Vorlesungsmitschrift

Lineare Algebra I+II

Prof. Franke

Wintersemester 1995/1996
Sommersemester 1996

Inhaltsverzeichnis

1 Naive Mengenlehre **1**
- 1.1 Mengen . . . 1
- 1.2 Abbildungen . . . 3

2 Vektorräume über einem Körper **6**
- 2.1 Gruppen und Körper . . . 6
- 2.2 Vektorräume . . . 11
- 2.3 Unterobjekte . . . 12
- 2.4 Erzeugendensysteme, lineare Unabhängigkeit und Basen . . . 13
- 2.5 Lineare Abbildungen . . . 18

3 Matrizen **26**
- 3.1 Matrizenrechnung . . . 26
- 3.2 Das Eliminationsverfahren von Gauß . . . 33

4 Determinanten und multilineare Algebra **39**
- 4.1 Einige Fakten über Permutationen . . . 39
- 4.2 Charakteristik eines Körpers . . . 42
- 4.3 Schiefsymmetrische n-Linearformen . . . 45
- 4.4 Der Determinantenbegriff . . . 49
- 4.5 Determinanten von Matrizen . . . 52
- 4.6 Orientierungen von reellen Vektorräumen . . . 58
- 4.7 Symmetrische n-Linearformen . . . 61

5 Quadratische, hermitesche und symplektische Formen **62**
- 5.1 Quadratische und symmetrische Bilinearformen . . . 62
- 5.2 Nicht ausgeartete Bilinearformen. Der Rang. . . . 66
- 5.3 Klassifikation der quadratischen Formen . . . 68
- 5.4 Quadratische Formen auf reellen Vektorräumen . . . 70
- 5.5 Die Gramsche Determinante, die Gruppen $\mathrm{O}(\beta)$ und $\mathrm{SO}(\beta)$ . . . 73
- 5.6 Euklidische Vektorräume . . . 75
  - 5.6.1 Das Vektorprodukt auf einem dreidimensionalen orientierten Euklidischen Raum . . . 80
- 5.7 Hermitesche Formen . . . 83
- 5.8 Symplektische Vektorräume . . . 88

6 Eigenwerttheorie **91**
- 6.1 Eigenwerttheorie für normale und selbstadjungierte Endomorphismen eines endlichdimensionalen Hilbertraums . . . 93
- 6.2 Eigenwerttheorie für normale und selbstadjungierte Endomorphismen eines Euklidischen Vektorraumes . . . 99
- 6.3 Das charakteristische Polynom . . . 105
- 6.4 Die Jordansche Normalform . . . 112

7 Affine Räume und Faktorräume **118**
- 7.1 Affine Räume . . . 118
- 7.2 Der Faktorraum V/W nach einem Unterraum . . . 121

8 Tensorprodukte **123**

1 Naive Mengenlehre

1.1 Mengen

Eine Menge M ist eine Zusammenfassung von Objekten (die Elemente von M) unseres Denkens zu einem neuen Objekt. Beispiel: $M = \{0; 5; 163\}$, $N = \{130, 0\}$, $L = \{3; 10; \{4; 5\}\}$. Man kann auch eine Serie von Elementen weglassen und durch drei Punkte andeuten, wenn das Gesetz der Serie klar ist.

Beispiel: $\mathbb{N} = \{0; 1; 2; 3; \ldots\}$ die Menge der natürlichen Zahlen
$\mathbb{Z} = \{0; 1; -1; 2; -2; 3; \ldots\}$ die Menge der ganzen Zahlen
$M = \{0; 1; 2; \ldots; 1000\}$

Wenn P eine in der mathematischen Sprache ausdrückbare Eigenschaft ist, so kann man (meistens) die Menge aller Objekte x bilden, die diese Eigenschaft haben. Beispiel: $M = \{x \mid x$ hat die Eigenschaft $P\}$. Ein Objekt x gehört genau dann zu dieser Menge, wenn es die Eigenschaft P hat.
Wenn x ein Element der Menge M ist, so schreiben wir $x \in M$. Wir sagen auch: x gehört zu M. Andernfalls schreiben wir $x \notin M$. Beispiel: $M = \{x \mid x$ ist eine gerade natürliche Zahl$\} = \{0; 2; 4; \ldots\}$ $x \in M$ genau dann wenn x eine gerade natürliche Zahl ist. $N = \{x \mid x$ ist eine Primzahl$\} = \{2; 3; 5; 7; 9; \ldots\}$ $R = \{x \mid x$ ist eine Menge und $x \notin x\}$. Wenn R existieren würde, gäbe es einen logischen Widerspruch. Für alle x gilt: $x \in R$ genau dann, wenn $x \notin x$. Setze $x = R$, $R \in R$ genau dann, wenn $R \notin R$ (Russelsche Antinomie).
Die leere Menge $\emptyset = \{\}$. Für alle x gilt: $x \notin \emptyset$. Generell gilt für zwei Mengen M und N: $M = N$ genau dann, wenn für alle x die Eigenschaften $x \in M$ und $x \in N$ äquivalent sind.

Definition 1:

1. Eine Menge N ist genau dann **Teilmenge** einer Menge M, wenn für alle $x \in N$ gilt: $x \in M$. Wenn das der Fall ist, so schreiben wir $N \subseteq M$ oder $M \supseteq N$ und nennen M auch **Obermenge** von N.

2. N ist eine **echte Teilmenge** von M, wenn $N \subseteq M$ und $N \neq M$ gilt. Wenn dies der Fall ist, so schreiben wir $N \subset M$ oder $M \supset N$. Es muß also $N \subseteq M$ gelten, und es gibt ein $x \in M$ für das gilt $x \notin N$.

Fakt 1:

1. Für alle Mengen M gilt: $\emptyset \subseteq M$,$M \subseteq M$
2. Für keine Menge M gilt: $M \subset M$
3. Aus $A \subseteq B$ und $B \subseteq C$ folgt $A \subseteq C$.
4. Unter der Voraussetzung von (3) gilt auch: Aus $A \subset B$ und $B \subset C$ folgt $A \subset C$. Aus $A \subseteq B$ und $B \subset C$ folgt $A \subset C$.
5. Aus $M \subseteq N$ und $N \subseteq M$ folgt $M = N$.

Beweis:

3. Sei $A \subseteq B$ und $B \subseteq C$, aus $x \in A$ folgt $x \in B$, daraus folgt $x \in C$. Also gilt $A \subseteq C$.

4. Fall 1: Sei $A \subset B$ und $B \subseteq C$. Dann gibt es ein $x \in B$ mit $x \notin A$. Es gilt wegen $B \subseteq C$ auch $x \in C$. Also ist $A \neq C$. Wegen $A \subseteq C$ gilt $A \subset C$.
   Fall 2: Sei $A \subseteq B$ und $B \subset C$. Dann gibt es ein $x \in C$ und $x \notin B$. Wegen $A \subseteq B$ folgt $x \notin A$. Also ist $A \neq C$ und wegen $A \subseteq C$ gilt $A \subset C$.

5. Aus den Voraussetzungen ergeben sich folgende Implikationen: Aus $x \in M$ folgt $x \in N$ und aus $x \in N$ folgt $x \in M$. Also ist $x \in M$ zu $x \in N$ äquivalent. Folglich enthalten M und N dieselben Elemente, und es gilt: $M = N$. □

Definition 2: Seien M,N Mengen.

1. Der **Durchschnitt** $M \cap N$ von M und N ist die Menge
   $M \cap N = \{x \mid x \in M \wedge x \in N\}$
2. Die **Vereinigung** $M \cup N$ von M und N ist die Menge
   $M \cup N = \{x \mid x \in M \vee x \in N\}$
3. Die **Differenz** $M \setminus N$ (auch $M - N$) von M und N ist die Menge
   $M \setminus N = \{x \mid x \in M \wedge x \notin N\}$

Fakt 2:

1. $M \cap N \subseteq M \subseteq M \cup N$
   $M \cap N \subseteq N \subseteq M \cup N$
2. $M \cap N = N \cap M$; $M \cup N = N \cup M$
   $(L \cap M) \cap N = L \cap (M \cap N)$; $(L \cup M) \cup N = L \cup (M \cup N)$ Aber es gilt im Allgemeinen: $M \setminus N \neq N \setminus M$.
3. $(M \setminus N) \subseteq M$
4. $(M \cup N) \cap L = (M \cap L) \cup (N \cap L)$
   $(M \cap N) \cup L = (M \cup L) \cap (N \cup L)$
5. $(M \setminus N) \cap L = (M \cap L) \setminus N = (M \cap L) \setminus (N \cap L)$
6. $(M \setminus A) \cup (M \setminus B) = M \setminus (A \cap B)$
   $(M \setminus A) \cap (M \setminus B) = M \setminus (A \cup B)$
7. $(M \setminus N) \cap N = \emptyset$
8. Es gilt, daß $M \setminus N = \emptyset$ genau dann, wenn $M \subseteq N$.
9. Es gilt, daß $M \cup N = \emptyset$ genau dann, wenn $M = \emptyset \wedge N = \emptyset$.

Beweis:

4. "⇒": Sei $x \in (M \cup N) \cap L$. Dann ist $x \in L$ und $x \in M \cup N$. Dann lassen sich die Fälle $x \in L \wedge x \in M$ und $x \in L \wedge x \in N$ unterscheiden. Es folgt $x \in L \cap M$ bzw. $x \in L \cap N$. Also gilt in beiden Fällen $x \in (M \cap L) \cup (N \cap L)$. Ferner gilt $(M \cup N) \cap L \subseteq (M \cap L) \cup (N \cap L)$.
   "⇐": Sei $x \in (M \cap L) \cup (N \cap L)$. Es lassen sich die Fälle $x \in M \wedge x \in L$ und $x \in N \wedge x \in L$ unterscheiden. In beiden Fällen gilt $x \in M \cup N \wedge x \in L$, damit folgt $x \in (M \cup N) \cap L$. Ferner gilt $(M \cap L) \cup (N \cap L) \subseteq (M \cup N) \cap L$. Nach Fakt 1 folgt $(M \cup N) \cap L = (M \cap L) \cup (N \cap L)$.
   □

Definition 3: Zwei Mengen M und N sind **disjunkt**, falls $M \cap N = \emptyset$. Es existiert kein gemeinsames Element.

Definition 4: Ein **geordnetes Paar** (a, b) ist die geordnete Aufzählung der beiden Objekte a und b. Das **Mengenprodukt** $M \times N$ ist als

$$M \times N = \{(m, n) \mid m \in M, n \in N\}$$

definiert.

Beispiel: Sei $\mathbb{R}$ die Menge aller reeller Zahlen, $\mathbb{R} \times \mathbb{R} = \{(a, b) \mid a, b \in \mathbb{R}\}$. Die geordneten Paare (a, b) mit $a, b \in \mathbb{R}$ lassen sich mit Punkten in der Ebene identifizieren. Also kann $\mathbb{R} \times \mathbb{R}$ mit der Menge der Punkte der euklidischen Ebene identifiziert werden. Wenn $I = [0; 1] = \{t \in \mathbb{R} \mid 0 \leq t \leq 1\} \subseteq \mathbb{R}$, dann ist $I \times I$ ein Quadrat.

Fakt 3:

1. Es gilt meistens, daß $M \times N \neq N \times M$.
2. Es gilt $(a, b) = (c, d)$ genau dann, wenn $a = c$ und $b = d$.
3. Aus $\tilde{M} \subseteq M$ und $\tilde{N} \subseteq N$ folgt $\tilde{M} \times \tilde{N} \subseteq M \times N$
4. Es gilt

$$\begin{aligned} (A \cap B) \times C &= (A \times C) \cap (B \times C) \\ (A \cup B) \times C &= (A \times C) \cup (B \times C) \\ (A \setminus B) \times C &= (A \times C) \setminus (B \times C) \end{aligned}$$

5. $M \times \emptyset = \emptyset$

Beweis der ersten Gleichung von (4): Sei $x \in (A \cap B) \times C$ und $x = (\alpha, \beta)$ mit $\alpha \in A \cap B$ und $\beta \in C$. Daraus folgt $x \in (A \times C) \cap (B \times C)$. □

Definition 5: Ein n-**Tupel** ist eine geordnete Aufzählung $(x_1, \ldots, x_n)$. Sei

$$X_1 \times \ldots \times X_n = \{(x_1, \ldots, x_n) \mid x_i \in X_i \text{ für } 1 \leq i \leq n\}.$$

Für eine Menge X sei $X^n = X \times \ldots \times X$ (n Faktoren).

Bemerkung: Für $n = 1$ wollen wir (x) mit x identifizieren, also $X^1 = X$. 2-Tupel sind Paare, 3-Tupel sind Tripel.

1.2 Abbildungen

Definition 1: Eine **Abbildung** f von einer Menge M in eine Menge N ist eine Zuordnung, welche jedem Element $m \in M$ ein Element $n \in N$ zuordnet. Schreibweise: $f : M \to N$; $M \xrightarrow{f} N$, wenn f eine Abbildung von M nach N ist. Sei N^M die Menge aller Abbildungen von M nach N.

Beispiel:

1. Sei X eine Menge, durch $\mathrm{Id}_X(x) = x$ für $x \in X$ ist eine Abbildung $X \xrightarrow{\mathrm{Id}_X} X$ definiert, die **identische Abbildung**.

2. Seien X, Y Mengen und $y_0 \in Y$. Durch $f(x) = y_0$ für alle $x \in X$ ist eine Abbildung $f : X \to Y$ definiert. Eine solche Abbildung nennt man **konstant**.

3. Abbildungen von $\mathbf{R}$ in $\mathbf{R}$ (oder allgemeiner Abbildungen von $\mathbf{R}^n$ in $\mathbf{R}$) nennt man auch **Funktionen**. Beispiel: $f(x) = x^2$; $f(x) = 2x$; $f(x) = \frac{x}{1+x^2}$ etc. Ein anderes Beispiel ist die Funktion

$$f(x) = \begin{cases} 0 & x \in \mathbb{Q} \subset \mathbf{R} \\ 1 & x \text{ irrational} \end{cases} \quad \text{oder } f(x) = \begin{cases} \sin x & x \text{ rational} \\ e^x & x \text{ irrational} \end{cases} .$$

   Ein Beispiel für eine Funktion in zwei Variablen ist die Addition:

$$\begin{array}{rcl} + : \mathbf{R}^2 & \to & \mathbf{R} \\ (x, y) & \mapsto & x + y \end{array}$$

4. Seien M und N Mengen, durch

$$\begin{array}{rcll} \mathrm{pr}_1 : M \times N & \to & M & \\ \mathrm{pr}_1(m, n) := \mathrm{pr}_1((m, n)) & \mapsto & m & m \in M;\ n \in N \end{array}$$

   sind Abbildungen von $M \times N$ nach M bzw. N definiert, die Projektion auf den ersten bzw. zweiten Faktor. Für $M = N = \mathbf{R}$ wird ein Punkt in der Ebene durch pr_1 bzw. pr_2 auf seine x bzw. y-Koordinate abgebildet.

5. Sei M eine Menge, dann enthält $M^\emptyset$ stets genau ein Element, die leere Abbildung. Dagegen ist für $M \neq \emptyset$ die Menge $\emptyset^M$ leer. Für $M = \emptyset$ besteht $\emptyset^\emptyset$ aus genau einem Element, nämlich $\mathrm{Id}_\emptyset$.

6. Seien A, B endliche Mengen mit a, b Elementen. Dann enthält die Produktmenge $A \times B$ genau $a \cdot b$ Elemente, und die Menge aller Abbildungen B^A enthält b^a Elemente.

7. Für $\tilde{N} \subseteq N$ gilt: $\tilde{N}^M \subseteq N^M$. Wenn $\tilde{M} \subseteq M$, $f : M \to N$, so ist $f|_{\tilde{M}} : \tilde{M} \to N$ durch $f|_{\tilde{M}}(m) = f(m)$ für $m \in \tilde{M}$ definiert, die **Einschränkung** von f auf $\tilde{M}$.

Definition 2: Sei $f : M \longrightarrow N$. Für Untermengen $X \subseteq M$ und $Y \subseteq N$ sei

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

das **Bild** von X unter f und

$$f^{-1}(Y) := \{m \in M \mid f(m) \in Y\}$$

nennt man das **Urbild** oder **inverse Bild** von Y bezüglich f.

Fakt 1: Mit den Bezeichnungen und unter den Voraussetzungen von Definition 2 gilt: $f^{-1}(N) = M$; $f(\emptyset) = \emptyset$; $f^{-1}(\emptyset) = \emptyset$; $f(X) \subseteq Y \Leftrightarrow X \subseteq f^{-1}(Y)$; $X \subseteq f^{-1}(f(X))$, $f(f^{-1}(Y)) \subseteq Y$.Im Allgemeinen gilt nicht: $f(M) = N$. □

Beispiel: Seinen A, B Mengen und $X \subseteq A$. Dann gilt $\mathrm{pr}_1^{-1}(X) = X \times B$, wobei $\mathrm{pr}_1 : A \times B \to A$ die Projektion auf den ersten Faktor ist.

Definition 3: Eine Abbildung $M \xrightarrow{f} N$ ist

1. **injektiv**, wenn für jedes $n \in N$ $f^{-1}(\{n\})$ höchstens ein Element enthält.
2. **surjektiv**, wenn für jedes $n \in N$ $f^{-1}(\{n\})$ mindestens ein Element enthält.
3. **bijektiv**, wenn f sowohl injektiv als auch surjektiv ist.

Definition 4: Seien $A \xrightarrow{f} B \xrightarrow{g} C$ Abbildungen. Die **Hintereinanderausführung** $g \circ f : A \to C$ ist durch $(g \circ f)(a) := g(f(a))$ definiert. Eine Abbildung $B \xrightarrow{g} A$ ist ein **Linksinverses** zu f, wenn $g \circ f = \mathrm{Id}_A$, ein **Rechtsinverses** zu f, wenn $f \circ g = \mathrm{Id}_B$. Ein Inverses zu f ist eine Abbildung $g : B \to A$, welche sowohl linksinvers als auch rechtsinvers zu f ist.

Lemma: Sei $f : A \to B$ eine Abbildung.

1. f ist genau dann bijektiv, wenn es ein Inverses zu f gibt. In diesem Fall ist die zu f inverse Abbildung eindeutig bestimmt, sie wird mit f^{-1} bezeichnet.
2. Sei $A \neq \emptyset$. Dann ist f genau dann injektiv, wenn ein Linksinverses zu f existiert.
3. f ist genau dann surjektiv, wenn ein Rechtsinverses zu f existiert.

Das Auswahlaxiom: Seien M und N Mengen. Sei 2^M die Menge aller Untermengen von M. Sei $G : N \to 2^M$ eine Abbildung, so daß $G(n) \neq \emptyset$ für alle $n \in N$ gilt. Dann gibt es eine Funktion $g : N \to M$ mit $g(n) \in G(n)$ für alle $n \in N$.
Äquivalente Formulierung: Sei $M \neq \emptyset$. Dann gibt es eine Abbildung $a : 2^M \setminus \{\emptyset\} \to M$ mit der Eigenschaft: $a(X) \in X$ für $X \subseteq M$, $X \neq \emptyset$.

Beweis des Lemmas:

1. Sei f bijektiv. Für $n \in B$ sei $f^{-1}(n)$ das einzige Element von $f^{-1}(\{n\})$. Man prüft leicht, daß f^{-1} zu f invers ist und auch das einzige Inverse zu f ist.
   Sei g Inverses zu f. Dann gilt $g(n) \in f^{-1}(\{n\})$. Umgekehrt folgt aus $m \in f^{-1}(\{n\})$ $m = g(f(m)) = g(n)$. Also gilt $f^{-1}(\{n\}) = \{g(n)\}$, und die Bijektivität von f ist gezeigt.
2. Sei g linksinvers zu f. Aus $n \in f^{-1}(\{m\})$ folgt $n = g(f(n))$ (g ist linksinvers). Also gilt: $f^{-1}(\{m\}) \subseteq \{g(m)\}$, damit ist f injektiv. Sei $A \neq \emptyset$, $a_0 \in A$. Wenn f injektiv ist, so ist durch
   $$g(b) = \begin{cases} \text{das einzige Element von } f^{-1}(\{b\}), \text{ falls } f^{-1}(\{b\}) \neq \emptyset \\ a_0, \text{ falls } f^{-1}(\{b\}) = \emptyset \end{cases}$$
   ein Linksinverses zu f definiert.
3. "$\Rightarrow$": Sei g ein Rechtsinverses zu f. Für alle $b \in B$ gilt: $b = f(g(b))$ mit $g(b) \in f^{-1}(\{b\})$, also $f^{-1}(\{n\}) \neq \emptyset$. Also ist f surjektiv.
   "$\Leftarrow$": Sei $f : A \to B$ surjektiv. Sei $G : B \to 2^A$ gegeben durch durch $G(b) = f^{-1}(\{b\})$. Wegen der Surjektivität von f gilt: $G(b) \neq \emptyset$ für alle $b \in B$. Nach dem Auswahlaxiom gibt es eine Abbildung $g : B \to A$ mit $g(b) \in G(b)$ für alle $b \in B$. Also ist $f(g(b)) = b$. Damit ist g ein Rechtsinverses zu f. □

Bemerkung: Es gilt für $f : M \to N$, $X \subseteq M$, $Y \subseteq N$:

1. Bild von X: $f(X) = \{f(m) \mid m \in X\} \subseteq N$;
2. Urbild von Y: $f^{-1}(Y) = \{m \in M \mid f(m) \in Y\} \subseteq M$;
3. Wenn f bijektiv ist, bezeichnet man f^{-1} auch als das **Inverse** zu f. In diesem Fall stimmen das Urbild von Y unter f und das Bild von Y unter f^{-1} überein:

$$f^{-1}(Y) := \{m \in M \mid f(m) \in Y\} = \{f^{-1}(n) \mid n \in Y\} =: f^{-1}(Y).$$

   Insbesondere gilt für $Y = \{n\}$: $f^{-1}(\{n\}) = \{f^{-1}(n)\}$.

Beispiel: Sei $1 \leq k \leq m$, dann gibt es folgende Bijektion:

$$\begin{array}{rcl} (M_1 \times M_2 \times \ldots \times M_k) \times (M_{k+1} \times \ldots \times M_m) & \to & M_1 \times \ldots \times M_m \\ ((m_1, \ldots, m_k), (m_{k+1}, \ldots, m_m)) & \mapsto & (m_1, \ldots, m_m) \text{ mit } m_i \in M_i,\ 1 \leq i \leq m \end{array}$$

Die inverse Abbildung dazu ist gegeben durch:

$$(m_1, \ldots, m_m) \mapsto ((m_1, \ldots, m_k), (m_{k+1}, \ldots, m_m))$$

Insbesondere gibt es Bijektionen

$$A \times (B \times C) \overset{\alpha}{\tilde{\to}} A \times B \times C \overset{\beta}{\tilde{\leftarrow}} (A \times B) \times C,$$

außerdem gilt $A \times (B \times C) \overset{\beta^{-1} \circ \alpha}{\tilde{\to}} (A \times B) \times C$. Bis auf Bijektion ist $\times$ auch kommutativ:

$$\begin{array}{rcl} A \times B & \overset{\sigma_{AB}}{\tilde{\to}} & B \times A \\ \sigma_{AB}((a,b)) & = & (b,a) \end{array}$$

σ_{AB} ist normalerweise nicht die Identität, selbst wenn $A = B$.

2 Vektorräume über einem Körper

2.1 Gruppen und Körper

Definition 1: Eine Gruppe ist ein Paar $(G, \circ)$, wobei G eine Menge und $\circ : G \times G \to G$ eine Abbildung mit $(g, h) \mapsto (k \circ h)$ ist, wobei die folgenden Bedingungen erfüllt sein müssen:

1. **Assoziativgesetz:** Für $a, b, c \in G$ gilt: $a \circ (b \circ c) = (a \circ b) \circ c$;
2. **Neutrales Element:** Es gibt ein Element $e \in G$ (die Rechtseinheit, das rechtsneutrale Element, Einselement) von G mit der Eigenschaft: $a \circ e = a$ für alle $a \in G$;
3. **Inverses Element:** Für alle $a \in G$ gibt es ein $b \in G$ (das rechtsinverse Element) mit $a \circ b = e$.

Beispiele:

1. Sei $\mathbb{R}^* = \mathbb{R} - \{0\}$ und sei $\circ$ die übliche Multiplikation reeller Zahlen. Dann ist $(\mathbb{R}^*, \circ)$ eine Gruppe, ebenso $(\mathbb{Q}^*, \circ)$.

2. Sei X eine Menge und sei $S(X)$ die Menge aller bijektiven Abbildungen von X auf sich selbst. Sei $\circ$ gegeben durch die Hintereinanderausführung $(f \circ g)(x) = f(g(x))$. Dann ist $(S(X), \circ)$ eine Gruppe.

   Assoziativität: $((f \circ g) \circ h)(x) = (f \circ g) \circ (h(x)) = f(g(h(x))) = f((g \circ h)(x)) = (f \circ (g \circ h))(x)$. Neutrales Element ist $e = \mathrm{Id}_X$ und ein inverses Element zu f ist durch f^{-1} gegeben.

3. Die Gruppe $(S(\{1, \ldots, n\}), \circ)$ wird auch S_n genannt, ihre Elemente sind die **Permutationen** der Zahlen $1 \ldots n$. Die Gruppen S_n nennt man deshalb auch Permutationsgruppen.

Satz 1: Sei G eine Gruppe mit dem rechtsneutralen Element e.

1. Ist b ein Rechtsinverses von a, so gilt auch $b \circ a = e$, also ist b auch Linksinverses.
2. e ist auch linksneutrales Element: $e \circ a = e$ für alle $a \in G$.
3. Ein Inverses für $a \in G$ ist eindeutig bestimmt. Wir bezeichnen es mit a^{-1}.
4. Seien $a, f \in G$ mit $a \circ f = a$. Dann folgt $f = e$, selbiges folgt auch aus $f \circ a = a$.
5. Für $x, y \in G$ ist die Gleichung $x \circ z = y$ eindeutig nach z auflösbar , nämlich $z = x^{-1} \circ y$, ebenso ist $z \circ x = y$ eindeutig nach z auflösbar, nämlich $z = y \circ x^{-1}$.

Beweis:

1. Es gilt $bab = be = b$, also ist $ba = bae = babc = bec = bc = e$ mit $c = b^{-1}$.
2. Sei b ein Rechtsinverses zu a. Wegen (1) ist b auch linksinvers zu a. Es folgt: $ea = aba = ae = a$.
3. Sei $a \in G$ und seinen $b, \tilde{b}$ invers zu a. Es folgt: $\tilde{b} = e\tilde{b} = ba\tilde{b} = be = b$.
4. Sei $af = a$. Es folgt: $f = ef = a^{-1}af = a^{-1}a = e$. Den anderen Teil zeigt man analog: $f = fe = faa^{-1} = aa^{-1} = e$.
5. Zunächst wird der erste Teil bewiesen: Aus $z = x^{-1}y$ folgt $xz = xx^{-1}y = ey = y$. Damit ist $x^{-1}y$ eine Lösung der betrachteten Gleichung. Aus $xz = y$ folgt $z = ez = x^{-1}xz = x^{-1}y$. Also ist $x^{-1}y$ die einzige Lösung der Gleichung. Der zweite Teil kann analog bewiesen werden. □

Bemerkung: Für $a \in G$ und $n \in \mathbb{Z}$ sei

$$a^n = \begin{cases} \underbrace{a \cdot a \cdot \ldots \cdot a}_{n \text{ Faktoren}} & n > 0 \\ 1 & n = 0 \\ (a^n)^{-1} & n < 0 \end{cases}$$

Es gelten die Regeln $(a^n)^m = a^{nm}$ und $a^n a^m = a^{n+m}$ für $n, m \in \mathbb{Z}$. Der Beweis ist einfach, zum Beispiel für $n > 0, m > 0$ gilt

$$a^n a^m = (\underbrace{a \cdot a \cdot \ldots \cdot a}_{n \text{ Faktoren}})(\underbrace{a \cdot a \cdot \ldots \cdot a}_{m \text{ Faktoren}}) = \underbrace{a \cdot a \cdot \ldots \cdot a}_{n+m \text{ Faktoren}} = a^{n+m}.$$

Für viele Gruppen gilt nicht $(ab)^m = a^m b^m$, denn Kommutativität kann nicht vorausgesetzt werden.

