \in A$. Also ist $\varphi$ eine affine Abbildung. □

**2.5 Definition.** Ein **affiner Isomorphismus** $\varphi\colon A \to B$ ist eine affine Abbildung, die eine affine Umkehrabbildung hat.

**2.6 Proposition.** *Eine affine Abbildung $\varphi\colon A \to B$ ist genau dann ein Isomorphismus, wenn $\varphi$ bijektiv ist.*

**Beweis:** „$\Longrightarrow$“ Falls $\varphi\colon A \to B$ ein affiner Isomorphismus ist, dann gibt es nach Definition eine affine Abbildung $\psi\colon B \to A$ mit $\varphi \circ \psi = \mathrm{id}_B$ und $\psi \circ \varphi = \mathrm{id}_A$. Damit muss $\varphi$ bijektiv sein.

„$\Longleftarrow$“ Sei $\varphi$ eine affine Abbildung, die bijektiv ist. Weil $\varphi$ bijektiv ist, gibt es eine Umkehrabbildung $\psi\colon B \to A$. Wir müssen zeigen, dass $\psi$ affin ist. Nach Proposition 2.3 ist $\hat{\varphi}$ eine bijektive lineare Abbildung und damit ein Vektorraum-Isomorphismus. Also gibt es eine lineare Umkehrabbildung $g\colon V_B \to V_A$. Für $p', q' \in B$ gilt

$$\overrightarrow{\psi(p')\psi(q')} = \underbrace{g \circ \hat{\varphi}}_{\mathrm{id}_A}\left(\overrightarrow{\psi(p')\psi(q')}\right) \underset{\varphi \text{ affin}}{=} g\left(\overrightarrow{\varphi(\psi(p'))\varphi(\psi(q'))}\right) \underset{\psi=\varphi^{-1}}{=} g(\overrightarrow{p'q'}). \qquad \square$$

**2.7 Proposition.** *Die Verknüpfung $\varphi \circ \psi$ von affinen Abbildungen ist wieder eine affine Abbildung. Weiter ist $\mathrm{id}_A$ eine affine Abbildung.*

**Beweis:** Übung 8.4. $\square$

**2.8.** Ein affiner Isomorphismus $\varphi\colon A \to A$ heißt **Affinität**. Wir haben in Kapitel II gesehen, dass die Menge der bijektiven Selbstabbildungen von $A$ bzgl. $\circ$ eine Gruppe bilden, die wir symmetrische Gruppe von $A$ nennen. Nach Proposition 2.6 und 2.7 bilden die Affinitäten von $A$ eine Untergruppe der symmetrischen Gruppe von $A$. Wir nennen sie die **affine Gruppe** von $A$.

**2.9 Beispiel.** Sei $v \in V_A$. Für $p \in A \Longrightarrow \exists! q \in A$ mit $\overrightarrow{pq} = v$. Wir definieren $\tau_v(p) := q$. Damit entsteht eine Abbildung $\tau_v\colon A \to A$, die man die **Translation mit dem Translationsvektor v** nennt. Wegen

$$\begin{aligned}\overrightarrow{\tau_v(p_1)\tau_v(p_2)} &= \overrightarrow{\tau_v(p_1)p_1} + \overrightarrow{p_1\tau_v(p_2)} = -\overrightarrow{p_1\tau_v(p_1)} + \overrightarrow{p_1p_2} + \overrightarrow{p_2\tau_v(p_2)}\\ &= -v + \overrightarrow{p_1p_2} + v = \overrightarrow{p_1p_2}\end{aligned}$$

ist $\tau_v$ eine affine Abbildung mit $\hat{\tau}_v = \mathrm{id}_{V_A}$. Die Umkehrabbildung ist offenbar die Translation mit $-v$ und somit ist jede Translation eine Affinität. Wegen

$$\tau_v \circ \tau_w = \tau_{v+w} = \tau_w \circ \tau_v$$

bilden die Translationen von $A$ eine abelsche Untergruppe der affinen Gruppe.

**2.10 Proposition.** *Wir fixieren $p_0 \in A$. Für eine Affinität $\varphi$ von $A$ gibt es genau eine Faktorisierung $\varphi = \tau_v \circ \psi$, wobei $\tau_v$ die Translation mit $v \in V_A$ und $\psi$ eine Affinität mit $\psi(p_0) = p_0$ ist.*

**Beweis:** Setzte $v := \overrightarrow{p_0\varphi(p_0)}$ und $\psi := \tau_v^{-1} \circ \varphi = \tau_{-v} \circ \varphi$. Offenbar ist $\psi$ eine Affinität, weil alles in der affinen Gruppe gerechnet wurde. Es gilt

$$\psi(p_0) = \tau_v^{-1}\big(\varphi(p_0)\big) = p_0.$$

Somit erhalten wir die gewünschte Faktorisierung. Wenn umgekehrt $\varphi = \tau_v \circ \psi$ eine solche Darstellung ist, dann gilt

$$\tau_v(p_0) = \tau_v(\psi(p_0)) = \varphi(p_0)$$

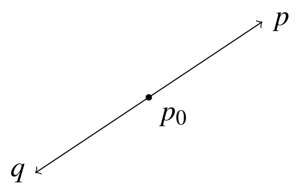
und damit gilt $v = \overrightarrow{p_0\tau_v(p_0)} = \overrightarrow{p_0\varphi(p_0)}$. Also ist $v$ eindeutig bestimmt und damit auch $\psi$, weil die Affinitäten eine Gruppe bilden. $\square$

**2.11 Bemerkung.** Durch die Wahl von $p_0 \in A$ wird $A$ affin isomorph zu $V_A$ (mit der Standardstruktur als affiner Raum aus Beispiel 1.11). Der affine Isomorphismus ist gegeben durch die Abbildung

$$\rho : A \to V_A,\ p \mapsto \overrightarrow{p_0p}.$$

Nach dem Axiom (1) der affinen Räume ist $\rho$ tatsächlich bijektiv und man zeigt leicht, dass $\rho$ eine affine Abbildung mit $\hat{\rho} = \mathrm{id}_{V_A}$ ist. Wenn man also $A$ dadurch mit $V_A$ identifiziert, dann besagt Proposition 2.10, dass jede Affinität sich als Verknüpfung einer Translation mit einem linearen Automorphismus schreiben lässt.

**2.12 Beispiel.** Wir fixieren $p_0 \in A$. Für $p \in A$ $\exists! q \in A$ mit $\overrightarrow{p_0q} = -\overrightarrow{p_0p}$. Wir erhalten eine Abbildung $\rho : A \to A,\ p \mapsto q$.




Wir nennen $\rho$ die Punktspiegelung an $p_0$. Wegen $\hat{\rho} = -\mathrm{id}_{V_A}$ ist $\rho$ eine affine Abbildung. Wegen $\rho^2 = \mathrm{id}_A$ ist $\rho$ die Umkehrabbildung von $\rho$ selber und damit auch affin. Also ist $\rho$ eine Affinität.

## 3. Kongruenzen

In diesem Abschnitt betrachten wir euklidische (bzw. unitäre) affine Räume über dem Körper $\mathbb{R}$ (bzw. $\mathbb{C}$), d.h. $V_A$ ist mit einem Skalarprodukt versehen. Damit kann man einen Abstand auf $A$ definieren. Diejenigen Affinitäten, die diesen Abstand erhalten, nennen wir Kongruenzen. Wir werden in diesem Abschnitt Resultate über Kongruenzen beweisen mit Hilfe der Resultate aus Abschnitt VIII.6 über Isometrien.

Wie immer bezeichnet $\mathbb{K}$ einen Körper aus $\{\mathbb{R}, \mathbb{C}\}$.

**3.1 Definition.** Ein **euklidischer** (bzw. **unitärer**) **affiner Raum** ist ein reeller (bzw. komplexer) affiner Raum $A$ so, dass $V_A$ ein euklidischer (bzw. unitärer) Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$ ist. Der **Abstand** zweier Punkte $p, q \in A$ ist definiert durch

$$d(p,q) := \|\overrightarrow{pq}\|$$

wobei $\|\ \|$ die Norm zum Skalarprodukt $\langle\ ,\ \rangle$ ist.

Ein **orthonormiertes Koordinatensystem** eines euklidischen (bzw. unitären) affinen Raumes ist ein affines Koordinatensystem $o, p_1, \ldots, p_n$ so, dass $\overrightarrow{op_1}, \ldots, \overrightarrow{op_n}$ eine Orthonormalbasis von $\big(V_A, \langle\ ,\ \rangle\big)$ bilden.

Nach dem Gram-Schmidt'schen Verfahren gibt es zu jedem Ursprung $o \in A$ ein orthonormales Koordinatensystem $o, p_1, \ldots, p_n$. Das hat den Vorteil, dass wir mit Hilfe der Koordinatenabbildung den affinen Raum $A$ mit $\mathbb{K}^n$ und $\langle\ ,\ \rangle$ mit dem Standard-Skalarprodukt identifizieren können.

Ab jetzt sei $A$ ein euklidischer (bzw. unitärer) affiner Raum der Dimension $n$ über $\mathbb{K} = \mathbb{R}$ (bzw. $\mathbb{K} = \mathbb{C}$) mit Skalarprodukt $\langle\ ,\ \rangle$ auf $V_A$ und Abstand $d$ wie oben definiert.

**3.2 Definition.** Eine Affinität $\varphi$ von $A$ heißt **Kongruenz**, wenn

$$d\big(\varphi(p), \varphi(q)\big) = d(p,q)$$

gilt für alle $p, q \in A$.

**3.3 Bemerkung.** Eine Affinität $\varphi$ ist genau dann eine Kongruenz, wenn $\hat{\varphi}$ eine orthogonale (bzw. unitäre) Abbildung ist.

**Beweis:** Für $p,q \in A$ gilt

$$d(\varphi(p),\varphi(q)) = \left\|\overrightarrow{\varphi(p)\varphi(q)}\right\| = \|\hat{\varphi}(\overrightarrow{pq})\|$$

und

$$d(p,q) = \|\overrightarrow{pq}\|.$$

Wenn $p,q$ über ganz $A$ läuft, dann durchquert $\overrightarrow{pq}$ den ganzen Vektorraum $V_A$ nach Axiom 1) der affinen Räume. Also ist $\varphi$ genau dann eine Kongruenz, wenn $\|\hat{\varphi}(v)\| = \|v\|$ gilt für alle $v \in V_A$. Das ist äquivalent zu $\hat{\varphi}$ gleich eine Isometrie wie gewünscht. $\square$

**3.4 Proposition.** *Die Kongruenzen von A bilden eine Untergruppe der affinen Gruppe von A.*

**Beweis:** Nach Definition ist jede Kongruenz eine Affinität.

*Abgeschlossenheit bzgl.* $\circ$*:* Seien $\varphi,\psi$ Kongruenzen. Dann ist $\varphi\circ\psi$ eine Affinität und es gilt

$$d\big(\varphi\circ\psi(p),\varphi\circ\psi(q)\big) = d\big(\psi(p),\psi(q)\big) = d(p,q)$$

für $p,q \in A$. Also ist $\varphi\circ\psi$ eine Kongruenz.

*Neutralelement:* Das ist in der affinen Gruppe $\mathrm{id}_A$. Offenbar ist $\mathrm{id}_A$ auch eine Kongruenz.

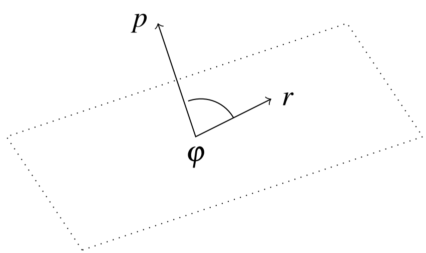
*Inverse:* Sei $\varphi$ eine Kongruenz. Dann gibt es $\varphi^{-1}$ in der affinen Gruppe und wir haben gesehen, dass $\varphi^{-1}$ die Umkehrabbildung ist. Wir müssen zeigen, dass $\varphi^{-1}$ den Abstand erhält. Seien $p,q \in A$. Definiere $p' := \varphi^{-1}(p), q' := \varphi^{-1}(q)$. Dann gilt

$$d\big(\varphi^{-1}(p),\varphi^{-1}(q)\big) = d(p',q') = d\big(\varphi(p'),\varphi(q')\big) = d(p,q)$$

und somit ist auch $\varphi^{-1}$ eine Kongruenz. Damit haben wir gezeigt, dass die Kongruenzen eine Untergruppe der affinen Gruppe bilden. $\square$

**3.5 Beispiel.** Sei $v \in V_A$. Dann hatten wir die Translation $\tau_v$ mit dem Vektor $v$ definiert und gesehen, dass sie eine Affinität von $A$ ist mit $\hat{\tau}_v = \mathrm{id}_{V_A}$ (siehe Beispiel 2.9). Nach Bemerkung 3.3 muss die Translation eine Kongruenz sein.

**3.6 Beispiel.** Ein affiner Unterraum $H$ von $A$ der Dimension $n-1$ heißt **Hyperebene** von $A$. Für $p \in A$ gibt es genau ein $q \in H$ so, dass $\overrightarrow{qp} \perp \overrightarrow{qr}$ für alle $r \in H$:




Wir nennen $q$ die **Orthogonalprojektion** von $p$ auf $A$. Um die Existenz und Eindeutigkeit von $q$ zu beweisen, legen wir den Ursprung $o$ von $A$ irgendwo nach $H$. Dann können wir $A$ mit einem euklidischen (bzw. unitären) Vektorraum identifizieren. Weil $o$ in $H$ liegt, wird dabei $H$ zu einem Unterraum des Vektorraums $A = V_A$ (siehe 2.11). Dann ist $q$ die übliche Orthogonalprojektion aus VIII.4 und damit folgt die Behauptung.

Wir definieren nun die **Spiegelung** $\sigma_H : A \to A$ an der Hyperebene $H$ folgendermaßen: Einem Punkt $p \in A$ ordnen wir den Punkt $s \in A$ zu, der durch $\overrightarrow{ps} = 2\overrightarrow{pq}$ charakterisiert wird, wobei $q$ die Orthogonalprojektion von $p$ auf $H$ ist. Dabei benutzen wir wieder Axiom 1) der affinen Raums. Offenbar gilt

$\sigma_H^2 = \mathrm{id}_A$. Wir behaupten, dass $\sigma_H$ eine Kongruenz ist. Wir wählen dazu ein orthonormales Koordinatensystem $o, p_1, \ldots, p_n$ von $A$ so, dass $o, p_1, \ldots, p_{n-1}$ ein Koordinatensystem von $H$ ist. Dann wird $\sigma_H$ zu einer linearen Abbildung, die durch die Matrix $\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & -1 \end{pmatrix}$ dargestellt wird. Also ist $\sigma_H$ eine bijektive affine Abbildung und $\widehat{\sigma_H}$ eine Isometrie. Mit Bemerkung 3.3 folgt, dass $\sigma_H$ eine Kongruenz ist. □

**3.7 Theorem.** *Jede Kongruenz von einem euklidischen affinen Raum $A$ ist Verknüpfung von maximal $n+2$ Spiegelungen.*

**Beweis:** Wir nehmen zuerst an, dass die Kongruenz $\varphi$ einen Fixpunkt $p \in A$ hat. Dann werden wir sogar zeigen, dass $\varphi$ Verknüpfung von maximal $n$ Spiegelungen ist: Wenn wir den Ursprung von $A$ nach $p$ legen, dann können wir $A$ mit $V_A$ identifizieren (siehe Bemerkung 2.11). Dabei wird $\varphi$ zu $\hat{\varphi}$ wegen $\varphi(p) = p$. Wegen Bemerkung 3.3 ist $\varphi = \hat{\varphi}$ dann eine orthogonale Abbildung von $A = V_A$. Nach Theorem VIII.6.11 gibt es eine Orthonormalbasis von $A = V_A$ so, dass $\varphi$ durch eine Matrix

$$C = \begin{pmatrix} E_r & & & & \\ & -E_s & & & \\ & & K_1 & & \\ & & & \ddots & \\ & & & & K_t \end{pmatrix}$$

in Kästchenform dargestellt wird mit $2 \times 2$ Kästchen $K_j = \begin{pmatrix} \cos(\alpha_j) & \sin(\alpha_j) \\ -\sin(\alpha_j) & \cos(\alpha_j) \end{pmatrix}$. Nach Aufgabe 8.1 lässt sich jede Drehung in $\mathbb{R}^2$ als Produkt von 2 Geradenspiegelungen schreiben. Also gilt

$$K_j = L_{j1} \cdot L_{j2},$$

wobei $L_{jk}$ die $2 \times 2$ Matrix ist, die die Matrix der Spiegelung an der Geraden $n_{jk}^{\perp}$ mit Normalenvektor $n_{jk} \in \mathbb{R}^2$. Wir ersetzen das $j$-te Kästchen $K_j$ in $C$ durch $L_{jk}$ und benutzen sonst die Einheitsmatrix. Dann erhalten wir die Matrix

$$S_{jk} = \begin{pmatrix} E_{r+s+2(j-1)} & 0 & 0 \\ 0 & L_{jk} & 0 \\ 0 & 0 & E_{n-r-s-2j} \end{pmatrix} \in M(n \times n, \mathbb{R}).$$

Dies ist die Matrix der Spiegelung an der Hyperebene $\begin{pmatrix} 0 \\ n_{jk} \\ 0 \end{pmatrix}^{\perp} \in \mathbb{R}^n$, wobei wir beim Vektor mit Null auffüllen für diejenigen Koordinaten, die nicht zum Kästchen $K_j$ beitragen. Weil wir blockweise rechnen können, folgt

$$C = \begin{pmatrix} E_r & & \\ & -E_s & \\ & & E_{n-r-s} \end{pmatrix} S_{11} S_{12} S_{21} S_{22} \cdots S_{t1} S_{t2}.$$

Weil der erste Faktor $s$-faches Produkt von Matrizen der Form $\begin{pmatrix} E_j & & \\ & -1 & \\ & & E_{n-1-j} \end{pmatrix}$ (für ein passendes $j$) ist, welche auch Spiegelungen repräsentieren, lässt sich $\varphi$ als Produkt von $s + 2t \leq n$ Spiegelungen schreiben. Das zeigt die Behauptung im Fall eines Fixpunktes. Im Allgemeinen lässt sich $\varphi$ schreiben als $\varphi = \tau_v \circ \psi$, wobei $\tau_v$ die Translation mit $v \in V_A$ und $\psi$ eine Affinität mit Fixpunkt ist. weil die Kongruenzen eine Untergruppe der affinen Gruppe bilden und weil $\tau_v$ auch eine Kongruenz ist, muss $\psi = \tau_v^{-1} \circ \varphi$ auch eine Kongruenz sein. Weil $\tau_v$ das Produkt von zwei Spiegelungen ist (siehe Proposition 3.8 unten), kann man $\varphi = \tau_v \circ \psi$ als Produkt von maximal $n+2$ Spiegelungen schreiben, denn $\psi$ ist ja nach dem 1. Fall Produkt von maximal $n$ Spiegelungen. □

**3.8 Beispiel.** Jede Translation ist das Produkt von 2 Spiegelungen. Um das zu sehen, identifizieren wir $A$ mit $V_A$ wie oben. Sei $\sigma_1$ die Spiegelung am linearen Unterraum $v^\perp$ und $\sigma_2$ die Spiegelung an der affinen Hyperebene $\{\frac{1}{2}v + w \mid w \in v^\perp\}$, dann gilt $\sigma_2 \circ \sigma_1 = \tau_v$.

## 4. Quadriken

Quadriken sind die Nullstellen einer quadratischen Gleichung in mehreren Variablen. Bei zwei Variablen läuft das im Wesentlichen auf die Kegelschnitte hinaus. Wir werden alle Quadriken bis auf Kongruenz klassifizieren mit Hilfe der Spektraltheorie für selbstadjungierte Abbildungen.

In diesem Abschnitt sei $A$ ein $n$-dimensionaler affiner Raum über $\mathbb{R}$.

**4.1 Definition.** Eine Teilmenge $Q \subseteq A$ heißt **Quadrik**, wenn es ein Koordinatensystem auf $A$ gibt so, dass $Q$ in diesem Koordinatensystem gegeben ist durch $\{\alpha \in \mathbb{R}^n \mid q(\alpha) = 0\}$ für ein Polynom $q(x_1, \ldots, x_n)$ mit Koeffizienten in $\mathbb{R}$ vom Grad 2, d.h.

$$q(x_1, \ldots, x_n) = \sum_{1 \leq i \leq j \leq n} b_{ij} x_i x_j + l_1 x_1 + \cdots + l_n x_n + c \tag{IX.1}$$

mit $b_{ij}, l_k, c \in \mathbb{R}$ und mindestens ein $b_{ij} \neq 0$.

**4.2 Proposition.** *Es sei $Q$ eine Quadrik und $o, p_1, \ldots, p_n$ ein beliebiges Koordinatensystem. Dann gibt es eine symmetrische Matrix $A \in M(n \times n, \mathbb{R}), l \in \mathbb{R}^n$ und $c \in \mathbb{R}$ so, dass $Q$ in diesen Koordinaten gegeben ist durch*

$$\{\alpha \in \mathbb{R}^n | \alpha^t \cdot A \cdot \alpha + l^t \cdot \alpha + c = 0\}. \tag{IX.2}$$

**Beweis:** Zuerst beweisen wir die Behauptung für das spezielle Koordinatensystem aus der Definition 4.1 der Quadrik $Q$. Dabei benutzt man die Form (IX.1) und setzt

$$a_{ij} := \begin{cases} \frac{1}{2} b_{ij} & i < j \\ b_{ij} & i = j \\ \frac{1}{2} b_{ji} & i > j. \end{cases}$$

Weiter wählen wir $l = \begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix}$, dann gilt

$$q(\alpha) = \alpha^t \cdot A \cdot \alpha + l^t \cdot \alpha + c.$$

Offenbar ist $A$ auch symmetrisch. Wir müssen jetzt zeigen, dass die Form (IX.2) unter affinen Koordinatenwechsel erhalten bleibt. Der Koordinatenwechsel ist gleich

$$\alpha' = T \cdot \alpha + \lambda$$

für $T \in \mathrm{GL}(n \times n, \mathbb{R})$ und $\lambda \in \mathbb{R}^n$ (siehe Proposition 1.15). Also gilt

$$\alpha = S \cdot \alpha' + \mu$$

mit $S = T^{-1}$ und $\mu = -T^{-1}\lambda$. Wegen

$$\begin{aligned} \alpha^t \cdot A \cdot \alpha &= (S \cdot \alpha' + \mu)^t \cdot A \cdot (S \cdot \alpha' + \mu) \\ &= \left((\alpha')^t \cdot S^t + \mu^t\right) \cdot A \cdot (S \cdot \alpha' + \mu) \\ &= (\alpha')^t \cdot S^t \cdot A \cdot S \cdot \alpha' + (\alpha')^t \cdot S^t \cdot A \cdot \mu + \mu^t \cdot A \cdot S \cdot \alpha' + \mu^t \cdot A \cdot \mu \end{aligned}$$

und wegen

$$l^t \cdot \alpha = l^t \cdot (S \cdot \alpha' + \mu) = l^t \cdot S \cdot \alpha' + l^t \cdot \mu$$

gilt, dass $Q$ in den neuen Koordinaten gegeben ist durch

$$\{\alpha' \in \mathbb{R}^n | (\alpha')^t \cdot A' \cdot \alpha' + (l')^t \cdot \alpha' + c' = 0\}$$

mit

$$\begin{aligned} A' &:= S^t \cdot A \cdot S, \\ l' &:= S^t \cdot l + 2 \cdot S^t \cdot A \cdot \mu, \\ c' &:= \mu^t \cdot A \cdot \mu + l^t \cdot \mu + c. \end{aligned}$$

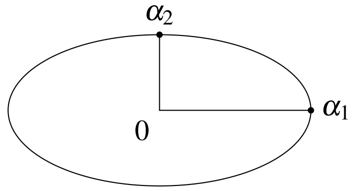
Es bleibt zu zeigen, dass $A'$ symmetrisch ist. Dies folgt aus

$$(A')^t = (S^t \cdot A \cdot S)^t = S^t \cdot A^t \cdot (S^t)^t = S^t \cdot A \cdot S = A'. \qquad \square$$

**4.3 Bemerkung.** Nach Beispiel 1.12 ist die Nullstellenmenge eines Polynoms vom Grad 1 gleich einer Hyperebene in $\mathbb{R}^n$ und jede affine Hyperebene ist von dieser Gestalt.

**4.4 Definition.** Ein Punkt $o \in A$ heißt **Zentrum** der Quadrik $Q$, wenn $\rho(Q) = Q$ gilt für die Punktspiegelung $\rho$ am Punkt $o$.

**4.5 Beispiel.** Eine Ellipse $\{\alpha \in \mathbb{R}^2 | \frac{\alpha_1^2}{v_1^2} + \frac{\alpha_2^2}{v_2^2} = 1\}$ ist eine Quadrik mit Zentrum $\binom{0}{0}$.




**4.6 Beispiel.** $Q := \{\alpha \in \mathbb{R}^3 | \alpha_1 \cdot \alpha_2 = 0\}$ ist eine Quadrik in $\mathbb{R}^3$. Die Zentren von $Q$ bilden die $x_3$-Achse.

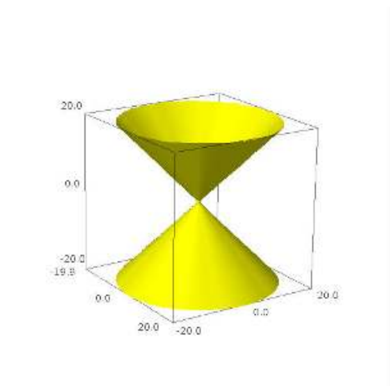
**4.7 Beispiel.** Die Parabel $\{\alpha \in \mathbb{R}^2 \mid \alpha_2 = \alpha_1^2\}$ ist eine Quadrik in $\mathbb{R}^2$ ohne Zentrum.

**4.8 Definition.** Eine Quadrik $Q$ heißt **Kegel**, wenn es einen Punkt $o \in Q$ gibt so, dass für alle $p \in Q$ und alle $q \in A$ mit $\overrightarrow{oq} \in \{\lambda \overrightarrow{op} \mid \lambda \geq 0\}$ folgt, dass $q \in Q$ ist. Der Punkt $o$ heißt **Spitze** von $Q$.

**4.9 Beispiel.** Die Beispiele 4.5, 4.7 sind keine Kegel, während Beispiel 4.6 ein Kegel mit Spitzen auf der $x_3$-Achse. Ein interessanteres Beispiel ist die Quadrik

$$Q := \{\alpha \in \mathbb{R}^3 \mid \alpha_1^2 + \alpha_2^2 = \alpha_3^2\}$$

in $\mathbb{R}^3$, die ein Kegel ist mit Spitze 0.



**4.10 Bemerkung.** Jede Spitze eines Kegels $Q$ ist ein Zentrum von $Q$, das in $Q$ liegt. Umgekehrt ist eine Quadrik $Q$ mit Zentrum $o \in Q$ ein Kegel mit Spitze $o$.

**Beweis:** „$\Longrightarrow$" Wenn $o$ eine Spitze von einem Kegel $Q = \{x \in \mathbb{R}^n \mid q(x) = 0\}$ ist, dann gilt für $p \in Q$, dass die Verbindungsgerade von $o$ und $p$ in $Q$ liegt. Um das zu sehen, wählt man ein Koordinatensystem $o, p_1 = p, p_2, \ldots, p_n$ entlang der Verbindungsgeraden. Dann ist $q(x_1, 0)$ ein Polynom vom Grad $\leq 2$ in $x_1$, dessen Nullstellen den Durchschnitt von $Q$ mit der Verbindungsgeraden beschreiben. Weil $Q$ ein Kegel ist, gibt es unendlich viele Nullstellen (welche den Punkten des Halbstrahls von $o$ in Richtung $p$ entsprechen) und damit ist $q$ das Nullpolynom. Das zeigt, dass die Verbindungsgerade von $o$ und $p$ wirklich in $Q$ liegt.

Spiegelt man $p$ am Punkt $o$, dann erhält man einen Punkt $\rho(p)$ auf dieser Verbindungsgeraden und deshalb gilt $\rho(p) \in Q$. Also ist $Q$ punktsymmetrisch bzgl. $o$ und damit ist $o$ ein Zentrum von $Q$. Nach Definition der Spitze gilt auch $o \in Q$.

„$\Longleftarrow$" Sei jetzt $o$ ein Zentrum von $Q$ mit $o \in Q$. Wir wählen $p \in Q \setminus \{o\}$ und müssen zeigen, dass der Halbstrahl von $o$ in Richtung $p$ in $Q$ enthalten ist. Wir wählen ein affines Koordinatensystem $o, p_1 = p, p_2, \ldots, p_n$ mit Koordinaten $x_1, \ldots, x_n$ wie oben. Dann ist $Q$ gegeben durch eine quadratische Gleichung $q(x_1, \ldots, x_n) = 0$ nach Proposition 4.2. Wir schränken $q$ ein auf die Verbindungsgerade von $o$ und $p$, d.h. wir müssen $x_2 = 0, \ldots, x_n = 0$ setzen. Wir erhalten aus $q(x_1, \ldots, x_n) = 0$ eine Gleichung der Form

$$ax_1^2 + bx_1 + c = 0$$

mit $a, b, c \in \mathbb{R}$. Da diese Gleichung mindestens drei Lösungen entsprechend zu den Punkten $p, o, \rho(p)$ hat, gilt

$$a = b = c = 0.$$

Damit ist die ganze Verbindungsgerade enthalten in $Q$, wie gewünscht. $\square$

**4.11 Beispiel.** Wir betrachten die Quadrik $Q = \{\alpha \in \mathbb{R}^3 \mid \alpha_1\alpha_2 + \alpha_1 + 1 = 0\}$ und wollen sie geometrisch verstehen, dazu machen wir einige Koordinatentransformationen. Wir schreiben

$$\alpha_1\alpha_2 + \alpha_1 + 1 = \alpha_1(\alpha_2 + 1) + 1 = \alpha_1'\alpha_2' + 1$$

mit neuen Koordinaten $\alpha_1' = \alpha_1$, $\alpha_2' = \alpha_2 + 1$. Dann erhalten wir die Quadrik

$$\{\alpha' \in \mathbb{R}^3 \mid \alpha_1' \cdot \alpha_2' = -1\}.$$

Setze $\alpha_1'' := -\alpha_1, \alpha_2'' := \alpha_2'$. In den neuen Koordinaten gilt

$$\{\alpha'' \in \mathbb{R}^3 \mid \alpha_1'' \cdot \alpha_2'' = 1\}.$$

Dies ist ein hyperbolischer Zylinder. In den ursprünglichen Koordinaten liegen die Zentren auf der Geraden $0 = x_1'' = -x_1$, $0 = x_2'' = x_2 + 1$ d.h. gegeben durch $x_1 = 0$, $x_2 = -1$.

**4.12 Theorem.** *Sei $Q$ eine nicht leere Quadrik im euklidischen affinen Raum $A$. Dann gibt es ein orthonormales Koordinatensystem $o, p_1, \ldots, p_n$ so, dass $Q$ in diesen Koordinaten gegeben ist durch eine der folgenden* ***Normalformen****:*

*(a)* $\dfrac{x_1^2}{v_1^2} + \cdots + \dfrac{x_s^2}{v_s^2} - \dfrac{x_{s+1}^2}{v_{s+1}^2} - \cdots - \dfrac{x_{s+t}^2}{v_{s+t}^2} = 0, \quad (1 \leq s, 0 \leq t \leq s, s+t \leq n);$

*(b)* $\dfrac{x_1^2}{v_1^2} + \cdots + \dfrac{x_s^2}{v_s^2} - \dfrac{x_{s+1}^2}{v_{s+1}^2} - \cdots - \dfrac{x_{s+t}^2}{v_{s+t}^2} = 1, \quad (1 \leq s, 0 \leq t, s+t \leq n);$

*(c)* $\dfrac{x_1^2}{v_1^2} + \cdots + \dfrac{x_s^2}{v_s^2} - \dfrac{x_{s+1}^2}{v_{s+1}^2} - \cdots - \dfrac{x_{s+t}^2}{v_{s+t}^2} + x_{s+t+1} = 0, \quad (1 \leq s, 0 \leq t \leq s, s+t < n).$

*mit allen* $\nu_i > 0$.

**Beweis:** Wir wählen ein orthonormales Koordinatensystem $o, p_1, \dots, p_n$ von $A$ so, dass wir $A$ mit $\mathbb{R}^n$ und dem Standard-Skalarprodukt identifizieren können. Nach Proposition 4.2 gibt es eine symmetrische $n \times n$-Matrix $A$, $l \in \mathbb{R}^n$, $c \in \mathbb{R}$ so, dass

$$Q = \{\alpha \in \mathbb{R}^n \mid \alpha^t \cdot A \cdot \alpha + l^t \cdot \alpha + c = 0\}.$$

Mit der Hauptachsentransformation (Korollar VIII.7.2) gibt es eine Orthonormalbasis $v_1, \dots, v_n$ in $\mathbb{R}^n$ und $\lambda_1, \dots, \lambda_n \in K$ so, dass $\alpha^t \cdot A \cdot \alpha = \lambda_1(\alpha_1')^2 + \cdots + \lambda_n(\alpha_n')^2$ gilt in den neuen Koordinaten $\alpha'$. Der Koordinatenwechsel von den Koordinaten bzgl. der Orthonormalbasis $v_1, \dots, v_n$ in die Standardkoordinaten wird durch eine orthogonale Basiswechselmatrix $T$ beschrieben (Proposition VIII.6.7), d.h. $\alpha = T \cdot \alpha'$. Damit haben wir die gemischten Terme $\alpha_i \cdot \alpha_j$ $(i \neq j)$ in dem quadratischen Teil der Gleichung weggebracht und es gilt

$$Q = \{\alpha' \in \mathbb{R}^n \mid (\alpha')^t \cdot A' \cdot \alpha' + (l')^t \cdot \alpha' + c' = 0\}$$

mit $A' = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, $l' = T^t \cdot l$, $c' = c$. Nun machen wir einen affinen Koordinatenwechsel der Form

$$\alpha'' = \alpha' + \mu'$$

mit $\mu' \in \mathbb{R}^n$. Wir erhalten

$$\begin{aligned}
(\alpha')^t \cdot A' \cdot \alpha' + (l')^t \cdot \alpha' + c' &= (\alpha'' - \mu')^t \cdot A' \cdot (\alpha'' - \mu') + (l')^t \cdot (\alpha'' - \mu') + c' \\
&= (\alpha'')^t \cdot A' \cdot \alpha'' - (\mu')^t \cdot A' \cdot \alpha'' - (\alpha'')^t \cdot A' \cdot \mu' \\
&\qquad + (\mu')^t \cdot A' \cdot \mu' + (l')^t \cdot \alpha'' - (l')^t \cdot \mu' + c' \\
&= (\alpha'')^t \cdot A' \cdot \alpha'' + (l'')^t \cdot \alpha'' + c'' \qquad \text{(IX.3)}
\end{aligned}$$

mit

$$l'' := l' - \left((\mu')^t \cdot A'\right)^t - A' \cdot \mu' \underset{A'=(A')^t}{=} l' - 2A'\mu' \qquad c'' := c' - (l')^t\mu' + (\mu')^t \cdot A' \cdot \mu'.$$

**1.Fall:** $l'$ ist im Bild von $A'$.

Das bedeutet präziser, dass das Gleichungssystem $A' \cdot x = l'$ lösbar ist. Dann ist auch $A' \cdot x = \frac{1}{2}l'$ lösbar. Also gibt es ein $\mu' \in \mathbb{R}^n$ mit $2 \cdot A' \cdot \mu' = l'$. Mit dem Koordinatenwechsel $\alpha'' = \alpha' + \mu'$ von oben ergibt sich

$$Q = \{\alpha'' \in \mathbb{R}^n \mid \lambda_1(\alpha_1'')^2 + \ldots + \lambda_n(\alpha_n'')^2 + c'' = 0\}$$

**1.1.** $c'' = 0$. Durch Umordnen der Koordinaten erreichen wir

$$\lambda_1 > 0, \dots, \lambda_s > 0,\ \lambda_{s+1} < 0, \dots, \lambda_{s+t} < 0,\ \lambda_{s+t+1} = 0, \dots, \lambda_n = 0.$$

Dabei dürfen wir auch annehmen, dass $s \geq 1$, $0 \leq t \leq s$ und $s + t \leq n$. Wir setzen

$$\nu_i := \begin{cases} \sqrt{\frac{1}{\lambda_i}} & \text{für } 1 \leq i \leq s, \\ \sqrt{-\frac{1}{\lambda_i}} & \text{für } s+1 \leq i \leq s+t, \end{cases}$$

und erhalten a).

**1.2.** $c'' \neq 0$. Durch Umordnen der Koordinaten erreichen wir

$$\frac{\lambda_1}{c''} < 0, \dots, \frac{\lambda_s}{c''} < 0,\ \frac{\lambda_{s+1}}{c''} > 0, \dots, \frac{\lambda_{s+t}}{c''} > 0\, \lambda_{s+t+1} = 0, \dots, \lambda_n = 0$$

mit $0 \le s,t$ und $1 \le s+t \le n$. Wir setzen

$$v_i := \begin{cases} \sqrt{-\frac{c''}{\lambda_i}} & \text{für } 1 \le i \le s \\ \sqrt{\frac{c''}{\lambda_i}} & \text{für } s+1 \le i \le s+t \end{cases}$$

und erhalten b).

**2.Fall.** $l'$ ist nicht im Bild von $A'$.
Weil $A'$ symmetrisch ist, muss die lineare Abbildung

$$\varphi_{A'}: \mathbb{R}^n \to \mathbb{R}^n,\ \alpha' \mapsto A' \cdot \alpha'$$

selbstadjungiert sein. Nach Korollar VIII.4.7 gilt

$$\mathbb{R}^n = \ker(\varphi_{A'}) \oplus \ker(\varphi_{A'})^\perp$$

und nach Proposition VIII.5.16 gilt

$$\ker(\varphi_{A'})^\perp = \varphi_{A'}(\mathbb{R}^n).$$

Wir können zwar nicht den linearen Teil der Gleichung durch eine Translation wegbringen, aber wir können

$$l' = l'_- + l'_+$$

schreiben mit $l'_- \in \ker(\varphi_{A'})$, $l'_+ \in \varphi_{A'}(\mathbb{R}^n)$. Also existiert ein $\mu' \in \mathbb{R}^n$ mit $l'_+ = 2 \cdot A'\mu'$. Mit der affinen Koordinatentransformation $\alpha'' = \alpha' + \mu'$ folgt wie im 1.Fall

$$Q = \{\alpha'' \in \mathbb{R}^n \mid \lambda_1(\alpha''_1)^2 + \ldots + \lambda_n(\alpha''_n)^2 + (l'_-)^t \cdot \alpha'' + c'' = 0\}.$$

Weil $\varphi_{A'}(\mathbb{R}^n) \neq \mathbb{R}^n$, muss $r := \mathrm{Rang}(A') < n$ gelten. Der Rang $r$ ist bei der Diagonalmatrix $A'$ gleich $|\{j \mid \lambda_j \neq 0\}|$. Durch Umordnen der Koordinaten können wir annehmen, dass

$$\lambda_1 \neq 0, \ldots, \lambda_r \neq 0,\ \lambda_{r+1} = 0, \ldots, \lambda_n = 0.$$

Wir ergänzen $b_{r+1} := \frac{l'_-}{\|l'_-\|}$ zu einer Orthonormalbasis $b_{r+1}, \ldots, b_n$ von $\varphi_{A'}(\mathbb{R}^n)^\perp = \ker(\varphi_{A'})$ und betrachten die Matrix

$$S := \begin{pmatrix} \begin{array}{c|c} E_r & 0 \\ \hline \multicolumn{2}{c}{(b_{r+1})^t} \\ \hline \multicolumn{2}{c}{\vdots} \\ \hline \multicolumn{2}{c}{(b_n)^t} \end{array} \end{pmatrix} \in M(n \times n, \mathbb{R}).$$

Die Matrix $S$ ist orthogonal, weil die Zeilenvektoren orthonormal sind. Um dies zu sehen, bemerken wir, dass $e_1, \ldots, e_r$ ($e_i$ ist hier der $i$-te Standardvektor) im Bild von $\varphi_{A'}$ sind wegen $\varphi_{A'}(e_j) = \lambda_j e_j$ und unter Ausnutzung von $\varphi_{A'}(\mathbb{R}^n)^\perp = \ker(\varphi_{A'})$ folgt die Orthogonalität der Zeilen. Durch den orthogonalen Basiswechsel

$$\alpha''' = S \cdot \alpha''$$

gilt $\alpha'''_i = \alpha''_i$ für $1 \le r$ und $\alpha'''_{r+1} = b^t_{r+1} \cdot \alpha'' = (l'_-/\|l'_-\|)^t \cdot \alpha''$. Also folgt

$$Q = \{\alpha''' \in \mathbb{R}^n \mid \lambda_1(\alpha'''_1)^2 + \cdots + \lambda_r(\alpha'''_r)^2 + \|l'_-\|\alpha'''_{r+1} + c'' = 0\}. \tag{IX.4}$$

Mit einer Translation $\alpha^{(IV)}_{r+1} = \alpha'''_{r+1} + \frac{c''}{\|l'_-\|}$ bringt man noch den konstanten Term $c''$ weg. Durch eine einfache Normierung erhält man (c). $\square$

**4.13 Bemerkung.** Geometrisch ist der Fall (a) dadurch charakterisiert, dass die Quadrik ein Kegel ist. Die Menge der Spitzen ist durch die Gleichungen $x_1 = 0, \ldots, x_{s+t} = 0$ gegeben.

Der Fall (b) ist geometrisch charakterisiert, dass die Quadrik $Q$ einen Mittelpunkt hat, aber kein Kegel ist. Die Menge der Mittelpunkte ist wieder durch die Gleichungen $x_1 = 0, \ldots, x_{s+t} = 0$ gegeben.

Im Fall (c) liegt kein Mittelpunkt vor und man spricht von einem **Paraboloid**.

**4.14 Korollar.** *Jede nicht leere Quadrik im $\mathbb{R}^n$ ist kongruent zu einer Quadrik in Normalform (a)-(c) (siehe Theorem 4.12).*

**Beweis:** Jeder affine Koordinatenwechsel entspricht genau einer Abbildung $\alpha \mapsto T \cdot \alpha + \lambda$ (siehe Proposition 1.15). Es ist leicht zu sehen, dass die neue Basis dann orthonormal ist, wenn $T$ orthogonal (benutzte Proposition VIII.6.7). Also entspricht dem Koordinatenwechsel aus Theorem 4.12 eine Kongruenz und damit folgt die Behauptung. $\square$

**4.15 Bemerkung.** Man kann mit einigem Aufwand zeigen, dass eine Quadrik $Q$, die in keiner affinen Hyperebene enthalten ist, zu *genau einer* Quadrik aus der Liste (a)-(c) kongruent ist (siehe [Wa]).

## 5. Quadriken im $\mathbb{R}^3$

In diesem Abschnitt werden wir alle auftretenden nicht leeren Quadriken im $\mathbb{R}^3$ mit Hilfe von Theorem 4.12 beschreiben. Nach dem Korollar 4.14 muss jede nicht leere Quadrik $Q$ zu einer Normalform aus der Liste (a)-(c) von Theorem 4.12 kongruent sein. Wir studieren zunächst die Beispiele aus dem Fall (a). Das sind die möglichen Kegel in $\mathbb{R}^3$.

**5.1 Beispiel.** $s = 3, t = 0$. Dann ist die Normalform gegeben durch

$$\frac{x_1^2}{v_1^2} + \frac{x_2^2}{v_2^2} + \frac{x_3^2}{v_3^2} = 0$$

und damit liegt ein **Punkt** vor (Nullpunkt).

**5.2 Beispiel.** $s = 2, t = 1$.

$$\frac{x_1^2}{v_1^2} + \frac{x_2^2}{v_2^2} - \frac{x_3^2}{v_3^2} = 0.$$

Dies ist ein elliptischer Kegel mit Spitze in $\left(\begin{smallmatrix}0\\0\\0\end{smallmatrix}\right)$ (siehe Abbildung IX.1).

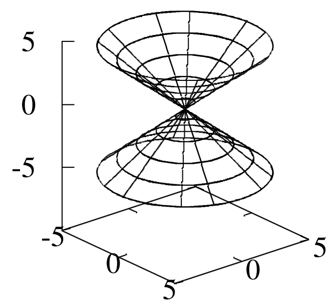



Abbildung IX.1.: elliptischer Kegel

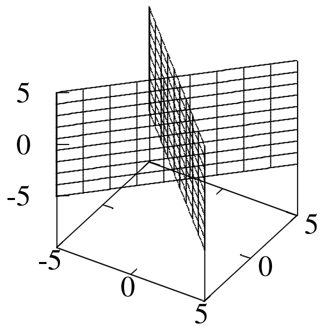



Abbildung IX.2.: schneidendes Ebenenpaar

**5.3 Beispiel.** $s = 2, t = 0$.

$$\frac{x_1^2}{v_1^2} + \frac{x_2^2}{v_2^2} = 0$$

Das ist eine Gerade ($x_3$-Achse).

**5.4 Beispiel.** $s = 1, t = 1$.

$$\frac{x_1^2}{v_1^2} - \frac{x_2^2}{v_2^2} = \left(\frac{x_1}{v_1} + \frac{x_2}{v_2}\right) \cdot \left(\frac{x_1}{v_1} - \frac{x_2}{v_2}\right) = 0$$

Dies ist ein sich **schneidendes Ebenenpaar** (siehe Abbildung IX.2).

**5.5 Beispiel.** $s = 1, t = 0$.

$$\frac{x_1^2}{v_1^2} = 0$$

Dies ist eine **Ebene** ($x_2x_3$-Ebene).

Nun kommen wir zu den Normalformen bei einer Quadrik, die ein Zentrum haben, aber keine Kegel sind. Dies ist der Fall (b) von Theorem 4.12.

**5.6 Beispiel.** $s = 3, t = 0$.

$$\frac{x_1^2}{v_1^2} + \frac{x_2^2}{v_2^2} + \frac{x_3^2}{v_3^2} = 1$$

Dies ist ein **Ellipsoid** (siehe Abbildung IX.3).

**5.7 Beispiel.** $s = 2, t = 1$.

$$\frac{x_1^2}{v_1^2} + \frac{x_2^2}{v_2^2} - \frac{x_3^2}{v_3^2} = 1$$

Dies ist ein **einschaliges Hyperboloid** (siehe Abbildung IX.4).

**5.8 Beispiel.** $s = 2, t = 0$.

$$\frac{x_1^2}{v_1^2} + \frac{x_2^2}{v_2^2} = 1$$

Dies ist ein **elliptischer Zylinder** (siehe Abbildung IX.5).

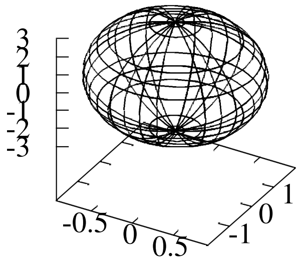



Abbildung IX.3.: Ellipsoid

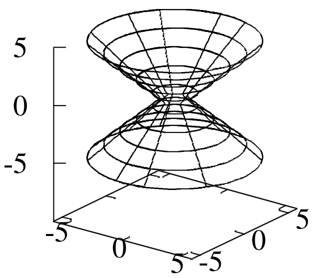



Abbildung IX.4.: einschaliges Hyperboloid

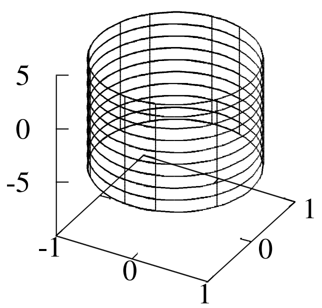



Abbildung IX.5.: elliptischer zylinder

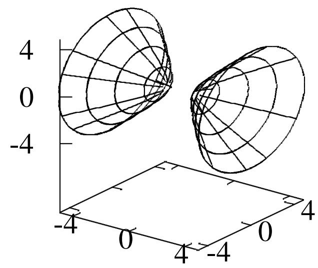



Abbildung IX.6.: zweischaliges Hyperboloid

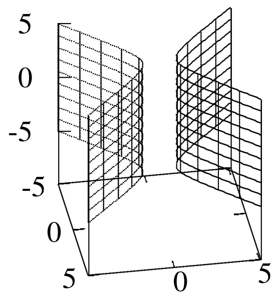



Abbildung IX.7.: hyperbolischer Zylinder

**5.9 Beispiel.** $s = 1, t = 2$.

$$\frac{x_1^2}{v_1^2} - \frac{x_2^2}{v_2^2} - \frac{x_3^2}{v_3^2} = 1$$

Dies ist ein **zweischaliges Hyperboloid** (siehe Abbildung IX.6).

**5.10 Beispiel.** $s = 1, t = 1$.

$$\frac{x_1^2}{v_1^2} - \frac{x_2^2}{v_2^2} = 1$$

Dies ist ein **hyperbolischer Zylinder** (siehe Abbildung IX.7).

**5.11 Beispiel.** $s = 1, t = 0$.

$$\frac{x_1^2}{v_1^2} = 1$$

Dies ist ein **paralleles Ebenenpaar** (siehe Abbildung IX.8).

Nun kommen wir zu den **Paraboloiden** in $\mathbb{R}^3$, d.h. den Normalformen aus Fall (c) von Theorem 4.12.

**5.12 Beispiel.** $s = 2, t = 0$.

$$\frac{x_1^2}{v_1^2} + \frac{x_2^2}{v_2^2} + x_3 = 0$$

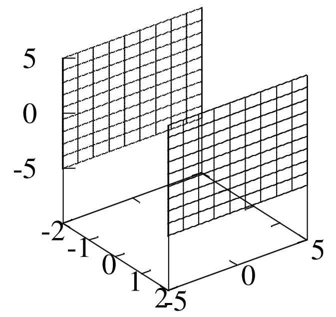



Abbildung IX.8.: paralleles Ebenenpaar

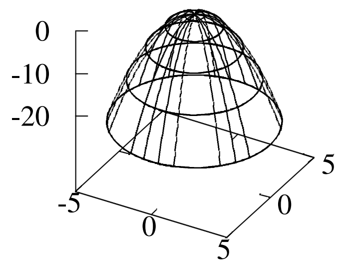



Abbildung IX.9.: elliptisches Paraboloid

Dies ist ein **einschaliges** bzw. **elliptisches Paraboloid** (siehe Abbildung IX.9).