Definition: Eine Gruppe $(G, \circ)$ ist **kommutativ** oder **abelsch**, falls für alle $a, b \in G$ gilt: $a \circ b = b \circ a$.

Beispiele:

1. $(\mathbb{Q}^*, \cdot)$ und $(\mathbb{R}^*, \cdot)$ sind kommutativ.
2. Wenn M wenigstens drei Elemente enthält, so ist $S(M)$ nicht kommutativ.
3. Die Gruppen der Bewegungen in der euklidischen Ebene bzw. im euklidischen Raum sind nicht kommutativ.

Für abelsche Gruppen gilt die Rechenregel $a^n b^n = (ab)^n$.

Bemerkung: Eine additiv geschriebene Gruppe ist im Allgemeinen kommutativ und wir definieren:

$$a \cdot n = \begin{cases} \underbrace{a + a + \ldots + a}_{n \text{ Summanden}} & n > 0 \\ 0 & n = 0 \\ (-a) \cdot (-n) & n < 0 \end{cases}$$

Es gilt dann $n(mx) = (nm)x$, $nx + mx = (n+m)x$, $n(x+y) = nx + ny$ (G ist abelsch).

Beispiel: $(\mathbb{Z}, +)$, $(\mathbb{R}, +)$, $(\mathbb{Q}, +)$ sind abelsche Gruppen.

Definition: Ein **Körper** ist ein Tripel $(K, +, \cdot)$ bestehend aus einer Menge K und zwei Abbildungen

$$\begin{array}{rcl} + : K \times K & \to & K \\ (x, y) & \mapsto & x + y \end{array} \quad \text{und} \quad \begin{array}{rcl} \cdot : K \times K & \to & K \\ (x, y) & \mapsto & x \cdot y \end{array},$$

so daß die folgenden Axiome gelten:

1. $(K, +)$ ist eine abelsche Gruppe mit dem neutralen Element 0, $K^* := K - \{0\}$.
2. $(K^*, \cdot)$ ist eine abelsche Gruppe mit dem neutralen Element 1.
3. Es gilt das Distributivgesetz: $a \cdot (b + c) = a \cdot b + a \cdot c$.

Beispiele:

1. $(\mathbb{R}, +, \cdot)$ und $(\mathbb{Q}, +, \cdot)$ sind Körper.
2. $(\mathbb{Z}, +, \cdot)$ ist kein Körper (es gibt keine multiplikativ Inversen).

Fakt 1: Sei $(K, +, \cdot)$ ein Körper. Dann gilt:

1. $0 \neq 1$.
2. Für alle $x \in K$ gilt $0 \cdot x = 0$.
3. Für alle $x \in K$ gilt $1 \cdot x = x$.

Beweis:

1. Wegen $1 \in K - \{0\}$ gilt $1 \neq 0$.
   Nebenbemerkung: Wenn $(G, \cdot)$ eine Gruppe ist, dann folgt $G \neq \emptyset$, da die Gruppenaxiome die Existenz des neutralen Elements fordern.

2. Es gilt $0 \cdot x + 0 \cdot x = (0+0) \cdot x = 0 \cdot x = 0 \cdot x + 0$. Aus der Eindeutigkeit folgt die Behauptung.

3. $(K^*, \cdot)$ ist eine Gruppe mit neutralem Element 1.

□

Beispiel: Sei $\mathbb{F}_2 = \{0, 1\}$ mit folgenden Operationen:

$+$	0	1
0	0	1
1	1	0

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

Diese Operationen entsprechen der logischen XOR bzw. AND-Operation, $(\mathbb{F}_2, XOR, AND)$ ist ein Körper.

Definition 4: Sei $n > 0$ eine natürliche Zahl. Eine **Restklasse modulo** n ist eine Menge $M \subseteq \mathbb{Z}$ mit folgenden Eigenschaften:

1. Aus $a, b \in M$ folgt $n \mid a - b$ (n teilt $a - b$).
2. Aus $a \in M, b \in \mathbb{Z}$ und $n \mid a - b$ folgt, daß auch $b \in M$.
3. $M \neq \emptyset$

Die Elemente von M nennen wir auch die Vertreter der Restklasse.

Restklassen werden addiert und multipliziert, indem man ihre Vertreter addiert bzw. multipliziert und die Restklasse bildet, die das Resultat enthält; $M + N = \{a + b \mid a \in M, b \in N\}$, $MN = \{ab + kn \mid a \in M, b \in N, k \in \mathbb{Z}\}$.

Satz 2:

1. Jede ganze Zahl x ist in genau einer Restklasse modulo n enthalten, nämlich in $\{x + kn \mid k \in \mathbb{Z}\}$.
2. Es gibt genau n Restklassen modulo n.
3. Die oben beschriebenen Mengen $M + N$ und MN sind Restklassen modulo n , falls M und N Restklassen modulo n sind.
4. Sei $\mathbb{Z}/n\mathbb{Z}$ ($\mathbb{Z}$ nach $n\mathbb{Z}$) die Menge aller Restklassen modulo n. Dann ist $(\mathbb{Z}/n\mathbb{Z}, +)$ eine abelsche Gruppe.
5. Sei n eine Primzahl, dann ist $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ ein Körper.

Beweis:

1. Sei $M = \{x + kn \mid k \in \mathbb{Z}\}$. M ist eine Restklasse. $M \neq \emptyset$, denn $x = x + 0n \in M$. Sei $a = x + kn$ und $b \in \mathbb{Z}$ und $n \mid a - b$. Sei $\delta = \frac{a-b}{n} \in \mathbb{Z}$. Dann folgt $b = a - (a - b) = x + kn - \delta n = x + (k - \delta) \cdot n \in M$. Ferner gilt: $n \mid ((x + kn) - (x + ln)) = (k - l)n$.
   Sei N eine andere Restklasse, die x enthält. Weil N Definition 4.2 erfüllt, gilt auch $x + kn \in N$ für alle $k \in \mathbb{Z}$, also ist $M \subseteq N$. Sei $y \in N$. Weil N Definition 4.1 erfüllt, gilt $n \mid x - y$, also auch $y = x + kn$ mit $k \in \mathbb{Z}$, also ist $N \subseteq M$, damit gilt $N = M$.

2. Wir behaupten, daß jede Restklasse M genau einen Vertreter x mit $0 \leq x \leq n$ enthält. Wenn diese Behauptung gezeigt ist, gibt es eine Bijektion:

$$\begin{aligned} \text{Restklasse modulo } n &\to \{x \in \mathbb{Z} \mid 0 \leq x \leq n\} \\ M &\mapsto \text{das einzige } x \in M \text{ mit } 0 \leq x \leq n \end{aligned}$$

(Es gibt genau n Restklassen $\Leftrightarrow$ es gibt genau n Elemente)
Beweis der Behauptung: Existenz: Da $M \neq \emptyset$ gibt es ein $a \in M$. Dividiere $\frac{a}{n}$ mit Rest x für $a = kn + x$ für $0 \leq x \leq n$. Es folgt $x \in M$. Also leistet x das Gewünschte.
Eindeutigkeit: Aus $0 \leq x, y \leq n$ und $x \neq y$ folgt $0 <\mid x - y \mid < n$, also teilt n nicht $x - y$. Also können x und y nicht zur selben Restklasse gehören.

3. Sei $x \in M$, $y \in N$, dann ist

$$\begin{aligned} M + N &= \{a + b \mid a \in M, b \in N\} \qquad a = x + kn,\ b = y + ln \\ &= \{x + y + (k + l)n \mid k, l \in \mathbb{Z}\} \\ &= \{x + y + mn \mid m \in \mathbb{Z}\} \qquad m := k + l \end{aligned}$$

$$\begin{aligned} MN &= \{ab + kn \mid a \in M, b \in Nk \in \mathbb{Z}\} \qquad a = x + \alpha n,\ b = y + \beta n \\ &= \{xy + (k + \alpha y + \beta x + \alpha\beta n)n \mid \alpha, \beta, k, \in \mathbb{Z}\} \\ &= \{xy + ln \mid l \in \mathbb{Z}\} \qquad m := k + \alpha y + \beta x + \alpha\beta n \end{aligned}$$

Mit Satz 2.1 folgt die Behauptung.

4. Kommutativität ist offensichtlich. Assoziativität: Sei $x \in M$, $y \in N$, $z \in O$. Dann ist $x + y \in M + N$ und $y + z \in N + O$, also folgt $x + (y + z) \in M + (N + O)$ und $(x + y) + z \in (M + N) + O$, aus (3) folgt $M + (N + O) = (M + N) + O$. Neutrales Element ist die Restklasse, welche die Zahl 0 enthält. Das inverse Element zu M ist $\{-m \mid m \in M\}$.

5. Die Kommutativität und Assoziativität der Multiplikation zeigt man wie bei (4) des Beweises, ebenso das Distributivgesetz. Das neutrale Element der Multiplikation von Restklassen modulo n ist die Restklasse von $1 = \{1 + kn \mid k \in \mathbb{Z}\}$. Sei M ein von 0 verschiedenes Element von $\mathbb{Z}/n\mathbb{Z}$ und sei n eine Primzahl. Wir möchten ein inverses Element zu M konstruieren. Wir zeigen zunächst, daß die Abbildung

$$\begin{aligned} \mathbb{Z}/n\mathbb{Z} &\to \mathbb{Z}/n\mathbb{Z} \\ N &\mapsto MN \end{aligned}$$

injektiv ist. Seien $MN = M\tilde{N}$, $x \in M, y \in N$ und $\tilde{y} \in \tilde{N}$. Dann gilt: $MN = \{xy + kn \mid k \in \mathbb{Z}\}$ und $M\tilde{N} = \{x\tilde{y} + kn \mid k \in \mathbb{Z}\}$. Aus $MN = M\tilde{N}$ folgt daher, daß $n \mid x(y - \tilde{y})$. Weil n eine Primzahl ist, gilt $n \nmid x$, also $n \mid (y - \tilde{y})$, was zu zeigen war.

Lemma: Sei X eine *endliche* Menge, und sei $f : X \to X$ eine injektive Abbildung. Dann ist f surjektiv, folglich auch bijektiv. □

Es folgt die Surjektivität der Abbildung

$$\begin{aligned} \mathbb{Z}/n\mathbb{Z} &\to \mathbb{Z}/n\mathbb{Z} \\ N &\longmapsto MN \end{aligned}$$

Insbesondere gibt es ein $N \in \mathbb{Z}/n\mathbb{Z}$ mit $NM = \{1 + kn \mid k \in \mathbb{Z}\}$. N ist also das Inverse zu M.
Für eine Primzahl n ist $\mathbb{Z}/n\mathbb{Z} = F_n$ ein Körper mit n Elementen. □

Definition 5: Sei $\mathbb{C}$ die Menge aller geordneten Paare (a, b) mit $a, b \in \mathbb{R}$ versehen mit den Operationen der Addition und der Multiplikation:

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

Das Paar (a, b) bezeichnen wir in diesem Zusammenhang als $a + ib$. Die Elemente von $\mathbb{C}$ nennt man komplexe Zahlen. Jede reelle Zahl t kann mit der komplexen Zahl $t + i0$ identifiziert werden, man betrachtet daher die reellen Zahlen als Teilmenge der komplexen Zahlen. Ferner gilt: $i = (0, 1) \Rightarrow i^2 = -1$.

Nachweis der Körperaxiome: Kommutativität und Assoziativität von + und · rechnet man leicht nach, ebenso folgt die Distributivität. Das Nullelement für + ist die reelle Zahl 0, d.h. (0,0). Das Einselement für · ist die reelle Zahl 1, d.h. $(1, 0)$. Das inverse Element zu $a + ib$ ist $\frac{a-ib}{a^2+b^2}$. □

2.2 Vektorräume

Definition: Sei K ein Körper. Eine Menge $(V, +, \cdot)$ (oft auch einfach V) mit der Addition + und der skalaren Multiplikation · heißt ein **Vektorraum** über K, wenn folgende Axiome erfüllt sind:

1. $(V, +)$ ist eine abelsche Gruppe
2. Für alle $v \in V$, $k, l \in K$ gilt: $k \cdot (l \cdot v) = (kl) \cdot v$
3. Für alle $v, w \in V$, $k, l \in K$ gilt:
   (a) $k \cdot (v + w) = k \cdot v + k \cdot w$
   (b) $(k + l) \cdot v = k \cdot v + l \cdot v$
4. Das multiplikativ Neutrale 1_K von K operiert trivial, d.h. für alle $v \in V$ gilt $1_K \cdot v = v$

Bemerkung: Das additiv Neutrale 0_K des Körpers operiert als $0_K \cdot v = 0_V$, denn $0_K \cdot v = (k - k) \cdot v = k \cdot v - k \cdot v = 0_V$.

Beispiel:

1. Der Vektorraum K^n aller $n-$Tupel $(x_1, x_2, \ldots, x_n)$ mit $x_1, x_2, \ldots, x_n \in K$ wird mit folgender Addition und skalarer Multiplikation definiert:
   (a) $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$ für alle $x_1, \ldots x_n, y_1 \ldots, y_n \in K$;
   (b) $k(x_1, \ldots, x_n) = (kx_1, \ldots, kx_n)$ für alle $k, x_1, \ldots, x_n \in K$.
2. Sei X eine Menge und K^X die Menge der Abbildungen von X nach K zusammen mit der Addition bzw. skalaren Multiplikation

$$\begin{array}{rcl} + : K^X \times K^X & \to & K^X \\ (f, g) & \mapsto & f(x) + g(x) \end{array} \quad \text{und} \quad \begin{array}{rcl} \cdot : K \times K^X & \longrightarrow & K^X \\ (k, g) & \longmapsto & k \cdot g(x) \end{array}.$$

   Dann ist $(K^X, +, \cdot)$ ein K-Vektorraum.
3. $V = \{0\}$, der Nullvektorraum.

4. Seien V, W zwei Vektorräume über einem Körper K, dann ist die Summe

$$V \oplus W := \{(u, w) \mid v \in V, w \in W\}$$

wieder ein Vektorraum mit der Addition $(v, w)+(\tilde{v}, \tilde{w}) = (v+\tilde{v}, w+\tilde{w})$, $v, \tilde{v} \in V$, $w, \tilde{w} \in W$ und der skalaren Multiplikation $k(v, w) = (kv, kw)$, $k \in K$, $v \in V, w \in W$.

2.3 Unterobjekte

Definition: Eine Teilmenge H einer Gruppe $(G, \circ)$ heißt **Untergruppe** von G, wenn für alle $g, h \in H$ auch $g \circ h^{-1} \in H$ gilt.

Bemerkung: Das neutrale Element e_G von G liegt in H und ist auch dort das neutrale Element, H ist auch eine Gruppe mit der eingeschränkten Operation von G.
Für $h \in H$ ist $h \circ h^{-1} = e_G$ das neutrale Element, also $e_G \in H$. Die Gruppenaxiome folgen aus der Einschränkung der Rechenregeln für G.

Beispiel: Die symmetrische Gruppe S_{n-1} zu $n-1$ Elementen ist eine Untergruppe der symmetrischen Gruppe S_n zu n Elementen, betrachte dazu nur jene Permutationen von n Elementen, die das letzte Element an seinem Platz lassen.

Definition: Sei K ein Körper und $L \subset K$ eine Teilmenge von K. L heißt **Unterkörper** von K, wenn

1. L mit der eingeschränkten Addition eine Untergruppe von $(K, +)$ und
2. $L - \{0\}$ mit der eingeschränkten Multiplikation eine Untergruppe von $(K - \{0\}, \cdot)$ ist.

Das heißt, für alle $h, l \in L$ gilt $k - l \in L$ und für alle $k, l \in L$, $l \neq 0$ gilt $k \cdot l^{-1} \in L$.

Beispiel: Ist L Unterkörper von K, so ist K ein Vektorraum über L.

Beweis:

1. $(K, +)$ ist eine kommutative Gruppe.
2. Für alle $l_1, l_2 \in L$, $k \in K$ gilt $l_1(l_2 k) = (l_1 l_2)k$ nach dem Assoziativgesetz in K.
3. Für alle $l \in L$, $k_1, k_2 \in K$ gilt $l(k_1 + k_2) = lk_1 + lk_2$ und für alle $l_1, l_2 \in L$, $k \in K$ gilt $(l_1 + l_2)k = l_1 k + l_2 k$ (Distributivgesetz in K).
4. Für alle $k \in K$ gilt $1_L k = 1_K k = k$. □

Allgemeiner ist jeder Vektorraum V über K ein Vektorraum über L.

Beweis: Axiome für K eingeschränkt auf L, zum Beispiel Axiom 2: für alle $l_1, l_2 \in L$, $v \in V$ gilt $l_1(l_2 v) = (l_1 l_2)v$, weil $l_1, l_2 \in K$ und die Regel über K gilt. Axiom 4: für alle $v \in V$ gilt $1_L v = v$, weil $1_L = 1_K$. □

Definition: Eine Teilmenge U eines Vektorraumes $(V, +, \cdot)$ über einem Körper K heißt **Untervektorraum**, falls für alle $k \in K, v, w \in U$ auch $v + w \in U$ und $kv \in U$.

Bemerkung: U ist mit der eingeschränkten Addition von V eine Untergruppe von $(V, +)$, denn für $v, w \in U$ ist $v - w = v + (-1_K)w \in U$.

Bemerkung: Wenn U ein Untervektorraum von V ist und $u_1, \ldots, u_n \in U$, $k_1, \ldots, k_n \in K$, $n \in \mathbf{N} - \{0\}$, dann ist

$$\sum_{i=1}^{n} k_i u_i = k_1 u_1 + \ldots + k_n u_n \in U.$$

Beweis durch vollständige Induktion: $n = 2$: folgt aus der Definition eines Untervektorraumes, gilt die Behauptung für n, so gilt $\sum_{i=1}^{n+1} k_i u_i = \sum_{i=1}^{n} k_i u_i + k_{n+1} u_{n+1} \in U$. □

Definition: Seien $U_1, U_2, \ldots, U_n$ Untervektorräume von V, dann ist die Summe der Untervektorräume U_i zu schreiben als:

$$\sum_{i=1}^{n} U_i = U_1 + U_2 + \ldots + U_n := \left\{ \sum_{i=1}^{n} k_i u_i \mid k_i \in K; u_i \in U_i \right\}.$$

$\sum_{i=1}^{n} U_i$ ist wieder ein Vektorraum.

Lemma: Seien $U_1, U_2, \ldots, U_n$ Untervektorräume von V, dann sind folgende Aussagen äquivalent:

1. Für alle $u \in \sum_{i=1}^{n} U_i$ ist die Darstellung $w = \sum_{i=1}^{n} v_i$, $v_i \in U_i$, eindeutig.
2. Für alle $1 \leq i \leq n$ gilt: $U_i \cap \sum_{\substack{j=1 \\ j \neq i}}^{n} U_j = \{0\}$

Beweis:

$$\begin{aligned}
\text{(1) gilt nicht} \quad &\Leftrightarrow \quad \exists\, \emptyset \neq I \subseteq \{1, \ldots, n\} : \forall\, i \in I\, \exists\, v_i, \tilde{v}_i \in U_i : v_i \neq \tilde{v}_i, \sum_{i=1}^{n} v_i = \sum_{i=1}^{n} \tilde{v}_i \\
&\Leftrightarrow \quad \forall\, i \in I : \hat{v}_i = v_i - \tilde{v}_i \in U_i - \{0\} : \sum_{j \in I} \hat{v}_j = 0 \\
&\Leftrightarrow \quad \exists\, i \in I : -\sum_{j \in I - \{i\}} \hat{v}_j = \hat{v}_i \in U_i \cap \sum_{j \in I - \{i\}} U_j \subseteq U_i \cap \sum_{\substack{j=1 \\ j \neq i}} U_j \\
&\Leftrightarrow \quad \text{(2) gilt nicht}
\end{aligned}$$

□

Definition: Falls (1) und (2) gelten, heißt die Summe der U_i **direkte Summe der Untervektorräume** U_i, geschrieben: $\bigoplus_{i=1}^{n} U_i$.

2.4 Erzeugendensysteme, lineare Unabhängigkeit und Basen

Sei V ein Vektorraum über einen Körper K.

Definition: Der von einer Teilmenge $S \subset V$ **erzeugte** Unterraum von V ist

$$\mathcal{L}(S) := \{ \sum_{i=1}^{n} k_i v_i \mid n \in \mathbf{N}, k_i \in K, v_i \in S \}.$$

S heißt **Erzeugendensytem**, falls $\mathcal{L}(S) = V$; anders ausgedrückt: S erzeugt V. $\mathcal{L}(\emptyset) := \{0\}$.

Bemerkung: Sei $S \subseteq V$, $W_S := \{$ sämtliche Unterräume U von $V \mid S \subseteq U\}$. Dann gilt $\mathcal{L}(S) = \bigcap_{U \in W_S} U$, d.h. $\mathcal{L}(S)$ ist der kleinste Unterraum, der S enthält.

Beweis: Einerseits gilt $\mathcal{L}(S) \supseteq \bigcap_{U \in W_S} U$, weil $\mathcal{L}(S) \in W_S$. Andererseits gilt für alle $k_i \in K$, $v_i \in S$: $\sum_{i=1}^n k_i v_i \in U$ für alle $U \in W_S$, also ist $\mathcal{L}(S) \subseteq \bigcap_{U \in W_S} U$. □

Bemerkungen:

1. Falls $T \subseteq S$, dann gilt $\mathcal{L}(T) \subseteq \mathcal{L}(S)$, da für alle $k_i \in K, v_i \in T$ gilt: $\sum_{i=1}^n k_i v_i \in \mathcal{L}(S)$.
2. $\mathcal{L}(\mathcal{L}(S)) = \mathcal{L}(S)$, denn $\mathcal{L}(S) = \bigcap_{\substack{W_S \\ S \subseteq U}} U = \bigcap_{\substack{W_S \\ \mathcal{L}(S) \subseteq U}} U = \mathcal{L}(\mathcal{L}(S))$.
3. Für beliebige Teilmengen $T, S \subseteq U$ gilt: $\mathcal{L}(T \cup S) = \{\sum_{i=1}^n k_i v_i + \sum_{j=1}^m \tilde{k}_j \tilde{v}_j \mid k_i, \tilde{k}_i \in K; v_i \in S; \tilde{v}_j \in T\} = \mathcal{L}(T) + \mathcal{L}(S)$.

Definition: $v \in V$ heißt **linear abhängig** von $S \subseteq V$, falls $v \in \mathcal{L}(S)$; $v \in V$ heißt **linear unabhängig** von S, wenn v nicht linear abhängig von S ist. S heißt linear unabhängig, wenn alle $v \in S$ von von $S - \{v\}$ linear unabhängig sind. S heißt **Basis**, wenn S ein linear unabhängiges Erzeugendensystem ist.

Beispiel: $\{v\}$ für $v \in V$ ist linear unabhängig genau dann, wenn $v \neq 0$.

Lemma 1: Die folgenden Aussagen für eine Teilmenge $S \subseteq V$ sind äquivalent:

1. S ist linear unabhängig;
2. für alle $v_i \in S, v_i \neq v_j$ für $i \neq j$, $k_i \in K$ folgt aus $\sum_{i=1}^n k_i v_i = 0$, daß $k_1 = \ldots = k_n = 0$ (wichtiges Kriterium!);
3. Für jedes $w \in \mathcal{L}(S)$ ist die Darstellung $w = \sum_{i=1}^n k_i v_i, n \in \mathbf{N}$, $k_i \in K^*$, $v_i \in S$, $v_i \neq v_j$ für $i \neq j$ in folgendem Sinne bis auf Vertauschung eindeutig: sei $w = \sum_{i=1}^m \tilde{k}_i \tilde{v}_i$ eine weitere derartige Darstellung, dann ist $m = n$ und es gibt eine Permutation σ von $\{1, \ldots, n\}$ mit $\tilde{k}_{\sigma(i)} = k_i, \tilde{v}_{\sigma(i)} = v_i$, $1 \leq i \leq n$

Beweis: (2) gilt nicht $\Leftrightarrow \exists\, v_i \in S, k_i \in K^*$, $1 \leq i \leq n : \sum_{i=1}^n k_i v_i = 0 \Leftrightarrow \exists\, v_i \in S, k_i \in K^*$, $1 \leq i \leq n$, $j : k_j \neq 0$, $v_j = -\sum_{\substack{i=1 \\ i \neq j}}^n \frac{k_i}{k_j} v_i \Leftrightarrow \exists\, v_i \in S, k_i \in K^*$, $1 \leq i \leq n$, $j : v_j \in \mathcal{L}(S - \{v_j\}) \Leftrightarrow$ (1) gilt nicht.
(3) gilt nicht $\Leftrightarrow \exists\, k_i, \tilde{k}_i \in K^*$, $v_i, \tilde{v}_i \in S : 0 = \sum_{i=1}^n k_i v_i - \sum_{i=1}^m \tilde{k}_i \tilde{v}_i$, $\exists\, i : k_i v_i \neq \tilde{k}_j \tilde{v}_j$, $1 \leq j \leq m$ $\Leftrightarrow \exists\, \tilde{\tilde{v}}_i \in K^*$, $\tilde{\tilde{v}}_i \in S : \sum_{i=1}^{\max(m,n)} \tilde{\tilde{k}}_i \tilde{\tilde{v}}_i = 0 \Leftrightarrow$ (2) gilt nicht. □

Fakt 2: Eine Untermenge $S \subseteq V$ ist genau dann eine Basis, wenn sich jedes Element $w \in V$ als Summe $w = \sum_{i=1}^n k_i v_i$ mit $k_i \in K$, $v_i \in S$ schreiben läßt, und diese Darstellung bis auf Vertauschung der Indizes eindeutig ist.