**5.13 Beispiel.** $s = 1, t = 1$.

$$\frac{x_1^2}{v_1^2} - \frac{x_2^2}{v_2^2} + x_3 = 0$$

Dies ist ein **hyperbolisches Paraboloid** (siehe Abbildung IX.10).

**5.14 Beispiel.** $s = 1, t = 0$.

$$\frac{x_1^2}{v_1^2} + x_2 = 0$$

Dies ist ein **parabolischer Zylinder** (siehe Abbildung IX.11).

Dies schließt die Klassifikation der Quadriken in $\mathbb{R}^3$ ab. Zum Abschluss zeigen wir, wie man das in der Praxis umsetzen kann.

**5.15 Beispiel.** Wir betrachten die Quadrik $Q$ in $\mathbb{R}^3$ gegeben durch

$$x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_2x_3 + 2x_1 - 2x_3 + 1 = 0.$$

Unser Ziel ist es $Q$ geometrisch zu beschreiben. Wir bestimmen dazu die Normalform von $Q$ nach dem im Beweis von Theorem 4.12 angegebenen Verfahren und können dann entscheiden, welcher der

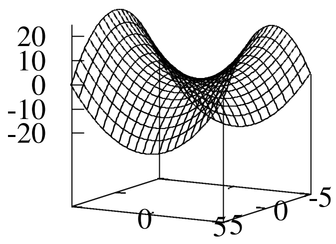



Abbildung IX.10.: hyperbolisches Paraboloid

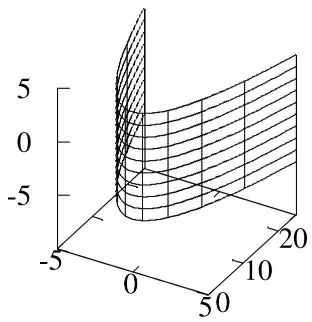



Abbildung IX.11.: parabolischer Zylinder obigen 14 Typen (aus Beispiel 1-14) vorliegt. Weiter wollen wir alle existierenden Zentren bestimmen.

*1.Schritt:* Bestimme die symmetrische Matrix $A$ und $l, c$ so, dass

$$Q = \{\alpha \in \mathbb{R}^3 \mid \alpha^t A \alpha + l^t \alpha + c = 0\}.$$

Die symmetrische Matrix $A$ ist die Matrix Bilinearform zur quadratischen Form $x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_2x_3$ und damit gilt

$$A = \begin{pmatrix} 1 & -1/2 & 0 \\ -1/2 & 1 & -1/2 \\ 0 & -1/2 & 1 \end{pmatrix}, \quad l = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}, \quad c = 1$$

(analog zu Proposition 4.2).

*2.Schritt:* Hauptachsentransformation von $A$.
Das bedeutet, dass wir die Eigenwerte und Eigenräume von $A$ finden müssen. Das kann man wie in Beispiel VII.2.10 machen. Da wurde dies für die Matrix $A$ gemacht. Also findet man, dass $A$ die Eigenwerte $\lambda_1 = 1$, $\lambda_{2/3} = 1 \pm \frac{\sqrt{2}}{2}$ hat mit zugehörigen Eigenvektoren $\left(\begin{smallmatrix} -1 \\ 0 \\ 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 \\ \mp\sqrt{2} \\ 1 \end{smallmatrix}\right)$. Weil $A$ symmetrisch ist, müssen die Eigenvektoren orthogonal sein und damit finden wir die Orthonormalbasis

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \frac{1}{2} \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}, \quad v_3 = \frac{1}{2} \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}.$$

Wir wollen die Transformationsmatrix $S$ für den Basiswechsel von $e_1, e_2, e_3$ (Standardbasis von $\mathbb{R}^3$) in die neue Orthonormalbasis $v_1, v_2, v_3$ angeben. Dazu betrachten wir zuerst den umgekehrten Basiswechsel der durch die Matrix

$$T = \begin{pmatrix} -1/\sqrt{2} & 1/2 & 1/2 \\ 0 & -\sqrt{2}/2 & \sqrt{2}/2 \\ 1/\sqrt{2} & 1/2 & 1/2 \end{pmatrix}$$

gegeben wird (nach V.4.3). Wegen Korollar V.4.5 gilt $S = T^{-1}$. Bei einem Basiswechsel von einer Orthonormalbasis in eine andere Orthonormalbasis ist die Transformationsmatrix orthogonal (Proposition VIII.6.7). Somit gilt $S = T^{-1} = T^t$. Wir betrachten die Koordinatentransformation $\alpha' = S \cdot \alpha$. Nach Konstruktion gilt

$$\alpha^t \cdot A \cdot \alpha = \lambda_1 (\alpha_1')^2 + \lambda_2 (\alpha_2')^2 + \lambda_3 (\alpha_3')^2.$$

Weiter haben wir

$$l^t \cdot \alpha = l^t \cdot T \cdot \alpha' = \begin{pmatrix} -2\sqrt{2} & 0 & 0 \end{pmatrix} \cdot \alpha' = -2\sqrt{2}\alpha_1'$$

Also gilt

$$Q = \{\alpha' \in \mathbb{R}^3 \mid (\alpha')^t \cdot A' \cdot \alpha' + (l')^t \alpha' + 1 = 0\}$$

wobei

$$A' = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \quad l' = \begin{pmatrix} -2\sqrt{2} \\ 0 \\ 0 \end{pmatrix}.$$

*3. Schritt:* Ist $l'$ im Bild von $A'$? Da $\operatorname{Rang}(A') = 3$, ja! Somit kann man direkt nach Zentren suchen. Sonst müsste man mit Fall 2 im Beweis von Theorem 4.12 weitermachen. Wir müssen das Gleichungssystem

$$A' \cdot x = \frac{1}{2} \cdot l'$$

lösen. Als einzige Lösung ergibt sich $\mu' = \left(\begin{smallmatrix} -\sqrt{2} \\ 0 \\ 0 \end{smallmatrix}\right)$. Wir machen den Koordinatenwechsel

$$\alpha'' = \alpha' + \mu'$$

und erhalten für $Q$ die Gleichung

$$
\begin{aligned}
0 &= \lambda_1(\alpha_1')^2 + \lambda_2(\alpha_2')^2 + \lambda_3(\alpha_3')^2 - 2\sqrt{2}\alpha_1' + 1 \\
&= \lambda_1(\alpha_1'' + \sqrt{2})^2 + \lambda_2(\alpha_2'')^2 + \lambda_3(\alpha_3'') - 2\sqrt{2}(\alpha_1'' + \sqrt{2}) + 1 \\
&= \lambda_1(\alpha_1'')^2 + \lambda_2(\alpha_2'')^2 + \lambda_2(\alpha_3'')^2 - 1.
\end{aligned}
$$

Als Normalform haben wir

$$
\underbrace{1}_{=\lambda_1} \cdot (\alpha_1'')^2 + \underbrace{\left(1 + \frac{\sqrt{2}}{2}\right)}_{\lambda_2} (\alpha_2'')^2 + \underbrace{\left(1 - \frac{\sqrt{2}}{2}\right)}_{\lambda_3} (\alpha_3'')^2 = 1
$$

und wegen $\lambda_2 > \lambda_1 > \lambda_3 > 0$ liegt ein **Ellipsoid** vor (Beispiel 5.6). Das einzige Zentrum ist $\alpha'' = 0$. Wir müssen dies zurück transformieren. Zunächst ist das Zentrum $\alpha' = -\mu'$ und damit in den ursprünglichen Koordinaten gegeben durch

$$
Z = T \cdot (-\mu') = \begin{pmatrix} -\sqrt{2}/2 & 1/2 & 1/2 \\ 0 & -\sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} \sqrt{2} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.
$$

# 10 Kapitel X. Normalformen

## 1. Polynomfaktorisierung

Wir brauchen ein paar zusätzliche Grundlagen für Polynome in der Variablen $x$ mit Koeffizienten im Körper $K$. Deren Ring $K[x]$ weißt erstaunliche Parallelen zu $\mathbb{Z}$ auf. Das beruht darauf, dass es für beide Ringe eine Division mit Rest gibt. Dabei spielt der Grad deg der Polynome die Rolle des Betrages der ganzen Zahlen. Wir werden sehen, dass die beiden Ringe $\mathbb{Z}$ und $K[x]$ eine analoge Teilbarkeitstheorie haben. Insbesondere werden wir ein analoges Resultat für $K[x]$ zur Primfaktorisierung in $\mathbb{N}$ beweisen. Wir fassen die bekannten Resultate für $K[x]$ zusammen.

**1.1 Proposition.** *Es seinen $a(x), b(x), c(x) \in K[x]$. Dann gilt*

*a)* $\deg\big(a(x)\cdot b(x)\big) = \deg\big(a(x)\big) + \deg\big(b(x)\big)$; *(Gradformel)*

*b)* $a(x)b(x) = a(x)c(x) \Longrightarrow b(x) = c(x)$ *falls* $a(x) \neq 0$; *(Kürzungsregel)*

*c)* $a(x)b(x) = 0 \Longrightarrow a(x) = 0$ *oder* $b(x) = 0$; *(Nullteilerfreiheit)*

*d) Falls* $b(x) \neq 0$ *dann* $\exists! q(x), r(x) \in K[x]$ *mit* *(Division mit Rest)*

$$a(x) = q(x)b(x) + r(x) \quad \wedge \quad \deg\big(r(x)\big) < \deg\big(b(x)\big);$$

*e) Für* $\alpha \in K$ *mit* $a(\alpha) = 0 \Longrightarrow \exists! q(x) \in K[x]$ *mit* *(Abspalten einer Nullstelle)*

$$a(x) = (x-\alpha)q(x).$$

**Beweis:** Abschnitt II.3. □

**1.2 Korollar.** *Ein Polynom $p(x) \in K[x]$ ist genau dann invertierbar bzgl. $\cdot$, wenn es ein $\alpha \in K\setminus\{0\}$ gibt mit $p(x) = \alpha$.*

**Beweis:** „$\Longrightarrow$“ Es gelte $p(x)q(x) = 1$ für ein $q(x) \in K[x]$. Nach der Gradformel folgt $\deg\big(p(x)\big) = 0 = \deg\big(q(x)\big)$ und damit $p(x) = a_0$ mit $a_0 \in K\setminus\{0\}$.

„$\Longleftarrow$“ Falls $p(x) = \alpha \in K\setminus\{0\} \Longrightarrow \exists\beta \in K\setminus\{0\}$ mit $p(x)\cdot\beta = \alpha\cdot\beta = 1$ aufgrund der Körperaxiome. Also ist $p(x)$ invertierbar. □

**1.3.** Wir übertragen jetzt die Teilbarkeitstheorie auf $K[x]$. Die Definitionen sind vollkommen analog zu $\mathbb{Z}$: Zuerst definieren wir einen **Teiler** $a(x)$ von $b(x) \in K[x]$.

$$a(x) \mid b(x) :\Longleftrightarrow \exists c(x) \in K[x] \text{ mit } b(x) = a(x)\cdot c(x).$$

Nun definieren wir etwas ähnliches wie eine Primzahl. Ein Polynom $p(x) \in K[x]$ heißt **irreduzibel** genau dann, wenn $\deg\big(p(x)\big) \geq 1$ und

$$p(x) = a(x)\cdot b(x) \Longrightarrow \deg\big(a(x)\big) = 0 \quad \text{oder} \quad \deg\big(b(x)\big) = 0$$

gilt.

**1.4 Beispiel.** $p(x) = x^2 + 1$ ist irreduzibel für $K = \mathbb{R}$, aber nicht irreduzibel für $K = \mathbb{C}$, weil es dann Teiler $x + \mathrm{i}$ und $x - \mathrm{i}$ hat.

**1.5 Bemerkung.** Wir erinnern daran, dass ein Polynom **normiert** ist, wenn es von der Form

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1x^1 + a_0$$

ist. In der Teilbarkeitstheorie kann man sich auf normierte Polynome beschränken, weil sich die Teilbarkeitseigenschaften beim Normieren $\big($d.h. Multiplikation und einem Element aus $K \setminus \{0\}\big)$ nicht ändern.

Wir definieren den **größten gemeinsamen Teiler** von $a(x), b(x) \in K[x] \setminus \{0\}$ als denjenigen normierten Teiler von $a(x)$ und $b(x)$ von maximalem Grad. Da $a(x)$ und $b(x)$ den gemeinsamen Teiler 1 haben und da $\deg\big(g(x)\big) \leq \min\big\{\deg\big(a(x)\big), \deg\big(b(x)\big)\big\}$ für jeden gemeinsamen Teiler $g(x)$ gilt, existiert der größte gemeinsame Teiler von $a(x)$ und $b(x)$ tatsächlich und ist eindeutig. Wir bezeichnen ihn mit $\mathrm{ggT}\big(a(x), b(x)\big)$.

**1.6 Proposition.** *Für einen gemeinsamen normierten Teiler $g(x)$ von $a(x), b(x) \in K[x] \setminus \{0\}$ sind die folgenden Aussagen äquivalent:*

*(i)* $g(x) = \mathrm{ggT}\big(a(x), b(x)\big)$;

*(ii)* $\exists p(x), q(x) \in K[x]$ *mit* $a(x)p(x) + b(x)q(x) = g(x)$;

*(iii) Jeder gemeinsamer Teiler von $a(x)$ und $b(x)$ teilt $g(x)$.*

**Beweis:** „(i)$\Longrightarrow$(ii)“: Zuerst betrachten wir den Spezialfall $b(x) \mid a(x)$. Es sei $b(x) = b_m x^m + \ldots + b_0$ mit $b_m \neq 0$. Dann gilt

$$b_m^{-1} b(x) = \mathrm{ggT}\big(a(x), b(x)\big) = g(x)$$

und damit folgt (ii) mit $p(x) = 0$, $q(x) = b_m^{-1}$. Jetzt argumentieren wir mit verbesserter vollständiger Induktion nach $n := \min\big\{\deg\big(a(x)\big), \deg\big(b(x)\big)\big\}$. $n = 0$: Dann gilt $a(x) \in K$ oder $b(x) \in K$ und die Behauptung folgt aus obigem Spezialfall. Sei nun $n \geq 1$ und die Aussage bekannt für $a'(x), b'(x) \in K[x]$ mit $\min\big\{\deg\big(a'(x)\big), \deg\big(b'(x)\big)\big\} < n$. Durch Vertauschen gilt OBdA $\deg\big(a(x)\big) \geq \deg\big(b(x)\big)$. Jetzt machen wir Division mit Rest

$$a(x) = q(x) \cdot b(x) + r(x) \quad \wedge \quad \deg\big(b(x)\big) > \deg\big(r(x)\big).$$

Wie in Proposition I.2.10 beweist man, dass die beiden Polynome $a(x)$ und $b(x)$ dieselben gemeinsamen Teiler haben wie die Polynome $b(x)$ und $r(x)$, also gilt $g(x) = \mathrm{ggT}\big(b(x), r(x)\big)$. Falls $r(x) = 0$, dann gilt $b(x) \mid a(x)$ und (ii) folgt aus obigem Spezialfall. Falls $r(x) \neq 0$, dann gilt

$$\min\big\{\deg\big(b(x)\big), \deg\big(r(x)\big)\big\} = \deg\big(r(x)\big) < \min\big\{\deg\big(a(x)\big), \deg\big(b(x)\big)\big\} = n$$

Nach Induktionsannahme gilt

$$b(x) \cdot e(x) + r(x) \cdot f(x) = g(x)$$

für geeignete $e(x), f(x) \in K[x]$ und damit folgt

$$g(x) = b(x) \cdot e(x) + \big(a(x) - q(x) \cdot b(x)\big) \cdot f(x) = a(x) \cdot f(x) + b(x)\big(e(x) - q(x) \cdot f(x)\big)$$

wie in (ii) gewünscht. Das Verfahren entsprach dem **euklidischen Algorithmus** (siehe [Gu], §1.3).

„(ii)$\Longrightarrow$(iii)“: Falls $d(x) \mid a(x) \ \wedge \ d(x) \mid b(x)$ dann folgt

$$d(x) \mid a(x)p(x) + b(x)q(x) = g(x).$$

„(iii)$\Longrightarrow$(i)“: Sei $d(x)$ ein normierter Teiler von $a(x), b(x)$. Dann gilt nach (iii) $d(x) \mid g(x)$ und somit $\deg\big(d(x)\big) \leq \deg\big(g(x)\big)$. Wenden wir das für $d(x) = \mathrm{ggT}\big(a(x), b(x)\big)$ an, dann folgt aus der Maximalität von $\deg\big(d(x)\big)$, dass $\deg\big(d(x)\big) = \deg\big(g(x)\big)$. Wegen $d(x) \mid g(x)$ und der Gradformel folgt, dass $g(x) = \alpha \cdot d(x)$ für ein $\alpha \in K$. Weil $d(x)$ und $g(x)$ normiert sind, muss $\alpha = 1$ und damit folgt (i). □

**1.7 Korollar.** *Falls $a(x) \mid b(x)c(x)$ in $K[x] \setminus \{0\}$ und $\mathrm{ggT}\big(a(x), b(x)\big) = 1$. Dann gilt $a(x) \mid c(x)$.*

**Beweis:** Nach Proposition 1.6 gibt es $p(x), q(x) \in K[x]$ mit

$$a(x)p(x) + b(x)q(x) = 1.$$

Also folgt mit $a(x)e(x) = b(x)c(x)$ die Behauptung aus

$$\begin{aligned} c(x) &= \big(a(x)p(x) + b(x)q(x)\big)c(x) = a(x)p(x)c(x) + b(x)c(x)q(x) \\ &= a(x)\big(p(x)c(x) + q(x)e(x)\big). \end{aligned}$$

□

**1.8 Theorem.** *Sei $p(x) \in K[x]$ ein normiertes Polynom mit $\deg\big(p(x)\big) \geq 1$.*

*a) Es gibt irreduzible normierte Polynome $p_1(x), \ldots, p_r(x) \in K[x]$ mit $p(x) = p_1(x) \cdots p_r(x)$.*

*b) Diese Darstellung ist bis auf Vertauschungen der irreduziblen Faktoren eindeutig.*

**Beweis:** a) Wir behandeln zuerst den Spezialfall $p(x)$ irreduzibel. Dann setzen wir $p_1(x) := p(x)$ und das ist die gewünschte „Faktorisierung“. Jetzt argumentieren wir mit verbesserter vollständiger Induktion nach $n := \deg\big(p(x)\big)$. $n = 1$: Dann ist $p(x)$ irreduzibel nach der Gradformel und somit die Faktorisierung aus obigem Spezialfall. $n \geq 2$ und die Behauptung bekannt für alle $p'(x) \in K[x] \setminus \{K\}$ mit $\deg\big(p'(x)\big) < n$. Nach obigem Spezialfall dürfen wir annehmen, dass $p(x)$ nicht irreduzibel ist. Also $\exists p_1(x), p_2(x) \in K[x]$ mit $p(x) = p_1(x) \cdot p_2(x)$ und $\deg\big(p_1(x)\big) \neq 0 \neq \deg\big(p_2(x)\big)$. Nach der Gradformel folgt $\deg\big(p_i(x)\big) < n$ $(i = 1, 2)$. Nach Induktionsannahme gibt es irreduzible normierte Polynome $q_1(x), \ldots, q_s(x), q_{s+1}(x), \ldots, q_{s+t}(x) \in K[x]$ mit

$$p_1(x) = q_1(x) \cdots q_s(x) \quad \wedge \quad p_2(x) = q_{s+1}(x) \cdots q_{s+t}(x).$$

Dann ist $p(x) = p_1(x)p_2(x) = q_1(x) \cdots q_{s+t}(x)$ die gewünschte Faktorisierung.

b) Wir betrachten zwei Faktorisierungen

$$p(x) = q_1(x) \cdots q_r(x) = q_1'(x) \cdots q_s'(x) \tag{X.1}$$

von $p(x)$ in irreduzible normierte Polynome. Wir dürfen $r \leq s$ annehmen. Wir argumentieren mit Induktion nach $r$. $r = 1$ $q_1(x) = q_1'(x) \cdots q_s'(x)$. Weil $q_1(x)$ irreduzibel ist, folgt $s = 1$ und $q_1(x) = q_1'(x)$. Sei $r \geq 2$. Aus (X.1) folgt $q_r(x) \mid q_1'(x) \cdots q_s'(x)$. Nach Korollar 1.7 folgt $q_r(x) \mid q_j'(x)$ für ein $j \in \{1, \ldots, s\}$. Durch Vertauschen gilt OBdA $j = s$. Weil $q_j'(x)$ irreduzibel ist, muss $q_r(x) = q_s'(x)$. Mit der Kürzungsregel folgt aus (X.1)

$$q_1(x) \cdots q_{r-1}(x) = q_1'(x) \cdots q_{s-1}'(x).$$

Mit Induktion folgt $r - 1 = s - 1$ und $q_1(x) = q_1'(x), \ldots, q_{r-1}(x) = q_{s-1}'(x)$ nach eventueller Tauschung. Dies zeigt die Eindeutigkeit. □

**1.9 Beispiel.** Gesucht ist die Faktorisierung von $p(x) = x^4 - 1$ in irreduzible Faktoren in $\mathbb{Q}[x]$. Die Lösung ergibt sich aus

$$p(x) = x^4 - 1 = (x^2 - 1) \cdot (x^2 + 1) = (x + 1) \cdot (x - 1) \cdot (x^2 + 1)$$

Dann sind $x \pm 1$ irreduzible normierte Faktoren, weil sie Grad 1 haben. Um zu zeigen, dass $x^2 + 1$ irreduzibel ist, nehmen wir an, dass $x^2 + 1 = q(x)r(x)$ eine Faktorisierung von $x^2 + 1$ ist mit $\deg(q(x)) \neq 0 \neq \deg(r(x))$. Nach der Gradformel gilt: $\deg(r(x)) = \deg(q(x)) = 1$. Durch Normieren könne wir annehmen, dass $r(x) = x - \alpha \wedge q(x) = x - \beta$ mit $\alpha, \beta \in \mathbb{Q}$. Dann sind $\alpha, \beta \in \mathbb{Q}$ Nullstellen von $x^2 + 1$, was wegen $\alpha^2 + 1 > 0$ unmöglich ist. Das Argument zeigt, dass

$$x^4 - 1 = (x - 1)(x + 1)(x - \mathrm{i})(x + \mathrm{i})$$

die Faktorisierung in normierte Polynome in $\mathbb{C}[x]$ wäre.

## 2. Primärzerlegung

In diesem Abschnitt betrachten wir $\varphi \in \mathrm{End}_K(V)$ für einen $K$-Vektorraum $V$ mit $\dim(V) = n < \infty$. Wir studieren die $\varphi$-invarianten Unterräume von $V$, d.h. Unterräume $U \subseteq V$ mit $\varphi(U) \subseteq U$. Wir werden zuerst das Minimalpolynom $m_\varphi(x) \in K[x]$ studieren, das minimal ist mit der Eigenschaft $m_\varphi(\varphi) = 0$ und damit ein Teiler des charakteristischen Polynoms $\chi_\varphi(x)$ ist. Die Faktorisierung in irreduzible Faktoren aus Theorem 1.8 wird dann eine kanonische Zerlegung $V = W_1 \oplus \cdots \oplus W_r$ in $\varphi$-invariante Unterräume, die Primärzerlegung heißt, liefern.

**2.1 Proposition.** *$\exists m(x) \in K[x] \setminus \{0\}$ mit $m(\varphi) = 0$.*

**Beweis:** Nach Korollar V.1.16 ist $\mathrm{End}_K(V)$ ein $n^2$-dimensionaler Vektorraum. Also sind $\mathrm{id}_V, \varphi, \ldots, \varphi^{n^2}$ $K$-linear abhängig, d.h. $\exists a_0, \ldots, a_{n^2} \in K$, nicht alle 0, mit

$$a_{n^2}\varphi^{n^2} + \cdots + a_1\varphi + a_0\,\mathrm{id}_V = 0 \in \mathrm{End}_K(V).$$

Mit $m_\varphi(x) := \sum_{j=0}^{n^2} a_j x^j$ folgt die Behauptung. $\square$

**2.2.** Es gibt also ein eindeutiges normiertes Polynom $m_\varphi(x) \in K[x] \setminus \{0\}$ mit $m_\varphi(\varphi) = 0$ von minimalen Grad. Wir nennen $m_\varphi(x)$ das **Minimalpolynom von $\varphi$**. Es gilt $m_\varphi(x) = 1 \Leftrightarrow V = 0$. In den anderen Fällen ist $\deg m_\varphi(x) \geq 1$.