Beweis: S ist Erzeugendensystem $\Rightarrow \exists\, k_i \in K, v_i \in S : w = \sum_{i=1}^n k_i v_i$. S ist linear unabhängig $\Leftrightarrow$ Diese Darstellung ist eindeutig nach Lemma 1, d.h. $V = \bigoplus_{v_i \in S} \mathcal{L}(\{v_i\}) = \bigoplus_{v_i \in S} k_i v_i$. □

Beispiel: Sei $n \in \mathbf{N}, v \in K^n$. Für $1 \leq i \leq n$ sei $e_i := (0, \ldots, 0, 1, 0, \ldots, 0) \in K^n$ der Vektor, dessen Einträge überall null sind, außer an der i-ten Stelle, wo sich der Eintrag 1 befindet. Dann ist die Menge $\{e_i \mid 1 \leq i \leq n\}$ eine Basis von K^n, die **kanonische Basis** oder auch Standardbasis.
Beweis: $S = \{e_i \mid 1 \leq i \leq n\}$ erzeugt K^n, weil $\sum_{i=1}^n k_i e_i = (k_1, \ldots, k_n)$ für alle $k_1, \ldots, k_n \in K$. S ist linear unabhängig, weil $\mathcal{L}(S - \{e_i\}) = \{\sum_{\substack{j=1 \\ j \neq i}}^n k_j e_j \mid k_j \in k\} = \{(k_1, \ldots, k_{i-1}, 0, k_{i+1}, \ldots, k_n) \mid k_j \in K\}$, also $e_i \notin \mathcal{L}(S - \{e_i\})$. □

Beispiel 1: $\emptyset$ ist eine Basis von $\{0\}$. $\mathcal{L}(M) = \{\sum_{i=1}^{n} \lambda_i x_i \mid \lambda_i \in K, x_i \in M\}$, $M \subset V$, $\mathcal{L}(\emptyset) = \{0\}$, also folgt die Behauptung. Für jeden Vektorraum V ist $\emptyset \subset V$ linear unabhängig.

Fakt 3: Sei $\{x_1, \ldots, x_n\}$ eine Basis des Vektorraums V. Dann kann jedes Element $x \in V$ auf genau eine Weise in der Form $x = \sum_{i=1}^{n} \lambda_i x_i$ mit $\lambda_i \in K$ dargestellt werden.
Das ist offensichtlich aus Lemma 1 und Fakt 2. □

Lemma 2: Sei V ein Vektorraum und $S \subseteq V$ eine endliche Untermenge von V, welche V erzeugt. Dann gibt es eine Untermenge $B \subseteq S$, welche eine Basis von V ist.

Beweis: Sei M die Menge aller Untermengen von S, welche V erzeugen. M ist eine nichtleere Menge, denn $S \in M$. Die Elemente dieser Menge sind endliche Mengen. Also gibt es ein Element $B \in M$ mit der kleinstmöglichen Anzahl an Elementen, d.h. für alle $T \in M$ gilt $|T| \geq |B|$ ($|M|$:=Anzahl der Elemente von M). Wir behaupten, daß B linear unabhängig und daher eine Basis von V ist. Wenn B nicht linear unabhängig wäre, so gäbe es ein $b \in B$ mit $b \in \mathcal{L}(B - \{b\})$. Sei $B' = B - \{b\}$, wir behaupten, daß $B' \in M$. Tatsächlich gilt

$$\mathcal{L}(B') = \mathcal{L}(\mathcal{L}(B')) = \mathcal{L}(\mathcal{L}(B') \cup \{b\}) \supseteq \mathcal{L}(B' \cup \{b\}) = \mathcal{L}(B) = V.$$

Also $\mathcal{L}(B') = V$, damit $B' \in M$ im Widerspruch zur Definition von B, denn $|B'| < |B|$. □

Bemerkung: Wir hätten auch eine linear unabhängige Untermenge von S mit der größtmöglichen Anzahl an Elementen betrachten können.

Bemerkung: Mit Hilfe des Auswahlaxioms kann der Beweis von Lemma 2 auch für unendliches S geführt werden. Insbesondere hat jeder Vektorraum eine Basis.

Lemma 3: Sei $M \subset V$ linear unabhängig, sei $x \notin \mathcal{L}(M)$. Dann ist auch $N := M \cup \{x\}$ linear unabhängig.

Beweis: Andernfalls gäbe es ein $y \in N$ mit $y \in \mathcal{L}(N - \{y\}) = \mathcal{L}((M \cup \{x\}) - \{y\})$ und damit $x_1, \ldots, x_n \in M - \{y\}$ und $\vartheta, \lambda_1, \ldots, \lambda_n \in K$ mit $y = \vartheta x + \lambda_1 x_1 + \ldots + \lambda_n x_n$. Aus $\vartheta = 0$ würde $y \in \mathcal{L}(M - \{y\})$ folgern, was nicht sein soll (M linear unabhängig). Also $\vartheta \neq 0$ und es folgt $x = \frac{1}{\vartheta}(y - \lambda_1 x_1 - \ldots - \lambda_n x_n) \in \mathcal{L}(M)$, was auch nicht sein soll. □

Folgerung: Jeder Vektorraum, der eine endliche Erzeugendenmenge besitzt, besitzt auch eine endliche Basis.

Definition 3: Einen Vektorraum nennt man genau dann **endlichdimensional**, wenn er eine endliche Basis hat.

Satz 1: (Austauschsatz von Steinitz) Sei V ein Vektorraum über einem Körper K, T eine endliche Untermenge von V, welche V erzeugt, und $S \subset V$ eine linear unabhängige Menge. Dann gilt:

1. $|S| \leq |T|$
2. S kann durch $|T| - |S|$ Elemente von T zu einer Erzeugendenmenge für V ergänzt werden.

Beweis: Durch vollständige Induktion nach $|S|$. Für $|S| = 0$, also $S = \emptyset$, sind beide Behauptungen trivial. Seien beide Behauptungen wahr für Mengen mit $|S| - 1$ Elementen. Sei $b \in S$ und sei $\tilde{S} = S - \{b\}$. Sei $l = |T| - |\tilde{S}|$, es gilt $l \geq 0$ und es gibt $t_1, \ldots, t_l \in T$, so daß $\tilde{S} \cup \{t_1, \ldots, t_l\}$ V erzeugt. Insbesondere gilt $b \in \mathcal{L}(\tilde{S} \cup \{t_1, \ldots, t_l\})$, also gibt es ein $s \in \mathcal{L}(\tilde{S})$ und $\lambda_1, \ldots, \lambda_l \in K$

mit $b = s + \sum_{i=1}^{l} \lambda_i t_i$. Für $l = 0$ würde dies $b = s \in \mathcal{L}(\tilde{S})$ bedeuten, im Widerspruch zur linearen Unabhängigkeit von S. Also gilt $l \geq 1$ und (1) ist für S bewiesen. Aus demselben Grund gibt es ein i mit $\lambda_i \neq 0$, ohne Beschränkung der Allgemeinheit $i = 1$. Dann wird V durch $S \cup \{t_2, \ldots, t_l\} = R$ erzeugt, woraus (2) folgt, denn es gilt $t_1 = \frac{1}{\lambda_1}(b - s - \sum_{i=2}^{l} \lambda_i t_i) \in \mathcal{L}(R)$, also $\tilde{S} \cup \{t_1, \ldots, t_l\} \subseteq \mathcal{L}(R)$, also $\mathcal{L}(R) = V$ nach den wohlbekannten Eigenschaften von $\mathcal{L}$. □

Folgerung 2: Zwei Basen eines endlichdimensionalen Vektorraumes haben dieselbe Anzahl von Elementen (was die Endlichkeit einschließt).

Definition 5: Die Anzahl der Elemente einer Basis eines endlichdimensionalen Vektorraumes V ist die **Dimension** von V, und wir schreiben $\dim V$.

Beweis von Folgerung 2: Seien B, B' endliche Basen. Wende Satz 1 an mit:

$$\begin{aligned} S = B, T = B' &\Rightarrow |B| \leq |B'| \\ S = B', T = B &\Rightarrow |B'| \leq |B| \end{aligned}$$

Daraus folgt $|B| = |B'|$. Angenommen, es gäbe eine unendliche Basis B von V. Sei T eine endliche Basis von V, sei $S \subset B$ eine Untermenge mit $|T| + 1$ Elementen. Dann ist S linear unabhängig, Widerspruch zu Satz 1. □

Folgerung 3: Sei V ein endlichdimensionaler K-Vektorraum, $S \subset V$ eine linear unabhängige Menge und $T \subset V$ eine Erzeugendenmenge. Dann gilt:

1. $|S| \leq \dim V$
2. S ist genau dann eine Basis, wenn $|S| = \dim V$.
3. $|T| \geq \dim V$
4. T ist genau dann eine Basis, wenn die Gleichheit bei (3) eintritt.
5. Jede linear unabhängige Untermenge von V kann zu einer Basis ergänzt werden.

Beweis: Sei B eine Basis von V. Zu:

1. Wende Steinitz (1) auf S und B an: $|S| \leq |B| = \dim V$.
2. Sei $|S| = \dim V = |B|$. Wende Steinitz (2) an. S kann durch $|B| - |S| = 0$ Elemente von B zu einer Erzeugendenmenge ergänzt werden, also ist S selbst eine Erzeugendenmenge.
3. Wende Steinitz (1) auf B und T an: $\dim V = |B| \leq |T|$.
4. Nach Lemma 2 gibt es ein $M \subseteq T$ welches eine Basis von V ist. Wenn $M \subset T$ wäre, so wäre $\dim V = |M| < |T|$. Das ist ein Widerspruch zur Voraussetzung. Also gilt: $T = M$ und T selbst ist eine Basis von V.
5. Sei $S \subset V$ linear unabhängig. Wende Steinitz (2) auf S und B an: S kann durch $|B| - |S|$ Elemente von B zu einer Erzeugendenmenge S' ergänzt werden. Es gilt: $|S'| \leq |S| + |B| - |S| = \dim V$. Aus (3) folgt: $|S'| = \dim V$. Aus (4) folgt, daß S' eine Basis von V ist. Eine andere Möglichkeit des Beweises von (5) ist es, mit Hilfe von Lemma 3 zu S solange Elemente hinzuzufügen, bis man eine linear unabhängige Erzeugendenmenge erhält. □

Fakt 4: Sei $W \subseteq V$ ein Unterraum eines endlichdimensionalen Vektorraumes. Dann ist auch W endlichdimensional, und es gilt: $\dim W \leq \dim V$. Dabei gilt die Gleichheit genau dann, wenn $W = V$ gilt.

Beweis: Wir nehmen zunächst an, daß W endlichdimensional ist. Sei $d = \dim W$ und sei $B = \{w_1, \ldots, w_d\}$ eine Basis von W. Dann ist B linear unabhängige Teilmenge von W, also auch von V. Aus Folgerung 3.1 folgt $\dim W \leq d \leq \dim V$. Wenn die Gleichheit eintreten sollte, so folgt aus 3.2), daß B eine Basis von V ist, also $V = W$.
Beweis der Endlichdimensionalität von W: Jede linear unabhängige Menge $S \subseteq W$ ist auch in V linear unabhängig, also gilt $|S| \leq \dim V$ (3.1). Unter Anwendung von Fakt 5 folgt die Behauptung. □

Fakt 5: Sei W ein K-Vektorraum und $d \in \mathbb{N}$, so daß $|S| \leq d$ für jede linear unabhängige Untermenge $S \subset V$ gilt. Dann ist dieser Vektorraum endlichdimensional (d ist obere Schranke).

Beweis: Sei S eine linear unabhängige Untermenge von W mit der größtmöglichen Anzahl an Elementen. Dann ist S auch eine Erzeugendenmenge, denn andernfalls gäbe es ein $v \in W - \mathcal{L}(S)$ und nach Lemma 3 wäre $S \cup \{v\}$ linear unabhängig, im Widerspruch zu Auswahl von S. □

Satz 2: Seien U und V endlichdimensionale Unterräume eines Vektorraumes X. Dann ist auch $U + V$ endlichdimensional, und es gilt:

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

Beweis: Sei $a = \dim(U \cap V)$, $b = \dim U - \dim(U \cap V)$, $c = \dim V - \dim(U \cap V)$. Sei $B = \{w_1, \ldots, w_a\}$ eine Basis von $U \cap V$. Nach Folgerung 3.5 kann diese Basis von $U \cap V$ zu einer Basis $\{w_1, \ldots, w_a, u_1, \ldots, u_b\}$ von U und zu einer Basis $\{w_1, \ldots, w_a, v_1, \ldots, v_c\}$ von V ergänzt werden.
Behauptung: Die Menge $\{u_1, \ldots, u_b, v_1, \ldots, v_c, w_1, \ldots, w_a\}$ ist eine Basis von $U + V$.
Beweis von $\mathcal{L}(B) = U + V$: Sei $x \in U + V$. Dann kann x in der Form $x = u + v$ mit $v \in V$ und $u \in U$ dargestellt werden. Wir können u und v als $u = \sum_{i=1}^{a} k_i w_i + \sum_{j=1}^{b} l_j u_j$, $v = \sum_{i=1}^{a} k'_i w_i + \sum_{l=1}^{c} m_l v_l$ mit $k_i, k'_i, l_i, m_i \in K$ darstellen. Also gilt

$$u + v = \sum_{i=1}^{a} (k_i + k'_i) w_i + \sum_{j=1}^{b} l_j u_j + \sum_{l=1}^{c} m_l v_l \in \mathcal{L}(B).$$

Also erzeugt B $U + V$, damit ist $U + V$ endlichdimensional.
Beweis der linearen Unabhängigkeit von B: Wäre B nicht linear unabhängig, so gäbe es $\alpha_1, \ldots, \alpha_a$, $\beta_1, \ldots, \beta_b$, $\gamma_1, \ldots, \gamma_c \in K$, die nicht alle verschwinden, mit

$$\sum_{i=1}^{a} \alpha_i w_i + \sum_{j=1}^{b} \beta_j u_j + \sum_{l=1}^{c} \gamma_l v_l = 0,$$

also $\sum_{i=1}^{a} \alpha_i w_i + \sum_{j=1}^{b} \beta_j u_j = -\sum_{l=1}^{c} \gamma_l v_l$. Die linke Seite gehört zu U, die rechte zu V, also gehören beide Seiten zu $U \cap V$. Die rechte Seite kann als Linearkombination der w_i geschrieben werden: $\sum_{l=1}^{c} \gamma_l v_l = \sum_{i=1}^{a} \delta_i w_i$ $(\delta \in K) = \sum_{i=1}^{a} 0 w_i + \sum_{l=1}^{c} \gamma_l v_l = \sum_{i=1}^{a} \delta_i w_i + \sum_{i=1}^{c} 0 v_k$. Weil $\{w_1, \ldots, w_a, v_1, \ldots, v_c\}$ eine Basis von V ist, folgt nach Fakt 3: $\gamma_1 = \ldots = \gamma_c = 0$. Wir erhalten $\sum_{i=1}^{a} \alpha_i w_i + \sum_{j=1}^{b} \beta_j u_j = 0$. Weil $\{w_1, \ldots, w_a, u_1, \ldots, u_b\}$ eine Basis von U ist, folgt nach Fakt 3 $\alpha_i = \beta_i = 0$. Das ist aber ein Widerspruch zur Annahme, daß B linear abhängig ist. Folglich ist B unabhängig, und es gilt:

1. $\dim(U+V) = |B| = a+b+c$
2. $\dim U = a+b$; $\dim V = a+c$; $\dim(U \cap V) = a$
3. $\dim(U \cap V) + \dim(U+V) = 2a+b+c = \dim U + \dim V$ □

2.5 Lineare Abbildungen

Definition 1: Seien G und H multiplikativ geschriebene Gruppen. Ein **Gruppenhomomorphismus** ist eine Abbildung $\varphi : G \to H$ mit der Eigenschaft $\varphi(g_1 \cdot g_2) = \varphi(g_1) \cdot \varphi(g_2)$ für $g_1, g_2 \in G$. Sei $\mathrm{Hom}\,(G,H)$ die Menge aller Homomorphismen von G nach H.

Fakt 1:

1. Sei $\varphi \in \mathrm{Hom}\,(G,H)$. Dann gilt: $\varphi(1) = 1$. Für $g \in G$ gilt: $\varphi(g^{-1}) = \varphi(g)^{-1}$
2. Sei H abelsch und sei $\varphi, \psi \in \mathrm{Hom}\,(G,H)$. Sei $\varphi \cdot \psi$ durch $(\varphi \cdot \psi)(g) = \varphi(g)\psi(g)$ definiert. Dann gilt auch $\varphi \cdot \psi \in \mathrm{Hom}\,(G,H)$, und $\mathrm{Hom}\,(G,H)$ wird mit dieser Verknüpfung eine abelsche Gruppe.

Beweis:

1. Sei $f = \varphi(1) \in H$, dann gilt $f \cdot f = \varphi(1) \cdot \varphi(1) = \varphi(1 \cdot 1) = \varphi(1) = f$. Aus den grundlegenden Eigenschaften von Gruppen folgt $f = 1$.
2. φ, ψ sind Gruppenhomomorphismen: Für $g, h \in G$ gilt

$$\begin{aligned}(\varphi \cdot \psi)(gh) &= \varphi(gh)\psi(gh) = \varphi(g) \cdot \varphi(h)\psi(g)\psi(h) = \varphi(g)\psi(g)\varphi(h)\psi(h) \\ &= (\varphi \cdot \psi)(g)(\varphi \cdot \psi)(h).\end{aligned}$$

   Gruppenaxiome für $\mathrm{Hom}\,(G,H)$: Seien $\alpha, \beta, \gamma \in \mathrm{Hom}\,(G,H)$. Für $g \in G$ gilt

$$((\alpha \cdot \beta) \cdot \gamma)(g) = (\alpha \cdot \beta)(g)\gamma(g) = \alpha(g)\beta(g)\gamma(g) = \alpha(g)(\beta \cdot \gamma)(g) = (\alpha \cdot (\beta \cdot \gamma))(g).$$

   Es sei $e(g) = 1$ für alle $e \in G$, denn $\varphi e = e\varphi = \varphi$. Sei $\varphi \in \mathrm{Hom}\,(G,H)$ und $\psi(g) = \varphi(g)^{-1}$. Mit Hilfe der Kommutativität von H prüft man leicht $\psi \in \mathrm{Hom}\,(G,H)$, ψ ist invers zu φ. Kommutativität prüft man wie die Assoziativität. Damit sind alle geforderten Eigenschaften nachgewiesen. □

Bemerkung:

1. Die identische Abbildung Id_G ist für jede Gruppe G ein Homomorphismus von G nach G.
2. Seien G, H Gruppen, dann ist durch $e(g) = 1 \in H$ für alle $g \in G$ ein Homomorphismus von G nach H definiert.
3. Seien $G \xrightarrow{l} H \xrightarrow{m} I$ Gruppenhomomorphismen, dann ist auch $m \circ l$ ein Gruppenhomomorphismus: $(m \circ l)(gh) = m(l(gh)) = m(l(g)l(h)) = m(l(g))m(l(h)) = (m \circ l)(g)(m \circ l)(h)$.
4. Wenn H additiv geschrieben ist, wird auch $\mathrm{Hom}\,(G,H)$ additiv geschrieben.

Definition 2: Seien $(K, +, \cdot)$ und $(L, +, \cdot)$ Körper. Eine Abbildung $f : K \to L$ ist ein **Körperhomomorphismus** genau dann, wenn folgende Eigenschaften erfüllt sind:

1. f ist ein Gruppenhomomorphismus von $(K, +)$ und $(L, +)$;
2. $f(K^*) \subseteq L^*$ und $f : (K^*, \cdot) \to (L^*, \cdot)$ ist ein Gruppenhomomorphismus.

Bemerkung: Ist f ein Körperhomomorphismus, so gilt $f(xy) = f(x)f(y)$ für alle $x, y \in K$, denn es gibt nur die zwei Möglichkeiten

1. $x, y \in K^*$: wende die Tatsache an, daß $f : (K^*, \cdot) \to (L^*, \cdot)$ ein Gruppenhomomorphismus ist.
2. $x = 0$ oder $y = 0$: Sei $x = 0$, dann ist $f(x) = 0$ und $f(xy) = f(x) \cdot f(y) = 0 \cdot f(y) = 0 = f(0 \cdot y) = f(xy)$.

Die Hintereinanderschaltung von Körperhomomorphismen ist wieder ein Körperhomomorphismus.

Beispiel: Die identische Abbildung ist ein Körperhomomorphismus von K nach K. Sei $f : \mathbb{Q} \to \mathbb{R}$ durch $f(x) = x$ für $x \in \mathbb{R}$ definiert. Dann ist f ein Körperhomomorphismus, ebenso $j : \mathbb{R} \to \mathbb{C}$ mit $j(x) = x + i0$ für alle $x \in \mathbb{R}$.

Definition 3: Seien X, Y Vektorräume über einem Körper K. Eine **K-lineare Abbildung** (oder ein **K-Vektorraumhomomorphismus**) ist eine Abbildung $f : X \to Y$ mit folgenden Eigenschaften:

1. $f : (X, +) \to (Y, +)$ ist ein Gruppenhomomorphismus.
2. Sei $\lambda \in K$, $x \in X$, dann gilt $f(\lambda x) = \lambda f(x)$.

Die Menge der K-linearen Abbildungen von X nach Y bezeichnen wir mit $L(X, Y)$.

In den folgenden Betrachtungen ist der Grundkörper K meist vorgegeben, und wir sprechen von linearen Abbildungen statt K-linearen Abbildungen.

Bemerkung: $\mathbb{C}$ ist sowohl $\mathbb{R}$-, als auch $\mathbb{C}$-Vektorraum. Die Abbildung $f : \mathbb{C} \to \mathbb{C}$ mit $f(z) = \overline{z}$ ist $\mathbb{R}$-linear, aber nicht $\mathbb{C}$-linear.

Fakt 2:

1. Sei f linear, dann gilt
$$f\left(\sum_{i=1}^{n} k_i x_i\right) = \sum_{i=1}^{n} k_i f(x_i),\ k_i \in K, x_i \in X, 1 \leq i \leq n;$$
2. Die identische Abbildung ist linear;
3. Die Hintereinanderausführung zweier linearer Abbildungen ist linear;
4. Die Einschränkung eine linearen Abbildung auf einen Untervektorraum ist linear.

Beweis:

1. Induktion nach n:

$$n = 1: \qquad f(\sum_{i=1}^{1} k_i x_i) = f(k_1 x_1) = k_1 f(x_1) = \sum_{i=1}^{1} k_1 f(x_1)$$

$$n \to n+1: \qquad f(\sum_{i=1}^{n+1} k_i x_i) = f(\sum_{i=1}^{n} k_i x_i + k_{n+1} x_{n+1}) = f(\sum_{i=1}^{n} k_i x_i) +$$
$$f(k_{n+1} x_{n+1}) = \sum_{i=1}^{n} k_i f(x_i) + k_{n+1} f(x_{n+1}) = \sum_{i=1}^{n+1} k_i f(x_i).$$

2. ist trivial.

3. Seien $X \xrightarrow{f} Y \xrightarrow{g} Z$ linear, dann gilt für $x_1, x_2 \in X, k \in K$:

$$\begin{aligned}(g \circ f)(x_1 + y_1) &= g(f(x_1 + x_2)) = g(f(x_1) + f(x_2)) = g(f(x_1)) + g(f(x_2)) \\ &= (g \circ f)(x_1) + (gf)(x_2),\end{aligned}$$

$$(g \circ f)(kx_1) = g(f(kx_1)) = g(kf(x_1)) = kg(f(x_1)) = k(g \circ f)(x_1).$$

4. ist trivial. □

Fakt 3: Wenn $f, g : X \to Y$ lineare Abbildungen sind, so sind auch die durch

1. $(f+g)(x) = f(x) + g(x)$, $x \in X, k \in K$,
2. $(kf)(x) = kf(x)$

definierten Abbildung $f + g$ und kf linear. Auf diese Weise wird $L(X, Y)$ zu einem K-Vektorraum.

Beweis: Die Gültigkeit der Behauptung von Definition 3.1 folgt aus Fakt 1.2. Wegen

$$(f+g)(\lambda x) = f(\lambda x) + g(\lambda x) = \lambda f(x) + \lambda g(x) = \lambda(f+g)(x)$$

ist auch Definition 3.2 erfüllt.
Wegen

$$(kf)(x+y) = kf(x+y) = k(f(x) + f(y)) = kf(x) + kf(y) = (kf)(x) + (kf)(y)$$

ist Definition 3.1 erfüllt, wegen

$$(kf)(\lambda x) = kf(\lambda x) = k(\lambda f(x)) = (k\lambda)f(x) = (\lambda k)f(x) = \lambda(kf(x)) = \lambda(kf)(x)$$

auch Definition 3.2. Der Nachweis der Vektorraumaxiome für $L(X, Y)$ ist klar. □

Definition 4: Sei X ein Vektorraum über K. Wir definieren den **Raum der linearen Funktionale auf** X durch $X^* = L(X, K)$. Man nennt X^* auch den **Dualraum** von X.

Fakt 4: Sei $f : X \to Y$ linear, und seien $U \subseteq X$ und $V \subseteq Y$ Unterräume. Dann ist $f^{-1}(V)$ ein Unterraum von X und $f(U)$ ist ein Unterraum von Y.

Beweis: Seien $y_1, y_2 \in f(U)$ und $\lambda \in K$. Dann gibt es $x_i \in U$ mit $f(x_i) = y_i$ für $i \in \{1, 2\}$. Es folgt $\lambda y_1 = \lambda f(x_1) = f(\lambda x_1) \in f(U)$ und $y_1 + y_2 = f(x_1) + f(x_2) = f(x_1 + x_2) \in f(U)$, weiterhin $0 = f(0) \in f(U)$, also $f(U) \neq \emptyset$.
Seien $x_1, x_2 \in f^{-1}(V)$, dann gilt $f(x_1 + x_2) = f(x_1) + f(x_2) \in V$, also $x_1 + x_2 \in f^{-1}(V)$, $f(\lambda x) = \lambda f(x) \in V$, also $\lambda x \in f^{-1}(V)$, natürlich gilt auch $0 \in f^{-1}(V)$. □

Definition 5: Der **Kern** einer linearen Abbildung $f : X \to Y$ ist durch $\ker(f) = f^{-1}(\{0\}) = \{x \in X \mid f(x) = 0\}$ definiert. Das **Bild** von F ist durch $\text{Bild}\,(f) = f(X)$ definiert. Man schreibt auch oft $\text{Im}\,(f)$ statt $\text{Bild}\,(f)$.
Sei $\varphi : G \to H$ ein Gruppenhomomorphismus. Wir definieren $\text{Bild}\,(\varphi) = f(G)$ und den Kern von φ als das Urbild des neutralen Elementes von H (also 0 bzw. 1). Es handelt sich dabei um Untergruppen.

Bemerkung: f ist surjektiv $\Leftrightarrow \text{Bild}\,(f) = Y$; $f(\{0\}) = \{0\}$.

Fakt 5: f ist genau dann injektiv, wenn $\ker(f) = \{0\}$.

Beweis: f injektiv $\Rightarrow |f^{-1}(\{0\})| \leq 1$. Wegen $0 \in f^{-1}(\{0\})$ folgt $f^{-1}(\{0\}) = \{0\}$. Sei $\ker(f) = \{0\}$. Dann folgt aus $x_1, x_2 \in X$ und $f(x_1) = f(x_2)$, daß auch $f(x_1 - x_2) = 0$, also $x_1 - x_2 \in \ker(f)$, damit $x_1 - x_2 = 0$, folglich $x_1 = x_2$, was die Injektivität von f zeigt. □

Bemerkung: Ein Gruppenhomomorphismus ist genau dann injektiv, wenn sein Kern aus dem neutralen Element besteht.