**2.3 Proposition.** *Für jedes $p(x) \in K[x]$ mit $p(\varphi) = 0$ gilt $m_\varphi(x) \mid p(x)$.*

**Beweis:** Die Division mit Rest liefert:

$$p(x) = q(x) \cdot m_\varphi(x) + r(x) \text{ und } \deg\big(r(x)\big) < \deg\big(m_\varphi(x)\big)$$

Wir wissen von VII.1.18, dass das Einsetzen von $\varphi$ in Polynome ein Ringhomomorphismus ist und damit folgt: $r(\varphi) = p(\varphi) - q(\varphi) \circ m_\varphi(\varphi) = 0 - 0 = 0$. Wegen der Minimalität von $\deg\big(m_\varphi(x)\big)$ folgt $r(x) = 0$ und damit folgt $m_\varphi(x) \mid p(x)$. $\square$

**2.4 Korollar.** *Das Minimalpolynom $m_\varphi(x)$ teilt das charakteristische Polynom $\chi_\varphi(x)$. Insbesondere gilt $\deg(m_\varphi(x)) \leq \deg(\chi_\varphi(x)) = n$.*

**Beweis:** Nach dem Satz von Cayley-Hamilton (siehe VII.4.11) gilt $\chi_\varphi(\varphi) = 0$. Mit Proposition 2.3 folgt $m_\varphi(x) \mid \chi_\varphi(x)$. Nach der Gradformel folgt weiter $\deg\big(m_\varphi(x)\big) \leq \deg\big(\chi_\varphi(x)\big) = n$. $\square$

**2.5 Proposition.** *Es sei $V = W_1 \oplus W_2$ für $\varphi$-invariante Unterräume $W_i$ und $\varphi_i \colon W_i \to W_i, w \mapsto \varphi(w)$ (für $i = 1, 2$). Dann gilt $m_{\varphi_i}(x) \mid m_\varphi(x)$ für $i = 1, 2$ und $m_\varphi(x) \mid m_{\varphi_1}(x) \cdot m_{\varphi_2}(x)$.*

**Beweis:** Wir zeigen zuerst, dass $m(\varphi) = 0$ mit $m(x) := m_{\varphi_1}(x) \cdot m_{\varphi_2}(x)$.

Wir müssen also zeigen, dass $\big(m(\varphi)\big)(w) = 0$ für alle $w \in V$. Weil $w = w_1 + w_2$ und $w_i \in W_i$ und weil $m(\varphi)$ linear, genügt es $\big(m(\varphi)\big)(w_i) = 0$ für $w_i \in W_i$ zu zeigen. Nun gilt

$$\big(m(\varphi)\big)(w_2) = \big(m_{\varphi_1}(\varphi) \circ m_{\varphi_2}(\varphi)\big)(w_2) = m_{\varphi_1}(\varphi)\big((m_{\varphi_2}(\varphi)(w_2)\big).$$

Weil $W_2$ $\varphi$-invariant ist gilt

$$\big(m_{\varphi_2}(\varphi)\big)(w_2) = \big(m_{\varphi_2}(\varphi_2)\big)(w_2) = 0 \in W_2$$

und somit $\big(m(\varphi)\big)(w_1) = 0$ für $w_1 \in W_1$. Dies zeigt $\big(m(\varphi)\big)(w) = 0$ also $m(\varphi) = 0$. Aus Proposition 2.3 folgt $m_\varphi(x) \mid m(x)$.

Andererseits folgt wie oben $\big(m_{\varphi_i}(\varphi)\big)(w_i) = \big(m_\varphi(\varphi)\big)(w_i) = 0$ und damit $m_{\varphi_i}(x) \mid m_\varphi(x)$ nach Proposition 2.3. $\square$

**2.6 Korollar.** *Mit denselben Bezeichnungen wie in Proposition 2.5 gilt* $m_\varphi(x) = m_{\varphi_1}(x) \cdot m_{\varphi_2}(x)$, *falls* $\mathrm{ggT}\big(m_{\varphi_1}(x), m_{\varphi_2}(x)\big) = 1$.

**Beweis:** Seien $m_{\varphi_1}(x) = p_1(x) \cdots p_r(x)$, $m_{\varphi_2}(x) = p_{r+1}(x) \cdots p_{r+s}(x)$ die Faktorisierungen in irreduzible normierte Polynome aus Theorem 1.8. Wenn dabei $W_i = 0$ ist, können wir $m_{\varphi_i}(x) = 1$ setzen. Weil $\mathrm{ggT}\big(m_{\varphi_1}(x), m_{\varphi_2}(x)\big) = 1$ gilt, müssen die irreduzible Faktoren von $m_{\varphi_1}$ verschieden von denjenigen von $m_{\varphi_2}$ sein. Nach Proposition 2.5 gilt:

$$m_\varphi(x) \mid m_{\varphi_1}(x) m_{\varphi_2}(x) = p_1(x) \cdots p_{r+s}(x).$$

Wegen Theorem 1.8 folgt, dass es $J \subseteq \{1, \ldots, r+s\}$ gibt mit:

$$m_{\varphi(x)} = \prod_{j \in J} p_j(x).$$

Andererseits gilt nach Proposition 2.5 $m_{\varphi_1}(x) \mid m_\varphi(x)$ und damit $\{1, \ldots, r\} \subseteq J$ wieder nach Theorem 1.8. Analog folgt $\{r+1, \ldots, r+s\} \subseteq J$ und damit $J = \{1, \ldots, r+s\}$. Das heißt

$$m_\varphi(x) = p_1(x) \cdots p_{r+s}(x) = m_{\varphi_1}(x) \cdot m_{\varphi_2}(x)$$

□

Jetzt betrachten wir die folgende Umkehrung:

**2.7 Theorem.** *Sei* $m_\varphi(x) = m_1(x) \cdots m_r(x)$ *eine Faktorisierung des Minimalpolynoms in paarweise teilerfremde Polynome* $m_i(x) \in K[x]$, *d.h.* $\mathrm{ggT}\big(m_i(x), m_j(x)\big) = 1$ *für* $i \neq j$. *Dann sind die Unterräume* $V_i := \ker\big(m_i(\varphi)\big)$ $\varphi$*-invariant und es gilt* $V = V_1 \oplus \cdots \oplus V_r$.

**Beweis:** Die $\varphi$-Invarianz der Unterräume $V_i$ folgt ohne Bedingung an die Polynome $m_i(\varphi)$ (analog zu Übung 11.2). Betrachte $\hat{m}_i(x) = \frac{m_\varphi(x)}{m_i(x)} = m_1(x) \cdots m_{i-1}(x) \cdot m_{i+1}(x) \cdots m_r(x)$.

*1.Schritt:* $\mathrm{ggT}\big(\hat{m}_1(x), \ldots, \hat{m}_r(x)\big) = 1$.

⌜Indirekt: Wir nehmen an, dass es einen gemeinsamen Teiler von $\hat{m}_1(x), \ldots, \hat{m}_r(x)$ von Grad $\geq 1$ gibt. Nach Theorem 1.8 gibt es dann auch eine irreduziblen Teiler $p(x)$. Wegen $p(x) \mid \hat{m}_1(x) = m_2(x) \cdots m_r(x)$ und weil die $m_i$'s paarweise teilerfremd sind, ist $p(x)$ ein irreduzibler Faktor von genau einem $m_i(x)$ mit $i \in \{2, \ldots, r\}$ (wieder Theorem 1.8). Weil $p(x) \mid \hat{m}_i(x)$ folgt mit demselben Argument, dass $p(x)$ ein irreduzibler Faktor von genau einem $m_j(x)$ ist mit $i \neq j$. Dies widerspricht der Annahme $\mathrm{ggT}\big(m_i(x), m_j(x)\big) = 1$. ⌟

*2.Schritt:* $\exists q_1(x), \ldots, q_r(x) \in K[x]$ mit

$$\hat{m}_1(x) q_1(x) + \cdots + \hat{m}_r(x) q_r(x) = 1.$$

Dies folgt aus Übung 10.4. Wir setzen $p_i(x) := \hat{m}_i(x) q_i(x)$, $P_i := p_i(\varphi) \in \mathrm{End}_K(V)$ und $W_i := P_i(V)$.

*3.Schritt:* $P_1 + \cdots + P_r = \mathrm{id}_V$.

Wegen dem 2. Schritt gilt $p_1(x) + \cdots + p_r(x) = 1$. Weil das Einsetzen von $\varphi$ ein Ringhomomorphismus $K[x] \to \mathrm{End}_K(V)$ induziert (siehe VII1.18), folgt

$$P_1 + \cdots + P_r = p_1(\varphi) + \cdots + p_r(\varphi) = \mathrm{id}_V.$$

*4.Schritt:* $W_1 + \cdots + W_r = V$.

Sei $v \in V$, dann gilt nach dem 3. Schritt $v = P_1(v) + \cdots + P_r(v)$ und wegen $P_i(v) \in W_i$ folgt die Behauptung.

*5.Schritt:* Für $i \neq j$ gilt $P_i \circ P_j = 0$.

⌜Wieder benutzen wir den Einsetzungshomomorphismus mit $\varphi$:

$$P_i \circ P_j = p_i(\varphi) \circ p_j(\varphi) = (p_i p_j)(\varphi) = (q_i q_j \hat{m}_i \hat{m}_j)(\varphi) = q_i(\varphi) \circ q_j(\varphi) \circ (\hat{m}_i \hat{m}_j)(\varphi).$$

Es gilt

$$\hat{m}_i\hat{m}_j = m_\varphi \cdot m_{ij} \quad \text{mit} \quad m_{ij} := \prod_{k\neq i,j} m_k.$$

Also folgt

$$(\hat{m}_i, \hat{m}_j)(\varphi) = m_\varphi(\varphi) \circ m_{ij}(\varphi) = 0$$

und damit $P_i \circ P_j = 0$ wie gewünscht. ⌟

*6.Schritt:* Jedes $P_i$ ist eine Projektion auf $W_i$, d.h. $P_i(V) = W_i$ und $P_i^2 = P_i$.
$P_i(V) = W_i$ nach Definition. Die zweite Behauptung folgt aus

$$P_i^2 \overset{3.Schritt}{=} P_i \circ \Big(\mathrm{id}_V - \sum_{j\neq i} P_j\Big) = P_i - \sum_{j\neq i} P_i \circ P_j \overset{5.Schritt}{=} P_i.$$

*7.Schritt:* $V = W_1 \oplus \cdots \oplus W_r$.

Nach dem 4. Schritt hat jedes $v \in V$ eine Darstellung $v = w_1 + \cdots + w_r$ mit $w_i \in W_i$. Wir müssen zeigen, dass diese Darstellung eindeutig ist. Aus der Definition folgt $w_j = P(v_j)$ für ein $v_j \in V$. Die Eindeutigkeit der Darstellung folgt aus

$$P_i(v) = P_i\Big(\sum_j w_j\Big) = P_i\Big(\sum_j P_j(v_j)\Big) = \sum_j (P_i \circ P_j)(v_j) = P_i(v_i) = w_i.$$

*8.Schritt:* $W_i \subseteq \ker\big(m_i(\varphi)\big)$.

Sei $w_i \in W_i$, das heißt $w_i = P_i(v_i)$ für $v_i \in V$. Zu zeigen ist, dass $\big(m_i(\varphi)\big)(w_i) = \big(m_i(\varphi) \circ P_i\big)(v_i)$ der Nullvektor ist. Das folgt aus

$$m_i(\varphi) \circ P_i = m_i(\varphi) \circ p_i(\varphi) = (m_i p_i)(\varphi) = (m_i \hat{m}_i q_i)(\varphi) = (q_i m_\varphi)(\varphi) = q_i(\varphi) \circ m_\varphi(\varphi) = 0.$$

*9.Schritt:* $\ker\big(m_i(\varphi)\big) \subseteq W_i$.

Sei $u_i \in \ker\big(m_i(\varphi)\big)$, d.h. $\big(m_i(\varphi)\big)(u_i) = 0$. Es gilt:

$$\mathrm{id}_V - P_i \overset{3.\text{Schritt}}{=} \sum_{j\neq i} P_j = \sum_{j\neq i} p_j(\varphi) = \Big(\sum_{j\neq i} \hat{m}_j q_j\Big)(\varphi).$$

Nun ist $m_i$ ein Teiler von $\hat{m}_j q_j$ mit $j \neq i$ und damit gilt

$$\sum_{j\neq i} \hat{m}_j q_j = r_i m_i$$

für ein $r_i \in K[x]$. Also folgt

$$\mathrm{id}_V - P_i = (r_i m_i)(\varphi) = r_i(\varphi) \circ m_i(\varphi).$$

Wir setzen $u_i \in \ker\big(m_i(\varphi)\big)$ und erhalten 0, d.h. $u_i = P_i(u_i) \in W_i$ wie gewünscht.

Aus dem 8. und 9.Schritt ergibt sich $W_i = \ker\big(m_i(\varphi)\big)$ und mit dem 7.Schritt folgt die Behauptung. □

**2.8 Bemerkung.** Die Zerlegung $V = V_1 \oplus \cdots \oplus V_r$ aus Theorem 2.7 ist kanonisch assoziiert zu der Faktorisierung $m_\varphi(x) = m_1(x) \cdots m_r(x)$ in paarweise teilerfremde normierte Polynome. Da $V_i$ $\varphi$-invariant ist, induziert $\varphi$ einen Endomorphismus $\varphi_i \in \mathrm{End}_K(V_i)$ durch

$$\varphi_i \colon V_i \to V_i, \quad v_i \mapsto \varphi(v_i).$$

**2.9 Korollar.** *Es gilt* $m_{\varphi_i}(x) = m_i(x)$. *Insbesondere gilt* $V_i = 0 \iff m_i(x) = 1$.

**Beweis:** Wie in Proposition 2.5 folgt

$$m_\varphi(x) \mid m_{\varphi_1}(x) \cdots m_{\varphi_r}(x).$$

Aus der Definition von $V_i$ ergibt sich $m_i(\varphi_i) = m_i(\varphi)|_{V_i} = 0$ und somit gilt $m_{\varphi_i} \mid m_i$ nach Proposition 2.3. Wegen

$$m_\varphi(x) = m_1(x) \cdots m_r(x)$$

folgt $m_{\varphi_i}(x) = m_i(x)$ wie gewünscht. Die letzte Behauptung ergibt sich aus 2.2. □

**2.10.** Wir haben eine eindeutige Faktorisierung $m_\varphi(x) = p_1(x)^{k_1} \cdots p_r(x)^{k_r}$ in irreduzible normierte und paarweise verschiedene Polynome $p_1(x), \ldots, p_r(x)$ und $k \geq 1$ nach Theorem 1.8. Wir erhalten damit paarweise teilerfremde Polynome $m_i(x) := p_i(x)^{k_i}$. Dann liefert uns Theorem 2.7 eine kanonische Zerlegung $V = V_1 \oplus \cdots \oplus V_r$ in $\varphi$-invariante Unterräume, die wir **Primärzerlegung von V bzgl.** $\varphi$ nennen.

**2.11 Korollar.** *Die Primärzerlegung $V = V_1 \oplus \cdots \oplus V_r$ von $V$ bzgl. $\varphi$ hat die folgenden Eigenschaften:*

*a) $V_i$ ist ein $\varphi$-invarianter Unterraum $\neq \{0\}$;*

*b) $V_i = \ker\big(p_i(\varphi)^{k_i}\big)$;*

*c) $p_i(x)^{k_i}$ ist das Minimalpolynom von $\varphi\colon V_i \to V_i,\ v_i \mapsto \varphi(v_i)$.*

**Beweis:** Folgt direkt aus Theorem 2.7 und Korollar 2.9 für die in c) gewählte Faktorisierung. □

## 3. Jordan-Zerlegung

Wir betrachten wieder einen $n$-dimensionalen Vektorraum $V$ über dem Körper $K$. Wenn das Minimalpolynom $m_\varphi(x)$ von $\varphi \in \mathrm{End}_K(V)$ vollständig in Linearfaktoren zerfällt, dann gibt es eine kanonische Zerlegung von $\varphi$ in eine Summe von einem diagonalisierbaren und einem nilpotenten Endomorphismus so, dass die beiden Endomorphismus kommutieren. Das ist die Jordansche Zerlegung, die in der Theorie der Lie-Algebren eine Rolle spielt.

**3.1 Proposition.** *Sei $\varphi \in \mathrm{End}_K(V)$ mit $m_\varphi(x) = (x-\lambda_1)^{k_1} \cdots (x-\lambda_r)^{k_r}$ für $\lambda_i \in K$ und $k_i \geq 1$. Dann gibt es eine Basis $b_1, \ldots, b_n$ von $V$ so, dass $\varphi$ bzgl. $b_1, \ldots, b_n$ durch eine Matrix der Form*

$$J = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_r \end{pmatrix}$$

*dargestellt wird mit oberen Dreiecksmatrizen $J_i$ der Form*

$$J_i = \begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix} \in M(n_i \times n_i, K).$$

*Weiter sind $\lambda_1, \ldots, \lambda_r$ alle Eigenwerte von $\varphi$ und $n_1, \ldots, n_r$ ihre algebraischen Vielfachheiten.*

**Beweis:** Wir betrachten die Primärzerlegung

$$V = V_1 \oplus \cdots \oplus V_r$$

von $V$ bzgl. $\varphi$ aus 2.10, dabei ist $p_i(x) = x - \lambda_i$. Wir betrachten die Einschränkung $\varphi_i \in \mathrm{End}_K(V_i)$ von $\varphi$ auf $V_i$. Nach Korollar 2.11 gilt $m_{\varphi_i}(x) = (x-\lambda_i)^{k_i}$. Es gilt somit

$$0 = m_{\varphi_i}(\varphi_i) = (\varphi_i - \lambda_i \,\mathrm{id}_{V_i})^{k_i},$$

das heißt $\psi_i := \varphi_i - \lambda_i \operatorname{id}_{V_i}$ ist nilpotent. Nach Korollar VII.4.10 gibt es eine Basis von $V_i$ so, dass $\psi_i$ durch eine $n_i \times n_i$-Matrix der Form $\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$ dargestellt wird, wobei $n_i := \dim(V_i)$. Also wird $\varphi_i = \psi_i + \lambda_i \cdot \operatorname{id}_{V_i}$ durch eine Matrix der Form $J_i = \begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix}$ dargestellt. Setzen wir alle diese Basen zusammen erhalten wir eine Basis von $V$ bzgl. der $\varphi$ die gemischte Form $J$ hat. Weil $J$ eine obere Dreiecksmatrix ist, gilt für das charakteristische Polynom von $\varphi$:

$$\chi_\varphi(x) = \chi_J(x) = \det(x \cdot E_n - J) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_r)^{n_r}$$

nach Proposition VI.3.3. Dies zeigt auch die letzte Behauptung. □

**3.2 Proposition.** *Wenn $\varphi$ diagonalisierbar und $W$ ein $\varphi$-invarianter Unterraum von $V$ ist, dann ist $\varphi_W : W \to W,\ w \to \varphi(w)$ auch diagonalisierbar.*

**Beweis:** Weil $\varphi$ diagonalisierbar ist, gilt $V = \bigoplus_\lambda V_\lambda$, wobei $V_\lambda$ der Eigenraum zum Eigenwert $\lambda$ ist (siehe Theorem VII.2.9). Setze $W_\lambda := V_\lambda \cap W$. Es ist klar, dass $W_\lambda$ der Eigenraum von $\varphi_W$ zu $\lambda$ ist. Nach Theorem VII.2.9 genügt es zu zeigen, dass $W = \sum_\lambda W_\lambda$ gilt. Dies geschieht indirekt. Annahme: $\exists w \in W \setminus \sum_\lambda W_\lambda$. Dann existiert eine eindeutige Darstellung

$$w = \sum_\lambda v_\lambda$$

mit $v_\lambda \in V_\lambda$. Wir wählen $w$ so, dass $|\{\lambda \mid v_\lambda \neq 0\}|$ minimal ist. Es gilt

$$\varphi(w) = \sum_\lambda \varphi(v_\lambda) = \sum_\lambda \lambda v_\lambda.$$

Nun wählen wir einen Eigenwert $\mu$ mit $v_\mu \neq 0$. Wegen $\varphi(W) \subseteq W$ folgt

$$w' := \sum_\lambda (\lambda - \mu) v_\lambda = \varphi(w) - \mu \cdot w \in W.$$

Es gilt $v'_\lambda := (\lambda - \mu) v_\lambda \in V_\lambda$. Wegen $|\{\lambda \mid v'_\lambda \neq 0\}| = |\{\lambda \mid v_\lambda \neq 0\}| - 1$ (da $v'_\mu = 0$), muss $w' \in \sum_\lambda W_\lambda$ aufgrund der Minimalität von $w$ gelten. Weil die Darstellung eindeutig ist, muss $v'_\lambda \in W_\lambda$ und damit $v_\lambda \in W_\lambda$ für alle Eigenwerte $\lambda \neq \mu$. Es folgt

$$v_\mu = w - \sum_{\lambda \neq \mu} v_\lambda \in W$$

und damit auch $v_\mu \in W \cap V_\mu = W_\mu$ im Widerspruch zur Wahl von $w$. Also gilt $W = \sum_\lambda W_\lambda$ und damit ist $\varphi_w$ diagonalisierbar. □

**3.3 Korollar.** *Unter Voraussetzung von Proposition 3.2 gilt $W = \bigoplus_\lambda (W \cap V_\lambda)$, wobei $\lambda$ über alle Eigenwerte läuft.*

**Beweis:** Folgt direkt aus dem Beweis von Proposition 3.2. □

**3.4 Lemma.** *Seien $\psi, \rho \in \operatorname{End}_K(V)$ mit $\psi \circ \rho = \rho \circ \psi$. Wenn $\psi$ diagonalisierbar und $\rho$ trigonalisierbar, dann gibt es eine Basis von $V$ bzgl. der $\psi$ durch eine Diagonalmatrix und $\rho$ durch eine obere Dreiecksmatrix dargestellt wird.*

**Beweis:** Weil $\psi$ diagonalisierbar ist, gilt nach Theorem VII.2.9

$$V = \bigoplus_\lambda V_\lambda \tag{X.2}$$

für alle Eigenräume $V_\lambda$ von $\psi$ zu den Eigenwerten $\lambda$.

*1.Schritt:* $V_\lambda$ ist $\rho$-invariant.

⌜Für $v \in V_\lambda$ gilt $\rho(v) = \sum_\mu w_\mu$ und $w_\mu \in V_\mu$.

$$\sum_\mu \mu w_\mu = \sum_\mu \psi(w_\mu) = \psi \circ \rho(v) = \rho \circ \psi(v) = \rho(\lambda \cdot v) = \sum_\mu \lambda \cdot w_\mu$$

und wegen der Eindeutigkeit der Darstellung folgt $w_\mu = 0$ für $\lambda \neq \mu$ und somit $\rho(v) = w_\lambda \in V_\lambda$. ⌟

*2.Schritt:* $\rho_\lambda : V_\lambda \to V_\lambda,\ v \mapsto \rho(v)$ trigonalisierbar.

⌜Nach Proposition VII.4.5 gilt $\chi_{\rho_\lambda} \mid \chi_\rho$ für die charakteristischen Polynome. Weil $\rho$ trigonalisierbar ist, muss $\chi_\rho$ vollständig in Linearfaktoren aus $K[x]$ zerfallen und damit gilt dies auch für $\chi_{\rho_\lambda}$, was wiederum $\rho_\lambda$ trigonalisierbar impliziert (zwei mal Theorem VII.4.7). ⌟

Wir wählen in jedem $V_\lambda$ eine Basis so, dass $\rho_\lambda$ durch eine obere Dreiecksmatrix dargestellt wird. Wegen (X.2) ergeben diese Basen zusammen eine Basis von $V$. Nach Konstruktion wird $\rho$ durch eine obere Dreiecksmatrix dargestellt und weil alle Basisvektoren Eigenvektoren von $\psi$ sind, muss $\psi$ durch eine Diagonalmatrix dargestellt werden. □

**3.5 Korollar.** *Es seien $\psi, \rho \in \mathrm{End}_K(V)$ mit $\psi \circ \rho = \rho \circ \psi$. Wenn $\psi$ diagonalisierbar ist und $\rho$ nilpotent, dann ist $\psi + \rho$ trigonalisierbar und $\chi_{\psi+\rho}(x) = \chi_\psi(x)$.*

**Beweis:** Nach Korollar VII.4.10 ist $\rho$ trigonalisierbar und alle Eigenwerte sind 0. Mit Lemma 3.4 folgt, dass es eine Basis von $V$ gibt mit Darstellungsmatrizen

$$A_\psi = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad A_\rho = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}, \quad A_{\psi+\rho} = A_\psi + A_\rho = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Die Behauptung folgt aus

$$\chi_{\psi+\rho}(x) = (x - \lambda_1) \cdots (x - \lambda_n) = \chi_\psi(x)$$ □

**3.6 Theorem (Jordan-Zerlegung).** *Sei $\varphi \in \mathrm{End}_K(V)$ mit $m_\varphi(x) = (x - \lambda_1)^{k_1} \cdots (x - \lambda_r)^{k_r}$ für verschiedene $\lambda_1, \ldots, \lambda_r \in K$ und $k_i \geq 1$. Dann gibt es genau ein Paar $(\psi, \rho) \in \mathrm{End}_K(V)^2$ mit den folgenden Eigenschaften*

*a)* $\varphi = \psi + \rho$*;*

*b)* $\psi$ *ist diagonalisierbar;*

*c)* $\rho$ *ist nilpotent;*

*d)* $\psi \circ \rho = \rho \circ \psi$*.*

**Beweis:** Wir wählen die Basis $b_1, \ldots, b_n$ von $V$ aus Proposition 3.1. Dann wird $\varphi$ dargestellt durch die Matrix

$$J = \begin{pmatrix} J_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & J_r \end{pmatrix},$$

wobei

$$J_i = \lambda_i E_{n_i} + S_i, \quad S_i = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}.$$

Seien $\psi$ und $\rho$ die Endomorphismen von $V$, die bzgl. der Basis $b_1, \ldots, b_n$ durch die Matrizen

$$D = \begin{pmatrix} \lambda_1 E_{n_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_r E_{n_r} \end{pmatrix} \quad \wedge \quad S = \begin{pmatrix} S_1 & & 0 \\ & \ddots & \\ 0 & & S_r \end{pmatrix}$$

dargestellt werden. Dann ist $\psi$ offensichtlich diagonalisierbar. Nach Korollar VII.4.10 ist $\rho$ nilpotent. Wegen $J = D + S$ gilt auch $\varphi = \psi + \rho$ und damit folgt a). Weiter gilt

$$(\lambda_i E_{n_i}) \cdot S_i = S_i \cdot (\lambda_i E_{n_i}).$$

Weil man blockweise rechnen kann, folgt $D \cdot S = S \cdot D$ und damit auch d).

Jetzt zeigen wir die Eindeutigkeit. Wir betrachten die folgende Primärzerlegung:

$$V = V_1 \oplus \cdots \oplus V_r$$

von $V$ bzgl. $\varphi$ aus 2.10. Dabei ist

$$V_i := \ker\big((\varphi - \lambda_i \operatorname{id}_V)^{k_i}\big)$$

ein $\varphi$-invarianter Unterraum. Wir betrachten jetzt $\psi, \rho \in \operatorname{End}_K(V)$ mit a)-d). Dann gilt

$$\psi \circ \varphi \underset{\text{a)}}{=} \psi \circ (\psi + \rho) = \psi \circ \psi + \psi \circ \rho \underset{\text{d)}}{=} \psi \circ \psi + \rho \circ \psi = (\psi + \rho) \circ \psi = \varphi \circ \psi$$

und damit auch

$$\psi \circ (\varphi - \lambda_i \cdot \operatorname{id}_V)^{k_i} = (\varphi - \lambda_i \cdot \operatorname{id}_V)^{k_i} \circ \psi \tag{X.3}$$

nach der binomischen Formel. Wir behaupten, dass $V_i$ auch $\psi$-invariant ist:

Wähle $v \in V_i$. Zu zeigen: $\psi(v) \in V_i$. Dies folgt aus

$$(\varphi - \lambda_i \cdot \operatorname{id}_V)^{k_i}\big(\psi(v)\big) \underset{\text{(X.3)}}{=} \psi \circ (\varphi - \lambda_i \cdot \operatorname{id}_V)^{k_i}(v) \underset{v \in V_i}{=} 0.$$

Weil $V_i$ sowohl $\varphi$-invariant wie auch $\psi$-invariant ist, muss $V_i$ auch $\rho = \varphi - \psi$-invariant sein. Wir können somit die Einschränkung $\varphi_i, \psi_i, \rho_i \in \operatorname{End}_K(V_i)$ definieren für $\varphi, \psi, \rho$. Offenbar ist $\rho_i$ auch nilpotent. Wegen Proposition 3.2 ist $\psi_i$ diagonalisierbar. Wegen $\psi_i \circ \rho_i = \rho_i \circ \psi_i$ können wir Korollar 3.5 anwenden und folgern, dass $\psi_i$ und $\varphi_i = \psi_i + \rho_i$ dieselben Eigenwerte haben. Da $m_{\varphi_i}(x) = (x - \lambda_i)^{k_i}$ nach Korollar 2.11, hat $\varphi_i$ den einzigen Eigenwert $\lambda_i$. Damit hat auch $\psi_i$ den einzigen Eigenwert $\lambda_i$. Da $\psi_i$ diagonalisierbar ist, folgt $\psi_i = \lambda_i \cdot \operatorname{id}_{V_i}$. Somit ist $\psi_i$ eindeutig durch $\varphi$ bestimmt und damit auch $\psi$ und $\rho = \varphi - \psi$. Also folgt die Eindeutigkeit. □

## 4. Jordansche Normalform

Wir betrachten weiter einen $n$-dimensionalen Vektorraum $V$ über Körper $K$. Wenn das Minimalpolynom $m_\varphi(x)$ von $\varphi \in \operatorname{End}_K(V)$ vollständig in Linearfaktoren aus $K[x]$ zerfällt, dann besitzt $V$ eine Basis bzgl. der $\varphi$ durch eine besonders einfache Matrix beschrieben wird, die wir Jordansche Normalform nennen. Wir werden das benutzten, um alle komplexen quadratischen Matrizen bis auf Ähnlichkeit zu klassifizieren.

**4.1 Definition.** Für $\lambda \in K$ und $k \geq 1$ sei

$$J_k(\lambda) := \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

Wir nennen $J_k(\lambda)$ den **Jordan-Block** zu $\lambda$ der Länge $k$.

**4.2 Theorem.** *Es sei $\varphi \in \operatorname{End}_K(V)$ so, dass $m_\varphi(x)$ vollständig in Linearfaktoren aus $K[x]$ zerfällt. Dann gibt es eine Basis $V$ so, dass $\varphi$ durch eine Matrix $J$ der Form*

$$J = \begin{pmatrix} J_{k_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{k_s}(\lambda_s) \end{pmatrix} \in M(n \times n, K)$$

*dargestellt wird mit Jordan-Blöcken $J_{k_i}(\lambda_i)$ wie in Definition 4.1.*

**Beweis:** Es gilt $m_\varphi(x) = (x-\lambda_1)^{k_1}\cdots(x-\lambda_r)^{k_r}$ für paarweise verschiedene $\lambda_1,\ldots,\lambda_r$. Wir betrachten dazu die Primärzerlegung

$$V = V_1 \oplus \cdots \oplus V_r$$

aus Korollar 2.11. Jeder Unterraum $V_i$ ist $\varphi$-invariant und die Einschränkung $\varphi_i$ von $\varphi$ auf $V_i$ hat das Minimalpolynom $m_{\varphi_i}(x) = (x-\lambda_i)^{k_i}$. Es genügt die Behauptung für $\varphi_i \in \mathrm{End}_K(V_i)$ zu beweisen, weil wir dann die gefundenen Basen für $V_i$ zu einer Basis von $V$ zusammensetzen können und damit die gewünschte Normalform erhalten. Also dürfen wir $V = V_i$ und $m_\varphi(x) = (x-\lambda)^k$ annehmen. Nach Proposition 3.1 gilt $\chi_\varphi(x) = (x-\lambda)^n$ für das charakteristische Polynom von $\varphi$. Wir ersetzen $\varphi$ durch $\psi = \varphi - \lambda\,\mathrm{id}_V$ und beweisen die Behauptung für $\psi$. Damit folgt die Behauptung auch für $\varphi$, weil wir zu der Normalform von $\psi$ einfach $\lambda \cdot E_n$ hinzu addieren können. Dabei ändert sich nur die Diagonaleinträge der Jordan-Blöcke (mit Addition von $\lambda$). Also dürfen wir annehmen, dass $m_\varphi(X) = x^k$ und damit $\chi_\varphi(x) = x^n$. Nach Korollar VII.4.10 ist $\varphi$ dann nilpotent. Wir beweisen nun die Behauptung für nilpotente Endomorphismen mit verbesserter vollständiger Induktion nach $n$:

$n = 1$: $\Longrightarrow \varphi = \lambda \cdot \mathrm{id}_V$ (sogar $\lambda = 0$, weil $\varphi$ nilpotent). Somit hat jede Darstellungsmatrix die gewünschte Form.

$n \geq 2$: Wir nehmen an, dass die Behauptung für alle Vektorräume der Dimension $m < n$ richtig ist. Offenbar ist $W := \varphi(V)$ ist $\varphi$-invariant. Wir betrachten die Einschränkung $\varphi_W : W \to W,\ w \mapsto \varphi(w)$. Es gilt $m := \dim(W) < \dim(V) = n$, denn sonst würde $\varphi(V) = V$ für alle $k \geq 1$ im Widerspruch zu $\varphi$ nilpotent. Wir können die Induktionsannahme also für $\varphi_W$ benutzen, weil $\varphi_W$ auch nilpotent ist. Nach Induktionsannahme gibt es eine Basis $w_1,\ldots,w_m$ von $W$ so, dass $\varphi_W$ durch eine Matrix der Form

$$J' = \begin{pmatrix} J_{k_1}(0) & & 0 \\ & \ddots & \\ 0 & & J_{k_t}(0) \end{pmatrix}, \quad J_{k_i} = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

dargestellt wird. Dabei treten nur Nullen auf der Diagonalen dieser oberen $m \times m$ Matrix auf, weil $\varphi_W$ nilpotent ist. Wegen der speziellen Gestalt von $J'$ sieht man sofort

$$\mathrm{Rang}(J') = m - t$$

und der Kern von $\varphi_W$ hat Basis $b_1 := w_1, b_2 := w_{k_1+1}, \ldots, b_t := w_{k_1+\cdots+w_{t-1}+1}$. Wir ergänzen dies zu einer Basis $b_1,\ldots,b_t,b_{t+1},\ldots,b_{n-m}$ von $\ker(\varphi)$, wobei wir

$$\dim\big(\ker(\varphi)\big) = \dim(V) - \dim\big(\varphi(V)\big) = n - m$$

benutzt haben (Korollar IV.3.17). Wir betrachten nun folgenden Teil

$$c_1 := w_{k_1},\ c_2 := w_{k_1+k_2}, \ldots,\ c_t := w_{k_1+\cdots k_t} = w_m$$

unserer Basis von $W$. Wegen $W = \varphi(V)$ gibt es $d_i \in V$ mit $c_i = \varphi(d_i)$ für $i = 1,\ldots,t$. Wir behaupten, dass

$$w_1,\ldots,w_m,b_{t+1},\ldots,b_{n-m},d_1,\ldots,d_t \tag{X.4}$$

eine Basis von $V$ bilden. Weil das $n$ Elemente sind, müssen wir nur die lineare Unabhängigkeit prüfen. Aus

$$\sum_{i=1}^{m} \alpha_i w_i + \sum_{j=t+1}^{n-m} \beta_j b_j + \sum_{l=1}^{t} \gamma_l d_l = 0 \tag{X.5}$$

folgt durch Anwenden von $\varphi$

$$\sum_{i \notin \{k_1, k_1+k_2, \ldots, k_1+\cdots+k_t\}} \alpha_{i+1} w_i + \sum_{l=1}^{t} \gamma_l c_l = 0.$$

Wegen $c_l = w_{k_1+\cdots+k_l}$ durchlaufen die auftretenden Vektoren die ganze Basis $w_1,\ldots,w_n$ und damit gilt $\alpha_{i+1} = 0$ für $i \notin \{k_1,\ldots,k_1+\cdots+k_t\}$ und alle $\gamma_l = 0$. Setzt man dies in (X.5) ein, durchlaufen die übrig gebliebenen Vektoren die Basis $b_1,\ldots,b_n$ von $\ker(\varphi)$ und somit $\alpha_i = 0$, $\beta_j = 0$, $\gamma_l = 0$ für alle $i,j,l$. Dies zeigt, dass (X.4) eine Basis von $V$ ist.

Wenn wir nun die Basis (X.4) umordnen und $d_i$ direkt hinter $c_i$ einordnen, dann hat $\varphi$ bzgl. dieser Basis die Darstellungsmatrix

$$J = \left(\begin{array}{ccc|c} J_{k_1+1}(0) & & 0 & 0 \\ & \ddots & & \\ 0 & & J_{k_t+1}(0) & \\ \hline & 0 & & 0 \end{array}\right)$$

wie gewünscht. $\square$

**4.3.** Sei $\varphi \in \mathrm{End}_K(V)$ so, dass $m_\varphi(x)$ vollständig in Linearfaktoren aus $K[x]$ zerfällt. Dann haben wir in Theorem 4.2 gesehen, dass es eine Basis von $V$ gibt so, dass $\varphi$ durch eine Matrix $J$ der Form

$$J = \begin{pmatrix} J_{k_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{k_s}(\lambda_s) \end{pmatrix} \in M(n\times n, K)$$

dargestellt wird, wobei

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \in M(k\times k, K).$$

Wir nennen $J$ die **Jordansche Normalform** von $\varphi$ und $J_{k_1}(\lambda_1),\ldots,J_{k_s}(\lambda_s)$ heißen die **Jordan-Blöcke**.

**4.4 Theorem.** *Die Jordan-Normalform aus Theorem 4.2 ist bis auf Permutation der Jordan-Blöcke eindeutig durch $\varphi$ bestimmt, d.h. unabhängig von der Wahl der Basis.*

**Beweis:** Für $l \geq 0$ und $\mu \in K\setminus\{0\}$ gilt

$$J_k(\mu)^l = \begin{pmatrix} \mu^l & & * \\ & \ddots & \\ 0 & & \mu^l \end{pmatrix} \quad \wedge \quad J_k(0)^l = \begin{pmatrix} 0 & \cdots & 0 & 1 & & * \\ & \ddots & & \ddots & \ddots & \\ & & \ddots & & \ddots & 1 \\ & & & \ddots & & 0 \\ & & & & \ddots & \vdots \\ 0 & & & & & 0 \end{pmatrix}$$

(wobei in der ersten Zeile $l$-mal die 0 und in der letzten Spalte $l$-mal die 0 auftauchen) nach Übung 11.4. Also gilt

$$\mathrm{Rang}\big(J_k(\mu)^l\big) = k \quad \wedge \quad \mathrm{Rang}\big(J_k(0)^l\big) = \max\{k-l, 0\}.$$

Sei $\lambda \in K$. Mit der Blockform der Matrix

$$(J-\lambda\cdot E_n)^l = \begin{pmatrix} J_{k_s}(\lambda_1-\lambda)^l & & 0 \\ & \ddots & \\ 0 & & J_{k_s}(\lambda_s-\lambda)^l \end{pmatrix}$$

von $(\varphi - \lambda \cdot \mathrm{id}_V)^l$ erhalten wir

$$\mathrm{Rang}\big((\varphi - \lambda\, \mathrm{id}_V)^l\big) = \sum_{j=1}^{s} \mathrm{Rang}\big(J_{k_j}(\lambda_j - \lambda)^l\big) = \sum_{\lambda_i \neq \lambda} k_i + \sum_{\lambda_i = \lambda} \max\{k_i - l, 0\}. \qquad \text{(X.6)}$$

Es sei jetzt $l \geq 1$ und wir vergleichen diese Formel (X.6) mit derjenigen für $l-1$. Es gilt

$$\max\{k_i - (l-1), 0\} - \max\{k_i - l, 0\} = \begin{cases} 0 & \text{für } k_i < l, \\ 1 & \text{für } k_i \geq l. \end{cases}$$

Somit folgt aus (X.6)

$$\mathrm{Rang}\big((\varphi - \lambda\, \mathrm{id}_V)^{l-1}\big) - \mathrm{Rang}\big((\varphi - \lambda\, \mathrm{id}_V)^{l}\big) = \sum_{\lambda_i = \lambda,\, k_i \geq l} 1 = |\{i = 1, \ldots, s \mid \lambda_i = \lambda,\, k_i \geq l\}|.$$

Die linke Seite ist eine Invariante, die durch $\varphi$ allein bestimmt wird. Somit bestimmt sie die Anzahl Jordan-Blöcke $J_{k_i}(\lambda_i)$ mit $\lambda_i = \lambda$ und $k_i \geq l$. Weil dies für alle $l$ gilt, ist auch die Anzahl Jordan-Blöcke $J_{k_i}(\lambda_i)$ mit $k_i = l$ und $\lambda_i = \lambda$ eindeutig durch $\varphi$ bestimmt. Sie ist gleich

$$\begin{aligned} &\mathrm{Rang}\big((\varphi - \lambda\, \mathrm{id}_V)^{l-1}\big) - \mathrm{Rang}\big((\varphi - \lambda\, \mathrm{id}_V)^{l}\big) - \mathrm{Rang}\big((\varphi - \lambda\, \mathrm{id}_V)^{l}\big) + \mathrm{Rang}\big((\varphi - \lambda\, \mathrm{id}_V)^{l+1}\big) \\ &\qquad = \mathrm{Rang}\big((\varphi - \lambda\, \mathrm{id}_V)^{l-1}\big) - 2\,\mathrm{Rang}\big((\varphi - \lambda\, \mathrm{id}_V)^{l}\big) + \mathrm{Rang}\big((\varphi - \lambda\, \mathrm{id}_V)^{l+1}\big). \end{aligned}$$

Dies zeigt, wenn man $\lambda$ und $l$ variiert, dass die Jordan-Blöcke von $\varphi$ bis auf Permutation eindeutig durch $\varphi$ bestimmt sind. □

**4.5 Proposition.** *Sei $J$ die Jordan-Normalform von $\varphi$ aus Theorem 4.2. Dann gilt*

*a) $\lambda_1, \ldots, \lambda_s$ sind die Eigenwerte von $\varphi$ (nicht notwendig verschieden);*

*b) $\chi_\varphi(x) = (x - \lambda_1)^{k_1} \cdots (x - \lambda_s)^{k_s}$;*

*c) $m_\varphi(x) = \prod_{\lambda \in \{\lambda_1, \ldots, \lambda_s\}} (x - \lambda)^{\max\{k_i \mid \lambda_i = \lambda\}}$;*

*d) Anzahl Jordan-Blöcke $J_{k_i}(\lambda_i)$ von $J$ mit $k_i \geq l$ und $\lambda_i = \lambda$ ist gleich*

$$\dim\big(\ker((\varphi - \lambda \cdot \mathrm{id}_V)^l)\big) - \dim\big(\ker((\varphi - \lambda \cdot \mathrm{id}_V)^{l-1})\big)$$

*für jedes $\lambda \in K$ und $l \geq 1$.*

**Beweis:** Weil $J$ eine obere Dreiecksmatrix ist, gilt

$$\chi_\varphi(x) = \chi_J(x) = \det(x \cdot E_n - J) = (x - \lambda_1)^{k_1} \cdots (x - \lambda_s)^{k_s}.$$

Dies zeigt b). Weil die Eigenwerte von $\varphi$ die Nullstellen von $\chi_\varphi(x)$ sind, folgt a) aus b). Wir werden c) in Übung 12.4 beweisen. Aus dem Beweis von Theorem 4.4 folgt, dass die fragliche Anzahl in b) gleich

$$\begin{aligned} &\mathrm{Rang}\big((\varphi - \lambda \cdot \mathrm{id}_V)^{l-1}\big) - \mathrm{Rang}\big((\varphi - \lambda \cdot \mathrm{id}_V)^{l}\big) \\ &= n - \dim\big(\ker(\varphi - \lambda \cdot \mathrm{id}_V)^{l-1}\big) - \Big(n - \dim\Big(\ker((\varphi - \lambda \cdot \mathrm{id}_V)^l)\Big)\Big) \\ &= \dim\big(\ker((\varphi - \lambda \cdot \mathrm{id}_V)^l)\big) - \dim\big(\ker((\varphi - \lambda \cdot \mathrm{id}_V)^{l-1})\big), \end{aligned}$$

wobei wir die Dimensionsformel aus Korollar IV.3.17 benutzt haben. □

**4.6 Korollar.** *Es sei $\varphi \in \mathrm{End}_K(V)$. Dann ist $\varphi$ genau dann diagonalisierbar, wenn das Minimalpolynom $m_\varphi(x)$ in ein Produkt paarweise verschiedener Linearfaktoren aus $K[x]$ zerfällt.*

**Beweis:** Wenn $\varphi$ diagonalisierbar ist, dann zerfällt das Minimalpolynom als Teiler des charakteristischen Polynoms (Korollar 2.4) in Linearfaktoren aus $K[x]$. Weil $\varphi$ nach Voraussetzung durch eine Diagonalmatrix dargestellt werden kann und weil jede Diagonalmatrix auch Jordan-Normalform ist mit Jordan-Blöcken der Länge 1, folgt aus Proposition 4.5 c), dass $m_\varphi(x)$ das Produkt von paarweise verschiedene Faktoren ist. Dies zeigt „$\Longrightarrow$“.

„$\Longleftarrow$“: Wenn $m_\varphi(x)$ in ein Produkt von paarweise verschiedenen Linearfaktoren aus $K[x]$ zerfällt, dann gibt es eine Jordan-Normalform $J$ nach Theorem 4.2. Wir benutzen Proposition 4.5 c) und es folgt

$$m_\varphi(x) = \prod_{\lambda \in \{\lambda_1,\ldots,\lambda_s\}} (x-\lambda)^{\max\{k_i \mid \lambda_i = \lambda\}}.$$

Andererseits wissen wir nach Voraussetzung, dass $m_\varphi(x)$ ein Produkt von paarweise verschiedenen Linearfaktoren ist. Weil die Faktorisierung eines Polynoms in irreduzible Faktoren bis auf Permutation eindeutig ist, müssen deshalb alle Exponenten in der obigen Formel gleich 1 sein, d.h. $k_i = 1$ für $i = 1,\ldots,s$. Also hat $J$ nur Jordan-Blöcke der Länge 1 und ist damit eine Diagonalmatrix. Es folgt, dass $\varphi$ diagonalisierbar ist. □

**4.7 Bemerkung.** Die **Jordan-Normalform** von $A \in M(n \times n, K)$ ist definiert als die Jordan-Normalform von $\varphi_A \colon K^n \to K^n,\ \alpha \mapsto A \cdot \alpha$. Sie existiert natürlich nur, wenn das Minimalpolynom $m_A(x)$ vollständig in Linearfaktoren aus $K[x]$ zerfällt. Nach Korollar 3.1 ist die äquivalent dazu, dass $\chi_A(x)$ vollständig in Linearfaktoren aus $K[x]$ zerfällt. Nach dem Fundamentalsatz der Algebra ist diese Bedingung immer erfüllt, falls $K = \mathbb{C}$.

Zur Erinnerung: $A, B \in M(n \times n, K)$ heißen **ähnlich** $:\Longleftrightarrow \exists G \in \mathrm{GL}(n \times n, K)$ mit $\beta = G \cdot A \cdot G^{-1}$.

**4.8 Theorem.** *Zwei komplexe $n \times n$ Matrizen sind genau dann ähnlich, wenn sie dieselbe Jordan-Normalform haben bis auf Vertauschen der Jordan-Blöcke.*

**Beweis:** „$\Longrightarrow$“: Seien $A, B \in M(n \times n, \mathbb{C})$ mit $A \sim B$. Dann gibt es einen $n$-dimensionalen komplexen Vektorraum $V$, $\varphi \in \mathrm{End}_{\mathbb{C}}(V)$ und zwei Basen von $V$ so, dass $\varphi$ die Darstellungsmatrix $A$ bzgl. der einen Basis und die Darstellungsmatrix $B$ bzgl. der anderen Basis hat (siehe Korollar V.4.8). Es seien $J_A$ und $J_B$ die Jordan-Normalformen von $A$ und $B$. Damit sind $J_A$ und $J_B$ auch Jordan-Normalformen von $\varphi$. Nach Theorem 4.4 ist die Jordan-Normalform von $\varphi$ bis auf Permutation der Jordan-Blöcke eindeutig. Das zeigt „$\Longrightarrow$“.

„$\Longleftarrow$“: Es seien $A, B \in M(n \times n, \mathbb{C})$ mit Jordan-Normalformen $J_A$ und $J_B$, die bis auf Permutation der Jordan-Blöcke gleich sind. Letztere entsteht aber aus einer entsprechenden Permutation der Koordinaten, d.h. aus einem (besonders einfachen) Basiswechsel. Mit Korollar V.4.8 folgt $J_A \sim J_B$. Also folgt $A \sim B$ aus

$$A \sim J_A \sim J_B \sim B,$$

weil die Ähnlichkeit eine Äquivalenzrelation ist (Proposition V.2.7). □

# 11 Kapitel XI. Multilineare Algebra

## 1. Direkte Summen

Zuerst wollen wir den Begriff der direkten Summe, den wir in Abschnitt III.2 eingeführt hatten, auf beliebige Familien von Vektorräumen verallgemeinern.

In diesem Abschnitt sei $K$ ein Körper.

**1.1.** Sei $(V_i)_{i\in I}$ eine Familie von $K$-Vektorräumen. Wir haben in III.2.1 gesehen, dass das kartesische Produkt $\prod_{i\in I} V_i = \{(v_i)_{i\in I} \mid v_i \in V_i\}$ einen $K$-Vektorraum bildet unter komponentenweiser Addition und skalarer Multiplikation. Es ist leicht zu sehen, dass

$$\bigoplus_{i\in I} V_i := \{(v_i)_{i\in I} \in \prod_{i\in I} V_i \mid v_i \neq 0 \text{ für nur endlich viele } i \in I\}$$

ein Unterraum ist und damit auf natürliche Art ein $K$-Vektorraum bildet. Wir nennen $\bigoplus_{i\in I} V_i$ die **direkte Summe** der $V_i$.

**1.2.** Für $j \in I$ ist die Abbildung $V_j \to \bigoplus_{i\in I} V_i$, $v_j \mapsto (\dots, 0, v_j, 0 \dots)$ injektiv und linear. Wir identifizieren ab jetzt $V_j$ mit seinem Bild in $\bigoplus_{i\in I} V_i$ und betrachten somit $V_j$ als Unterraum von $\bigoplus_{i\in I} V_i$.