Satz 1: Sei $f : X \to Y$ eine lineare Abbildung, und sei X endlichdimensional. Dann ist $\text{Bild}\,(f)$ endlichdimensional und $\dim(\text{Bild}\,(f)) = \dim X - \dim(\ker(f))$. Insbesondere ist f genau dann injektiv, wenn $\dim(\text{Bild}\,(f)) = \dim(X)$ gilt.

Beweis: Sei $\{x_1, \ldots, x_d\}$ eine Basis für $\ker(f)$ und $\{x_1, \ldots x_d, x_{d+1}, \ldots, x_n\}$ eine Basis von X ($d = \dim(\ker(f))$; $n = \dim X$).
Behauptung: $\{f(x_{d+1}), \ldots, f(x_n)\}$ hat $n - d$ Elemente und ist eine Basis von $\text{Bild}\,(f)$. Aus der Behauptung folgt: $\dim(\text{Bild}\,(f)) = n - d = \dim X - \dim(\ker(f))$ und der Rest ist trivial.
Beweis der Behauptung: Angenommen, $f(x_i) = f(x_j)$ für $d + 1 \leq i < j \leq n$. Dann gilt: $x_i - x_j \in \ker(f)$, $x_i = x_j + \sum_{k=1}^{d} \lambda_k x_k$ mit geeigneten $x_k \in K$. Dann ist $x_i - x_j = \sum_{k=1}^{d} \lambda_k x_k$ im Widerspruch zur linearen Unabhängigkeit von $\{x_1, \ldots, x_n\}$. $\mathcal{L}(\{f(x_{d+1}), .., f(x_n)\}) = \text{Bild}\,(f)$: Sei $y \in \text{Bild}\,(f)$, also $y = f(x)$ mit $x \in X$. Dann kann x als $x = \sum_{k=1}^{n} \lambda_k x_k$ mit $x_k \in K$ dargestellt werden. Es folgt

$$f(x) = f(\sum_{k=1}^{n} \lambda_k x_k) = \sum_{k=1}^{n} \lambda_k f(x_k) = \sum_{k=d+1}^{n} \lambda_k f(x_k) \in \mathcal{L}(\{f(x_{d+1}), \ldots, f(x_n)\}).$$

Lineare Unabhängigkeit: Angenommen $\sum_{k=d+1}^{n} \lambda_k f(x_k) = 0$. Dann folgt $\sum_{k=d+1}^{n} \lambda_k x_k \in \ker(f)$. Weil die $\{x_1, \ldots, x_d\}$ eine Basis von $\ker(f)$ bilden, gibt es $\lambda_1, \ldots, \lambda_d \in K$ mit $\sum_{k=d+1}^{n} \lambda_k x_k = -\sum_{k=1}^{d} \lambda_k x_k$. Es folgt $\sum_{k=1}^{n} \lambda_k x_k = 0$, also folgt $\lambda_1 = \ldots = \lambda_n = 0$ wegen der linearen Unabhängigkeit der x_n, also gilt auch $\lambda_{d+1} = \ldots = \lambda_n = 0$, und die lineare Unabhängigkeit von $\{f(x_{d+1}), \ldots, f(x_n)\}$ ist bewiesen. □

Definition 6:

- Ein **Monomorphismus** ist ein Homomorphismus von Gruppen bzw. von K-Vektorräumen, welcher injektiv ist.
- Ein **Epimorphismus** ist ein Homomorphismus von Gruppen, Körpern bzw. K-Vektorräumen, welcher surjektiv ist.
- Ein **Isomorphismus** von Gruppen bzw. Körpern bzw. K-Vektorräumen ist ein bijektiver Homomorphismus.
- Ein **Endomorphismus** einer Gruppe, eines Körpers bzw. eines K-Vektorraumes X ist ein Homomorphismus, welcher X nach sich selbst abbildet.
- Ein **Automorphismus** ist ein Endomorphismus welcher zugleich ein Isomorphismus ist.
- Zwei Vektorräume, Gruppen oder Körper sind **isomorph**, wenn zwischen ihnen ein Isomorphismus existiert. Zwei endlichdimensionale K-Vektorräume sind genau dann isomorph, wenn ihre Dimensionen übereinstimmen.

Bemerkung: Wenn f ein Isomorphismus ist, dann ist die inverse Abbildung ebenfalls ein Homomorphismus und damit ein Isomorphismus.

Folgerung 1: (aus Satz 1) Sei f ein Endomorphismus eines endlichdimensionalen Vektorraumes X. Dann sind folgende Aussagen äquivalent:

1. f ist ein Monomorphismus.
2. f ist ein Epimorphismus.
3. f ist ein Automorphismus.

Beweis: (3)$\Leftrightarrow$(1) $\wedge$ (2): Ist gerade die Definition des Begriffs Automorphismus. (1) $\Rightarrow$ (2): Aus $\ker(f) = \{0\}$ folgt $\dim(\text{Bild}\,(f)) = \dim X$. Aus Fakt 4 in 2.3 folgt $\text{Bild}\,(f) = X$ und f ist ein Epimorphismus. (2) $\Rightarrow$ (1): Dies folgt aus Satz 1. □

Warnung: Für Epimorphismen von unendlichdimensionalen Vektorräumen gilt dieser Sachverhalt nicht.

Fakt 6: Eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen $f : X \to Y$ ist genau dann ein Isomorphismus, wenn folgende Bedingungen erfüllt sind:

1. $\dim X = \dim Y$
2. f ist ein Monomorphismus oder ein Epimorphismus.

Beweis: Wie bei Folgerung 1.

Fakt 7: Zwei endlichdimensionale Vektorräume über K sind genau dann isomorph, wenn sie dieselbe Dimension haben. Mit anderen Worten: für jede natürliche Zahl n gibt es bis auf Isomorphie genau einen K-Vektorraum der Dimension n.

Beweis: Seien V und W endlichdimensionale K-Vektorräume. Aus Satz 1 folgt $\dim(W) = \dim(\text{Bild}\,(\varphi)) = \dim V - \dim(\ker(\varphi)) = \dim V$.
Sei weiterhin B eine geordnete Basis von V, $B = (b_1, \ldots, b_n), b_i \in V$. Dann ist durch $i_B : K^n \to$

V, $i_B(\lambda_1, \ldots, \lambda_n) = \sum_{k=1}^{\infty} \lambda_k b_k$, ein Isomorphismus von Vektorräumen definiert. $V \simeq K^{\dim V} = K^{\dim W} \simeq W$, also $V \simeq W$ nach dem nachfolgenden Fakt. □

Bemerkung: Ein Isomorphismus $\varphi : V \to W$ überführt eine Basis B von V in die Basis $\varphi(B)$ von W.

Fakt: Wenn zwei Vektorräume beide zu einem dritten Vektorraum isomorph sind, so sind sie auch untereinander isomorph. □

Satz 2:

1. Seien X und Y endlichdimensionale K-Vektorräume mit geordneten Basen $(x_1, \ldots, x_m)$ und $(y_1, \ldots, y_n)$, $m = \dim(X)$, $n = \dim(Y)$. Sei $A_{ij} : X \to Y$ definiert durch
$$A_{ij}(\sum_{k=1}^{m} \lambda_k x_k) = \lambda_i y_j,\ 1 \leq i \leq m,\ 1 \leq j \leq n.$$
Dann sind die A_{ij} lineare Abbildungen und bilden eine Basis von $L(X,Y)$. Es folgt $\dim(L(X,Y)) = \dim(X) \cdot \dim(Y)$

2. Für $Y = K$ und eine Basis $(x_1, \ldots, x_m)$ von X bilden die linearen Funktionale $\xi_i \in X^*$ mit $\xi_i(\sum_{k=1}^{m} \lambda_k x_k) = \lambda_i$ eine Basis von X^*, die zu $(x_1, \ldots, x_m)$ duale Basis. Insbesondere folgt: $\dim X^* = \dim X$.

Beweis: (2) ist offenbar ein Spezialfall von (1), es genügt also, (1) zu beweisen.
Linearität der A_{ij}: Dies ist leicht zu beweisen, denn es gilt für $x = \sum_{k=1}^{m} \lambda_k x_k$, $\tilde{x} = \sum_{k=1}^{m} \tilde{\lambda}_k x_k$

$$A_{ij}(x + \tilde{x}) = A_{ij}(\sum_{k=1}^{m} (\lambda + \tilde{\lambda}) x_k) = (\lambda_i + \tilde{\lambda}_i) y_i = A_{ij}(x) + A_{ij}(\tilde{x}).$$

Den anderen Teil der Linearitätsbedingung prüft man analog.
Lineare Unabhängigkeit der A_{ij}: Sei $\sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} A_{ij} = 0$ mit $\alpha_{ij} \in K$. Dann gilt für $1 \leq k \leq m$:
$$0 = (\sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} A_{ij})(x_k) = \sum_{j=1}^{n} \alpha_{kj} y_j.$$
Daraus und aus der linearen Unabhängigkeit der y_j folgt $\alpha_{ij} = 0$ für $1 \leq j \leq m$. Also sind die A_{ij} linear unabhängig.
Die A_{ij} erzeugen $L(X,Y)$: Sei $f \in L(X,Y)$. Für $1 \leq i \leq n$ gibt es $\alpha_{ij} \in K$, $1 \leq j \leq n$, mit $f(x_i) = \sum_{j=1}^{n} \alpha_{ij} y_j$. Es folgt

$$f(\sum_{k=1}^{m} \lambda_k x_k) = \sum_{j=1}^{n} \sum_{k=1}^{m} \alpha_{kj} \lambda_k y_j = \sum_{j=1}^{n} \sum_{k=1}^{m} a_{kj} A_{kj} (\sum_{i=1}^{m} \lambda_i x_i),$$

also $f = \sum_{i=1}^{n} \sum_{k=1}^{m} \alpha_{kj} A_{kj}$. □

Insbesondere $X \simeq X^*$ und $X^* \simeq X^{**}$, also $X \simeq X^{**}$. Ein Isomorphismus $X \simeq X^{**}$ läßt sich auf kanonische Weise beschreiben:

Folgerung 2: Für jeden K-Vektorraum X sei die K-lineare Abbildung $i_X : X \to X^{**}$ durch $i_X(x)(l) = l(x)$ mit $x \in X$, $l \in X^*$ definiert. Für endlichdimensionale Vektorräume X ist i_X ein Isomorphismus. Dieser Isomorphismus bildet die Basis von X auf die duale Basis ihrer dualen Basis ab.

Beweis: Für alle $x \in X$ ist $i_X(x)$ ein lineares Funktional auf X^*: $i_X : X \to X^{**}$ ist K-linear. Sei $n = \dim(X)$, $(x_1, \ldots, x_n)$ eine Basis von X, $(\xi_1, \ldots, \xi_n)$ die dazu duale Basis von X^*, $(\tilde{x}_1, \ldots, \tilde{x}_n)$ die dazu duale Basis von X^{**}. Die dualen Basen sind durch

$$\xi_i(x_j) = \delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases} \quad \text{und } \tilde{x}_i(\xi_j) = \delta_{ij}$$

gegeben. Es gilt dann $i_X(x_k)(\xi_l) = \delta_{kl} = \tilde{x}_k(\xi_l)$ für alle l, also $i_X(x_k) = \tilde{x}_k$. Damit ist die letzte Behauptung gezeigt. Daraus folgt auch, daß i_X ein Isomorphismus ist, denn es gilt folgender

Fakt: Sei $f : X \to Y$ eine lineare Abbildung und $(b_1, \ldots, b_n)$ eine geordnete Basis von X. Wenn $(f(b_1), \ldots, f(b_n))$ eine geordnete Basis von Y ist, so ist f ein Isomorphismus. □

Folgerung 3: Sei $U \subseteq X$ ein Unterraum eines endlichdimensionalen Vektorraumes X. Wenn Y endlichdimensional ist, so auch die lineare Abbildung $L(X,Y) \to L(U,Y)$, $f \mapsto f|_U$ surjektiv. Mit anderen Worten, zu jeder linearen Abbildung $g : U \to Y$ gibt es eine lineare Abbildung $G : X \to Y$, welche g fortsetzt, d.h. $G|_U = g$.

Beweis: Sei $a = \dim U$, $b = \dim X \geq \dim U$, $(x_1, \ldots, x_a)$ eine Basis von U, $(x_1, \ldots, x_b)$ eine Basis von X, $c = \dim Y$, $(y_1, \ldots, y_c)$ eine Basis von Y. Für $1 \leq i \leq a$, $1 \leq j \leq c$ sei $A_{ij} \in L(U,Y)$ durch $A_{ij}(\sum_{k=1}^{a} \lambda_k x_k) = \lambda_i y_j$ definiert. Für $1 \leq i \leq b$, $1 \leq j \leq c$ sei $\tilde{A}_{ij} \in L(X,Y)$ durch $\tilde{A}_{ij}(\sum_{k=1}^{b} \lambda_k x_k) = \lambda_i y_j$ definiert. Es folgt $\tilde{A}_{ij}|_U = A_{ij}$ für $1 \leq i \leq a$, $1 \leq j \leq c$. Da die A_{ij} $L(U,Y)$ erzeugen, folgt nach Satz 2 die Behauptung.

Fakt 8: Für eine lineare Abbildung $X \xrightarrow{f} Y$ sei die duale Abbildung $Y^* \xrightarrow{f^*} X^*$ durch $(f^*(l))(x)) = l(f(x))$ mit $x \in X$, $l \in Y^*$ definiert, d.h. $f^*(l) = l \circ f$. Dann ist f^* eine lineare Abbildung und es gilt $f^{**} i_X = i_Y f$, wobei i_X und i_Y in Folgerung 2 definiert sind. Unter Ausnutzung von Folgerung 2, um für endlichdimensionale X, Y X^{**} mit X und Y^{**} mit Y zu identifizieren, geht f^{**} in f über.

Beweis: Sei $l \in Y^*$ und $x \in X$. Es gilt

$$(f^{**}(i_X(x)))(l) = (i_X(x))(f^*(l)) = f^*(l)(x) = l(f(x)) = i_Y(f(x))(l).$$

Weil das für alle l gilt, folgt $f^{**}(i_X(x)) = i_Y(f(x))$. □

Definition 7: Sei X ein endlichdimensionaler Vektorraum. Für $U \subseteq X$ und $V \subseteq X^*$ seien

$$\begin{aligned} U^\perp &= \{l \in X^* \mid \forall\, x \in U : l(x) = 0\} \\ V^\perp &= \{x \in X \mid \forall\, l \in V : l(x) = 0\} \end{aligned}$$

die **orthogonalen Komplemente** von U und V. Dies sind Unterräume $U^\perp \subseteq X^*$ und $V^\perp \subseteq X$, selbst dann, wenn U und V nur Untermengen sind.

Bemerkung: Sei $V \subset X^*$ und

$$\begin{aligned} A = \{x \in X \mid \forall\, l \in V : l(x) = 0\} &= V^\perp \subseteq X \\ B = \{\lambda \in X^{**} \mid \forall\, l \in V : \lambda(l) = 0\} &= V^\perp \subseteq X^{**}. \end{aligned}$$

Dann ist $i_X : X \to X^{**}$ eine Abbildung, welche A isomorph auf B abbildet:

$$\begin{aligned} \forall\, l \in V : x \in A &\Leftrightarrow l(x) = 0 \\ \Leftrightarrow (i_X(x))(l) = 0 &\Leftrightarrow i_X(x) \in B. \end{aligned}$$

Fakt 9: Sei $U \subseteq X$ ein Unterraum. Dann gilt:

1. $\dim(U^\perp) = \dim(X) - \dim(U)$;
2. $U^{\perp\perp} = U$.

Beweis: Zum Beweis der ersten Aussage sei $(x_1, \ldots, x_d)$ eine geordnete Basis von U und sei $(x_1, \ldots, x_n)$, $n = \dim X \geq d = \dim U$, eine Basis von X. Sei $(\xi_1, \ldots, \xi_n)$ die dazu duale Basis von X^*. Behauptung: $(\xi_{d+1}, \ldots, \xi_n)$ ist eine Basis von $U^\perp$.

- Für $d+1 \geq i \geq n$ und $U \ni x = \sum_{k=1}^{d} \lambda_k x_k$ mit $\lambda_k \in K$, folgt:

$$\xi_i(x) = \sum_{k=1}^{d} \lambda_k \xi_i(x_k) = \sum_{k=1}^{d} \lambda_k \delta_{ik} = 0$$

  Also $\xi_{d+1}, \ldots, \xi_n \in U^\perp$. Weil diese Elemente Teil einer Basis von X^* sind, sind sie linear unabhängig.

- Sei $\varphi \in U^\perp$. Für $d+1 \leq k \leq n$ sei $\lambda_k = \varphi(x_k)$. Dann gilt für $\varphi = \sum_{k=d+1}^{n} \lambda_k \xi_k \stackrel{!}{=} \tilde{\varphi} \in \mathcal{L}(\{\xi_{d+1}, \ldots, \xi_n\})$, denn es gilt:
  - für $1 \leq l \leq d$: $\varphi(x_l) = 0$ wegen $\varphi \in U^\perp$ und $x_l \in U$, weiterhin $\tilde{\varphi}(x_l) = 0$, also $\varphi(x_l) = \tilde{\varphi}(x_l)$.
  - für $d+1 \leq l \leq n$: $\varphi(x_l) = \lambda_l = \sum_{k=d+1}^{n} \lambda_k \delta_{kl} = \sum_{k=d+1}^{n} \lambda_k \xi_k(x_l) = \tilde{\varphi}(x_l)$.

Weil dies für alle Elemente der Basis $(x_1, \ldots, x_n)$ gilt, folgt $\varphi = \tilde{\varphi}$ und $U^\perp = \mathcal{L}(\{\xi_{d+1}, \ldots, \xi_n\})$. Also ist die Behauptung gezeigt. Es folgt $\dim U^\perp = n - d = \dim X - \dim U$.
Es gilt $U \subseteq U^{\perp\perp}$, denn für $u \in U$ und $\eta \in V^\perp$ gilt $\eta(u) = 0$ nach der Definition von $U^\perp$, also gilt $u \in U^{\perp\perp}$. Aber

$$\begin{aligned} \dim U^{\perp\perp} &= \dim X^* - \dim U^\perp \\ &= \dim X^* - (\dim X - \dim U) = \dim U, \end{aligned}$$

also $U^{\perp\perp} = U$. □

Satz 3: Sei $X \xrightarrow{f} Y$ ein lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Dann gilt:

1. $\text{Bild}\,(f) = \ker(f^*)^\perp$;
2. $\text{Bild}\,(f^*) = \ker(f)^\perp$.

Beweis:

1. Sei $y \in \text{Bild}\,(f)$. Dann gibt es $x \in X$ mit $y = f(x)$. Sei $\xi \in \ker(f^*)$. Dann gilt: $\xi(y) = \xi(f(x)) = (f^*(\xi))(x) = 0$, also $y \in \ker(f^*)^\perp$. Sei umgekehrt $y \notin \text{Bild}\,(f)$. Sei $V = \text{Bild}\,(f) + k \cdot y \subseteq Y$, $k \in K$. Durch $\varphi(v) = \lambda$ für $v = u + \lambda y$, $u \in \text{Bild}\,(f)$, $\lambda \in K$ (Diese Darstellung ist wegen $y \notin \text{Bild}\,(f)$ eindeutig.) ist ein Funktional $\varphi \in V^*$ definiert. Nach Folgerung 3 gibt es ein Funktional $\xi \in Y^*$ mit $\xi|_V = \varphi$. Dann gilt $\xi(y) \stackrel{y \in V}{=} \varphi(y) = \varphi(0 + 1 \cdot y) = 1$. Es gilt $\xi \in \ker(f^*)$, denn für alle $x \in X$ gilt

$$(f^*(\xi))(x) = \xi(f(x)) \stackrel{f(x) \in V}{=} \varphi(f(x)) = \varphi(f(x) + 0 \cdot y) = 0.$$

   Also $\xi \in \ker(f^*)$ und $\xi(y) \neq 0$. Also $y \notin \ker(f^*)^\perp$.

2. Die zweite Gleichung folgt durch Anwendung der ersten Gleichung auf f^* und durch Anwendung von Fakt 8:

$$\text{Bild}\,(f^*) = \ker(f^{**})^\perp = \ker(f)^\perp.$$ □

Folgerung 4: In der Situation von Satz 3 gilt:

$$\dim(\text{Bild}\,(f)) = \dim(\text{Bild}\,(f^*)).$$

Beweis:

$$\dim(\text{Bild}\,(f)) \overset{\text{Satz 3}}{=} \dim(\ker(f^*)^\perp) \overset{\text{Fakt 8}}{=} \dim(Y) - \dim(\ker(f^*)$$
$$\overset{\text{Satz 2.2}}{=} \dim(Y^*) - \dim(\ker(f^*)) \overset{\text{Satz 1}}{=} \dim(\text{Bild}\,(f^*))$$

□

Definition 8: Sei $X \xrightarrow{f} Y$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Die Dimension des Bildes von f nennt man auch den **Rang von** f, $\text{Rang}\,(f)$ oder $\text{rank}\,(f)$.

3 Matrizen

3.1 Matrizenrechnung

Sei K ein Körper.

Definition 1:

- Eine $m \times n$-**Matrix** A ist ein rechteckiges Schema mit m Zeilen, n Spalten und Einträgen aus K:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad \text{oder} \qquad A = (a_{ij})_{i=1\,j=1}^{m\;\;n}.$$

- $\text{Mat}\,(m, n, K)$ ist die Menge aller $m \times n$-Matrizen von Elementen aus K.
- $\text{Mat}\,(m, n, K)$ erhält die Struktur eines K-Vektorraumes wie folgt:
  - $\lambda(a_{ij})_{i=1\,j=1}^{m\;\;n} = (\lambda a_{ij})_{i=1\,j=1}^{m\;\;n}$;
  - $(a_{ij})_{i=1\,j=1}^{m\;\;n} + (\tilde{a}_{ij})_{i=1\,j=1}^{m\;\;n} = (a_{ij} + \tilde{a}_{ij})_{i=1\,j=1}^{m\;\;n}$
- Die a_{ij} nennt man **Matrixkoeffizienten**.
- Für $1 \leq i \leq m$ ist $(a_{ij})_{j=1}^{n} \in K^n$ der i**-te Zeilenvektor** der Matrix und für $1 \leq j \leq n$ ist $(a_{ij})_{i=1}^{m} \in K^m$ der j**-te Spaltenvektor**.

Beispiel:

$$\begin{aligned} A &= \begin{pmatrix} 2 & 0 & 5 \\ \frac{3}{2} & 1 & 7 \end{pmatrix},\ B = \begin{pmatrix} 3 & 1 & 3 \\ 0 & 8 & 7 \end{pmatrix},\ A, B \in \mathrm{Mat}\,(2,3,\mathbb{R}\,) \\ \frac{2}{3}A &= \begin{pmatrix} \frac{4}{3} & 0 & \frac{10}{3} \\ 1 & \frac{2}{3} & \frac{14}{3} \end{pmatrix},\ A+B = \begin{pmatrix} 5 & 1 & 8 \\ \frac{3}{2} & 9 & 14 \end{pmatrix} \end{aligned}$$

A hat zwei Zeilenvektoren, nämlich (2 0 5) und $(\frac{3}{2}\ 1\ 7)$ und drei Spaltenvektoren $\binom{2}{\frac{3}{2}}, \binom{0}{1}, \binom{5}{7}$.

Definition 2: Sei $A \in \mathrm{Mat}(m,n,K)$. Der **Operator der Multiplikation mit** A $A\cdot$ ist eine lineare Abbildung

$$\begin{aligned} A\cdot : K^n &\to K^m \\ A \cdot (x_i)_{i=1}^n &= (\textstyle\sum_{j=1}^n a_{ij}x_j)_{i=1}^m \end{aligned}$$

mit $A = (a_{ij})_{i=1\,j=1}^{m\ \ n}$, $a_{ij} \in K$, $x_i \in K$.

Fakt 1: Die Abbildung

$$\begin{aligned} \mathrm{Mat}(m,n,K) &\to L(K^n, K^m) \\ A &\mapsto A\cdot \end{aligned}$$

ist ein Isomorphismus von K-Vektorräumen.

Beweis: Für $1 \le i \le m$ und $1 \le j \le n$ sei $A^{(ij)}$ die Matrix $(a_{kl}^{(ij)})_{k,l=1}^{m,n}$, wobei

$$a_{kl}^{(ij)} = \begin{cases} 1 & \text{falls } i=k \text{ und } j=l \\ 0 & \text{sonst} \end{cases},\ \text{d.h. } A^{i,j} = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \leftarrow \begin{matrix} i\text{-te Zeile} \\ j\text{-te Spalte} \end{matrix}.$$

Offenbar bilden die $A^{(ij)}$ eine Basis von $\mathrm{Mat}(m,n,K)$, denn jede $m \times n$-Matrix $A = (a_{ij})_{i=1\,j=1}^{m\ \ n}$ kann als

$$A = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}A^{(i,j)} \qquad (*)$$

dargestellt werden, und umgekehrt ist diese Darstellung eindeutig, weil aus $(*)$ folgt, daß die a_{ij} mit den Matrixkoeffizienten von A übereinstimmen. Also bilden die $m \times n$ Matrizen $A^{(ij)}$, $1 \le i \le m$, $1 \le j \le n$, eine Basis von $\mathrm{Mat}\,(m,n,K)$ und $\dim(\mathrm{Mat}\,(m,n,K)) = m \cdot n = \dim(L(K^n,K^m))$. Daher genügt es, die Injektivität der betrachteten Abbildung zu zeigen. Aus $A\cdot = 0$ folgt bei Anwendung auf den Vektor $e_j = (\delta_{ij})_{i=1}^n \in K^n$ $0 = A \cdot e_j = (a_{ij})_{i=1}^m$ ($=j$-te Spalte von A). Also gilt $a_{1j} = \ldots = a_{mj} = 0$. Weil dies für alle j mit $1 \le j \le n$ gilt, folgt $A = 0$, die betrachtete Abbildung ist also injektiv und damit auch bijektiv. □

Fakt 2: Sie $A \in \mathrm{Mat}\,(m,n,K)$ und sei $e_j \in K^n$, $1 \le j \le n$, wie beim Beweis von Fakt 1 definiert. Wir haben gesehen, daß $A \cdot e_j$ der j-te Spaltenvektor von A ist. Der von diesen Vektoren erzeugte Unterraum von K^m (das Bild von A) ist der **Spaltenraum** von A. (Analog dazu ist der Zeilenraum Unterraum von K^n.) Daher stimmt das Bild von $A\cdot$ mit der linearen Hülle der Spaltenvektoren von A überein. Die Dimension dieses Unterraumes von K^m nennen wir den **Spaltenrang** von A, sie stimmt mit der maximalen Anzahl linear unabhängiger Spaltenvektoren überein. □

Definition 3: Sei $A \in \mathrm{Mat}\,(m,n,K)$ und $B \in \mathrm{Mat}\,(n,o,K)$, $A = (a_{ij})_{i=1\,j=1}^{m\;\;n}$, $B = (b_{jk})_{j=1\,k=1}^{n\;\;o}$. Das **Produkt** $A \cdot B$ ist durch die Matrix

$$A \cdot B := \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right)_{i=1}^{m} {}_{k=1}^{\;\;o} \in \mathrm{Mat}\,(m,o,K)$$

definiert.

Bemerkung 1: Wenn man in Definition 2 den Vektor x als $n \times 1$- Matrix auffaßt, so umfaßt Definition 3 Definition 2:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} x_j \end{pmatrix} .$$

Beispiel 1: Sei

$$\mathrm{Mat}\,(2,3,\mathbb{Q}) \ni A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad \mathrm{Mat}\,(3,3,\mathbb{Q}) \ni B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix}$$

„ Schema von Falk “

$$\begin{array}{cc} & \begin{pmatrix} 1 & 2 & 3 \\ 4 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} & \begin{matrix} 9 & 17 & 3 \\ 24 & 38 & 12 \end{matrix} \end{array}$$

$$A \cdot B = \begin{pmatrix} 9 & 17 & 3 \\ 24 & 38 & 12 \end{pmatrix},$$

$B \cdot A$ ist nicht definiert.

Fakt 3: Die Matrizenmultiplikation entspricht der Hintereinanderausführung linearer Operatoren, d.h. für $A \in \mathrm{Mat}\,(m,n,K)$, $B \in \mathrm{Mat}\,(n,o,K)$, $x \in K^o$, so gilt $A \cdot (B \cdot x) = (A \cdot B) \cdot x$, denn

$$\begin{aligned} A \cdot (B \cdot x) &= \left(\sum_{j=1}^{n} a_{ij} (Bx)_j \right)_{i=1}^{m} = \left(\sum_{j=1}^{n} a_{ij} \sum_{k=1}^{o} b_{jk} x_k \right)_{i=1}^{m} \\ &= \left(\sum_{k=1}^{o} \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right) x_k \right)_{i=1}^{m} = (AB) \cdot x \end{aligned}$$

Weil die Hintereinanderausführung von Abbildungen assoziativ ist, ist es auch die Matrixmultiplikation: aus $A \in \mathrm{Mat}\,(m,n,K)$, $B \in \mathrm{Mat}\,(n,o,K)$, $C \in \mathrm{Mat}\,(o,p,K)$ folgt $A \cdot (B \cdot C) = (A \cdot B) \cdot C \in \mathrm{Mat}\,(m,p,K)$. Es gelten für die Matrizenmultiplikation auch die beiden Distributivgesetze

$$\begin{aligned} (A + \tilde{A})B &= AB + \tilde{A}B \\ A \cdot (B + \tilde{B}) &= AB + A\tilde{B} \end{aligned}$$

und

$$(\alpha A)(\beta B) = (\alpha\beta)(AB),\ \alpha, \beta \in K.$$

Dagegen gilt das Kommutativgesetz für Matrizen nicht, selbst wenn beide Seiten wohldefiniert und vom gleichen Typ sein sollten:

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ aber } \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Definition 4: Sei $n = \dim(V)$, $m = \dim(W)$, $B = (b_1, \ldots, b_n)$ eine geordnete Basis von V , $\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_m)$ eine geordnete Basis von W. Sei $V \stackrel{f}{\to} W$ eine lineare Abbildung, und sei $(a_{ij})_{i=1}^{m}{}_{j=1}^{n}$, $a_{ij} \in K$ durch

$$f(b_j) = \sum_{i=1}^{m} a_{ij}\tilde{b}_i$$

definiert. Die Matrix $(a_{ij})_{i=1}^{m}{}_{j=1}^{n} = \text{Mat}_{B,\tilde{B}}(f) \in \text{Mat}\,(m, n, K)$ nennen wir die **Matrixdarstellung von** f bezüglich B und $\tilde{B}$.

Beispiel 2: $\text{Mat}_{(e_1,\ldots,e_n)(e_1,\ldots,e_m)}(A\cdot) = A$ für $A \in \text{Mat}\,(m, n, K)$. Die Matrixdarstellung von $A\cdot : K^n \to K^m$ bezüglich der Standardbasen ist A. (Auf diese Weise erhält man eine inverse Abbildung zu dem in Fakt 1 betrachteten Isomorphismus.)

Fakt 4:

1. Für eine geordnete Basis $B = (b_1, \ldots, b_n)$ von V sei $i_B(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i b_i$ definiert. In den Bezeichnungen von Definition 4 gilt: $fi_B = i_{\tilde{B}} \cdot (\text{Mat}_{B,\tilde{B}}(f)\cdot)$.

$$\begin{array}{ccc} K^n & \stackrel{A}{\longrightarrow} & K^m \\ {\scriptstyle i_B}\downarrow & & \downarrow{\scriptstyle i_{\tilde{B}}} \\ V & \stackrel{f}{\longrightarrow} & W \end{array}$$

   $A = \text{Mat}_{B,\tilde{B}}(f)$ kommutiert, d.h. $i_{\tilde{B}}A = fi_B$.

2. Die Abbildung

$$\begin{array}{rcl} L(V, W) & \to & \text{Mat}\,(m, n, K) \\ f & \mapsto & \text{Mat}_{B,\tilde{B}}(f) \end{array}$$

   ist ein Vektorraumisomorphismus.

3. Sei $U \stackrel{g}{\to} V$ eine lineare Abbildung, $\hat{B} = (\hat{b}_1, \ldots, \hat{b}_o)$ eine Basis von U, dann gilt

$$\text{Mat}_{\hat{B},\tilde{B}}(f \circ g) = \text{Mat}_{B,\tilde{B}}(f) \cdot \text{Mat}_{\hat{B},B}(g).$$

Beweis:

1. Weil i_B und $i_{\tilde{B}}$ Isomorphismen sind, ist der Operator der Multiplikation mit $\text{Mat}_{B,\tilde{B}}(f)$ durch die Bedingung $f(i_B(x)) = i_{\tilde{B}}(\text{Mat}_{B,\tilde{B}}(f) \cdot x)$ bestimmt, nämlich

$$\text{Mat}_{B,\tilde{B}}(f) \cdot x = i_{\tilde{B}}^{-1}(f(i_B(x))).$$

Nach Fakt 1 ist daher auch $\text{Mat}_{B,\bar{B}}(f)$ selbst eindeutig durch die Bedingung $f(i_B(x)) = i_{\bar{B}}(\text{Mat}_{B,\bar{B}}(f) \cdot x)$ bestimmt. Es verbleibt, diese Gleichung nachzurechnen:
sei $\text{Mat}_{B,\bar{B}}(f) = (a_{ij})_{i=1}^{n}{}_{j=1}^{m}$. Dann gilt:

$$\begin{aligned} f(i_B(x_1, \ldots, x_m)) &= f\left(\sum_{j=1}^{m} x_j b_j\right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij} x_j\right) \tilde{b}_i \\ &= i_{\bar{B}}\left(\text{Mat}_{B,\bar{B}}(f) \cdot x\right) \end{aligned}$$

2. Durch

$$\begin{aligned} L(V,W) &\xrightarrow{I} L((K^m, K^n) \\ I(f) &= i_{\bar{B}}^{-1} \circ f \circ i_B \end{aligned}$$

ist ein Isomorphismus I von Vektorräumen, denn die Linearität von I ist leicht nachprüfbar, und eine inverse Abbildung $K : L(K^m, K^n) \to L(V,W)$ ist durch $K(g) = i_{\bar{B}} \circ g \circ i_B^{-1}$ gegeben. Nach (1) ist die Abbildung $f \mapsto \text{Mat}_{B,\bar{B}}(f)$ die Verknüpfung von I mit dem Inversen des in Fakt 1 betrachteten Isomorphismus $\text{Mat}(n,m,K) \to L(K^m, K^n)$. Die Behauptung folgt, denn die Hintereinanderausführung von Isomorphismen ist ein Isomorphismus.

3. Es gilt für $y \in K^o$

$$\begin{aligned} (f \circ g)(i_{\hat{B}}(y)) = f(g(i_{\hat{B}}(y))) &= f(i_B(\text{Mat}_{\hat{B},B}(f) \cdot y)) \\ &= i_{\bar{B}}(\text{Mat}_{B,\bar{B}}(f) \cdot (\text{Mat}_{\hat{B},B}(g) \cdot y)) \\ &= i_{\bar{B}}((\text{Mat}_{B,\bar{B}}(f) \cdot \text{Mat}_{\hat{B},B}(g))y). \end{aligned}$$

Mit anderen Worten, $\text{Mat}_{B,\bar{B}}(f) \cdot \text{Mat}_{\hat{B},B}(g)$ erfüllt die unter (1) angegebene Bedingung für die Matrixdarstellung von $f \circ g$, die Behauptung folgt nun aus (1). □

Fakt 5: Sei $A \in \text{Mat}(m,m,K)$. Dann sind folgende Bedingungen äquivalent:

1. Es gibt $B \in \text{Mat}(m,m,K)$ mit

$$AB = \mathbf{1}_m = \begin{pmatrix} 1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 1 \end{pmatrix} = (\delta_{ij})_{i=1}^{m}{}_{j=1}^{m}.$$

2. Es gibt $C \in \text{Mat}(m,m,K)$ mit $CA = \mathbf{1}_m$.
3. Die lineare Abbildung $K^m \to K^m$, $x \mapsto Ax$, ist bijektiv.

Beweis: Es gilt $\mathbf{1}_m x = x$ für $x \in K^m$, denn $\mathbf{1}_m x = \left(\sum_{j=1}^{m} \delta_{ij} x_j\right)_{i=1}^{m} = (x_i)_{i=1}^{m}$. Allgemein gilt auch

$$\mathbf{1}_m X = X \text{ und } Y \mathbf{1}_m = Y, \; X \in \text{Mat}(m,k,K), Y \in \text{Mat}(n,m,K).$$

(3)⇒(1) und (2): Sei $x \to f(x) = A \cdot x$ bijektiv und sei g die inverse Abbildung. Nach Fakt 1 gibt es eine Matrix B mit $g(x) = B \cdot x$. Dann gilt nach Fakt 3:

$$(AB) \cdot x = A(B \cdot x) = A(g(x)) = f(g(x)) = x = \mathbf{1}_m x.$$

Aus Fakt 1 folgt $A \cdot B = \mathbf{1}_m$. Genauso leitet man $B \cdot A = \mathbf{1}_m$ aus der Tatsache her, daß g linksinvers zu f ist.
(1)⇒(3): Sei $AB = \mathbf{1}_m$. Für alle $x \in K^m$ folgt $x = \mathbf{1}_m x = (AB)x = A(Bx) \in \text{Bild}\,(A\cdot)$. Also ist der Endomorphismus $A\cdot$ des endlichdimensionalen Vektorraumes K^m ein Epimorphismus und damit auch ein Isomorphismus und (3) folgt.
(2)⇒(3) Praktisch genauso: aus $CA = \mathbf{1}_m$ folgt, daß $A\cdot$ ein Monomorphismus und damit ein Isomorphismus ist. □

Definition 5: Sei K ein Körper.

1. Eine Matrix $A \in \text{Mat}\,(m, m, K)$ nennen wir **invertierbar**, falls sie die äquivalenten Bedingungen von Fakt 5 erfüllt.
2. Sei $GL_m(K) \subset \text{Mat}\,(m, m, K)$ die Menge aller invertierbaren $m \times m$-Matrizen, diese wird, versehen mit dem Matrizenprodukt, eine Gruppe.
3. Sei V ein K-Vektorraum. $\text{Aut}\,(V)$ sei die Menge der Automorphismen von V, versehen mit der Hintereinanderausführung von Automorphismen als Gruppenoperation eine Gruppe.

Bemerkung: Die Gruppenaxiome für GL_m rechnet man leicht nach, das Einselement ist $\mathbf{1}_m$.

Fakt 6: Sei B eine Basis des m-dimensionalen K-Vektorraumes V, dann ist

$$\begin{array}{rcl} \text{Aut}\,(V) & \to & GL_m(K) \\ f & \mapsto & \text{Mat}_{B,B}(f) \end{array}$$

ein Gruppenhomomorphismus.

Beweis: folgt leicht aus Fakt 4. □

Bemerkung: Invertierbare $m \times m$-Matrizen können auch beim Basiswechsel entstehen. Sei nämlich V m-dimensional und seien $B, \tilde{B}$ Basen von V. Aus Fakt 4.3 folgt

$$\begin{array}{rcl} \text{Mat}_{B,\tilde{B}}(\text{Id}_V) \cdot \text{Mat}_{\tilde{B},B}(\text{Id}_V) & = & \text{Mat}_{\tilde{B},\tilde{B}}(\text{Id}_V \text{Id}_V) \\ & = & \text{Mat}_{\tilde{B},\tilde{B}}(\text{Id}_V) = \mathbf{1}_m. \end{array}$$

Es folgt $\text{Mat}_{B,\tilde{B}}(\text{Id}_V) \in GL_m$. Seien $U \xrightarrow{g} V \xrightarrow{f} W$ lineare Abbildungen, B' und B'' geordnete Basen von U bzw. W. Aus Fakt 4.3 folgt

$$\begin{array}{rcl} \text{Mat}_{B',\tilde{B}}(g) & = & \text{Mat}_{B,\tilde{B}}(\text{Id}_V) \cdot \text{Mat}_{B',B}(g); \\ \text{Mat}_{\tilde{B},B''}(g) & = & \text{Mat}_{B,B''}(f) \cdot \text{Mat}_{\tilde{B},B}(\text{Id}_V). \end{array}$$

Fakt 7: Sei A eine $m \times n$-Matrix, $B \in GL_m$, $C \in GL_n$. Dann stimmen die Spaltenräume von A und AB bzw. A und CAB überein.

Beweis: Es gilt wegen der Bijektivität von B

$$\begin{array}{rcl} \text{Spaltenraum}\,(AB) & = & \text{Bild}\,(AB)\cdot = \text{Bild}\,(AB\cdot) \\ & = & \text{Bild}\,(A\cdot) = \text{Spaltenraum}(A). \end{array}$$

Weil $C\cdot$ ein Isomorphismus ist, definiert er einen Isomorphismus

$$\text{Bild}\,(AB)\cdot \xrightarrow{\sim} \text{Bild}\,(C \cdot (AB)\cdot) = \text{Bild}\,((CAB)\cdot),$$

wobei $\text{Bild}\,(AB)\cdot$ der Spaltenraum von AB und $\text{Bild}\,((CAB)\cdot)$ der von CAB ist. Es folgt Spaltenrang(CAB) =Spaltenrang(AB) =Spaltenrang(A). □

Definition 6: Die **transponierte** Matrix einer $m \times n$ Matrix $A = (a_{ij})_{i=1\ j=1}^{m\ n}$ ist die $n \times m$ Matrix $A^T = (a_{ji})_{j=1\ i=1}^{n\ m}$.

Beispiel:

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

Fakt 8: Sei $V \xrightarrow{f} W$ eine lineare Abbildung endlichdimensionaler Vektorräume, $B = (b_1, \ldots, b_n)$ eine geordnete Basis von V, $\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_m)$ eine Basis von W, $B^* = (\beta_1, \ldots, \beta_n)$ die zu B duale Basis von V^*, $\tilde{B}^* = (\tilde{\beta}_1, \ldots, \tilde{\beta}_m)$ die zu $\tilde{B}$ duale Basis von W^*. Dann gilt

$$\text{Mat}_{\tilde{B}^*, B^*}(f^*) = \text{Mat}_{B,\tilde{B}}(f)^T.$$

Beweis: Sei $A = \text{Mat}_{B,\tilde{B}}(f), a = (a_{ij})_{i=1\ j=1}^{n\ m}, A^T = B, B = (b_{ij})_{i=1\ j=1}^{m\ n}, b_{ij} = a_{ji}$. Es gilt

$$\begin{aligned} (f^*(\tilde{\beta}_j))(\sum_{i=1}^{m} \lambda_i b_i) &= \tilde{\beta}_j(f(\sum_{i=1}^{m} \lambda_i b_i)) = \tilde{\beta}_j(\sum_{i=1}^{m}\sum_{l=1}^{n} a_{li}\lambda_i \tilde{b}_l) \\ &= \sum_{i=1}^{m}\sum_{l=1}^{n} a_{li}\lambda_i \delta_{jl} = \sum_{i=1}^{m} a_{ji}\lambda_i \\ &= \sum_{i=1}^{m} a_{ji}\beta_i(\sum_{k=1}^{m} \lambda_k b_k) \end{aligned}$$

Es folgt:

$$f^*(\tilde{\beta}_j) = \sum_{i=1}^{m} a_{ji}\beta_i = \sum_{i=1}^{m} b_{ij}\beta_i, \text{ also } \text{Mat}_{\tilde{B}^*, B^*}(f^*) = B = A^T. \qquad \square$$

Folgerung 1: Aus Fakt 2.4.8 ($(fg)^* = g^* f^*$), Fakt 3 und Fakt 8 folgt:

$$(AB)^T = B^T \cdot A^T.$$

Dies könnte man leicht direkt nachrechnen!

Folgerung 2: Die transponierte Matrix einer invertierbaren Matrix A ist invertierbar, und $(A^T)^{-1} = (A^{-1})^T$. Aus Folgerung 1 folgt, daß $(A^{-1})^T$ tatsächlich die inverse Matrix zu A^T ist.

Definition: Der Zeilenraum bzw. -rang von A ist der Spaltenraum bzw. -rang von A^T.

Bemerkung: Aus Folgerung 1, Folgerung 2 und Fakt 7 folgt, daß sich der Zeilenrang einer Matrix bei Multiplikation mit einer invertierbaren Matrix nicht ändert.

Folgerung 3: Der Spalten- und der Zeilenrang einer Matrix A stimmen überein.

Beweis: Es gibt Vektorräume V und W mit Basen B und $\tilde{B}$ und eine lineare Abbildung $f : V \to W$ mit $\text{Mat}_{B,\tilde{B}}(f) = A$. Dann gilt:

$$f(i_B(x)) = i_{\tilde{B}}(A \cdot x),\ x \in K^{\dim V}.$$

Weil i_B und $i_{\tilde{B}}$ Isomorphismen sind, definiert $i_{\tilde{B}}$ einen Isomorphismus des Bildes von $A\cdot$ auf das Bild von f. Also gilt: Spaltenrang(A) = dim(Bild $(A\cdot)$) = dim(Bild (f)). Wenn man denselben Sachverhalt aus f^* und die Basen $\tilde{B}^*$ und B^* anwendet, so folgt:

$$\text{Zeilenrang}(A) = \text{Spaltenrang}(A^T) = \dim(\text{Bild}\,(A^T)) = \dim(\text{Bild}\,(f^*)) = \dim(\text{Bild}\,(f))$$

nach Folgerung 2.5.4, damit folgt die Behauptung. □

Definition 8: Man nennt den Spaltenrang (=Zeilenrang) einer Matrix A den **Rang** von A und schreibt Rang (A) oder rank (A).

Bemerkung 1: Sei $V \xrightarrow{f} W$ einer lineare Abbildung zwischen endlichdimensionalen Vektorräumen, dann gilt:

$$\text{rang}\,(f) \overset{\text{Def } 2.4.8}{=} \dim(\text{Bild}\,(f)) = \text{rang}\,(\text{Mat}_{B,\tilde{B}}(f)).$$

Dies folgt aus dem Beweis von Folgerung 3.
Eine $n \times n$ Matrix A ist genau dann invertierbar, wenn ihr Rang gleich n ist. In der Tat gilt:

$$\begin{aligned} A \text{ invertierbar} &\Leftrightarrow A\cdot : K^n \to K^n \text{ ist bijektiv} \\ &\Leftrightarrow A\cdot : K^n \to K^n \text{ ist surjektiv} \\ &\Leftrightarrow \dim(\text{Bild}\,(A\cdot)) = n \\ &\Leftrightarrow \text{rang}\,(A) = n \end{aligned}$$

Bemerkung 2: Die Abbildung

$$\begin{aligned} \text{Mat}\,(m,n,K) &\longrightarrow \text{Mat}\,(n,m,K) \\ A &\longmapsto A^T \end{aligned}$$

ist ein Isomorphismus von Vektorräumen. Es gilt $(A^T)^T = A$.

3.2 Das Eliminationsverfahren von Gauß

Dieses Verfahren dient der Bestimmung des Ranges einer Matrix, gegebenenfalls der Berechnung ihrer inversen Matrix oder der Auflösung linearer Gleichungssysteme. Sei $A = (a_{ij})_{i=1}^{m}{}_{j=1}^{n}$ und sei $b = (b_i)_{i=1}^{m}$. Dann ist die Gleichung $A \cdot x = b$ für $x \in K^n$ äquivalent zu den m Gleichungen in K: $\sum_{i=1}^{n} a_{ij}x_j = b_j$ für $1 \le i \le m$, d.h.

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & + & \vdots & + & + & & \vdots & = & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

Die Menge der $b \in K^n$, für die diese Gleichungen eine gemeinsame Lösung haben, stimmt mit dem Bild von A (=Spaltenraum von A) überein.

Definition 1: Eine $m \times n$-Matrix A ist in **Zeilenstufenform**, falls eine Zahl $r \ge 0$ und Zahlen $0 < j_1 < j_2 < \ldots < j_r \le n$ existieren, so daß folgende Bedingungen erfüllt sind:

1. $a_{ij} = 0$ für $i > r$;
2. Für $i \le r$ gilt $a_{ij_i} \neq 0$ und $a_{ij} = 0$ für $j < j_i$.

Beispiel: Die folgenden Matrizen sind in Zeilenstufenform:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 7 \end{pmatrix} \qquad r = 2, j_1 = 1, j_2 = 3$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \end{pmatrix} \qquad r = 2, j_1 = 1, j_2 = 3$$

$$\begin{pmatrix} 1 & 4 & 7 \\ 0 & 5 & 8 \\ 0 & 0 & 9 \end{pmatrix} \qquad r = 3, j_1 = 1, j_2 = 2, j_3 = 3$$

Die folgenden Matrizen sind nicht in Zeilenstufenform:

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

Fakt 1: Wenn A eine Matrix in Zeilenstufenform ist, so sind ihre von 0 verschiedenen Zeilenvektoren linear unabhängig. Insbesondere bilden diese Vektoren eine Basis des Zeilenraumes, und der Zeilenrang ist r, wobei r dieselbe Bedeutung wie in Definition 1 hat.

Beweis: Die Zahlen $1 \leq j_1 < j_2 < \ldots < j_r \leq n$ seien wie in Definition 1 gewählt. Angenommen, $0 = \sum_{k=1}^{r} \lambda_k (a_{ki})_{i=1}^{n}$ und nicht alle $\lambda_k \in K$ sind 0. Seien $\lambda_1 = \ldots = \lambda_{k_0 - 1} = 0$ und $\lambda_{k_0} \neq 0$. Dann ist der j_{k_0}-te Koeffizient der untersuchten Summe

$$\sum_{k=1}^{r} \lambda_k a_{k j_{k_0}} = \sum_{k=k_0}^{r} \lambda_k a_{k j_{k_0}} = \lambda_{k_0} a_{k_0 j_{k_0}} \neq 0$$

im Widerspruch zu unserer Annahme. Die von 0 verschiedenen Zeilenvektoren sind also linear unabhängig. Sie bilden daher eine Basis des Zeilenraumes, und der Zeilenrang ist r.

Fakt 2: Sei A eine $m \times n$ -Matrix in Zeilenstufenform und $b = (b_1, \ldots, b_m) \in K^m$. Das lineare Gleichungssystem $A \cdot x = b$ hat genau dann eine Lösung $x \in K^n$, wenn $b_i = 0$ für $i > r$ gilt. Sei $M = \{j_1, \ldots, j_r\}$. Die Variablen x_i mit $i \notin M$ nennen wir **frei**. Wenn eine Lösung des untersuchten Gleichungssystems existiert, so ist diese durch Vorgabe der freien Variablen eindeutig bestimmt. Die gebundenen Variablen können aus den freien wie folgt berechnet werden:

1. $x_{j_r} = \frac{1}{a_{r j_r}} \cdot \left(b_r - \sum_{j=j_r+1}^{n} a_{rj} x_j \right)$ $\qquad$ x_j frei.
2. Seien $x_{j_{k+1}}, \ldots, x_{j_r}$ bereits bestimmt, dann gilt: $x_{j_k} = \frac{1}{a_{r j_k}} \cdot \left(b_k - \sum_{j=j_k+1}^{n} a_{kj} x_j \right)$
   x_j entweder frei oder gebunden und bereits berechnet.

Insbesondere gilt: wenn eine Lösung existiert, so ist die genau dann eindeutig, wenn keine freien Variablen existieren, d.h. , wenn $r = n$ gilt.

Beispiel: (rationale Koeffizienten)

$$\begin{pmatrix} 7 & 6 & 5 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \; m = n = 3, r = 2, j_n = 1, j_2 = 3.$$

Sei $b = (b_1, b_2, b_3)$ vorgegeben. Das Gleichungssystem hat dann die Gestalt:

$$\begin{array}{rcrcrcl} 7x_1 & + & 6x_2 & + & 5x_3 & = & b_1 \\ 0x_1 & + & 0x_2 & + & x_3 & = & b_2 \\ 0x_1 & + & 0x_2 & + & 0x_3 & = & b_3 \end{array} \quad \Rightarrow \quad \begin{array}{rcrcrcl} 7x_1 & + & 6x_2 & + & 5x_3 & = & b_1 \\ & & & & x_3 & = & b_2 \\ & & & & 0 & = & b_3 \end{array}$$

Es kann nur lösbar sein, wenn $b_3 = 0$. Dann ist genau die Variable x_2 frei, und unsere Formeln lauten:

$$\begin{aligned} x_3 &= b_2 \\ x_1 &= \frac{b_1 - 6x_2 - 5x_3}{7} = \frac{b_1 - 6x_2 - 5b_2}{7}. \end{aligned}$$

Auf diese Weise erhält man in der Tat Lösungen des Gleichungssystems (falls $b_3 = 0$).

Beweis von Fakt 2: Die k-te Gleichung des betrachteten Systems lautet:

$$\sum_{j=1}^{n} a_{kj} x_j = b_k.$$

Für $k > r$ verschwinden alle a_{kj} , also muß $b_k = 0$ sein, wenn eine Lösung existieren soll. Die Notwendigkeit unserer Bedingung für die Lösbarkeit ist somit bewiesen.
Wenn sie erfüllt ist, dann sind die letzten $m-r$ Gleichungen (die k-te mit $r < k \leq m$) automatisch erfüllt. Für $1 \leq k \leq r$ ist die k-te Gleichung $b_k = \sum_{j=1}^{n} a_{kj} x_j = \sum_{j=j_k}^{n} a_{kj} x_j$ (denn $a_{kj} = 0$ für $j < j_k$) zu $x_{j_k} = \frac{1}{a_{kj_k}} \cdot \left(b_k - \sum_{j=j_{k+1}}^{n} a_{kj} x_j \right)$ äquivalent, d.h. zu unserer Vorschrift zur Bestimmung der gebundenen Variablen aus der freien. Damit sind sowohl die Hinlänglichkeit als auch unsere Beschreibung der Lösungsmenge verifiziert. Die letzte Behauptung ist klar. □

Bemerkung 1: Der Spaltenraum von A stimmt mit der Menge aller b, für die $A \cdot x = b$ lösbar ist, überein. Nach Fakt 2 ist seine Dimension r. Insbesondere haben wir einen neuen Beweis für Zeilenrang = Spaltenrang für Zeilenstufenmatrizen gefunden.

Fakt 3: Eine $m \times m$-Matrix A in Zeilenstufenform kann genau dann invertiert werden, wenn $r = m$ gilt. In diesem Fall kann die Bestimmung der inversen Matrix dann durch Auflösung von

$$A \cdot s_i = e_i, 1 \leq i \leq n, e_j = (\delta_{ij})_{i=1}^{m}$$

nach s_i erfolgen. s_i ist dann der i-te Spaltenvektor von A.