**1.3 Proposition.** *Jedes $v \in \bigoplus_{i\in I} V_i$ hat genau eine Darstellung $v = \sum_{i\in I} v_i$, wobei $v_i \in V_i$ $\forall i \in I$ und $v_i \neq 0$ höchstens für endlich viele $i \in I$.*

**Beweis:** Nach Definition gilt $v = (v_i)_{i\in I}$ mit $v_i \in V_i$ und $v_i \neq 0$ höchstens für endlich viele $i \in I$.

$$v = (v_i)_{i\in I} \underset{\text{Def. von }+}{=} \textstyle\sum_{i\in I}(\dots, 0, v_i, 0, \dots) \underset{\text{Identifikation 1.2}}{=} \displaystyle\sum_{i\in I} v_i.$$

Das zeigt die Existenz. Für die Eindeutigkeit betrachten wir eine weitere solche Zerlegung $v = \sum_{i\in I} v_i'$. Dann gilt

$$v = \sum_{i\in I} v_i' \underset{1.2}{=} \sum_{i\in I}(\dots, 0, v_i', 0, \dots) = (v_i')_{i\in I}$$

Wegen der Eindeutigkeit der Komponenten in $\prod_{i\in I} V_i$ folgt $v_i = v_i'$ $\forall i \in I$. □

**1.4 Bemerkung.** Wenn $I$ endlich ist, dann zeigt Proposition 1.3, dass die in 1.1 definierte direkte Summe mit der in III.2.14 definierten direkten Summe übereinstimmt.

**1.5 Proposition.** *Für $i \in I$ sei $(e_{ij})_{j\in J_i}$ eine Basis von $V_i$. Dann ist $(e_{ij})_{(i,j)\in\{(i,j)\mid i\in I,\, j\in J_i\}}$ eine Basis von $\bigoplus_{i\in I} V_i$.*

**Beweis:** Sei $v \in \bigoplus_{i\in I} V_i$. Nach Proposition 1.3 gilt $v = \sum_{i\in I} v_i$ mit $v_i \in V_i$ und $v_i \neq 0$ für höchstens endlich viele $i \in I$. Für $i \in I \Longrightarrow v_i = \sum_{j\in J_i} \lambda_{ij} e_{ij}$ mit $\lambda_{ij} \in K$ und $\lambda_{ij} = 0$ bis auf endlich viele $j$.

$$\Longrightarrow v = \sum_{i\in I} v_i = \sum_{i\in I} \sum_{j\in J_i} \lambda_{ij} e_{ij}.$$

Also bilden $\{e_{ij} \mid i \in I, j \in J_i\}$ ein Erzeugendensystem von $\bigoplus_{i\in I} V_i$. Es bleibt die lineare Unabhängigkeit zu prüfen. Sei also

$$0 = \sum_{i\in I} \sum_{j\in J_i} \lambda_{ij} e_{ij},$$

wobei $\lambda_{ij} \neq 0$ für höchstens endlich viele $i, j$. Setze

$$v_i := \sum_{j \in J_i} \lambda_{ij} e_{ij}.$$

Dann gilt $v_i \in V_i$ und $0 = \sum_{i \in I} v_i$. Mit Proposition 1.3 gilt $v_i = 0\ \forall i \in I$. Weil nun $(e_{ij})_{j \in J_i}$ eine Basis ist von $V_i$ und $0 = v_i = \sum_{j \in J_i} \lambda_{ij} e_{ij}$, müssen alle $\lambda_{ij} = 0$ sein $\Longrightarrow (e_{ij})_{(i,j) \in \{(i,j) | i \in I, j \in J_i\}}$ ist auch linear unabhängig. $\square$

**1.6 Bemerkung.** Insbesondere besitzt $\bigoplus_{i \in I} K$ die Basis $(e_i)_{i \in I}$ mit $e_i = (\ldots, 0, 1, 0, \ldots)$, wobei der einzige Eintrag 1 an der $i$-ten Stelle ist. Wir sprechen von der Standardbasis von $\bigoplus_{i \in I} K$. Sie verallgemeinert den Spezialfall $K^n = \bigoplus_{i=1}^n K$.

**1.7 Proposition (Universelle Eigenschaft der direkten Summe).** *Für alle $i \in I$ seien $\varphi_i\colon V_i \to W$ lineare Abbildungen in den festen $K$-Vektorraum $W$. Dann existiert genau eine lineare Abbildung*

$$\varphi\colon \bigoplus_{i \in I} V_i \to W$$

*mit $\varphi|_{V_i} = \varphi_i$ für alle $i \in I$.*

**Beweis:** Für $v \in \bigoplus_{i \in I} V_i$ existiert genau eine Darstellung $v = \sum_{i \in I} v_i$ mit $v_i \in V_i$ und $v_i \neq 0$ für höchstens endlich viele $i \in I$ (Proposition 1.3).

$$\varphi(v) := \sum_{i \in I} \varphi_i(v_i).$$

Die ist wohldefiniert, weil höchstens endlich viele Summanden $\neq 0$ sind. Um zu zeigen, dass $\varphi$ linear ist, wählen wir $\lambda, \lambda' \in K$, $v = (v_i)_{i \in I}$, $v' = (v_i')_{i \in I} \in \bigoplus_{i \in I} V_i$.

$$\begin{aligned}
\varphi(\lambda v + \lambda' v') &\underset{1.2}{=} \varphi\Big(\lambda \sum_{i \in I} v_i + \lambda' \sum_{i \in I} v_i'\Big) = \varphi\Big(\sum_{i \in I} \lambda v_i + \lambda' v_i'\Big) \\
&\underset{\text{Def.}\varphi}{=} \sum_{i \in I} \varphi_i(\lambda v_i + \lambda' v_i') \\
&\underset{\varphi_i \text{ linear}}{=} \sum_{i \in I} \lambda \varphi_i(v_i) + \lambda' \varphi_i(v_i') = \lambda \sum_{i \in I} \varphi_i(v_i) + \lambda' \sum_{i \in I} \varphi_i(v_i') \\
&\underset{\text{Def. }\varphi}{=} \lambda \varphi(v) + \lambda' \varphi(v').
\end{aligned}$$

Also ist $\varphi$ linear. Für $v \in V_j$ gilt $v = \sum_{i \in I} v_i$ mit $v_i = \begin{cases} v & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$. Also folgt $\varphi(v) = \sum_{i \in I} \varphi_i(v_i) = \varphi_j(v_j) = \varphi_j(v)$. Also gilt $\varphi|_{V_j} = \varphi_j$. Die Eindeutigkeit von $\varphi$ ergibt sich daraus, dass der Vektorraum $\bigoplus_{i \in I} V_i$ erzeugt wird von den Unterräumen $V_i$. $\square$

## 2. Tensorprodukte

Das Tensorprodukte ist eine „bilineare Konstruktion“ die zu zwei $K$-Vektorräumen $V, W$ einen neuen $K$-Vektorraum $V \otimes W$ konstruiert. Neben der Algebra spielen Tensorprodukte in der Differentialgeometrie und sogar in der Physik eine wichtige Rolle. Wir werden hier das Tensorprodukt vom algebraischen Standpunkt betrachten.

In diesem Abschnitt sei $K$ ein beliebiger Körper.

**2.1 Definition.** Eine **bilineare Abbildung** $f\colon V \times W \to X$ von $K$-Vektorräumen $V, W, X$ erfüllt die Axiome:

i) $f(v + v', w) = f(v, w) + f(v', w)$;

ii) $f(v, w + w') = f(v, w) + f(v, w')$;

iii) $f(\lambda v, w) = \lambda f(v,w) = f(v, \lambda w)$;

für alle $\lambda \in K, v, v' \in V, w, w' \in W$.

**2.2.** Es seien nun im Folgenden $V$ und $W$ $K$-Vektorräume. Wir betrachten $S := V \times W$ als Menge und ignorieren die Vektorraumstruktur. Dann bilden wir die direkte Summe

$$F := \bigoplus_{s \in S} K.$$

Dies ist ein $K$-Vektorraum mit Standardbasis $(e_s)_{s\in S}$, wobei $e_s = (\dots, 0, 1, 0, \dots)$ eine 1 in der $s$-Koordinate hat. In $F$ betrachten wir den Unterraum $E$, der erzeugt wird von allen Elementen der Form

$(*_1)$ $e_{(a,b)} + e_{(a',b)} - e_{(a+a',b)}$

$(*_2)$ $e_{(a,b)} + e_{(a,b')} - e_{(a,b+b')}$

$(*_3)$ $e_{\lambda(a,b)} - e_{(\lambda a,b)}$

$(*_4)$ $e_{\lambda(a,b)} - e_{(a,\lambda b)}$

wobei $a, a' \in V$, $b, b' \in W$, $\lambda \in K$. Das **Tensorprodukt** von $V$ und $W$ wird definiert durch

$$V \otimes W := F/E.$$

Nach Konstruktion ist $V \otimes W$ ein $K$-Vektorraum. Das Bild von $e_{(a,b)}$ unter der Quotientenabbildung $\pi\colon F \to F/E$ wird mit $a \otimes b$ bezeichnet und heißt das **Tensorprodukt von a $\in$ V und b $\in$ W**.

**2.3 Proposition.** *Für $v, v' \in V$, $w, w' \in W$ und $\lambda \in K$ gelten die folgenden Rechenregeln:*

*1)* $(v + v') \otimes w = v \otimes w + v' \otimes w$;

*2)* $v \otimes (w + w') = v \otimes w + v \otimes w'$;

*3)* $\lambda \cdot (v \otimes w) = (\lambda \cdot v) \otimes w$;

*4)* $\lambda \cdot (v \otimes w) = v \otimes (\lambda \cdot w)$.

**Beweis:** Wir zeigen nur 1), die anderen Eigenschaften beweist man analog. Es gilt $\ker(\pi) = E$. Es liegt $e_{(v,w)} + e_{(v',w)} - e_{(v+v',w)}$ in $E$ und damit auch in $\ker(\pi)$.

$$\begin{aligned} \Longrightarrow 0 &= \pi(e_{(v,w)} + e_{(v',w)} - e_{(v+v',w)}) = \pi(e_{(v,w)}) + \pi(e_{(v',w)}) - \pi(e_{(v+v',w)}) \\ &= v \otimes w + v' \otimes w - (v + v') \otimes w. \end{aligned}$$

Dies zeigt 1). □

**2.4 Korollar.** *Die Abbildung $\tau\colon V \times W \to V \otimes W$, $(v, w) \mapsto v \otimes w$, ist bilinear.*

**Beweis:** Dies ist eine Umformulierung von Proposition 2.3. □

**2.5 Bemerkung.** Es lässt sich nicht jedes Element in $V \otimes W$ schreiben als $v \otimes w, v \in V, w \in W$. Zur Erinnerung: $V \otimes W = F/E$. Jedes Element in $F$ lässt sich schreiben als Summe $\sum_i \lambda_i e_{(v_i,w_i)}$ mit endlich vielen $v_i \in V$, $w_i \in W$ und $\lambda_i \in K$. Damit kann man durch Anwendung von $\pi$ nur schließen, dass jedes Element $\omega \in V \otimes W$ die Form $\omega = \sum_i \lambda_i v_i \otimes w_i$ hat. Wegen der Relation 4) in Proposition 2.3 folgt

$$\omega = \sum_{i \in I} v_i' \otimes w_i$$

mit $v_i' = \lambda_i v_i$. Die Elemente in $V \otimes W$ heißen **Tensoren**. Wenn sie die Form $v \otimes w$ haben mit $v \in V$, $w \in W$, dann spricht man von **reinen Tensoren**. Die obige Diskussion zeigt, dass die reinen Tensoren den Vektorraum $V \otimes W$ erzeugen (aber bilden keine Basis).

**2.6 Proposition.** *Das Tensorprodukt $V \otimes W$ zusammen mit der bilinearen Abbildung $\tau$ aus Korollar 2.4 erfüllt folgende universelle Eigenschaft:*

*a) Für jede bilineare Abbildung $f\colon V \times W \to X$ von $K$-Vektorräumen existiert genau eine lineare Abbildung $\varphi\colon V \otimes W \to X$ so, dass das Diagramm*

$$\begin{tikzcd} V \times W \arrow[r, "f"] \arrow[d, "\tau"'] & X \\ V \otimes W \arrow[ur, "\exists!\varphi"'] & \end{tikzcd}$$

*kommutiert, d.h. $f = \varphi \circ \tau$.*

*b) Wenn $\tau'\colon V \times W \to T$ eine bilineare Abbildung ist so, dass $(T, \tau')$ die universelle Eigenschaft a) erfüllt anstelle von $(V \otimes W, \tau)$, dann $\exists!$ Isomorphismus $\theta\colon V \otimes W \to T$ mit $\tau' = \theta \circ \tau$.*

**Beweis:** a) Nach Proposition V.1.1 kann man bei einer linearen Abbildung die Bilder einer Basis frei wählen und damit ist die lineare Abbildung auch eindeutig bestimmt. Wir benutzen nun, dass $(e_s)_{s \in S}$ (mit $S = V \times W$) eine Basis von $F$ ist und wenden das an. Dann gibt es also genau eine lineare Abbildung $\tilde{\varphi}\colon F \to X$ mit $\tilde{\varphi}(e_s) = f(a,b)$ für alle $(a,b) = s \in S$.

Wir zeigen nun, dass $E$ im Kern von $\tilde{\varphi}$ ist:

$$\begin{aligned} \tilde{\varphi}(e_{(a,b)} + e_{(a',b)} - e_{(a+a',b)}) &= \tilde{\varphi}(e_{(a,b)}) + \tilde{\varphi}(e_{(a',b)}) - \tilde{\varphi}(e_{(a+a',b)}) \\ &= f(a,b) + f(a',b) - f(a+a',b) = 0. \end{aligned}$$

Die Elemente vom Typ $(*_2)$ bis $(*_4)$ sind analog im Kern. Da die Erzeuger von $E$ im Kern liegen, liegt $E$ im Kern. Nach dem Homomorphiesatz (Theorem III.3.5) gibt es genau eine lineare Abbildung $\varphi\colon V \otimes W = F/E \to X$ mit $\tilde{\varphi} = \varphi \circ \pi$, wobei $\pi$ der Quotientenhomomorphismus ist. Es gilt $f(a,b) = \tilde{\varphi}(e_{(a,b)}) = \varphi(a \otimes b)$. Weil die reinen Tensoren $a \otimes b$ den Vektorraum $V \otimes W$ erzeugen, folgt $f = \varphi \circ \tau$.
*Eindeutigkeit:* Sei $\varphi'$ eine andere solche Abbildung. Sie stimmt dann insbesondere auf den reinen Tensoren mit $\varphi$ überein und weil die ein Erzeugendensystem bilden, folgt $\varphi' = \varphi$.

b) Wende die universelle Eigenschaft von $(V \otimes W, \tau)$ auf $\tau'$ an und erhalte $\varphi$. Wende danach die universelle Eigenschaft von $(T, \tau')$ auf $\tau$ an und erhalte $\psi$.

$$\begin{tikzcd} & V \times W \arrow[dl, "\tau"'] \arrow[d, "\tau'"] \arrow[dr, "\tau"] & \\ V \otimes W \arrow[r, "\exists!\varphi"] & T \arrow[r, "\exists!\psi"] & V \otimes W \end{tikzcd}$$

Wenn wir nun die universelle Eigenschaft von $(V \otimes W, \tau)$ auf $\tau$ anwenden, dann muss das „$\varphi$“ von a) offensichtlich gleich $\mathrm{id}_{V \otimes W}$ sein. Andererseits lässt auch $\psi \circ \varphi$ das Diagramm kommutieren und wegen der Eindeutigkeit in a) muss $\psi \circ \varphi = \mathrm{id}_{V \otimes W}$ gelten. Analog zeigt man $\psi \circ \varphi = \mathrm{id}_T$. □

**2.7 Bemerkung.** Die bilinearen Abbildung $f\colon V \times W \to X$ bilden bei festem $V, W, X$ einen $K$-Vektorraum unter der punktweisen Addition und skalaren Multiplikation (analog zu Beispiel III.1.1). Aus Proposition 2.6 folgert man sofort, dass dieser Vektorraum isomorph zu $\mathrm{Hom}(V \otimes W, X)$ ist. Der Isomorphismus ist gegeben durch $f \mapsto \varphi$ aus Proposition 2.6. Insbesondere ist der Raum der Bilinearformen $\{b\colon V \times W \to K \mid \text{ bilinear } \}$ isomorph zum Dualraum $(V \otimes W)^*$.

**2.8 Proposition.** *Es seien $\varphi_1\colon V_1 \to W_1$, $\varphi_2\colon V_2 \to W_2$ lineare Abbildungen von $K$-Vektorräumen. Dann existiert genau eine lineare Abbildung $\varphi\colon V_1 \otimes V_2 \to W_1 \otimes W_2$ mit*

$$\varphi(v_1 \otimes v_2) = \varphi_1(v_1) \otimes \varphi_2(v_2)$$

*für alle $v_i \in V_i$. Wir bezeichnen jetzt $\varphi$ mit $\varphi_1 \otimes \varphi_2$.*

**Beweis:** Wir betrachten die Abbildung

$$f\colon V_1 \times V_2 \to W_1 \otimes W_2,\ (v_1, v_2) \mapsto \varphi_1(v_1) \otimes \varphi_2(v_2).$$

Diese Abbildung ist bilinear:

$$\begin{aligned} f(v_1 + v_1', v_2) = \varphi_1(v_1 + v_1') \otimes \varphi_2(v_2) &\underset{\varphi_1 \text{ linear}}{=} (\varphi_1(v_1) + \varphi_1(v_1')) \otimes \varphi_2(v_2) \\ &\underset{2.3,\ 1}{=} \varphi_1(v_1) \otimes \varphi_2(v_2) + \varphi_1(v_1') \otimes \varphi_2(v_2) = f(v_1, v_2) + f(v_1', v_2). \end{aligned}$$

Analog zeigt man die anderen Axiome der Bilinearität. Nach der universellen Eigenschaft 2.6 gilt

$$\begin{tikzcd} V_1 \times V_2 \arrow[d, "\tau"'] \arrow[dr, "f"] & \\ V_1 \otimes V_2 \arrow[r, "\exists! \varphi"] & W_1 \otimes W_2. \end{tikzcd}$$

Für $v_1 \in V_1, v_2 \in V_2$ gilt

$$\varphi(v_1 \otimes v_2) \underset{\text{Def. } \tau}{=} \varphi \circ \tau(v_1, v_2) \underset{\text{Dreieck}}{=} f(v_1, v_2) = \varphi_1(v_1) \otimes \varphi_2(v_2).$$

Also hat $\varphi$ die gewünschte Eigenschaft. Weil die reinen Tensoren $v_1 \otimes v_2$ das Tensorprodukt $V_1 \otimes V_2$ erzeugen, ist $\varphi$ durch diese Eigenschaft eindeutig bestimmt. $\square$

**2.9 Proposition.** *Es sei $V$ ein $K$-Vektorraum. Dann gibt es kanonische Isomorphismen $K \otimes V \cong V \cong V \otimes K$.*

**Beweis:** Wir betrachten die bilineare Abbildung

$$f\colon K \times V \to V,\ (\lambda, v) \mapsto \lambda v.$$

Nach der universellen Eigenschaft in 2.6 gibt es genau eine lineare Abbildung $\varphi\colon K \otimes V \to V,\ \lambda \otimes v \mapsto \lambda v$. Sei $\psi\colon V \to K \otimes V,\ v \mapsto 1 \otimes v$. Nach den Rechenregeln in 2.3 ist $\psi$ wieder eine lineare Abbildung. Es gilt für $v \in V, \lambda \in K$:

$$\varphi \circ \psi(v) = \varphi(1 \otimes v) = 1 \cdot v = v$$

$$\psi \circ \varphi(\lambda \otimes v) = \psi(\lambda v) = 1 \otimes (\lambda v) \underset{2.3}{=} (\lambda \cdot 1) \otimes v = \lambda \otimes v.$$

Weil die reinen Tensoren $\lambda \otimes v$ das Tensorprodukt $K \otimes V$ erzeugen, muss $\varphi$ ein Isomorphismus sein mit Umkehrabbildung $\psi$. Analog zeigt man, dass $V \cong V \otimes K$ gilt. $\square$

**2.10 Proposition.** *Es seien $(V_i)_{i \in I}$ und $(W_j)_{j \in J}$ zwei Familien von $K$-Vektorräumen. dann gibt es genau einen Isomorphismus*

$$\left(\bigoplus_{i \in I} V_i\right) \otimes \left(\bigoplus_{j \in J} W_J\right) \xrightarrow{\sim} \bigoplus_{(i,j) \in I \times J} (V_i \otimes W_j),$$

*mit der Eigenschaft*

$$(v_i)_{i \in I} \otimes (w_j)_{j \in J} \mapsto (v_i \otimes w_j)_{(i,j) \in I \times J}.$$

**Beweis:** Sei $V := \bigoplus_{i \in I} V_i$, $W := \bigoplus_{j \in J} W_j$. Wir zeigen, dass

$$T' := \bigoplus_{(i,j) \in I \times J} V_i \otimes W_j$$

zusammen mit der bilinearen Abbildung

$$\tau'\colon V \times W \to T', \quad ((v_i)_{i \in I}, (w_j)_{j \in J}) \mapsto (v_i \otimes w_j)_{(i,j) \in I \times J},$$

die universelle Eigenschaft in Proposition 2.6 erfüllt. Wir müssen für eine bilineare Abbildung Abbildung $f\colon V\times W\to X$ von $K$-Vektorräumen zeigen, dass es genau eine lineare Abbildung $\varphi\colon T'\to X$ gibt so, dass das Dreieck

$$\begin{tikzcd} V\times W \arrow[d, "\tau'"'] \arrow[dr, "f"] & \\ T' \arrow[r, "\varphi"] & X \end{tikzcd}$$

kommutiert. Durch Einschränkung von $f$ auf den Unterraum $V_i\times W_j$ erhalten wir eine bilineare Abbildung $f_{ij}\colon V_i\times W_j\to X$.

$$\begin{tikzcd} V_i\times W_j \arrow[d, "\tau'_{ij}"'] \arrow[dr, "f_{ij}"] & \\ V_i\otimes W_j \arrow[r, "\exists!\varphi_{ij}"] & X. \end{tikzcd}$$

Hier haben wir die universelle Eigenschaft in Proposition 2.6 für $(V_i\otimes W_j,\tau)$ benutzt. Aufgrund der universellen Eigenschaft von direkten Summen (Proposition XI.1.7) gibt es genau eine lineare Abbildung $\varphi\colon T'\to X$ so, dass $\varphi|_{V_i\otimes W_j}=\varphi_{ij}$ ist. Mit Linearitätsargumenten folgt leicht, dass das erste Dreieck kommutiert und eindeutig ist. Wir haben also gezeigt, dass $T',\tau'$ die universelle Eigenschaft des Tensorprodukts $V\otimes W,\tau$ erfüllt. Nach dem Zusatz in Proposition 2.6 folgt, dass $V\otimes W\cong T$ ist. Man zeigt leicht, dass dies der gewünschte Isomorphismus ist und dass er eindeutig ist. □

**2.11 Proposition.** *Wenn $V$ die Basis $(v_i)_{i\in I}$ und $W$ die Basis $(w_j)_{j\in J}$ hat, dann ist $(v_i\otimes w_j)_{(i,j)\in I\times J}$ eine Basis von $V\otimes W$. Insbesondere gilt für endlich dimensionale $V$ und $W$, dass*

$$\dim(V\otimes W)=\dim(V)\cdot\dim(W).$$

**Beweis:** Es gilt

$$V\otimes W\underset{\text{Basis}}{=}\Big(\bigoplus_{i\in I}Kv_i\Big)\otimes\Big(\bigoplus_{j\in J}Kw_j\Big)\underset{2.10}{\cong}\bigoplus_{(i,j)\in I\times J}(Kv_i)\otimes(Kw_j)$$
$$\underset{Kv_i\cong K\cong Kw_j}{\cong}\bigoplus_{(i,j)\in I\times J}K\otimes K\underset{2.9}{\cong}\bigoplus_{(i,j)\in I\times J}K.$$

Die Standardbasis $\big(e_{(i,j)}\big)_{(i,j)\in I\times J}$ der rechten Seite entsteht unter dem Isomorphismus aus der Familie $(v_i\otimes w_j)_{(i,j)\in I\times J}$ und damit folgt die Behauptung. □

**2.12 Proposition.** *Sei $\varphi\colon V_1\to V_2$ eine lineare Abbildung von $K$-Vektorräumen und sei $W$ ein $K$-Vektorraum. dann gilt:*

*a) $\varphi$ surjektiv $\Longrightarrow \varphi\otimes\mathrm{id}_W\colon V_1\otimes W\to V_2\otimes W$ surjektiv;*

*b) $\varphi$ injektiv $\Longrightarrow \varphi\otimes\mathrm{id}_W\colon V_1\otimes W\to V_2\otimes W$ injektiv.*

**Beweis:** a) $V_2\otimes W$ wird von den reinen Tensoren $v_2\otimes w$ mit $v_2\in V_2$, $w\in W$ erzeugt. Weil $\varphi$ surjektiv ist, gibt es $v_1\in V_1$ mit $\varphi(v_1)=v_2$ und es folgt

$$(\varphi\otimes\mathrm{id}_W)(v_1\otimes w)=\varphi(v_1)\otimes\mathrm{id}_W(w)=\varphi(v_1)\otimes w=v_2\otimes w.$$

Also umfasst das Bild der linearen Abbildung $\varphi\otimes\mathrm{id}_W$ das Erzeugendensystem $(v_2\otimes w)_{v_2\in V_2,w\in W}$. Es folgt, dass $\varphi\otimes\mathrm{id}_W$ surjektiv ist.

b) Eine lineare Abbildung ist genau dann injektiv, wenn sie eine Basis in eine linear unabhängige Familie abbildet. Sei $(b_i)_{i\in I}$ eine Basis von $V_1$. Dann ist $\big(\varphi(b_i)\big)_{i\in I}$ linear unabhängig in $V_2$, weil wir $\varphi$ als injektiv voraussetzen. Somit kann man diese Familie nach dem Basisergänzungssatz zu einer Basis ergänzen. Es sei $(c_j)_{j\in J}$ eine Basis von $W$. Dann folgt aus Proposition 2.11, dass

$(\varphi(b_i)\otimes c_j)_{(i,j)\in I\times J}$ ein Teil einer Basis von $V_2\otimes W$ ist. Insbesondere sind die $\varphi(b_i)\otimes c_j$ linear unabhängig. Weil $(b_i\otimes c_j)_{(i,j)\in I\times J}$ eine Basis von $V_1\otimes W$ ist und weil

$$(\varphi\otimes \mathrm{id}_W)(b_i\otimes c_j)=\varphi(b_i)\otimes c_j$$

gilt, muss mit obigem Fakt auch $\varphi\otimes \mathrm{id}_W$ injektiv sein. □

**2.13.** Sei $L$ eine Körpererweiterung von $K$, d.h. $K\subseteq L$ Teilkörper mit denselben $+,\cdot$ auf $K$. Für einen $K$-Vektorraum $V$ heißt $V_L:=L\otimes V$ die **Skalarerweiterung von V nach L**. Hier benutzen wir, dass $L$ ein $K$-Vektorraum ist, indem wir die skalare Multiplikation $K\times L\to L,\ (\alpha,\beta)\mapsto\alpha\cdot\beta$ benutzen, die durch die Multiplikation von $L$ induziert wird. Somit ist $V_L$ ein $K$-Vektorraum. Wir behaupten, dass es auf $V_L$ genau eine $L$-Vektorraum-Struktur gibt mit demselben $+$ so, dass

$$\lambda\cdot(\alpha\otimes v)=(\lambda\alpha)\otimes v \tag{XI.1}$$

gilt für alle $\alpha,\lambda\in L$, $v\in V$.