Beweis: Sei A eine $m \times m$-Matrix und sei s_i der i-te Spaltenvektor von B, dann bilden die $A \cdot s_i$, $1 \leq i \leq m$, die Spaltenvektoren von $A \cdot B$. Daher ist $A \cdot B = \mathbf{1}_m$ äquivalent zu $A \cdot s_i = e_i$ (denn der i-te Spaltenvektor von $\mathbf{1}_m$ ist e_i). Für $i > r$ wird diese Gleichung nach Fakt 2 unlösbar, also muß $r = m$ sein, wenn A invertierbar sein soll. In diesem Fall sind die Gleichungen für die s_i nach Fakt 2 eindeutig lösbar.

Beispiel:

$$\begin{pmatrix} 1 & 0 & 80 \\ 0 & 23 & -46 \\ 0 & 0 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 1 & 0 & -40 \\ 0 & \frac{1}{23} & 1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

$$\begin{array}{rcrcl} 1 & + & 80g & = & -1 \\ 23d & - & 46g & = & 0 \\ & & 2e & = & 0 \end{array} \quad \Longrightarrow \quad \begin{array}{rcll} g & = & 0 & \text{(3.Gl)} \\ d & = & 0 & \text{(2.Gl)} \\ a & = & 1 & \text{(1.Gl)} \end{array}$$

$$\begin{array}{rcrcl} b & + & 80h & = & 0 \\ 23e & - & 46h & = & 1 \\ & & 2h & = & 0 \end{array} \quad \Longrightarrow \quad \begin{array}{rcll} h & = & 0 & \text{(3.Gl)} \\ e & = & \frac{1}{23} & \text{(2.Gl)} \\ b & = & 0 & \text{(1.Gl)} \end{array}$$

$$\begin{array}{rcrcl} c & + & 80i & = & 0 \\ 23f & - & 46i & = & 0 \\ & & 2i & = & 1 \end{array} \quad \Longrightarrow \quad \begin{array}{rcll} i & = & \frac{1}{2} & \text{(3.Gl)} \\ f & = & 1 & \text{(2.Gl)} \\ c & = & -40 & \text{(1.Gl)} \end{array}$$

Definition 2: Eine **elementare Zeilentransformation** einer Matrix A bzw. des Gleichungssystems $Ax = b$ ist eine der folgenden Transformationen:

1. Vertauschung zweier Zeilen in A bzw. in $Ax = b$.
2. Addition des Vielfachen einer Zeile zu einer anderen Zeile.

Fakt 4: Eine elementare Zeilentransformation ist die linke Multiplikation von A oder $Ax = b$ mit einer der invertierbaren Matrizen
Zeilenvertauschung:

$$T_{ij} = \begin{pmatrix} 1 & 0 & & & \dots & & & 0 & 0 \\ 0 & \ddots & & & & & & & 0 \\ & & 1 & & & & & & \\ & & & 0 & & & 1 & & \\ & & & & 1 & \dots & 0 & & \\ \vdots & & & & \vdots & \ddots & \vdots & & \vdots \\ & & & & 0 & \dots & 1 & & \\ & & & 1 & & & & 0 & \\ & & & & & & & & 1 \\ 0 & & & & & & & \ddots & 0 \\ 0 & 0 & & & \dots & & & 0 & 1 \end{pmatrix} = (t_{kl}^{(i,j)})_{k,l=1}^{m} \text{ mit } t_{kl}^{(i,j)} = \begin{cases} 1 & k = l \neq i, j \\ 1 & k = i, l = j \\ 1 & k = j, l = i \\ 0 & \text{sonst} \end{cases}$$

Addition des λ-fachen der i-ten Zeile auf die j-te, $\lambda \in K$, $i \neq j$:

$$E_{ij}(\lambda) = \begin{pmatrix} 1 & \dots & & & & \dots & 0 \\ \vdots & \ddots & & & & & \vdots \\ & & 1 & \dots & \lambda & & \\ & & \vdots & \ddots & \vdots & & \\ & & 0 & \dots & 1 & & \\ \vdots & & & & & \ddots & \vdots \\ 0 & \dots & & & & \dots & 1 \end{pmatrix} = (e_{kl}^{(i,j)}(\lambda))_{k,l=1}^{m} \text{ mit } e_{kl}^{(i,j)}(\lambda) = \begin{cases} 1 & k = l \\ \lambda & k = i, l = j \\ 0 & \text{sonst} \end{cases}$$

λ steht in der i-ten Zeile, j-ten Spalte. Weiterhin gilt

$$T_{ij}^2 = \mathbf{1}_m \text{ und } E_{ij}(\lambda) \cdot E_{ij}(-\lambda) = \mathbf{1}_m.$$

Insbesondere ändert eine solche Umformung weder die Lösungsmenge eines linearen Gleichungssystems noch den Zeilenrang bzw. Spaltenrang oder den Zeilenraum.

Beweis: Es ist leicht zu sehen, daß das Ersetzen von A (bzw. von $Ax = b$) durch $T_{ij}A$ oder $E_{ij}(\lambda)A$ (bzw. durch $(T_{ij}A)x = T_{ij}b$ oder $(E_{ij}(\lambda)A)x = E_{ij}(\lambda)b$) in der Tat das behauptete bewirkt.
Die Behauptung über das Inverse von T_{ij} und $E_{ij}(\lambda)$ sind ebenfalls leicht nachzurechnen.
Sei B invertierbar mit inverser Matrix B^{-1}. Aus $Ax = b$ folgt $(BA)x = B(Ax) = Bb$ und aus $(BA)x = Bb$ folgt $Ax = \mathbf{1}_m Ax = (B^{-1}B)Ax = B^{-1}((BA)x) = B^{-1}(Bb) = b$ Damit ändert sich die Lösungsmenge nicht.
Die Behauptung über den Zeilenraum und den Zeilenrang folgt aus Fakt 3.1.7 und Folgerung 3.1.1, ebenso die Behauptung über den Spaltenrang. □

Definition 3: Eine $m \times m$-Matrix $A = (a_{ij})_{i=1\,j=1}^{m\,n}$ ist in k-**Zeilenstufenform**, $0 \leq k \leq m$, falls Zahlen $0 < j_1 < j_2 < \ldots < j_k \leq m$ existieren, so daß folgende Bedingungen erfüllt sind:

1. Für $1 \leq i \leq k$ gilt $a_{ij_i} \neq 0$ und $a_{ij} = 0$ für $j < j_i$.
2. Für $k < i \leq m$ und $1 \leq j \leq j_k$ gilt $a_{ij} = 0$.

Beispiel 3:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 8 & 9 \\ 0 & 7 & 10 \end{pmatrix} \text{ und } \begin{pmatrix} 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 8 & 9 & 0 \\ 0 & 0 & 7 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

sind in 1-Zeilenstufenform ($j_1 = 1$ bzw. $j_1 = 2$). Jede Matrix liegt in 0-Zeilenstufenform vor. Eine Matrix A ist in m-Zeilenstufenform genau dann, wenn A ist in Zeilenstufenform und $r = m$ (r gemäß Definition 1).
A ist in Zeilenstufenform vom Rang r genau dann, wenn A in r-Zeilenstufenform ist und für $r < k \leq m$ die k-te Zeile verschwindet.
Wir möchten nun den Eliminationsalgorithmus von Gauß beschreiben. Generell besteht ein Algorithmus aus einer Folge von Anweisungen, durch die bestimmten Variablen bestimmte Werte (in unserem Fall natürliche Zahlen oder Elemente von K) zugeordnet werden. Wir schreiben

$$x \leftarrow \mathsf{Ausdruck},$$

falls der Variablen x der Wert von Ausdruck zugeordnet wird, z.B. ordnet $x \leftarrow y/3$ x ein Drittel der Variablen y zu.

Überführung einer Matrix $A \in$ Mat (m, n, K) oder des Gleichungssystemes $Ax = b$ oder von $Ax^{(\alpha)} = b^{(\alpha)}$, $1 \leq \alpha \leq r$ in Zeilenstufenform:

1. $k \leftarrow 1$.
2. An dieser Stelle liegt die Matrix A in $(k-1)$-Zeilenstufenform vor. Wenn $a_{ij} = 0$ für $k \leq i \leq m$ und $1 \leq i \leq m$ gilt, so liegt die Matrix sogar in vollständiger Zeilenstufenform mit $r = k - 1$ vor und das Verfahren bricht ab.
   Andernfalls sei $j_k = \min\{j \mid \exists\, i : k \leq i \leq m, a_{ij} \neq 0\}$. Da die Matrix bereits in $(k-1)$-Zeilenstufenform vorliegt, gilt $j_k > j_{k-1}$ ($= 0$ für $k = 1$). Sei $l \leftarrow$ (eine natürliche Zahl, so daß $k \leq l \leq m$ gilt und $a_{lj_k} \neq 0$). a_{lj_k} ist das k-**te Pivotelement**.
3. Indem wir die l-te mit der k-ten Zeile vertauschen (sowohl in A als auch, falls gegeben in b bzw. $b^{(\alpha)}$), können wir erreichen, daß $a_{kj_k} = k$-tes Pivotelement $\neq 0$.
4. Hier ist $a_{kj_k} \neq 0$. Setze $p \leftarrow a_{kj_k}^{-1}$. Für alle i mit $k < i \leq m$ führe man folgende elementare Zeilenumformung durch: Multiplikation der k-ten Zeile mit $-p \cdot a_{ij_k}$ und Addition des Ergebnisses zu der i-ten Zeile. Genauer gesagt:
   (a) Setze $q \leftarrow pa_{ij_k}$.
   (b) Für $j_k \leq j \leq n$ setze $a_{ij} \leftarrow a_{ij} - qa_{kj}$.
5. A ist nun in k-Zeilenstufenform. Wenn $k = m$ gilt, sind wir fertig. Andernfalls setzen wir $k \leftarrow k + 1$ und gehen zu Schritt 2 zurück.

Beispiel 4:

$$\begin{pmatrix} 1 & 1 & 2 \\ 30 & 30 & 63 \\ 14 & 16 & 37 \end{pmatrix} x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

(1) $k \leftarrow 1$
(2) $l \leftarrow 1, j_k \leftarrow 1$ Das Pivotelement ist $a_{11} = 1$
(3) bewirkt weiter nichts, denn $k = l$.
(4) $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 2 & 6 \end{pmatrix} x = \begin{pmatrix} 1 \\ -29 \\ -13 \end{pmatrix}$ (5) $k \leftarrow 2$, zurück zu 2
(2) $j_2 \leftarrow 2, l \leftarrow 3$
(3) Vertausche die 2. und 3. Zeile.

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix} x = \begin{pmatrix} 1 \\ -13 \\ -29 \end{pmatrix}$$

(4) In diesem Fall muß nur noch die dritte Zeile durch Subtrahieren des 0-fachen der 2. Zeile korrigiert werden, was überflüssig ist. Die Matrix ist in Zeilenstufenform und der Algorithmus bricht ab.

Bemerkung:

- Da für Zeilenstufenmatrizen Zeilen- und Stufenrang übereinstimmen und elementare Zeilentransformationen beide Ränge nicht ändern, haben wir auch einen neuen Beweis dafür gefunden, daß Zeilenrang = Stufenrang.
- *Bestimmung des Ranges einer Matrix:*
  Überführen in Zeilenstufenform mittels elementarer Zeilentransformationen. Ablesen des Ranges der dabei erhaltenen Zeilenstufenmatrix.
- *Auflösung linearer Gleichungssysteme:*
  Überführung in Zeilenstufenform, Anwendung des in Fakt 2 beschriebenen Verfahrens auf das umgeformte System.
- *Bestimmung der inversen Matrix:*
  Durch Auflösung von $As_i = e_i$. Die s_i bilden (falls sie existieren) die Spalten der inversen Matrix A^{-1}. Andernfalls ist die $m \times m$-Matrix A nicht invertierbar (vgl. Fakt 3).
- *Bestimmung einer Basis des Zeilenstufenraumes:*
  Überführung der Matrix durch elementare Zeilentransformationen in Zeilenstufenform. Die von 0 verschiedenen Zeilenvektoren des Ergebnisses bilden eine Basis des Zeilenraumes (Fakt 4 und Fakt 1).
- *Bestimmung einer Basis des Spaltenraumes:*
  Bestimme eine Basis des Zeilenraumes der transponierten Matrix.

Fakt 5: Die Überführung einer $m \times m$ Matrix in Zeilenstufenform nach dem beschriebenen Algorithmus kostet $\frac{m^3}{3} + O(m^2)$ Multiplikationen und Subtraktionen. □

Definition: Seien $f, g : \mathbf{N} \to \mathbf{R}$ Funktionen. Man sagt, daß f die Wachstumsordnung $O(g(n))$ hat, falls eine Konstante $c > 0$ und ein $n_0 \in \mathbf{N}$ existieren mit

$$|f(n)| \leq c \cdot g(n) \text{ für alle } n \geq n_0.$$

Bemerkung 3: In der Praxis kommt es beim Umgang mit reellen Matrizen sehr auf die geschickte Wahl der Pivotelemente an. (vgl. PRESS, FLANNERY, TEUKOLSKY, VETTERLING: *Numerical Recipes* — erhältlich für C, Fortran, Pascal).

4 Determinanten und multilineare Algebra

4.1 Einige Fakten über Permutationen

Wir erinnern daran, daß S_n die Gruppe der bijektiven Abbildungen von $\{1,\ldots,n\}$ auf sich selbst ist. Solche Abbildungen nennt man auch **Permutationen**. Man schreibt eine Permutation $\pi \in S_n$ oft als die zweizeilige Matrix:

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}.$$

Zum Beispiel ist $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ die Permutation mit $\pi(1)=3, \pi(2)=1, \pi(3)=2$.

Definition 1: Für zwei verschiedene Elemente $i,j \in \{1,\ldots,n\}$ sei (i,j) die Permutation, welche i und j vertauscht und alle anderen Zahlen auf sich selbst abbildet.

$$(i,j)(x) = \begin{cases} x & x \in \{1,\ldots,n\} - \{i,j\} \\ j & x = j \\ i & x = i \end{cases}.$$

Man nennt (i,j) die **Transposition** von i und j.

Fakt 1:

1. Jedes Element von S_n kann als Produkt von höchstens $n-1$ Transpositionen dargestellt werden.
2. $|S_n| = n! = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1$.

Beweis: Für $n = 1$ sind beide Behauptungen klar. Für $n > 1$ betrachten wir S_{n-1} als die Untergruppe

$$\{\pi \in S_n \mid \pi(n) = n\} \subset S_n,\ S_n = S_{n-1} \dot{\cup} \bigcup_{j=1}^{n-1} (j,n) S_{n-1}. \qquad (*)$$

Aus $(*)$ folgen beide Behauptungen durch Induktion nach n. $(*)$ impliziert

$$|S_n| = |S_{n-1}| + \sum_{j=1}^{n-1} |(j,n)S_{n-1}| = |S_{n-1}| + \sum_{j=1}^{n-1} |S_{n-1}| = n \cdot |S_{n-1}|,$$

was (2) beweist. Um (1) durch Induktion nach n zu zeigen, sei $\pi \in S_n$. Wenn sogar $\pi \in S_{n-1}$ gilt, so ist π nach Induktionsannahme ein Produkt von weniger als $n-2$ Transpositionen, andernfalls ist π nach $(*)$ von der Form $(j,n)\rho$ mit $1 \le j \le n$ und $\rho \in S_{n-1}$. Dann ist ρ nach Induktionsvoraussetzung Produkt von weniger als $n-2$ Transpositionen und π ist ein Produkt von weniger als $n-1$ Transpositionen.
Es verbleibt der Beweis von $(*)$. Die rechte Seite von $(*)$ ist sicher eine disjunkte Vereinigung, denn aus $\pi \in S_{n-1}$ folgt $\pi(n) = n$ und aus $\pi \in (j,n)S_{n-1}$ folgt $\pi(n) = j \neq n$. Sei $\pi \in S_n$ und

$j = \pi(n)$. Falls $j = n$ gilt, ist $\pi \in S_{n-1}$. Andernfalls gilt $(j,n)\pi \in S_{n-1}$. Nun betrachten wir $(j,n)^2 = \mathrm{Id}$. Es gilt

$$\pi = \mathrm{Id} \circ \pi = (j,n)((j,n)\pi) \in (j,n)S_{n-1}$$

Damit ist $(*)$ bewiesen. □

Bemerkung: Es gilt $(j,n)^2 = \mathrm{Id}$ und $(j,n) = (n,j)$.

Definition: Für $\pi \in S_n$ sei

$$F_\pi = \{(x,y) \mid x < y \text{ und } \pi(x) > \pi(y)\}$$

die Menge aller **Fehlstände** von π, $l(\pi) = |F_\pi|$ die **Länge** von π und $\mathrm{sign}\,(\pi) = (-1)^{l(\pi)}$ das **Signum** von π. Man nennt π **gerade** oder **ungerade**, je nachdem, ob $\mathrm{sign}\,(\pi) = 1$ oder $\mathrm{sign}\,(\pi) = -1$ gilt.

Bemerkung:

1. $\pi = \mathrm{Id}$ genau dann, wenn $l(\pi) = 0$.
2. Es gilt
$$F_{(i,j)} = \{(i,j)\} \dot\cup \{(i,k) \mid i < k < j\} \dot\cup \{(k,j) \mid i < k < j\}$$
für $i < j$. Die Vereinigung ist disjunkt, es folgt
$$l((i,j)) = 2(j-i) - 1 \text{ und } \mathrm{sign}\,((i,j)) = -1.$$

Satz 1: Für $\pi, \rho \in S_n$ gilt: $\mathrm{sign}\,(\pi\rho) = \mathrm{sign}\,(\pi)\mathrm{sign}\,(\rho)$, d.h. sign ist ein Gruppenhomorphismus von S_n nach der multiplikativen Gruppe $\{\pm 1\}$.

Beweis: Für $1 \le i < j \le n$ ist (i,j) genau dann ein Fehlstand von $\pi\rho$, wenn eine der beiden folgenden Aussagen gilt:

1. (i,j) ist kein Fehlstand von ρ, und $(\rho(i), \rho(j))$ ist ein Fehlstand von π.
2. (i,j) ist ein Fehlstand von ρ, und $(\rho(j), \rho(i))$ ist kein Fehlstand von π.

Diese Möglichkeiten schließen sich aus. Sei

$$N = \{(i,j) \mid (i,j) \notin F_\rho \wedge (\rho(i), \rho(j)) \in F_\pi) \vee ((i,j) \in F_\rho \wedge (\rho(j), \rho(i)) \in F_\pi)\}.$$

Es gilt $|N| = |F_\pi| = l(\pi)$. Also haben wir erhalten

$$F_{\pi\rho} = (F_\rho - N) \cup (N - F_\rho).$$

Dann folgt die Behautung aus folgendem

Lemma: Seine X und Y endliche Mengen, dann ist $|(X-Y) \cup (Y-X)| - |X| - |Y|$ gerade.

Es gilt

$$|F_{\pi\rho}| - |N| - |F_\rho| = |F_{\pi\rho}| - |F_\pi| - |F_\rho| = l(\pi\rho) - l(\pi) - l(\rho)$$

ist gerade. Man erhält

$$\mathrm{sign}\,(\pi\rho) = \mathrm{sign}\,(\pi)\mathrm{sign}\,(\rho) \cdot (-1)^{l(\pi\rho)-l(\pi)-l(\rho)} = \mathrm{sign}\,(\pi)\mathrm{sign}\,(\rho).$$ □

Beweis des Lemmas: Die Vereinigung $(X - Y) \cup (Y - X)$ ist eine disjunkte Vereinigung, also

$$|(X - Y) \cup (Y - X)| = |X - Y| + |Y - X|.$$

Es gilt $|X-Y| = |X|-|X\cap Y|$, denn $X = (X-Y)\dot\cup(X\cap Y)$. Ebenso gilt $|Y-X| = |Y|-|X\cap Y|$. Es folgt:

$$|(X - Y) \cup (Y - X)| = (|X| - |X \cap Y|) + (|Y| - |X \cap Y|)$$

und damit:

$$|(X - Y) \cup (Y - X)| - |X| - |Y| = -2|X \cap Y|. \qquad \square$$

Folgerung 1: Wenn eine Permutation π als Produkt von Transpositionen dargestellt ist, dann ist die Anzahl der Faktoren in dieser Darstellung genau dann gerade, wenn π gerade ist.

Beweis: Sei $\pi = \rho_1 \cdot \ldots \cdot \rho_k$ mit Transpositionen ρ_i, $1 \leq i \leq k$, dann gilt $\operatorname{sign}(\rho_i) = -1$, also ist $\operatorname{sign}(\pi) = \operatorname{sign}(\rho_1) \cdot \ldots \cdot \operatorname{sign}(\rho_k) = (-1)^k$. Also ist k genau dann gerade, wenn $\operatorname{sign}(\pi) = 1$. $\square$

Satz 2: Sei $\pi \in S_n$ und $1 \leq x < n$ und $\pi(x+1) < \pi(x)$. Dann gilt

$$l(\pi \cdot (x, x+1)) = l(\pi) - 1.$$

Beweis: Seien

$$\begin{aligned}
A &= \{y \mid y < x \wedge \pi(y) > \pi(x)\} \\
B &= \{y \mid y < x \wedge \pi(y) > \pi(x+1)\} \\
C &= \{y \mid y > x+1 \wedge \pi(y) < \pi(x)\} \\
D &= \{y \mid y > x+1 \wedge \pi(y) < \pi(x+1)\} \\
E &= \{(i,j) \in F_\pi \mid i \notin \{x, x+1\} \wedge j \notin \{x, x+1\}\}
\end{aligned}$$

Dann gilt

$$F_\pi = \underbrace{E}_{i,j \notin M := \{x,x+1\}} \cup \underbrace{A \times \{x\} \cup B \times \{x+1\}}_{i \notin M \wedge j \in M} \cup \underbrace{\{x\} \times C \cup \{x+1\} \times D}_{i \in M, j \notin M} \cup \underbrace{\{(x, x+1)\}}_{i,j \in M}$$

Zwischenüberlegungen: Sei $1 \leq i < j \leq n$.

1. Sei $\{(i,j)\} \cap \{(x, x+1)\} = \emptyset$. Dann läßt $(x, x+1)$ i und j fest, also gilt $(i,j) \in F_{\pi\cdot(x,x+1)}$.
2. Für $i \notin \{(x, x+1)\}, j \in \{(x, x+1)\}$ gilt
   (a) $j = x$: dann $(\pi \cdot (x, x+1))(x) = \pi(x+1)$. Also $(i,x) \in F_{\pi\cdot(x,x+1)} \Leftrightarrow i \in B$.
   (b) $j = x+1$: dann $(i, x+1) \in F_{\pi\cdot(x,x+1)} \Leftrightarrow i \in A$.
3. Für $i \in \{(x, x+1)\}, j \notin \{(x, x+1)\}$ gilt
   (a) $i = x$: dann $(x,j) \in F_{\pi\cdot(x,x+1)} \Leftrightarrow j \in D$.
   (b) $i = x+1$: dann $(x+1, j) \in F_{\pi\cdot(x,x+1)} \Leftrightarrow j \in C$.
4. Sei $\{(i,j)\} = \{(x, x+1)\}$. Dann: $(i,j) = (x, x+1) \notin F_{\pi\cdot(x,x+1)}$.

Damit folgt:

$$F_{\pi\cdot(x,x+1)} = \underbrace{E}_{1.} \cup \underbrace{B \times \{x\} \cup A \times \{x+1\}}_{2.} \cup \underbrace{\{x\} \times D \cup \{x+1\} \times C}_{3.}$$

und

$$\begin{aligned} l(\pi) &= |E| + |A| + |B| + |C| + |D| + 1 = 1 + |E| + |B| + |A| + |D| + |C| \\ &= 1 + l(\pi \cdot (x, x+1)) \end{aligned}$$

□

Bemerkung: Wenn $\pi(x) < \pi(x+1)$, so gilt $l(\pi \cdot (x, x+1)) = l(\pi) + 1$. Dies zeigt man genauso wie Satz 2 oder leitet es aus Satz 2 her (Übungsaufgabe).

Folgerung 2: Jede Permutation π kann als Produkt von $l(\pi)$ Transpositionen der Form $(x, x+1)$ geschrieben werden.

Beweis durch Induktion nach $l(\pi)$: Sei $l(\pi) = 0$. Dann ist $\pi = \mathrm{Id}$ zu zeigen. Wegen $l(\pi) = 0$ folgt $\pi(a) < \pi(b)$ für $1 \leq a < b \leq n$. Dann gilt $\pi(k) = k$ für $1 \leq k \leq n$: wäre $\pi(1) \neq 1$, so würde $\pi(1) > 1$ folgen, also $\pi(a) \geq \pi(1) > 1$ für ein $1 \leq a \leq n$, also $1 \notin \mathrm{Bild}\,(\pi)$ im Widerspruch zur Bijektion von π. Sei nun $\pi(1) = 1, \ldots, \pi(k-1) = k-1$. Aus $\pi(k) \neq k$ würde $\pi(k) > k$ und damit $\pi(a) \geq \pi(k) > k$ für ein $k \leq a \leq n$ folgen, also $k \notin \mathrm{Bild}\,(\pi)$ Widerspruch. Also gilt $\pi(k) = k$.
Die Behauptung sei für Permutationen ρ mit $l(\rho) < l(\pi)$ bewiesen, und es gelte $l(\pi) > 0$. Dann gibt es ein x mit $1 \leq x < n$ und $\pi(x+1) < \pi(x)$. (Andernfalls würde leicht $F_\pi = \emptyset$ folgen). Sei $\rho = \pi \cdot (x, x+1)$. Nach Satz 2 gilt $l(\rho) = l(\pi) - 1$, und ρ ist nach Induktionsvoraussetzung ein Produkt von $l(\pi) - 1$ Transpositionen benachbarter Zahlen. Es folgt, daß

$$\pi = \pi\mathrm{Id} = \pi(x, x+1) \cdot (x, x+1) = \rho(x, x+1)$$

ein Produkt von $l(\pi)$ Transpositionen benachbarter Zahlen ist. □

4.2 Charakteristik eines Körpers

Sei K ein Körper.

Definition 1: Für ganze Zahlen n sei $n_K \in K$ folgendermaßen definiert:

$$\begin{aligned} 0_K &= 0 \quad \text{(Nullelement der additiven Gruppe)} \\ 1_K &= 1 \quad \text{(Einselement der multiplikativen Gruppe)} \\ (n+1)_K &= n_K + 1 \text{ für } n \geq 0. \\ n_K &= -(-n)_K \text{ für } n < 0. \end{aligned}$$

Bemerkung: $n_k = n \cdot 1_K$ (Produkt der ganzen Zahl n mit dem Element 1 der additiven Gruppe)

$$n_K = \underbrace{1 + \ldots + 1}_{n \text{ Summanden}} \in K, \text{ für } n > 0.$$

Lemma 1:

1. $(n+m)_K = n_K + m_K$
2. $(n \cdot m)_K = n_K \cdot m_K$

Beweis:

1. (wurde bereits für beliebige additiv geschriebene Gruppen formuliert und begründet, wir wollen aber noch einen ordentlichen Induktionsbeweis anführen!)