**Beweis:** Für $\lambda\in L$ ist $L\times V\to V_L,\ (\alpha,v)\mapsto(\lambda\alpha)\otimes v$ $K$-bilinear und somit gibt es genau eine $K$-lineare Abbildung $T_\lambda\colon V_L\to V_L$ mit $T_\lambda(\alpha\otimes v)=(\lambda\alpha)\otimes v$ aufgrund der universellen Eigenschaft von $L\otimes V$. Wir definieren die skalare Multiplikation durch

$$L\times V_L\to V_L,\ (\lambda,\omega)\mapsto T_\lambda(\omega).$$

Mit Hilfe der Rechenregeln in 2.3 zeigt man leicht die Vektorraumaxiome. Die Identität (XI.1) und die Eindeutigkeit folgen aus der Konstruktion. □

**2.14 Proposition.** *Sei $\varphi\in\mathrm{Hom}_K(V_1,V_2)$. Dann gibt es genau eine L-lineare Abbildung $\varphi_L\colon(V_1)_L\to(V_2)_L$ mit $\varphi_L(\alpha\otimes v)=\alpha\otimes\varphi(v)\ \forall\alpha\in L,\ v\in V_1$.*

**Beweis:** Wir setzen $\varphi_L:=\mathrm{id}_L\otimes\varphi$. Dann ist $\varphi_L$ $K$-linear. Wir wählen $\lambda,\alpha\in L, v\in V_1$.

$$\varphi_L\big(\lambda(\alpha\otimes v)\big)=\varphi_L\big((\lambda\alpha)\otimes v\big)=(\lambda\alpha)\otimes\varphi(v)=\lambda\big(\alpha\otimes\varphi(v)\big)=\lambda\varphi(\alpha\otimes v).$$

Damit zeigt man sofort, dass $\varphi_L$ $L$-linear ist, Die Eindeutigkeit von $\varphi_L$ folgt, weil die reinen Tensoren $\alpha\otimes v$ das Tensorprodukt $(V_1)_L=L\otimes_K V_1$ erzeugen. □

**2.15 Korollar.** *a) $\varphi$ surjektiv $\Longrightarrow$ $\varphi_L$ surjektiv;*

*b) $\varphi$ injektiv $\Longrightarrow$ $\varphi_L$ injektiv.*

**Beweis:** Folgt aus Proposition 2.12. □

## 3. Die Tensoralgebra

Wir führen das Tensorprodukt von mehreren Vektorräumen ein. Damit machen Tensorpotenzen eines gegebenen Vektorraums Sinn und deren direkte Summe nennen wir die Tensoralgebra. Diese hat eine natürliche Ringstruktur, die wir in diesem Abschnitt studieren wollen.

In diesem Abschnitt seien alle auftretenden Vektorräume über dem Körper $K$ definiert. Es seien $V_1,\dots,V_m$ $K$-Vektorräume.

**3.1 Definition.** Eine Abbildung $f\colon V_1\times\cdots\times V_m\to W$ heißt **multilinear**, falls sie folgende Eigenschaft hat für alle $i\in\{1,\dots,m\}$:

$$f(v_1,\dots,v_{i-1},\lambda v_i+\lambda' v_i',v_{i+1},\dots,v_m)=$$
$$\lambda f(v_1,\dots,v_{i-1},v_i,v_{i+1},\dots,v_m)+\lambda' f(v_1,\dots,v_{i-1},v_i',v_{i+1},\dots,v_m)$$

für alle $v_1,\dots,v_m,v_i'\in V$ und $\lambda,\lambda'\in K$.

**3.2.** Für $m \geq 2$ definieren wir $V_1 \otimes \cdots \otimes V_m$ analog zum vorigen Abschnitt: Es sei $S := V_1 \times \cdots \times V_m$ als Menge. Wir betrachten $F := \bigoplus_{s \in S} K$. Dann ist $F$ ein $K$-Vektorraum mit Standardbasis $(e_s)_{s \in S}$, wobei $e_s = (\ldots, 0, 1, 0, \ldots)$ eine 1 in der $s$-Koordinate hat. Es sei $E$ der Unterraum von $F$, der erzeugt wird von allen Elementen der Form

$$e_{\lambda(v_1,\ldots,v_l,\ldots,v_m)} + e_{\lambda(v_1,\ldots,v_l',\ldots,v_m)} - e_{(v_1,\ldots,\lambda v_l + \lambda' v_l',\ldots,v_m)}$$

mit $l \in \{1, \ldots, m\}$, $v_i, v_i' \in V_i$, $\lambda, \lambda' \in K$. Wir definieren das **Tensorprodukt von $\mathbf{V_1}, \ldots, \mathbf{V_m}$** durch

$$V_1 \otimes \cdots \otimes V_m := F/E.$$

Dies ist wieder ein $K$-Vektorraum. Wir bezeichnen die Quotientenabbildung mit $\pi\colon F \to F/E = V_1 \otimes \cdots \otimes V_m$. Für $v_1 \in V_1, \ldots, v_m \in V_m$ definieren wir das *Tensorprodukt*

$$v_1 \otimes \ldots \otimes v_m := \pi(e_{(v_1,\ldots,v_m)}).$$

Nach Konstruktion erzeugen diese reinen Tensoren $v_1 \otimes \cdots \otimes v_m$ das Tensorprodukt $V_1 \otimes \cdots \otimes V_m$ als $K$-Vektorraum. Für $m = 2$ erhalten wir das alte Tensorprodukt $V_1 \otimes V_2$ aus Abschnitt 2.

**3.3 Proposition.** *Die Abbildung $\tau\colon V_1 \times \cdots \times V_m \to V_1 \otimes \cdots \otimes V_m,\ (v_1, \ldots, v_m) \mapsto v_1 \otimes \cdots \otimes v_m$ ist multilinear.*

**Beweis:** Analog wie Proposition 2.3. □

**3.4 Proposition.** *Das Tensorprodukt $V_1 \otimes \cdots \otimes V_m$ zusammen mit der multilinearen Abbildung $\tau$ aus Proposition 3.2 erfüllt folgende universelle Eigenschaft:*

a) *Für jede multilineare Abbildung $f\colon V_1 \times \cdots \times V_m \to W$ von $K$-Vektorräumen existiert genau eine lineare Abbildung $\varphi\colon V_1 \otimes \cdots \otimes V_m \to W$ so, dass das folgende Diagramm kommutiert:*

$$\begin{tikzcd} V_1 \times \ldots \times V_m \arrow[r, "f"] \arrow[d, "\tau"'] & X \\ V_1 \otimes \ldots \otimes V_m \arrow[ur, "\exists!\varphi"'] & \end{tikzcd}$$

b) *Wenn $\tau'\colon V_1 \times \cdots \times V_m \to T$ eine multilineare Abbildung von $K$-Vektorräumen ist so, dass $(T, \tau')$ die universelle Eigenschaft a) erfüllt anstelle von $(V_1 \otimes \cdots \otimes V_m, \tau)$, dann $\exists!$ Isomorphismus $\theta\colon V_1 \otimes \cdots \otimes V_m \to T$ so, dass $\tau' = \theta \circ \tau$.*

**Beweis:** Analog zu Proposition 2.6. □

**3.5 Proposition.** *Es gibt genau einen Isomorphismus $\psi\colon V_1 \otimes \cdots \otimes V_m \to (V_1 \otimes \cdots \otimes V_{m-1}) \otimes V_m$ mit $\psi(v_1 \otimes \cdots \otimes v_m) = (v_1 \otimes \cdots \otimes v_{m-1}) \otimes v_m$.*

**Beweis:** Die Abbildung $f\colon V_1 \times \cdots \times V_m \to (V_1 \otimes \cdots \otimes V_{m-1}) \otimes V_m, (v_1, \ldots, v_m) \mapsto (v_1 \otimes \cdots \otimes v_{m-1}) \otimes v_m$ ist multilinear und somit gibt es nach der universellen Eigenschaft 3.4 genau eine lineare Abbildung $\varphi\colon V_1 \otimes \cdots \otimes V_m \to W$ mit $\varphi(v_1 \otimes \cdots \otimes v_m) = (v_1 \otimes \cdots \otimes v_{m-1}) \otimes v_m$.

Es bleibt zu zeigen, dass $\varphi$ ein Isomorphismus ist. Dazu konstruieren wir die Umkehrabbildung. Für fest gewähltes $v_m \in V_m$ betrachten wir die multilineare Abbildung $V_1 \times \cdots \times V_{m-1} \to V_1 \otimes \cdots \otimes V_m$, $(v_1, \ldots, v_{m-1}) \to v_1 \otimes \cdots \otimes v_{m-1} \otimes v_m$. Nach der universellen Eigenschaft 3.4 gibt es genau eine lineare Abbildung $\psi_{v_m}\colon V_1 \otimes \cdots \otimes V_{m-1} \to V_1 \otimes \cdots \otimes V_m$ mit $\psi_{v_m}(v_1 \otimes \cdots \otimes v_{m-1}) = v_1 \otimes \cdots \otimes v_m$. Wir betrachten nun die bilineare Abbildung $(V_1 \otimes \cdots \otimes V_{m-1}) \times V_m \to V_1 \otimes \cdots \otimes V_m$, $(\omega, v_m) \mapsto \psi_{v_m}(\omega)$. Dabei ist die Linearität in $\omega \in V_1 \otimes \cdots \otimes V_{m-1}$ klar nach Konstruktion. Die Linearität in $v_m$ rechnet man für einen reinen Tensor $\omega$ direkt mit der Definition nach und der allgemeine Fall folgt aus Linearität in $\omega$. Nach der universellen Eigenschaft 2.6 gibt es genau eine lineare Abbildung $\psi\colon (V_1 \otimes \cdots \otimes V_{m-1}) \otimes V_m \to V_1 \otimes \cdots \otimes V_m$ mit $\psi\big((v_1 \otimes \cdots \otimes v_{m-1}) \otimes v_m\big) = \psi_{v_m}(v_1 \otimes \cdots \otimes v_{m-1}) = v_1 \otimes \cdots \otimes v_m$. Aus den Definitionen folgt $\varphi \circ \psi = \mathrm{id}$ und $\psi \circ \varphi = \mathrm{id}$ auf den reinen Tensoren und damit überall, weil die reinen Tensoren die Tensorprodukte erzeugen. Damit ist $\psi$ die Umkehrabbildung von $\varphi$. □

**3.6 Korollar.** *Seien $V, W, X$ $K$-Vektorräume. Dann gibt es genau einen Isomorphismus $\varphi: (V \otimes W) \otimes X \to V \otimes (W \otimes X)$ mit $\varphi((a \otimes b) \otimes c) = a \otimes (b \otimes c)$.*

**Beweis:** Dies folgt aus Proposition 3.5, weil beide Seiten isomorph zu $V \otimes W \otimes X$ sind. □

**3.7 Bemerkung.** Mit 3.5 kann man $m$-fache Tensorprodukte immer auf 2-fache Tensorprodukte zurückführen und nach 3.6 darf man beliebig umklammern. Man kann nun 2.8-2.11 auf mehrfache Tensorprodukte verallgemeinern, in dem man entweder analoge Beweise führt oder mit Hilfe von 3.5 und 3.6 auf den Fall von 2 Faktoren zurückführt.

**3.8.** Wir definieren die **k-te Tensorpotenz** $V^{\otimes k} := \underbrace{V \otimes \cdots \otimes V}_{k\text{-mal}}$ für einen $K$-Vektorraum $V$. Konvention: $V^{\otimes 0} := K, V^{\otimes 1} = V$. Aus Bemerkung 3.7 folgt, dass $(V^{\otimes k}) \otimes (V^{\otimes l})$ kanonisch isomorph zu $V^{\otimes (k+l)}$ ist. Wir werden im Folgenden diese Räume identifizieren. Wir definieren die **Tensoralgebra $\mathbf{T(V)}$ über V** als

$$T(V) := \bigoplus_{k=0}^{\infty} V^{\otimes k} = K \oplus V \oplus V^{\otimes 2} \oplus \cdots$$

Wir betrachten dabei die Summanden $K, V, V^{\otimes 2}, \ldots, V^{\otimes k}, \ldots$ als Unterräume des $K$-Vektorraums $T(V)$. Wir setzen $T_k(V) := V^{\otimes k}$. Die Elemente aus $T_k(V)$ heißen **homogene Elemente** von $T(V)$.

**3.9 Definition.** $A$ heißt $K$**-Algebra**, wenn $A$ ein $K$-Vektorraum und gleichzeitig ein Ring so, dass das Ringprodukt $K$-bilinear ist. Dabei soll das $+$ des Ringes gleich dem $+$ des Vektorraums sein. Ein Algebrahomomorphismus ist eine Abbildung zwischen $K$-Algebren, die ein Ringhomomorphismus ist, d.h. verträglich mit $+$ und $\cdot$ der Ringstruktur und die das Einselement auf das Einselement abbildet. Ein Algebraisomorphismus ist ein Algebrahomomorphismus, der eine Umkehrabbildung hat, die auch ein Algebrahomomorphismus ist.

**3.10 Beispiel.** Der Polynomring $K[x_1, \ldots, x_m]$ ist eine $K$-Algebra (siehe VII.1).

**3.11 Proposition.** *Es gibt genau ein Produkt $\cdot$ auf $T(V)$ so, dass der $K$-Vektorraum $T(V)$ zu einer $K$-Algebra wird mit*

$$(v_1 \otimes \cdots \otimes v_k) \cdot (w_1 \otimes \cdots \otimes w_l) = v_1 \otimes \cdots \otimes v_k \otimes w_1 \otimes \cdots \otimes w_l$$

*für alle $v_1, \ldots, v_k, w_1, \ldots, w_l \in V$.*

**Beweis:** Die Abbildung

$$V^{\otimes k} \times V^{\otimes l} \to V^{\otimes (k+l)}, \quad (\omega, v) \mapsto \omega \otimes v$$

ist bilinear aufgrund der Rechenregeln. Wir setzen

$$\left(\sum_k \omega_k\right) \cdot \left(\sum_l v_l\right) := \sum_{k,l} \omega_k \otimes v_l,$$

wobei $\omega_k \in V^{\otimes k}$, $v_l \in V^{\otimes l}$ sein soll und alle Summen endlich sein sollen. Man zeigt leicht, dass dieses Produkt das Gewünschte erfüllt. Weil die reinen Tensoren die Tensoralgebra erzeugen, folgt auch die Eindeutigkeit. □

**3.12 Bemerkung.** Weil das Produkt in Proposition 3.11 das Tensorprodukt $\otimes$ fortsetzt, werden wir es ab jetzt auch mit $\otimes$ bezeichnen.

**3.13 Proposition.** *Die Tensoralgebra $T(V)$ mit der kanonischen Inklusionsabbildung $V \to T(V)$ erfüllt folgende universelle Eigenschaft: Für jede $K$-Algebra $A$ und jede lineare Abbildung $\varphi: V \to A$*

*gibt es genau einen Algebrahomomorphismus $\tilde{\varphi}\colon T(V) \to A$ so, dass folgendes Diagramm kommutiert:*

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d] & A \\ T(V) \arrow[ur, dashed, "\exists!\tilde{\varphi}"'] & \end{tikzcd}$$

**Beweis:** Für $k \in \mathbb{N}$ ist $f : V^k \to A,\ (v_1, \ldots, v_k) \mapsto \varphi(v_1) \cdots \varphi(v_k)$ multilinear. Es existiert also genau ein lineare Abbildung $\tilde{\varphi}_k\colon V^{\otimes k} \to A$ so, dass das folgende Diagramm kommutiert:

$$\begin{tikzcd} V^k \arrow[r, "f"] \arrow[d, "\tau"'] & A \\ V^{\otimes k} \arrow[ur, "\exists!\tilde{\varphi}_k"'] & \end{tikzcd}$$

Mit Hilfe der universellen Eigenschaft der direkten Summe in Proposition 1.7 erhalten wir $\tilde{\varphi}$ aus den $\tilde{\varphi}_k$. Man rechnet nach, dass $\tilde{\varphi}$ das Gewünschte macht. Die Eindeutigkeit folgt daraus, dass die Tensoralgebra $T(V)$ als $K$-Algebra von $V$ erzeugt wird. Das wird im Rest dieses Abschnittes erklärt. □

**3.14.** Sei $A$ eine $K$-Algebra. Dann heißt eine Familie $(a_i)_{i\in I}$ ein *Erzeugendensystem* von $A \Longleftrightarrow \forall a \in A \exists\ r \in \mathbb{N}$, $p \in K[x_1, \ldots, x_r]$ und $a_{i_1}, \ldots, a_{i_r}$ aus $(a_i)_{i\in I}$ mit $a = p(a_{i_1}, \ldots, a_{i_r})$.

**3.15 Beispiel.** Die $K$-Algebra $K[x_1, \ldots, x_n]$ wird erzeugt von $x_1, \ldots, x_n$.

**3.16 Lemma.** *Seien $\varphi_1, \varphi_2 : A \to B$ Algebrahomomorphismen und $(a_i)_{i\in I}$ ein Erzeugendensystem von $A$. Dann gilt:*

$$\varphi_1 = \varphi_2 \Longleftrightarrow \varphi_1(a_i) = \varphi_2(a_i) \forall i \in I.$$

**Beweis:** „$\Longrightarrow$“ trivial.

„$\Longleftarrow$“:

$$\begin{aligned} \varphi_1(a) = \varphi_1\big(p(a_{i_1}, \ldots, a_{i_r})\big) &\overset{\text{Algebrahom.}}{=} p\big(\varphi_1(a_{i_1}), \ldots, \varphi_1(a_{i_r})\big) \\ &= p\big(\varphi_2(a_{i_1}), \ldots, \varphi_2(a_{i_r})\big) = \varphi_2\big(p(a_{i_1}, \ldots, a_{i_r})\big) = \varphi_2(a). \end{aligned}$$ □

**3.17 Proposition.** *$T(V)$ wird von jedem Erzeugendensystem von $V$ als $K$-Algebra erzeugt.*

**Beweis:** Nach Definition lässt sich jedes $\omega \in T(V)$ schreiben als Summe von reinen Tensoren aus gewissen $V^{\otimes k}$. Damit erhält man $\omega$ durch Einsetzen von Vektoren aus $V$ in ein geeignetes Polynom in $K[x_1, \ldots, x_m]$. Aufgrund der Bilinearität des Tensorproduktes gilt das auch für die Vektoren aus dem Erzeugendensystem von $V$. Dies zeigt die Behauptung. □

## 4. Die symmetrische Algebra

Die Tensoralgebra $T(V)$ ist nicht kommutativ, falls $\dim(V) \geq 2$. Wir konstruieren in diesem Abschnitt aus $T(V)$ die symmetrische Algebra $S(V)$, die kommutativ ist. Diese Algebra ist wichtig für die Algebra und die Differentialgeometrie.

In diesem Abschnitt sei $V$ ein Vektorraum über dem Körper $K$.

**4.1 Definition.** Sei $W$ der Unterraum von $T(V)$ erzeugt von

$$\{v_1 \otimes \cdots \otimes v_k - v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(k)} \mid k \in \mathbb{N},\ \sigma \in S_k,\ v_1, \ldots, v_k \in V\}.$$

Für $k \in \mathbb{N}$ sei $W_k$ der Unterraum von $V^{\otimes k}$, der erzeugt wird von

$$\{v_1 \otimes \cdots \otimes v_k - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)} \mid \sigma \in S_k, v_1, \ldots, v_k \in V\}.$$

Wir nennen $S(V) := T(V)/W$ die **symmetrische Algebra über V**. A priori ist $S(V)$ ein $K$-Vektorraum.

**4.2 Lemma.** *Für $\omega, \nu \in T(V)$ gelten:*

*a)* $W = \bigoplus_{k=0}^{\infty} W_k$, $W_0 := \{0\}$, $W_1 = \{0\}$;

*b)* $\omega \otimes \nu - \nu \otimes \omega \in W$;

*c) falls* $\nu \in W \Longrightarrow \omega \otimes \nu \in W$ *und* $\nu \otimes \omega \in W$.

**Beweis:** a) folgt direkt aus der Definition.

Wir zeigen b) zuerst für reine Tensoren $\omega = w_1 \otimes \cdots \otimes w_k$, $\nu = v_1 \otimes \cdots \otimes v_l$. Dann ist $w_1 \otimes \cdots \otimes w_k \otimes v_1 \otimes \cdots \otimes v_l$ bis auf eine Permutation der Faktoren gleich $v_1 \otimes \cdots \otimes v_l \otimes w_1 \otimes \cdots \otimes w_k$. Nach Definition folgt $\omega \otimes \nu - \nu \otimes \omega \in W$. Im allgemeinen sind $\omega$ und $\nu$ endliche Summen von reinen Tensoren und somit folgt b) aus der Bilinearität von $\otimes$.

c): Ab jetzt sei $\nu \in W$. Wir nehmen zuerst an, dass

$$\nu = v_1 \otimes \cdots \otimes v_k - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}, \ \omega = w_1 \otimes \cdots \otimes w_l$$

gilt. Dann folgt

$$\omega \otimes \nu = w_1 \otimes \cdots \otimes w_l \otimes v_1 \otimes \cdots \otimes v_k - w_1 \otimes \cdots \otimes w_l \otimes v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}.$$

Also gilt $\omega \otimes \nu \in W$ nach Definition. Im Allgemeinen sind $\nu$ und $\omega$ endliche Summen von solchen Tensoren und wieder folgt $\omega \otimes \nu \in W$ aus der Bilinearität von $\otimes$. Analog zeigt man $\nu \otimes \omega \in W$. □

**4.3 Proposition.** *Sei $S_k(V)$ das Bild von $V^{\otimes k}$ unter der Quotientenabbildung $\pi\colon T(V) \to T(V)/W = S(V)$. Dann gilt*

$$S(V) = \bigoplus_{k=0}^{\infty} S_k(V) \quad \textit{und} \quad S_k(V) \underset{\textit{kanonisch}}{\cong} V^{\otimes k}/W_k.$$

**Beweis:** Nach 4.2a) gilt $W = \bigoplus_{k=0}^{\infty} W_k$

$$\Longrightarrow S(V) = T(V)/W = \left(\bigoplus_{k=0}^{\infty} V^{\otimes k}\right) \Big/ \left(\bigoplus_{k=0}^{\infty} W_k\right) \underset{\text{kan.}}{\cong} \bigoplus_{k=0}^{\infty} V^{\otimes k}/W_k$$

und bei diesen kanonischen Isomorphismen wird $S_k(V)$ auf $V^{\otimes k}/W_k$ abgebildet. □

**4.4 Proposition.** *Es existiert genau ein Produkt $\cdot$ auf dem $K$-Vektorraum $S(V)$ so, dass $S(V)$ zu einer $K$-Algebra und die Quotientenabbildung $\pi\colon T(V) \to S(V)$ zu einem Algebrahomomorphismus wird.*

**Beweis:** Weil $\pi$ surjektiv ist, muss $\cdot$ eindeutig sein. Wir zeigen nun die Existenz. Für $a, b \in S(V)$ $\exists \omega, \nu \in T(V)$ mit $a = \pi(\omega)$ und $b = \pi(\nu)$. Wir definieren $a \cdot b = \pi(\omega \otimes \nu)$. Wir müssen zeigen, dass diese Definition unabhängig von der Wahl der Urbilder von $a, b$ ist. Wir zeigen zuerst, dass $a \cdot b$ unabhängig von der Wahl von $\omega$ ist. Sei also $\omega' \in T(V)$ mit $\pi(\omega') = a$. Es gilt also $\pi(\omega) = \pi(\omega')$ und damit $\omega' - \omega \in \ker(\pi) = W$. Nach Lemma 4.2 gilt $(\omega' - \omega) \otimes \nu \in W$. Also folgt

$$0 = \pi\big((\omega' - \omega) \otimes \nu\big) = \pi(\omega' \otimes \nu) - \pi(\omega \otimes \nu).$$

Es folgt $\pi(\omega' \otimes \nu) = \pi(\omega \otimes \nu)$ und damit ist die Definition von $a, b$ unabhängig von der Wahl des Urbildes $\omega$. Ähnlich argumentiert man für den zweiten Faktor. Die Algebraaxiome von $T(V)$ vererben sich auf $S(V)$ und damit wird $S(V)$ zu einer Algebra über $K$. Nach Konstruktion ist $\pi$ ein Algebrahomomorphismus. □

**4.5 Proposition.** *$S(V)$ ist eine kommutative $K$-Algebra.*

**Beweis:** Für $a, b \in S(V) \Longrightarrow \exists \omega, \nu \in T(V)$ mit $a = \pi(\omega)$, $b = \pi(\nu)$. Dann gilt

$$\begin{aligned} a \cdot b &= \pi(\omega) \cdot \pi(\nu) = \pi(\omega \otimes \nu) \\ &= \pi(\omega \otimes \nu - \nu \otimes \omega + \nu \otimes \omega) = \pi(\underbrace{\omega \otimes \nu - \nu \otimes \omega}_{\in W \text{ nach 4.2b}}) + \pi(\nu \otimes \omega) \\ &= 0 + \pi(\nu \otimes \omega) = b \cdot a. \end{aligned}$$

□

**4.6 Bemerkung.** Die Abbildung $\iota\colon V \to S(V)$, $v \mapsto \pi(v)$ ist eine injektive lineare Abbildung, da sie $V$ auf $S_1(V) \cong V/W_1 = V$ abbildet nach Proposition 4.3. Oft identifizieren wir deshalb $V$ mit seinem Bild $\iota(V)$. Wegen Proposition XI.3.17 und weil $\pi\colon T(V) \to S(V)$ ein surjektiver Algebrahomomorphismus ist, muss $V$ die Algebra $S(V)$ erzeugen.