   Fall 1: $n \geq 0$, $m \geq 0$
   Beweis durch Induktion nach m: Für $m = 0$ ist die Behauptung klar:

   $$(n+0)_K = n_K = n_K + 0 = n_k + 0_k$$

   Induktionsschluß von m auf $m+1$:

   $$\begin{array}{rcll} (n+(m+1))_K & = & (n+m)_K + 1 & \text{Definition 1} \\ & = & n_K + m_K + 1 & \text{Induktionsvoraussetzung} \\ & = & n_K + (m+1)_K & \text{Definition 1} \end{array}$$

   Fall 2: $n \geq 0, m < 0$:

   (i) $n \geq 0, m < 0, n+m \geq 0$: Es gilt $n_K = ((n+m)+(-m))_K = (n+m)_K + (-m)_K$ nach Fall 1. Aus Definition 1 folgt daher

   $$n_K + m_K = n_K - (-m)_K = (n+m)_K.$$

   (ii) $n \geq 0, m < 0, n+m < 0$: Dann gilt $(-m) = (-(n+m)) + n$, also $(-m)_K = (-(n+m))_K + n_K$ nach Fall 1 und damit $(n+m)_K = -(-(n+m))_K = -((-m)_K - n_K) = n_K - (-m)_K = n_K + m_K$ nach Definition 1.

   Fall 3: $n < 0, m > 0$: Wie Fall 2, denn die Aussage ist in n und m symmetrisch.
   Fall 4: $n < 0, m < 0$:

   $$\begin{array}{rcll} (n+m)_K & = & -((-n-m)_K) & \text{Definition 1} \\ & = & -((-n)_K + (-m)_K) & \text{Fall 1} \\ & = & -(-n)_K - (-m)_K & \\ & = & n_K + m_K & \text{Definition 1} \end{array}$$

2. Fall 1: $m \geq 0, n$ beliebig: Vollständige Induktion nach m.

   - Für $m = 0$ gilt $(n \cdot 0)_K = 0_K = 0 = n_K \cdot 0 = n_K m_K$
   - Schluß von m auf $m+1$:

   $$\begin{array}{rcl} n_K(m+1)_K & \overset{\text{Def.1}}{=} & n_K(m_K + 1) \\ & = & n_K \cdot m_K + n_K \\ & \overset{\text{IV}}{=} & (n \cdot m)_K + n_K \\ & \overset{(1)}{=} & (n \cdot m + n)_K \\ & = & (n \cdot (m+1))_K \end{array}$$

   Fall 2: $m < 0$,n beliebig: Es gilt $(mn)_K + (-mn)_K = 0_K = 0$ nach (1), also $-(mn)_K = (-mn)_K$. $n_K \cdot m_K \overset{\text{Def.1}}{=} n_K(-(-m)_K) = -n_K(-m)_K \overset{\text{Fall 1}}{=} -(-nm)_K = (nm)_K$. □

Lemma 2: Für einen Körper K gibt es zwei Möglichkeiten

1. $n_K \neq 0$ für alle $n \in \mathbb{Z} - \{0\}$;

2. andernfalls gibt es eine Primzahl p, so daß $n_K = 0$ genau dann gilt, wenn p n teilt.

Definition 2: Für ganze Zahlen n sei $n_K \in K$. In der Situation von Lemma 2 sagen wir, daß

1. K Charakteristik 0 hat;

2. K Charakteristik p hat.

Beweis von Lemma 2: Angenommen, Fall (1) trifft nicht zu. Dann gibt es $n \in \mathbb{Z} - \{0\}$ mit $n_K = 0$. Dann gilt auch $(-n)_K = 0$, also können wir $n > 0$ voraussetzen. Sei p die kleinste positive ganze Zahl mit $p_K = 0$.
Behauptung:

1. p ist Primzahl;

2. $n_K = 0$ genau dann, wenn $p \mid n$.

Aus (1) und (2) folgt, daß Fall (2) von Lemma 2 eintritt.
zu 1. Wenn p keine Primzahl ist, so gibt es $a, b \in \mathbb{Z}$, $a > 1, b > 1, p = a \cdot b$. Wegen $a < p$ und $b < p$ gilt $a_K \neq 0$ und $b_K \neq 0$ nach der Auswahl von p. Weil $K - \{0\}$ multiplikativ abgeschlossen ist, folgt $0 \neq a_K \cdot b_K = (a \cdot b)_K = p_K$ im Widerspruch zur Auswahl von p.
zu 2. Sei $n \in \mathbb{Z}$. Dividiere n mit Rest durch p: $n = a \cdot p + r$, $a, r \in \mathbb{Z}$, $0 \leq r < p$. Nach Lemma 1 gilt

$$n_K = a_K p_K + r_K = r_K.$$

Weil p die kleinste positive ganze Zahl mit $p_K = 0$ ist, verschwindet dieses Körperelement genau dann, wenn $r = 0$ gilt, d.h., wenn n durch p teilbar ist. □

Beispiel: $\mathrm{char}\,(\mathbb{Q}) = \mathrm{char}\,(\mathbb{R}) = \mathrm{char}\,(\mathbb{C}) = 0$. $\mathrm{char}\,(\mathbb{F}_p) = p$, denn

$$\begin{aligned} p_K &= \underbrace{1_K + \ldots + 1_K}_{p \text{ Summanden}} \\ &= \underbrace{1 + p\mathbb{Z} + \ldots + 1 + p\mathbb{Z}}_{p \text{ Summanden}} \\ &= p + p \cdot \mathbb{Z} \\ &= p \cdot \mathbb{Z} \\ &= 0_K. \end{aligned}$$

Bemerkung: Wenn K ein Körper und p eine Primzahl mit $p_K = 0$ ist, so hat K die Charakteristik p, denn p ist die einzige Primzahl, welche p teilt.
Die Bezeichnungsweise n_K wird normalerweise nicht verwendet.

Definition 3: Sei $\mathrm{char}\,(K) = p$, $x \in K$, und $r = \frac{z}{n} \in \mathbb{Q}$, wobei z und n teilerfremd sind und n nicht durch p teilbar ist (Für $p = 0$ bedeutet dies nur $n \neq 0$). Definiere

$$r \cdot x = \frac{z_K \cdot x}{n_K}.$$

Bemerkung: In Zukunft schreiben wir $n \cdot 1$ statt n_K.

Bemerkung: Für die Multiplikation von Körperelementen mit rationalen Zahlen gelten folgende Rechenregeln: sei $x, y \in \mathbb{Q}$, wobei die Nenner von x und y nicht durch $\mathrm{char}\,(K)$ teilbar sind. Dann gilt für $k, l \in K$:

$$\begin{aligned} x(yk) &= (xy)k \\ x(kl) &= (xk)l \\ (x+y)k &= xk + yk \\ x(k+l) &= xk + xl \end{aligned}$$

Der Beweis ist einfach.

4.3 Schiefsymmetrische n-Linearformen

Sei K ein Körper.

Definition 1: Seien $V_1, \ldots, V_n$ und W K-Vektorräume, eine Abbildung

$$f : V_1 \times \ldots \times V_n \longrightarrow W$$

ist n-**linear** (bilinear für $n = 2$; trilinear für $n = 3$), falls für alle i mit $1 \leq i \leq n$ und alle $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n, x_k \in V_k$ die Abbildung

$$V_i \longrightarrow W : V_i \ni x \mapsto f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)$$

eine lineare Abbildung ist. Sei $L(V_1, \ldots, V_n; W)$ die Menge aller n-linearen Abbildungen von $V_1, \ldots, V_n$ nach W, versehen mit der folgenden Struktur eines K-Vektorraumes:

$$\begin{aligned} (f+g)(x_1, \ldots, x_n) &= f(x_1, \ldots, x_n) + g(x_1, \ldots, x_n) \\ (\lambda f)(x_1, \ldots, x_n) &= \lambda \cdot f(x_1, \ldots, x_n) \end{aligned}$$

für $f, g \in L(V_1, \ldots, V_n; W)$ und $\lambda \in K$. Die Elemente von $L(\underbrace{V, \ldots, V}_{\text{n Stück}}; K)$ nennt man auch n-**Linearformen** auf V.

Fakt 1: Für eine n-lineare Abbildung gelten folgende Gesetze ($\lambda \in K$, $x_i \tilde{x}_i \in V_i$):

- $$\begin{aligned} f(x_1, \ldots, x_{i-1}, x_i + \tilde{x}_i, x_{i+1}, \ldots, x_n) &= f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) \\ &\quad + f(x_1, \ldots, x_{i-1}, \tilde{x}_i, x_{i+1}, \ldots, x_n); \end{aligned}$$
- $f(x_1, \ldots, x_{i-1}, \lambda x_i, x_{i+1}, \ldots, x_n) = \lambda f(x_1, \ldots, x_n)$.

Es gilt auch

$$f(\lambda_1 x_1, \ldots, \lambda_n x_n) = \lambda_1 \lambda_2 \cdot \ldots \cdot \lambda_n f(x_1, \ldots, x_n).$$

Insbesondere gilt $f(x_1, \ldots, x_n) = 0$, wenn eines der x_i verschwindet. □

Beispiel 1:

- Für alle n ist die Abbildung

$$\begin{aligned} K \times K \times \ldots \times K &\to K \\ (x_1, \ldots, x_n) &\mapsto x_1 \cdot \ldots \cdot x_n \end{aligned}$$

  eine n-Linearform auf K.

- linear entspricht 1-linear.

- Sei $f \in L(V_1, \ldots, V_n; W)$ und sei $\pi \in S_n$, dann ist die folgende Abbildung g ein Element von $L(V_{\pi(1)}, \ldots, V_{\pi(n)}; W)$;

$$g(y_1, \ldots, y_n) = f(y_{\pi^{-1}(1)}, \ldots, y_{\pi^{-1}(n)}).$$

  Zum Beispiel für $f \in L(V_1, V_2; W)$ ist g mit $g(y_1, y_2) = f(y_2, y_1)$ ein Element von $L(V_2, V_1, W)$.

- Für jeden Vektorraum V ist die Abbildung

$$\begin{aligned} f : K \times V &\to V \\ (\lambda, v) &\mapsto \lambda v \end{aligned}$$

  bilinear.

- Sei f eine k-Linearform und g eine l-Linearform auf V, dann ist durch

$$\begin{aligned} V \times V \times \ldots \times V &\to K \\ (v_1, \ldots, v_{k+l}) &\mapsto f(v_1, \ldots, v_k) \cdot g(v_{k+j}, \ldots, v_{k+1}) \end{aligned}$$

  eine $k + l$-Linearform definiert.

- Wenn $f \in L(V_1, \ldots, V_n; W)$ und $g \in L(W, \tilde{W})$, so ist $g \circ f \in L(V_1, \ldots, V_n; \tilde{W})$.

- Wenn $f \in L(V_1, \ldots, V_n; W)$ und $h_i \in L(\tilde{V}_i, V_i)$, so ist

$$\begin{aligned} f(h_1 \times \ldots \times h_n) : \tilde{V}_1 \times \ldots \times \tilde{V}_n &\to W \\ (\tilde{v}_1, \ldots, \tilde{v}_n) &\mapsto f(h_1(v_1), \ldots, h_n(v_n)) \end{aligned}$$

  n-linear.

- Sei $f \in L(V_1, \ldots, V_n; W)$ und sei für $1 \leq i \leq n$ $h_i \in L(U_{i,1}, \ldots, U_{i,m_i}; V_i)$, dann ist die Abbildung

$$\begin{aligned} f(h_1 \times \ldots \times h_n) \quad : \quad & L(U_{1,1} \times \ldots \times U_{1,m_1} \times \ldots \times U_{n,1} \times \ldots \times U_{n,m_n}) \to W \\ U_{ij} \quad \mapsto \quad & f(h_1(u_{1,1}, \ldots, u_{1,m_1}), \ldots, h_n(u_{n,1}, \ldots, u_{1,m_n})) \end{aligned}$$

  $m_1 + \ldots + m_n$-linear. Zum Beispiel ist für die Bilinearform f auf V und eine bilineare Abbildung $g : V \times V \to V$ die Abbildung $h : V \times V \times V \to K$, $h(v_1, v_2, v_3) = f(v_1, g(v_2, v_3))$ eine Trilinearform auf V.

- Sei $f : V_1 \times \ldots \times V_n \to W$ $(n-1)$-linear und sei $v_n \in V_n$. Dann ist

$$\begin{aligned} V_1 \times \ldots \times V_{n-1} &\to W \\ (v_1, \ldots, v_{n-1}) &\mapsto f(v_1, \ldots, v_{n-1}, v_n) \text{ } n\text{-linear} \end{aligned}$$

Lemma 1: Für einen Vektorraum V und eine n-lineare Abbildung $l : V \times \ldots \times V \to W$ sei

$$G = \{\pi \in S_n \mid \forall\, x_1, \ldots, x_n \in V : l(x_{\pi(1)}, \ldots, x_{\pi(n)}) = \operatorname{sign}(\pi) \cdot l(x_1, \ldots, x_n)\}$$

Dann ist G eine Untergruppe von S_n. Deshalb sind folgende Aussagen über l äquivalent:

1. $G = S_n$: l multipliziert sich bei Vertauschung seiner Argumente mit dem Vorzeichen der Permutation.
2. G enthält alle Transpositionen: l multipliziert sich mit -1 bei Vertauschung zweier Argumente.
3. G enthält $(i, i+1)$ für $1 \leq i < n$: l multipliziert sich mit -1 bei Vertauschung zweier benachbarter Argumente

Beweis: G ist eine Untergruppe:

- $\mathrm{Id} \in G$, denn $\operatorname{sign}(\mathrm{Id}) = 1$.
- $\sigma, \pi \in G \Rightarrow \sigma \cdot \pi \in G$: Sei $y_i = x_{\sigma(i)}$. Dann gilt $x_{\sigma\pi(i)} = y_{\pi(i)}$. Also

$$\begin{aligned} f(x_{(\sigma\pi)(1)}, \ldots, x_{(\sigma\pi)(n)}) &= f(y_{\pi(1)}, \ldots, y_{\pi(n)}) \\ &\overset{\pi \in G}{=} \operatorname{sign}(\pi) f(y_1, \ldots, y_n) \\ &= \operatorname{sign}(\pi) f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \\ &\overset{\sigma \in G}{=} \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\pi) f(x_1, \ldots, x_n) \\ &\overset{\text{Satz 4.1.1}}{=} \operatorname{sign}(\sigma\pi) f(x_1, \ldots, x_n) \end{aligned}$$

  Also $\sigma\pi \in G$.

- $\pi \in G \Rightarrow \pi^{-1} \in G$: Für $x_1, \ldots, x_n \in V$ sei $y_i = x_{\pi^{-1}(i)}$. Dann gilt:

$$\begin{aligned} f(x_1, \ldots, x_n) &= f(y_{\pi(1)}, \ldots, y_{\pi(n)}) \\ &\overset{\pi \in G}{=} \operatorname{sign}(\pi) f(y_1, \ldots, y_n) \\ &= \operatorname{sign}(\pi) f(x_{\pi^{-1}(1)}, \ldots, x_{\pi^{-1}(n)}) \\ \Rightarrow \quad f(x_{\pi^{-1}(1)}, \ldots, x_{\pi^{-1}(n)}) &= \operatorname{sign}(\pi^{-1}) f(x_1, \ldots, x_n). \end{aligned}$$

  Also $\pi^{-1} \in G$, damit ist G eine Untergruppe.

(1)⇒(2)⇒(3) ist klar. Angenommen (3) gilt. Dann $(i, i+1) \in G$ für $1 \leq i < n$. Nach Folgerung 4.1.2 ist aber jedes Element π von S_n Produkt solcher Transpositionen, also $\pi \in G$, damit gilt (1). □

Bemerkung: Wenn K ein Körper der Charakteristik 2 ist, so gilt $1 = -1$ in K. In solchen Körpern dürfen wir aus $x = -x$ noch nicht $x = 0$ folgern, denn $x = -x$ gilt für alle x aus einem K-Vektorraum. Sei $\operatorname{char}(K) \neq 2$, V ein K-Vektorraum, $x \in V$ und $x = -x$. Dann gilt nach 4.2:

$$x = \frac{1}{2} \cdot 2x = \frac{1}{2}(x + (-x)) = \frac{1}{2} \cdot 0 = 0 \Rightarrow x = 0$$

Lemma 2: Sei f wie l in Lemma 1. Dann sind folgende Aussagen äquivalent:

1. $f(x_1, \ldots, x_n) = 0$, falls ein i mit $1 \leq i < n$ und $x_{i+1} = x_i$ existiert;

2. $f(x_1, \ldots, x_n) = 0$, falls i, j mit $1 \leq i < j < n$ und $x_i = x_j$ existieren.

Wenn das der Fall ist, so gilt $G = S_n$ in Lemma 1. Wenn $\mathrm{char}\,(K) \neq 2$ gilt, so sind (1) und (2) zu $G = S_n$, d.h. zu jedem der Punkte (1),(2),(3) in Lemma 1 äquivalent.

Beweis:

- (2)⇒(1) ist klar.
- (1)⇒ $G = S_n$: Wir werden (3) aus Lemma 1 beweisen: sei $n = 2$. Aus (1) folgt:

$$\begin{aligned} 0 &= f(x_1 + x_2, x_1 + x_2) - f(x_1, x_1) - f(x_2, x_2) \\ &= f(x_1, x_1 + x_2) + f(x_2, x_1 + x_2) - f(x_1, x_1) - f(x_2, x_2) \\ &= f(x_1, x_1) + f(x_1, x_2) + f(x_2, x_1) + f(x_2, x_2) - f(x_1, x_1) - f(x_2, x_2) \\ &= f(x_1, x_2) + f(x_2, x_1) \end{aligned}$$

  Also $f(x_1, x_2) = -f(x_2, x_1)$.
  Sei $n > 2$, $1 \leq i < n$ und $x_1, \ldots, x_n \in V$. Wir haben

$$f(x_1, \ldots, x_n) = -f(x_1, \ldots, x_{i-1}, x_{i+1}, x_i, x_{i+2}, \ldots, x_n)$$

  zu zeigen. Sei $g(x, y) = f(x_1, \ldots, x_{i-1}, x, y, x_{i+2}, \ldots, x_n)$. Dann ist g bilinear und erfüllt (1), also $g(x_i, x_{i+1}) = -g(x_{i+1}, x_i)$ nach dem bewiesenen Fall $n = 2$, also erfüllt f (3) aus Lemma 1.

- (1)⇒(2): Wir wissen $G = S_n$. Sei $1 \leq i < j \leq n$ und sei $x_i = x_j$. Sei $\rho \in S_n$ wie folgt definiert:

$$\rho(k) = \begin{cases} k & 1 \leq k \leq i \vee j < k \leq 1 \\ j & k = i + 1 \\ k - 1 & i + 1 < k \leq j \end{cases}$$

  Sei $y_i = x_{\rho(i)}$. Dann gilt $y_i = y_{i+1}$, also gilt

$$f(x_1, \ldots, x_n) \overset{G=S_n}{=} \mathrm{sign}\,(\rho) f(y_1, \ldots, y_n) \overset{(1)}{=} 0.$$

- $G = S_n$ und $\mathrm{char}\,(K) \neq 2$ ⇒(1): Es gilt wegen $G = S_n$

$$a := f(x_1, \ldots, x_n) = -f(x_1, \ldots, x_{i-1}, x_{i+1}, x_i, x_{i+2}, \ldots, x_n).$$

  Falls $x_i = x_{i+1}$ sein sollte, so gilt

$$\begin{aligned} a &= -f(x_1, \ldots, x_{i-1}, x_{i+1}, x_i, x_{i+2}, \ldots, x_n) \\ &= f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, x_{i+2}, \ldots, x_n) = -a \end{aligned}$$

  Also $a = -a$. Wegen $\mathrm{char}\,(K) \neq 2$ folgt daraus $a = 0$. □

Definition 2: Eine n-lineare Abbildung $f : V \times \ldots \times V \to W$, die die äquivalenten Bedingungen aus Lemma 2 erfüllt, nennt man **schiefsymmetrisch** oder **antisymmetrisch**.

Fakt 2:

1. Wenn l schiefsymmetrisch ist, so ändert es seinen Wert nicht, wenn man ein Vielfaches eines Argumentes zu einem anderen Argument hinzuaddiert.

2. Wenn die $x_1, \ldots, x_n$ linear abhängig sein sollten, so gilt $f(x_1, \ldots, x_n) = 0$).

Beweis:

1.

$$\begin{aligned} f(x_1, \ldots, x_i + \lambda x_j, x_{i+1}, \ldots, x_n) &= f(x_1, \ldots, x_n) + \lambda f \underbrace{(x_1, \ldots, x_{i-1}, x_j, x_{i+1}, \ldots, x_n)}_{=0, \text{ da } j \text{ doppelt}} \\ &= f(x_1, \ldots, x_n) \end{aligned}$$

2. Wegen $G = S_n$ können wir o.B.d.A. annehmen, daß $x_n = \sum_{i=1}^{n-1} \lambda_i x_i$. Also gilt:

$$\begin{aligned} f(x_1, \ldots, x_n) &= f(x_1, \ldots, x_{n-1}, \sum_{i=1}^{n-1} \lambda_i x_i) \\ &= \sum_{i=1}^{n-1} \lambda_i f \underbrace{(x_1, \ldots, x_{n-1}, x_i)}_{x_i \text{ doppelt}} = 0. \end{aligned}$$

Definition 3: Sei l eine schiefsymmetrische n-Linearform auf V und sei $a : V' \to V$ linear. Sei a^*l die durch

$$(a^*l)(v'_1, \ldots, v'_n) = l(a(v'_1), \ldots, a(v'_n))$$

definierte schiefsymmetrische n-Linearform auf V'.

Fakt 3: Für $V \xrightarrow{b} V' \xrightarrow{a} V$ gilt

$$(ab)^*l = b^*a^*l \text{ und } \mathrm{id}^*l = l.$$

Wenn a ein Isomorphismus ist, so ist a^* ein Isomorphismus zwischen den Räumen der n-Formen auf V und V', und $(a^*)^{-1} = (a^{-1})^*$. □

4.4 Der Determinantenbegriff

Sei K ein Körper, V ein n-dimensionaler K-Vektorraum, und sei $B = (b_1, \ldots, b_n)$ eine geordnete Basis von V.

Lemma 1: Sei d eine schiefsymmetrische n-Linearform auf V. Wenn $d(b_1, \ldots, b_n) = 0$, so gilt $d = 0$. Zum Beweis benötigen wir

Fakt 1: Seien V, W Vektorräume, $V \xrightarrow{f} W$ eine n-lineare Abbildung, und sei $v_{i,j} \in V$ für $1 \leq i \leq n$ und $1 \leq j \leq m$, $m \in \mathbf{N}$. Dann gilt

$$f(\sum_{j=1}^{m} v_{i,j}, \ldots, \sum_{j=1}^{m} v_{n,j}) = \sum_{\Phi \in L} f(v_{1,\Phi(1)}, \ldots, v_{n,\Phi(n)}),$$

wobei L die Menge aller Abbildungen von $\{1, \ldots, n\}$ nach $\{1, \ldots, m\}$ ist.

Beweis durch vollständige Induktion nach n. Für $n = 1$ folgt die Behauptung aus der Vertauschbarkeit von Summen und linearen Abbildungen. Sei $n > 1$, und sei die Behauptung für

$(n-1)$-lineare Abbildungen bewiesen. Dann gilt:

$$\begin{aligned} f(\sum_{j=1}^{m} v_{1,j}, \ldots, \sum_{i=1}^{m} v_{n,i}) &= \sum_{i=1}^{m} f(\sum_{j=1}^{n} v_{1,j}, \ldots, \sum_{j=1}^{m} v_{n-1,j}, v_{n,i}) \\ &= \sum_{j=1}^{m} \sum_{\Psi \in L'} f(v_{1,\Psi(1)}, \ldots, v_{n-1,\Psi(n-1)}, v_{n,j}) \\ &= \sum_{j=1}^{m} \sum_{\Phi \in L} f(v_{1,\Phi(1)}, \ldots, v_{n-1,\Phi(n-1)}, v_{n,\Phi(n)}). \end{aligned}$$

Dabei ist $L' = \{1, \ldots, m\}^{\{1,\ldots,n-1\}}$. Durch $L \to L' \times \{1, \ldots, m\}$, $\Phi \mapsto (\Psi, i)$ mit $\Psi = \Phi|_{\{1,\ldots,n-1\}}$ und $i = \Phi(n)$ ist eine Bijektion definiert, daher die Gleichheit der beiden letzten Zeilen. □

Beweis von Lemma 1: Sei $v_i = \sum_{j=1}^{n} \lambda_{ij} b_j$. Aus Fakt 1 folgt

$$d(v_1, \ldots, v_n) = \sum_{\Phi \in L} \lambda_{1,\Phi(1)} \cdot \ldots \cdot \lambda_{n,\Phi(n)} \cdot d(b_{\Phi(1)}, \ldots, b_{\Phi(n)}),$$

wobei L die Menge aller Abbildungen von $\{1, \ldots, n\}$ nach $\{1, \ldots, n\}$ ist. Wenn $\Phi \in L$ nicht bijektiv ist, so gibt es $1 \leq i < j \leq n$ mit $\Phi(i) = \Phi(j)$, also taucht in $d(b_{\Phi(1)}, \ldots, b_{\Phi(n)})$ wenigstens ein Argument doppelt auf, und $d(b_{\Phi(1)}, \ldots, b_{\Phi(n)}) = 0$. Es folgt

$$\begin{aligned} d(v_1, \ldots, v_n) &= \sum_{\Phi \in S_n} \lambda_{1,\Phi(1)} \cdot \ldots \cdot \lambda_{n,\Phi(n)} \cdot d(b_{\Phi(1)}, \ldots, b_{\Phi(n)}) \\ &= \sum_{\Phi \in S_n} \lambda_{1,\Phi(1)} \cdot \ldots \cdot \lambda_{n,\Phi(n)} \cdot \operatorname{sign}(\Phi) d(b_1, \ldots, b_n), \end{aligned}$$

nach den Resultaten aus 4.3. □

Folgerung 1: Für einen n-dimensionalen Vektorraum V sei $D(V)$ der Raum aller schiefsymmetrischen n-Linearformen auf V. Dann ist $D(V)$ höchstens eindimensional. Wenn also $d \in D(V) - \{0\}$ ist, so ist jedes $\tilde{d} \in D(V)$ zu d proportional.