**4.7 Proposition.** *Die symmetrische Algebra $S(V)$ zusammen mit $\iota : V \to S(V)$ besitzt folgende universelle Eigenschaft: Wenn $\varphi$ eine lineare Abbildung $V \to A$ ist für eine kommutative $K$-Algebra $A$, dann gibt es genau einen $K$-Algebrahomomorphismus $\varphi'\colon S(V) \to A$ so, dass*

$$\begin{tikzcd} V \arrow[d, "\iota"] \arrow[dr, "\varphi"] & \\ S(V) \arrow[r, "\exists!\varphi'"'] & A \end{tikzcd}$$

*kommutiert.*

**Beweis:** Nach der universellen Eigenschaft der Tensoralgebra $T(V)$ zusammen mit der Inklusionsabbildung $i : V \to T(V)$ (Proposition XI.3.13) $\exists!$ Algebrahomomorphismus $\tilde{\varphi}\colon T(V) \to A$ so, dass das Diagramm

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "i"] & A \\ T(V) \arrow[ur, "\exists!\tilde{\varphi}"'] & \end{tikzcd}$$

kommutiert. Wir behaupten, dass $W \subseteq \ker(\tilde{\varphi})$ gilt: $W$ wird erzeugt von Elementen der Form $v_1 \otimes \cdots \otimes v_k - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}$. Weil $\tilde{\varphi}$ ein Algebrahomomorphismus ist, gilt

$$\tilde{\varphi}(v_1 \otimes \cdots \otimes v_k - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}) = \tilde{\varphi}(v_1) \cdots \tilde{\varphi}(v_k) - \tilde{\varphi}(v_{\sigma(1)}) \cdots \tilde{\varphi}(v_{\sigma(k)}) = 0.$$

Dabei haben wir im letzten Schritt die Kommutativität von $A$ benutzt. Das zeigt unsere Zwischenbehauptung.

Nach dem Homomorphiesatz (Theorem III.3.5) $\exists! \varphi'$ lineare Abbildung so, dass

$$\begin{tikzcd} T(V) \arrow[r, "\tilde{\varphi}"] \arrow[d, "\pi"] & A \\ S(V) \arrow[ur, "\exists!\varphi'"'] & \end{tikzcd}$$

kommutiert. Weil $\tilde{\varphi}$ ein Algebrahomomorphismus ist und $\pi$ ein surjektiver Algebrahomomorphismus ist, muss auch $\varphi'$ ein Algebrahomomorphismus sein. Nach Konstruktion macht $\varphi'$ das Gewünschte. Die Eindeutigkeit von $\varphi'$ ergibt sich daraus, dass die Algebra $S(V)$ von $\iota(V)$ erzeugt wird (siehe 4.6). □

**4.8 Beispiel.** Wir berechnen die symmetrische Algebra des $K$-Vektorraums $V = K$. Wir betrachten die lineare Abbildung $\varphi : K \to K[x]$, $\alpha \to \alpha \cdot x$. Der Polynomring $K[x]$ ist eine kommutative $K$-Algebra. Nach der universellen Eigenschaft von $(S(K), \iota)$ in Proposition 4.7 gibt es genau einen Algebrahomomorphismus $\varphi' \colon S(K) \to K[x]$ so, dass

$$\begin{tikzcd} K \arrow[r, "\varphi"] \arrow[d, "\iota"'] & K[x] \\ S(K) \arrow[ur, "\exists!\varphi'"'] & \end{tikzcd}$$

kommutiert.

Umgekehrt definieren wir eine Abbildung $\psi \colon K[x] \to S(K), p(x) \mapsto p\big(\iota(1)\big)$. Dies ist ein $K$-Algebrahomomorphismus und es gilt $\varphi' \circ \psi = \mathrm{id}_{K[x]}$, $\psi \circ \varphi' = \mathrm{id}_{S(K)}$. Fazit: $S(K)$ ist kanonisch isomorph zu $K[x]$ als $K$-Algebra.

**4.9 Beispiel.** Jetzt beschreiben wir die symmetrische Algebra von $V = K^n$. Wir gehen wie in Beispiel 4.8 vor. Wir betrachten die lineare Abbildung $\varphi : K^n \to K[x_1, \dots, x_n], \alpha \mapsto \alpha_1 x_1 + \dots \alpha_n x_n$. Nach der universellen Eigenschaft von $(S(K^n), \iota)$ in Proposition 4.7 gibt es genau einen Algebrahomomorphismus $\varphi' \colon S(K^n) \to K[x_1, \dots, x_n]$ so, dass

$$\begin{tikzcd} K^n \arrow[r, "\varphi"] \arrow[d, "\iota"'] & K[x_1, \dots, x_n] \\ S(K^n) \arrow[ur, "\exists!\varphi'"'] & \end{tikzcd}$$

kommutiert. Umgekehrt definieren wir $\psi \colon K[x_1, \dots, x_n] \to S(K^n), p(x_1, \dots, x_n) \mapsto p\big(\iota(e_1), \dots, \iota(e_n)\big)$ durch Einsetzen der Standardbasis $e_1, \dots, e_n$ in $p$ bzg. der Algebrastruktur von $S(K^n)$. Dies ist ein $K$-Algebrahomomorphismus und es gilt $\varphi' \circ \psi = \mathrm{id}_{K[x_1, \dots, x_n]}$, $\psi \circ \varphi' = \mathrm{id}_{S(K^n)}$. Fazit: $S(K^n)$ ist kanonisch isomorph zu $K[x_1, \dots, x_n]$ als $K$-Algebra.

**4.10 Korollar.** *Sei $V$ ein $K$-Vektorraum mit $n := \dim(V) < \infty$. Dann gilt $\dim\big(S_l(V)\big) = \binom{n+l-1}{l}$. Falls $v_1, \dots, v_n$ Basis von $V$ ist, dann ist $\{v_1^{k_1} \cdots v_n^{k_n} \mid k_1 + \dots + k_n = l\}$ Basis von $S_l(V)$.*

**Beweis:** $V \xrightarrow{\sim} K^n$, $v_i \mapsto e_i$ . OBdA $V = K^n$. Nach Beispiel 4.9 gilt $S(V) \cong K[x_1, \dots, x_n]$. Aus der Definition von $K[x_1, \dots, x_n]$ folgt, dass die Monome $\{x_1^{k_1} \cdots x_n^{k_n} \mid k_1, \dots, k_n \in \mathbb{N}\}$ eine Basis von $K[x_1, \dots, x_n]$ als $K$-Vektorraum bilden. Wenn wir diese Basis mit Hilfe des kanonischen Isomorphismus übertragen, erhalten wir leicht die zweite Behauptung. Die Länge der Basis ist gleich der Anzahl der Partitionen von $l$ in $n$ Teile. Kombinatorisches Problem: $n + l - 1$ Plätze und wir müssen $n - 1$ auswählen, das ergibt $\binom{n+l-1}{n-1} = \binom{n+l-1}{l}$ Möglichkeiten.
*Beispiel $l = 7$ und $n = 3$. Wir suchen die Anzahl Partitionen $k_1 + k_2 + k_3 = 7$:*
□ ■ □ □ □ □ ■ □ □ $\rightsquigarrow k_1 = 1, k_2 = 4, k_3 = 2$; oder
■ □ □ □ □ ■ □ □ □ $\rightsquigarrow k_1 = 0, k_2 = 4, k_3 = 3$.
Insgesamt gibt es $\binom{9}{2}$ Partitionen, weil man von 9 Plätzen 2 auswählen muss (oben schwarz). □

## 5. Äußere Algebra

Sei $V$ ein $K$-Vektorraum. In diesem Abschnitt werden wir analog zur symmetrischen Algebra die äußere Algebra konstruieren, deren Produkt alternierend (statt kommutativ) ist. Das hat eine Verbindung zu den Determinanten aus Kapitel VI.

**5.1 Definition.** Für $k \geq 2$ sei $U_k$ der Unterraum von $T_k(V) = V^{\otimes k}$ erzeugt von

$$\{v_1 \otimes \cdots \otimes v_k \mid v_i \in V \text{ und } v_j = v_l \text{ für ein Paar } (j,l) \text{ mit } j \neq l\}.$$

Weiter sei $U$ der Unterraum von $T(V)$ erzeugt von $\bigcup_{k\geq 2} U_k$. Wir nennen $\Lambda(V) := T(V)/U$ die **äußere Algebra über V**. $\pi\colon T(V) \to \Lambda(V)$ sei die Quotientenabbildung. Weiter setzen wir $U_0 := U_1 := \{0\}$.

**5.2 Lemma.** *a)* $U = \bigoplus_{k=0}^{\infty} U_k$*;*

*b)* $\omega \in T_k(V),\, \nu \in T_l(V) \Longrightarrow \omega \otimes \nu - (-1)^{kl} \nu \otimes \omega \in U_{k+l}$*;*

*c)* $\omega \in T(V),\, \nu \in U \Longrightarrow \omega \otimes \nu \in U$ *und* $\nu \otimes \omega \in U$.

**Beweis:** Ähnlich wie Lemma 4.2. □

**5.3 Proposition.** *Sei* $\Lambda^k(V) := \pi(V^{\otimes k})$*. Dann gilt*

$$\Lambda(V) = \bigoplus_{k=0}^{\infty} \Lambda^k(V) = K \oplus V \oplus \Lambda^2(V) \oplus \cdots$$

*und* $\Lambda^k(V) \underset{kan.}{\cong} V^{\otimes k}/U_k$.

**Beweis:** Ähnlich wie Proposition 4.3. □

**5.4 Proposition.** *Es existiert genau ein Produkt* $\wedge$ *auf* $\Lambda(V)$ *so, dass die Quotientenabbildung* $\pi\colon T(V) \to \Lambda(V)$ *ein Algebrahomomorphismus ist.*

**Beweis:** Analog zu Proposition 4.4. □

**5.5 Proposition.** *Für* $\omega \in \Lambda^k(V)$ *und* $\nu \in \Lambda^l(V)$ *gilt*

$$\omega \wedge \nu = (-1)^{kl} \nu \wedge \omega.$$

**Beweis:** Dies folgt aus Lemma 5.2b) und Proposition 5.4. □

**5.6 Bemerkung.** Wir haben eine kanonische injektive lineare Abbildung $V \to \Lambda(V),\, v \mapsto \pi(v)$. Wir identifizieren $V$ mit diesem Bild und damit mit dem direkten Summanden $V$ aus $\Lambda(V) = K \oplus V \oplus \Lambda^1(V) \oplus \cdots$. Beachte, dass $v \wedge v = 0$ für alle $v \in V$, da $v \otimes v \in U_2 \subset U$.

**5.7 Proposition (Universelle Eigenschaft).** *Die äußere Algebra* $\Lambda(V)$ *zusammen mit der kanonischen Inklusionsabbildung* $V \to \Lambda(V)$ *aus 5.6 hat folgende universelle Eigenschaft: Sei A eine K-Algebra und* $\varphi\colon V \to A$ *eine lineare Abbildung mit* $\varphi(v) \cdot \varphi(v) = 0$ *für alle* $v \in V$*. Dann existiert genau ein Algebrahomomorphismus* $\varphi'\colon \Lambda(V) \to A$ *so, dass das Diagramm*

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d] & A \\ \Lambda(V) \arrow[ur, "\exists!\varphi'"'] & \end{tikzcd}$$

*kommutiert.*

**Beweis:** Analog zu Proposition 4.7, wobei die obige Eigenschaft von $\varphi$ dazu führt, dass der Algebrahomomorphismus $\tilde{\varphi}\colon T(V) \to A$ durch $\Lambda(V)$ faktorisiert. □

**5.8 Definition.** Eine multilineare Abbildung $f\colon V_1 \times \cdots \times V_m \to W$ von $K$-Vektorräumen heißt **alternierend** $:\Longleftrightarrow f(v_1,\dots,v_m) = 0$ falls $v_i = v_j$ für ein $i \neq j$.

**5.9 Beispiel.** Die Abbildung $\rho\colon V^m \to \Lambda^m(V),\ (v_1,\dots,v_m) \mapsto v_1 \wedge \cdots \wedge v_m$ ist multilinear und alternierend.

**Beweis:** Multilinearität folgt, weil $\rho$ die Verknüpfung ist von der multilinearen Abbildung $\tau\colon V^m \to V^{\otimes m}$ und der linearen Abbildung $\pi\colon V^{\otimes m} \to \Lambda^m(V)$. Man kann die Multilinearität auch aus der Bilinearität des Algebraprodukts $\wedge$ gewinnen. Alternierend folgt sofort aus Proposition 5.5. $\square$

Für $m \in \mathbb{N}$ nennen wir $\Lambda^m(V)$ die **m-te äußere Potenz von V**. Zusammen mit der Abbildung $\rho : V^m \to \Lambda^m(V)$ besitzt sie folgende universelle Eigenschaft:

**5.10 Proposition.** *Für jede alternierende multilineare Abbildung $f\colon V^m \to W$ in einen $K$-Vektorraum $W$ existiert genau eine lineare Abbildung $\varphi\colon \Lambda^m(V) \to W$ mit $f = \varphi \circ \rho$.*

**Beweis:** Nach der universellen Eigenschaft von $(V^{\otimes m}, \tau)$ $\exists!\tilde{\varphi}\colon V^{\otimes m} \to W$ linear so, dass

$$\begin{tikzcd} V^m \arrow[r, "f"] \arrow[d, "\tau"'] & W \\ V^{\otimes m} \arrow[ur, "\exists!\tilde{\varphi}"'] & \end{tikzcd}$$

kommutiert. Weil $f$ alternierend ist, gilt $U_m \subseteq \ker(\tilde{\varphi})$: Um das zu zeigen, wählen wir $v_1, \ldots, v_n \in V$ mit $v_i = v_j$ für ein $i \neq j$. Dann gilt

$$\tilde{\varphi}(v_1 \otimes \cdots \otimes v_m) = \tilde{\varphi}(\tau(v_1, \ldots, v_m)) = f(v_1, \ldots, v_m) = 0.$$

Weil $U_m$ von solchen reinen Tensoren erzeugt wird, folgt $U_m \subseteq \ker(\tilde{\varphi})$.

Nach dem Homomorphiesatz (Theorem III.3.3) $\exists!\varphi'\colon \Lambda^m(V) \to W$ mit $\varphi' \circ \pi = \tilde{\varphi}$. Wegen $\rho = \pi \circ \tau$ macht $\varphi'$ das Gewünschte. Die Eindeutigkeit von $\varphi'$ ergibt sich daraus, dass $\{v_1 \wedge \cdots \wedge v_m \mid v_1, \ldots, v_m \in V\}$ ein Erzeugendensystem ist von $\Lambda^m(V)$. $\square$

**5.11 Proposition.** *Falls $V$ endlich dimensional ist mit Basis $e_1, \ldots, e_n$, dann ist*

$$(e_{i_1} \wedge \cdots \wedge e_{i_m})_{1 \leq i_i < \cdots < i_m \leq n}$$

*eine Basis von $\Lambda^m(V)$. Insbesondere gilt* $\dim\big(\Lambda^m(V)\big) = \binom{n}{m}$.

**Beweis:** Aufgrund der Multilinearität ist $(e_{i_1} \otimes \cdots \otimes e_{i_m})_{1 \leq i_1, \ldots, i_m \leq n}$ ein Erzeugendensystem von $V^{\otimes m}$ (sogar eine Basis). $\Longrightarrow$ Erzeugendensystem von $\Lambda^m(V)$, weil $\pi$ surjektiv. Wenn zwei Indizes gleich sind, gilt $e_{i_1} \wedge \cdots \wedge e_{i_m} = 0$, also können wir die weglassen. Permutationen der Faktoren führen auf Vorzeichenwechsel nach Proposition 5.5 und sind somit überflüssig. Also ist

$$(e_{i_1} \wedge \cdots \wedge e_{i_m})_{1 \leq i_1 < i_2 < \cdots < i_m \leq n}$$

ein Erzeugendensystem von $\Lambda^m(V)$. Wir müssen noch beweisen, dass diese Vektoren linear unabhängig sind. Wir betrachten die Dualbasis $(l_i)_{1 \leq i \leq n}$ in $V^*$ der Basis $e_1, \ldots, e_n$, d.h. $l_i(e_j) = \delta_{ij}$. Sei $1 \leq j_1 < \cdots < j_m \leq n$. Wir betrachten dazu folgende Abbildung:

$$f\colon V^m \to K,\ (v_1, \ldots, v_m) \mapsto \det\Big(\big(l_j(v_i)\big)_{\substack{i=1,\ldots,m \\ j=j_1,\ldots,j_m}}\Big)$$

Aufgrund der Eigenschaften der Determinanten ist diese Abbildung multilinear und alternierend. Nach Proposition 5.10 existiert genau eine lineare Abbildung $\varphi\colon \Lambda^m(V) \to K$ mit $f = \varphi \circ \rho$. Für $1 \leq i_1 < \cdots < i_m \leq n$ gilt

$$\begin{aligned} \varphi(e_{i_1} \wedge \cdots \wedge e_{i_m}) &= \varphi\big(\rho(e_{i_1}, \ldots, e_{i_m})\big) = f(e_{i_1}, \ldots, e_{i_m}) \\ &= \det\Big(\big(l_{j_h}(e_{i_g})\big)_{g,h=1,\ldots,m}\Big) = \begin{cases} 1 & \text{falls } i_1 = j_1, \ldots, i_m = j_m \text{ (Einheitsmatrix)} \\ 0 & \text{sonst (da Nullzeile)} \end{cases} \end{aligned}$$

Weil $\varphi$ linear ist, folgt sofort, dass $(e_{i_1} \wedge \cdots \wedge e_{i_m})_{1 \leq i_1 < \cdots < i_m \leq n}$ linear unabhängig sind. $\square$

**5.12 Bemerkung.** Sei $V = K^n$ und $e_1, \dots, e_n$ die Standardbasis. Für $v_1, \dots, v_n \in K^n$ gilt

$$v_1 \wedge \cdots \wedge v_n = \det(v_1, \dots, v_n) \cdot e_1 \wedge \cdots \wedge e_n.$$

**Beweis:** Schreibe $v_i = \sum_{k=1}^n \alpha_{ki} e_k$ mit Koordinaten $\alpha_{ki} \in K$ und rechne die linke Seite aus. Wenn man benutzt, dass $\wedge$ alternierend ist, folgt leicht die Behauptung mit Hilfe der Definition der Determinante (siehe Definition VI.2.9). $\square$

**5.13 Proposition.** *Sei $\varphi \colon V \to W$ eine lineare Abbildung von $K$-Vektorräumen und $m \in \mathbb{N}$. Dann $\exists!$ lineare Abbildung $\Lambda^m(\varphi) \colon \Lambda^m(V) \to \Lambda^m(W)$ mit $\Lambda^m(\varphi)(v_1 \wedge \cdots \wedge v_m) = \varphi(v_1) \wedge \cdots \wedge \varphi(v_m)$, $\forall v_1, \dots, v_m \in V$.*

**Beweis:** Betrachte $f \colon V^m \to \Lambda^m(W)$, $(v_1, \dots, v_m) \mapsto \varphi(v_1) \wedge \cdots \wedge \varphi(v_m)$. Dies ist eine multilineare Abbildung, weil $\varphi$ linear und $\rho$ multilinear ist. Weiter ist $f$ alternierend, weil $\rho$ alternierend ist (Beispiel 5.9). Nach der universellen Eigenschaft in Proposition 5.10 $\exists!$ lineare Abbildung $\Lambda^m(\varphi)$ so, dass das Diagramm

$$\begin{tikzcd} V^m \arrow[r, "f"] \arrow[d, "\rho"'] & \Lambda^m(W) \\ \Lambda^m(V) \arrow[ur, "\exists! \Lambda^m(\varphi)"'] & \end{tikzcd}$$

kommutiert. Nach Konstruktion erfüllt $\Lambda^m(\varphi)$ das Gewünschte. Die Eindeutigkeit von $\Lambda^m(\varphi)$ ergibt sich daraus, dass $\{v_1 \wedge \cdots \wedge v_m \mid v_1, \dots, v_m \in V\}$ ein Erzeugendensystem von $\Lambda^m(V)$ ist. $\square$

**5.14 Bemerkung.** Nach Proposition 5.3 und der universellen Eigenschaft der direkten Summe in Proposition 1.7 setzen sich $\Lambda^m(\varphi)$ zusammen zu einer linearen Abbildung $\Lambda(\varphi) \colon \Lambda(V) \to \Lambda(W)$. Man zeigt leicht, dass $\Lambda(\varphi)$ sogar ein Algebrahomomorphismus ist.

**5.15 Bemerkung.** Sei $\varphi \in \mathrm{End}(V)$ mit $n := \dim(V) < \infty$. Nach Proposition 5.13 erhalten wir einen linearen Endomorphismus $\Lambda^n(\varphi)$ von $\Lambda^n(V)$. Nach Proposition 5.11 ist $\dim\big(\Lambda^n(V)\big) = 1$. Also muss $\Lambda^n(\varphi) = \lambda \cdot \mathrm{id}_{\Lambda^n(V)}$ gelten für ein $\lambda \in K$. *Behauptung:* $\Lambda^n(\varphi) = \det(\varphi) \cdot \mathrm{id}_{\Lambda^n(V)}$

**Beweis:** Wir wählen eine Basis $v_1, \dots, v_n$ von $V$. Nach Proposition 5.11 ist $v_1 \wedge \ldots \wedge v_n$ eine Basis von $\Lambda^n(V)$. Es gilt

$$\Lambda^n(\varphi)(v_1 \wedge \ldots \wedge v_n) == \varphi(v_1) \wedge \cdots \wedge \varphi(v_n) = \det(\varphi) \cdot v_1 \wedge \ldots \wedge v_n,$$

wobei man für den letzten Schritt eine analoge Rechnung wie in Bemerkung 5.12 macht. Dies zeigt die Behauptung. $\square$

# Index

# Literaturverzeichnis


[Bo] S. Bosch: Lineare Algebra

[Fi] G. Fischer: Lineare Algebra

[Gr] Greub: Linear Algebra

[Ha] J. Hausen: Lineare Algebra I

[Jä] Jänich: Lineare Algebra

[KoMi] Kowalski-Michler: Lineare Algebra

[Gu] W. Gubler: Algebra und Zahlentheorie (2007); Skript Zahlentheorie und Algebra

[JN] P. Jordan, J. von Neumann: On inner products in linear, metric spaces. Ann. of Math. 36 (1935) [719 - 723]

[Wa] R. Walter: Lineare Algebra und Analytische Geometrie, Vieweg Verlag