Beweis: In der Tat, für $d \in D(V) - \{0\}$ und $\tilde{d} \in D(V)$ sei $d' = \tilde{d} - \frac{\tilde{d}(b_1,\ldots,b_n)}{d(b_1,\ldots,b_n)} \cdot d$. Dann gilt $d'(b_1, \ldots, b_n) = 0$, also $d' = 0$, also ist $\tilde{d}$ zu d proportional. Wir wollen zeigen, daß die Dimension von $D(V)$ auch genau eins ist. Sei d_B durch

$$d_B(v_1, \ldots, v_n) = \sum_{\Phi \in S_n} \operatorname{sign}(\Phi) \lambda_{1,\Phi(1)} \cdot \ldots \cdot \lambda_{n,\Phi(n)} \tag{1}$$

für $v_i = \sum_{j=1}^{n} \lambda_{ij} b_j$ definiert. Dann ist d_B eine n-Linearform. Angenommen, $v_i = v_{i+1}$. Sei M die Menge aller $\Phi \in S_n$ mit $\Phi(i) < \Phi(i+1)$. Dann kann jedes Element von S_n als Ψ oder als $\Psi(i, i+1)$ geschrieben werden, wobei $\Psi \in M$ ist. Also kann (1) zu

$$\begin{aligned} d_B(v_1, \ldots, v_n) = \sum_{\Psi \in M} \operatorname{sign}(\Psi) & \underbrace{\lambda_{1,\Psi(1)} \cdot \ldots \cdot \lambda_{i,\Psi(i)} \cdot \lambda_{i+1,\Psi(i+1)} \cdot \ldots \cdot \lambda_{n,\Psi(n)}}_{\text{Beitrag von } \Psi \text{ zu (1)}} \\ & \underbrace{- \lambda_{1,\Psi(1)} \cdot \ldots \cdot \lambda_{i,\Psi(i+1)} \cdot \ldots \cdot \lambda_{i+1,\Psi(i)} \cdot \ldots \cdot \lambda_{n,\Psi(n)}}_{\text{Beitrag von } \Psi(i,i+1) \text{ zu (1)}} = 0 \end{aligned}$$

umgeschrieben werden, denn $\lambda_{i,\Psi(i)} = \lambda_{i+1,\Psi(i)}$ und $\lambda_{i+1,\Psi(i+1)} = \lambda_{i,\Psi(i+1)}$, also

$$\lambda_{i,\Psi(i)} \lambda_{i+1,\Psi(i+1)} = \lambda_{i,\Psi(i+1)} \lambda_{i+1,\Psi(i)}.$$

Damit $d_B \in D(V)$.
Behauptung: $d_B(b_1, \ldots, b_n) = 1$. Es folgt $d_B \in D(V) - \{0\}$ und $D(V) \neq \{0\}$. In der Tat, für $v_i = b_i$ gilt $\lambda_{ij} = \delta_{ij}$ und (1) ergibt

$$d_B(b_1, \ldots, b_n) = \sum_{\Phi \in S_n} \operatorname{sign}(\Phi) \cdot \underbrace{\delta_{i,\Phi(1)} \cdot \ldots \cdot \delta_{i,\Phi(n)}}_{=0 \text{ für } \Phi \neq \mathrm{Id}} = 1,$$

denn für $\Phi \neq \mathrm{Id}$ gibt es j mit $\Phi(j) \neq j$, $1 \leq j \leq n$. □

Theorem 1:

1. $D(V)$ ist eindimensional
2. Für jeden Endomorphismus f von V gibt es genau ein $\det(f) \in K$ mit $f^*(d) = \det(f) \cdot d$ für alle $d \in D(V)$.

Beweis: (1) wurde soeben bewiesen. (2) folgt aus dem folgenden einfachen

Fakt: Sei g ein Endomorphismus eines eindimensionalen Vektorraumes D. Dann gibt es genau ein $y \in K$ mit $g(d) = yd$ für alle $d \in D$. □

Folgerung 2:

1. $\det(fg) = \det(f) \cdot \det(g)$ für $f, g \in \mathrm{End}(V)$.
2. $\det(\lambda f) = \lambda^n \det(f)$, $n = \dim(V)$.
3. $\det(0) = 0, \det(\mathrm{Id}_V) = 1$.
4. Sei $a : V \to W$ ein Isomorphismus und $f \in \mathrm{End}(V)$, dann gilt $\det(afa^{-1}) = \det(f)$.

Beweis:

1. Für $d \in D(V)$ gilt:

$$(fg)^* d \overset{\text{Fakt } 4.3.3}{=} g^*(f^*(d)) = \det(g) f^*(d) = \det(g) \det(f) d.$$

   Also $\det(fg) = \det(f) \det(g)$ nach der Eindeutigkeitsaussage in Theorem 1.

2. Zusammen mit dem nächsten Punkt.

3. Es genügt nach (1), $\det(\lambda \mathrm{Id}_V) = \lambda^n$ zu zeigen.

$$\begin{aligned} ((\lambda(\mathrm{Id}_V)^* d)(v_1, \ldots, v_n) &\overset{\text{Def } 4.3.3}{=} d((\lambda(\mathrm{Id}_V)(v_1), \ldots, (\lambda(\mathrm{Id}_V)(v_n)) \\ &= d(\lambda v_1, \ldots, \lambda v_n) \\ &= \lambda^n d(v_1, \ldots, v_n). \end{aligned}$$

   Damit sind (2) und (3) gezeigt.

4. Sei $d \in D(W)$. Es gilt

$$\begin{aligned}(afa^{-1})^*(d) &= (a^{-1})^*(f^*a^*(d)) \\ &= (a^{-1})^*(\det(f)a^*(d))) \\ &= \det(f)((a^*)^{-1})a^*(d)) \\ &= \det(f)d,\end{aligned}$$

wobei wir Fakt 4.3.3 verwendet haben. □

Satz 1: Sei $d \in D(V) - \{0\}$. Dann sind n Vektoren $v_1, \ldots, v_n \in V$ genau dann linear abhängig, wenn $d(v_1, \ldots, v_n) = 0$ gilt.

Beweis: Wenn die v_i linear unabhängig sind, so ist $C = (v_1, \ldots, v_n)$ eine geordnete Basis von V . Wegen Lemma 1 (angewendet auf C statt B) würde aus $d(v_1, \ldots, v_n) = 0$ $d = 0$ folgen, was nicht sein sollte. Wenn die v_i linear abhängig sind, so folgt $d(v_1, \ldots, v_n) = 0$ aus Fakt 4.3.2.

Folgerung 3: (Aus Theorem 1 und Satz 1)

1. Sei $d \in D(V) - \{0\}$, dann gilt

$$\det(f) = \frac{d(f(b_1), \ldots, f(b_n))}{d(b_1, \ldots, b_n)}.$$

2. Insbesondere: $\det(f) = d_B(f(b_1), \ldots, f(b_n))$.

3. $f \in \mathrm{End}\,(V)$ ist genau dann ein Automorphismus, wenn $\det(f) \neq 0$ gilt. In diesem Fall gilt $\det(f^{-1}) = \det(f)^{-1}$.

Beweis:

1. $d(f(b_1), \ldots, f(b_n)) = (f^*d)(b_1, \ldots, b_n) = \det(f)d(b_1, \ldots, b_n)$.

2. Wende (1) auf d_B an und beachte $d_B(b_1, \ldots, b_n) = 1$.

3. f ist Automorphismus $\Leftrightarrow$ $f(b_1), \ldots, f(b_n)$ linear unabhängig $\Leftrightarrow$ $d_B(f(b_1), \ldots, f(b_n)) \neq 0$. Die Gleichheit $\det(f^{-1}) = \det(f)^{-1}$ folgt aus Folgerung 2. □

4.5 Determinanten von Matrizen

Definition 1: Sei $A = (a_{ij})_{i,j=1}^n \in \mathrm{Mat}\,(n, n, K)$ und sei $A\cdot : K^n \to K^n$ der in 3.1 definierte Operator. Wir setzen $\det(A) = \det(A\cdot)$ und schreiben auch

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} := \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Folgerung 1: Für $A, B \in \mathrm{Mat}\,(n, n, K)$ gilt:

1. $\det(AB) = \det(A) \cdot \det(B)$;

2. $\det(\lambda A) = \lambda^n \det(A)$;

3. $\det(0) = 0$;

4. $\det(\mathbf{1}_n) = 1$.

Außerdem ist A invertierbar genau dann, wenn $\det(A) \neq 0$. In diesem Fall gilt, daß $\det(A^{-1}) = \det(A)^{-1}$

Beweis: Die Behauptungen folgen aus Folgerung 4.4.2 und 4.4.3 und den in 3.1. erklärten Zusammenhängen zwischen Matrizen und linearen Abbildungen. □

Satz 1: (LEIBNIZsche Determinantenformel)

$$\begin{aligned}\det((a_{ij})_{i,j=1}^n) &= \sum_{\pi\in S_n} \operatorname{sign}(\pi)\cdot a_{1,\pi(1)}\cdot a_{2,\pi(2)}\cdot\ldots\cdot a_{n,\pi(n)} \\ &= \sum_{\rho\in S_n} \operatorname{sign}(\rho)\cdot a_{\rho(1),1}\cdot a_{\rho(2),2}\cdot\ldots\cdot a_{\rho(n),n}\end{aligned}$$

Beweis: Sei $e_i = (\delta_{ij})_{j=1}^n$. Damit folgt mit Folgerung 4.4.3 und $A\cdot e_j = \sum_{i=1}^n a_{ij}e_j$

$$\begin{aligned}\det(A) &= d_{(e_1,\ldots,e_n)}(Ae_1,\ldots,Ae_n) \\ &= d_{(e_1,\ldots,e_n)}\left(\sum_{i=1}^n a_{i1}e_i,\ldots,\sum_{i=1}^n a_{in}e_i\right) \\ &= \sum_{\rho\in S_n} \operatorname{sign}(\rho)\cdot a_{\rho(1),1}\cdot a_{\rho(2),2}\cdot\ldots\cdot a_{\rho(n),n}\end{aligned}$$

mit Gleichung 4.4.1. Wenn wir $\pi = \rho^{-1}$ einsetzen, so gilt $\operatorname{sign}(\pi) = \operatorname{sign}(\rho)$, und der letzte Ausdruck wird zu

$$\begin{aligned}&\sum_{\pi\in S_n} \operatorname{sign}(\pi)\cdot a_{\pi^{-1}(1),1}\cdot a_{\pi^{-1}(2),2}\cdot\ldots\cdot a_{\pi^{-1}(n),n} \\ = &\sum_{\pi\in S_n} \operatorname{sign}(\pi)\cdot a_{1,\pi(1)}\cdot a_{2,\pi(2)}\cdot\ldots\cdot a_{n,\pi(n)}\end{aligned}$$

□

Folgerung 2: $\det(A^T) = \det(A)$

Beweis: Sei $A = (a_{ij})_{i,j=1}^n$ und $A^T = (a_{ij}^T)$. Es gilt $a_{ij}^T = a_{ji}$.

$$\begin{aligned}\det(A^T) &= \sum_{\pi\in S_n} \operatorname{sign}(\pi)\cdot a_{1,\pi(1)}^T\cdot a_{2,\pi(2)}^T\cdot\ldots\cdot a_{n,\pi(n)}^T \\ &= \sum_{\pi\in S_n} \operatorname{sign}(\pi)\cdot a_{\pi(1),1}\cdot a_{\pi(2),2}\cdot\ldots\cdot a_{\pi(n),n} \\ &= \det(A)\end{aligned}$$

□

Beispiel: Für $n = 2$ gilt $S_2 = \{\mathrm{Id}, (1,2)\}$, wobei $\operatorname{sign}(\mathrm{Id}) = 1$ und $\operatorname{sign}((1,2)) = -1$. Die rechte Seite der Leibnizschen Determinantenformel wird zu

$$\underbrace{a_{11}a_{22}}_{\pi=\mathrm{Id}} - \underbrace{a_{12}a_{21}}_{\pi=(1,2)} \text{ oder } \det\begin{pmatrix} a & b \\ c & d\end{pmatrix} = \begin{vmatrix} a & b \\ c & d\end{vmatrix} = ad - bc.$$

Für $n = 3$ gilt die **Regel von Sarrus**: Schreibe die beiden ersten Spalten hinter die Matrix. Bilde die Summe der drei Diagonalen von links oben nach rechts unten und subtrahiere von dieser die Summe der drei Diagonalen von links unten nach rechts oben.

$$\begin{pmatrix} a_{11} & & a_{12} & & a_{13} & & a_{11} & & a_{12} \\ & \searrow & & \searrow\!\!\!\searrow & & \searrow\!\!\!\searrow & & \nearrow & \\ a_{21} & & a_{22} & & a_{23} & & a_{21} & & a_{22} \\ & \nearrow & & \searrow\!\!\!\searrow & & \searrow\!\!\!\searrow & & \searrow & \\ a_{31} & & a_{32} & & a_{33} & & a_{31} & & a_{32} \end{pmatrix}$$

$$\det(A) = a_{12}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{13}a_{11} - a_{33}a_{21}a_{12}.$$

Für allgemeines n hat die Leibnizformel $n!$ Summanden, jeder Summand hat n Faktoren, man braucht also $n!\cdot(n-1)$ Multiplikationen. Für große n wird die Formel daher schnell unpraktikabel.

Definition 2: $\mathrm{Mat}_z(v_1, \ldots, v_m)$ ist die Matrix mit den Zeilenvektoren v_i:

$$\mathrm{Mat}_z((v_{11}, \ldots, v_{1n}), \ldots, (v_{m1}, \ldots, v_{mn})) = \begin{pmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{m1} & \cdots & v_{mn} \end{pmatrix}.$$

Für $w_1, \ldots, w_n \in K^m$ sei $\mathrm{Mat}_s(w_1, \ldots, w_n)$ die Matrix mit den Spaltenvektoren w_i:

$$\mathrm{Mat}_s(w_1, \ldots, w_n) = \mathrm{Mat}_z(w_1, \ldots, w_n)^T.$$

Satz 2: $\det(\mathrm{Mat}_s(v_1, \ldots, v_n)) = \det(\mathrm{Mat}_z(v_1, \ldots, v_n))$ ist eine schiefsymmetrische n-Linearform in den Argumenten $v_i \in K^n$.

Beweis: Es gilt nach Folgerung 4.4.3:

$$\begin{aligned} \det(\mathrm{Mat}_s(v_1, \ldots, v_n)) &= d_{(e_1,\ldots,e_n)}(\mathrm{Mat}_s(v_1, \ldots, v_n) \cdot e_1, \ldots, \mathrm{Mat}_s(v_1, \ldots, v_n) \cdot e_n) \\ &= d_{(e_1,\ldots,e_n)}(v_1, \ldots, v_n) \end{aligned}$$

denn $\mathrm{Mat}_s(v_1, \ldots, v_n)e_i = v_i$. Wegen Folgerung 2 und $\mathrm{Mat}_z(v_1, \ldots, v_n) = \mathrm{Mat}_s(v_1, \ldots, v_n)^T$ berechnet dieser Ausdruck auch die Determinante von $\mathrm{Mat}_z(v_1, \ldots, v_n)$. Weil $d_{(e_1,\ldots,e_n)}$ eine schiefsymmetrische n-Linearform ist, folgt die Behauptung. □

Folgerung 3: Die Determinate einer Matrix

1. ändert sich nicht, wenn man ein Vielfaches einer Zeile zu einer anderen Zeile addiert.
2. multipliziert sich mit λ, wenn man eine Zeile mit λ multipliziert.
3. multipliziert sich mit -1, wenn man zwei Zeilen vertauscht.

Analoges gilt für die entsprechenden Spaltenumformungen. □

Definition 3:

1. Für ein n-Tupel $x = (x_1, \ldots, x_n)$ und $1 \leq k \leq n$ sei

$$\begin{aligned} \mathrm{Min}_k(x_1, \ldots, x_k) &= (x_1, \ldots, \hat{x_k}, \ldots, x_n) \\ &:= \begin{cases} (x_2, \ldots, x_n) & k = 1 \\ (x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) & 1 < k < n \\ (x_1, \ldots, x_{n-1}) & k = n \end{cases} \end{aligned}$$

   das durch Streichen des k-ten Elements entstehende $(n-1)$-Tupel.

2. Für eine $m \times n$-Matrix $A = \mathrm{Mat}_z(a_1, \ldots, a_m)$ und $1 \leq i \leq m$, $1 \leq j \leq n$, sei

$$\mathrm{Min}_{ij}(A) = \mathrm{Mat}_z(\mathrm{Min}_j(a_1), \ldots, \widehat{\mathrm{Min}_j}(a_i), \ldots, \mathrm{Min}_j(a_m))$$

   die durch Streichen der i-ten Zeile und der j-ten Spalte entstehende Matrix.

Offenbar sind die Abbildungen $\mathrm{Min}_i : K^n \to K^{n-1}$ und $\mathrm{Min}_{ij} : \mathrm{Mat}(m, n, K) \to \mathrm{Mat}(n-1, m-1, K)$ linear.

Satz 3: (LAPLACE) Für jede $n \times n$-Matrix $A = (a_{ij})_{i,j=1}^n$ und $1 \leq k \leq n$ gelten die Entwicklungen nach der k-ten Zeile

$$\det(A) = \sum_{i=1}^{n} (-1)^{k+i} a_{ki} \cdot \det(\mathrm{Min}_{ki}(A)) \tag{1}$$

und die Entwicklung nach der k-ten Spalte

$$\det(A) = \sum_{i=1}^{n} (-1)^{k+i} a_{ik} \cdot \det(\mathrm{Min}_{ik}(A)). \tag{2}$$

Beispiel:

$$\begin{vmatrix} 3 & 2 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{vmatrix} = 3 \cdot \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} - 2 \cdot \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} + \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} = 1$$

Beweis von Satz 3: Die Entwicklung nach einer Spalte bekommt man durch Anwendung der Zeilenentwicklung auf die transponierte Matrix (Folgerung 2). Es genügt also, (1) zu beweisen. Wir fixieren dazu k und bezeichnen die rechte Seite von (1) mit $\mathrm{RS}(A)$. Für $x = (x_1, \ldots, x_n) \in K^n$ sei $(K^n)^* \ni l_i(x) = x_i$.

1. Wir behaupten, daß

$$\mathrm{RS}(\mathrm{Mat}_s(v_1, \ldots, v_n)) \tag{$*$}$$

   eine n-Linearform in $v_1, \ldots, v_n \in K^n$ ist. In der Tat, für $1 \leq i \leq n$ ist die i-te Summe in (1) für $A = \mathrm{Mat}_s(v_1, \ldots, v_n)$ durch

$$\mathrm{RS}_i(v_1, \ldots, v_n) = (-1)^{k+i} l_k(v_i) \cdot \det(\mathrm{Mat}_s(\mathrm{Min}_k(v_1), \ldots, \widehat{\mathrm{Min}_k}(v_i), \ldots, \mathrm{Min}_k(v_n))$$

gegeben. Dann ist RS_i eine n-Linearform in $v_1, \ldots, v_n$, denn $\det(\mathrm{Mat}_s(x_1, \ldots, \hat{x}_k, \ldots, x_n))$ ist $(n-1)$-linear und Min_k ist linear, also ist $\det(\mathrm{Mat}_s(\mathrm{Min}_k(v_1), \ldots, \widehat{\mathrm{Min}_k(v_i)}, \ldots, \mathrm{Min}_k(v_n))$ eine $(n-1)$-Linearform auf K^n, und die Abbildungen $K \times K \to K$, $(x, y) \mapsto \pm xy$ sind bilinear. Nach Beispiel 4.3.1 ist daher $\mathrm{RS}_i(v_1, \ldots, v_n)$ n-linear, und daher ist $\mathrm{RS}(\mathrm{Mat}_s(v_1, \ldots, v_n)) = \sum_{i=1}^n \mathrm{RS}_i(v_1, \ldots, v_n)$ n-linear.

2. $(*)$ ist eine schiefsymmetrische n-Linearform auf K^n: Das ist äquivalent dazu, daß $\mathrm{RS}(A)$ verschwindet, wenn in A zwei benachbarte Spalten übereinstimmen, etwa die l-te und die $(l+1)$-te für $1 \leq l < n$ (Lemma 4.3.2). Dann stimmen für $i \notin \{l, l+1\}$ auch in $\mathrm{Min}_{ki}(A)$ zwei Spalten überein, nämlich die l-te und die $(l+1)$-te für $i > l+1$ und die l-te und die $(l-1)$-te für $i < l$. Daher gilt $\forall\, i \neq \{l, l+1\} : \det(\mathrm{Min}_{ki}(A)) = 0$ und wir erhalten

$$\mathrm{RS}(A) = (-1)^{k+l}(a_{kl} \det(\mathrm{Min}_{kl}(A)) - a_{k,l+1} \det(\mathrm{Min}_{k,l+1}(A))). \qquad (\dagger)$$

   Aus unseren Vorraussetzungen an A folgt aber $a_{kl} = a_{k,l+1}$ und $\mathrm{Min}_{kl}(A) = \mathrm{Min}_{k,l+1}(A)$, also verschwindet $(\dagger)$ und die Behauptung aus Schritt 2 ist bewiesen.

3. Aus Schritt 2 und Theorem 4.4.1 folgt die Existenz eines $c \in K$ mit

$$\mathrm{RS}(A) = c \cdot \det(A) \text{ für alle } n \times n\text{-Matrizen} A.$$

   Wir haben noch $c = 1$ zu zeigen. Dazu genügt es, $\mathrm{RS}(\mathbf{1}_n) = 1$ zu zeigen. Da $\mathbf{1}_n = (\delta_{ij})_{i,j=1}^n$ gilt, ergibt sich

$$\begin{aligned} \mathrm{RS}(\mathbf{1}_n) &= \sum_{i=1}^n (-1)^{k+i} \delta_{ki} \det(\mathrm{Min}_{ki}(\mathbf{1}_n)) \\ &= (-1)^{k+k} \delta_{k,k} \det(\mathrm{Min}_{k,k}(\mathbf{1}_n)) \\ &= 1 \cdot 1 \det(\mathbf{1}_{n-1}) = 1 \end{aligned}$$

□

Fakt 1: Sei $A = (a_{ij})_{i,j=1}^n$ eine obere Dreiecksmatrix, d.h. $a_{ij} = 0$ für alle $j < i$. Dann ist $\det(A)$ das Produkt der Diagonalelemente, $\det(A) = a_{11} \cdot \ldots \cdot a_{nn}$.

Beweis: Entwicklung nach der ersten Spalte ergibt

$$\begin{vmatrix} a_{11} & a_{12} & \ldots \\ 0 & a_{22} & \\ 0 & 0 & \ddots \\ \vdots & & \ddots \end{vmatrix} = a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} & \ldots \\ 0 & a_{33} & \\ 0 & 0 & \ddots \\ \vdots & & \ddots \end{vmatrix} = \cdots = a_{11} \cdot \ldots \cdot a_{nn}. \qquad \square$$

Bemerkung: Insbesondere kann die Determinante einer Zeilenstufenmatrix (vgl. Abschnitt 3.2) nach dieser Vorschrift berechnet werden. Von den in 3.2 benutzten Zeilenumformungen ändert die Addition des Vielfachen einer Zeile zu einer anderen die Determinante nicht (Folgerung 3), und die Vertauschung zweier Zeilen ändert nur das Vorzeichen. Daher erhält man folgende Vorschrift zur Berechnung von $\det A$:

1. Bringe A durch Anwendung des in 3.2 beschriebenen Eliminationsverfahrens in Zeilenstufenform. Sei $\tilde{A}$ das Ergebnis, und sei l die Anzahl der Zeilenvertauschungen, die dabei benötigt wurden.

2. Es gilt dann mit $\tilde{A} = (\tilde{a}_{ij})_{i,j=1}^{n}$

$$\det(A) = (-1)^l \det(\tilde{A}) = (-1)^l \cdot \tilde{a}_{11} \cdot \ldots \cdot \tilde{a}_{nn}.$$

Dieser Algorithmus kostet $\frac{n^3}{3} + O(n^2)$ Multiplikationen. (Zum Vergleich die Kosten für die Anwendung der Leibniz-Formel: $(n-1) \cdot n!$).
Vom rechnerischen Standpunkt aus betrachtet ist die Berechnung einer Determinante durch wiederholte Anwendung des Entwicklungssatzes nur eine klügere Organisation der Berechnung der Produkte, die in der Leibniz-Formel auftauchen.

Satz 4: (VANDERMONDsche Determinante) Für $\lambda_1, \ldots, \lambda_n \in K$ gilt

$$\begin{vmatrix} 1 & \ldots & 1 \\ \lambda_1 & \ldots & \lambda_n \\ \lambda_1^2 & \ldots & \lambda_n^2 \\ \vdots & & \vdots \\ \lambda_1^{n-1} & \ldots & \lambda_n^{n-1} \end{vmatrix} = \prod_{1 \le i < j \le n} (\lambda_j - \lambda_i).$$

Beweis: Sei $z_i = (\lambda_1^{i-1}, \ldots, \lambda_n^{i-1})$ für $1 \le i \le n$ und $A = \mathrm{Mat}_z(z_1, \ldots, z_n)$. Sei $\tilde{z}_1 = z_1$ und $\tilde{z}_i = z_i - \lambda_1 z_{i-1}$ für $1 < i \le n$. Sei $\tilde{A} = \mathrm{Mat}_z(\tilde{z}_1, \ldots, \tilde{z}_n)$. Zu zeigen: $\det(A) = \det(\tilde{A})$.
Entwicklung von $\tilde{A}$ nach der ersten Spalte ergibt

$$\det(\tilde{A}) = (\lambda_2 - \lambda_1) \cdot \ldots \cdot (\lambda_n - \lambda_1)$$

(Siehe auch Übungsaufgaben). □

Beispiel:

$$\begin{vmatrix} 5 & 4 & 3 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 3 & 4 \\ 1 & 1 & 4 & 9 & 16 \\ 1 & 1 & 8 & 27 & 64 \end{vmatrix} = (5-4) \cdot \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \\ 1 & 8 & 27 & 64 \end{vmatrix}$$

$$= 1 \cdot (4-3)(4-2)(4-1)(3-2)(3-1)(2-1)$$

$$= 12$$

Definition 4: Die **Adjunkte** $\mathrm{Adj}\,(A)$ einer $n \times n$-Matrix A ist die Matrix $\mathrm{Adj}\,(A) = (\mathrm{adj}_{ij}(A))_{i,j=1}^{n}$ mit

$$\mathrm{adj}_{ij}(A) = (-1)^{i+j} \cdot \det(\mathrm{Min}_{ji}(A)).$$

Satz 5: (CRAMERsche Regel von 1750) Für alle $n \times n$-Matrizen A gilt:

$$A \cdot \mathrm{Adj}\,(A) = \mathrm{Adj}\,(A) \cdot A = \det(A) \cdot \mathbf{1}_n.$$

Insbesondere:

$$A^{-1} = \frac{\mathrm{Adj}\,(A)}{\det(A)} \text{ falls } \det(A) \neq 0.$$

Beispiel: Für $n = 2$ gilt:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Beweis von Satz 5: Wir zeigen $A \cdot \mathrm{Adj}\,(A) = (\gamma_{i,j})_{i,j=1}^{n} = (\delta_{ij} \det A)_{i,j=1}^{n}$ mit

$$\gamma_{ij} = \sum_{k=1}^{n} a_{ik} \cdot \mathrm{adj}_{\,kj}(A) \qquad (*)$$

mit $A = (a_{ij})_{i,j=1}^{n}$. Dazu fixieren wir i und j. Sei B die Matrix, die aus A durch Ersetzen der j-ten durch die i-te Zeile entsteht. Für $i \neq j$ stimmen die i-te und die j-te Zeile von B überein, also $\det(B) = 0$. Es folgt in beiden Fällen:

$$\det(B) = \det(A) \cdot \delta_{ij}. \qquad (\dagger)$$

Sei $B = (b_{kl})_{k,l=1}^{n}$. Dann gilt $b_{jk} = a_{ik}$ und $\mathrm{Min}_{\,jk}(B) = \mathrm{Min}_{\,jk}(A)$ für $1 \leq k \leq n$. Entwicklung von $\det(B)$ nach der j-ten Zeile gibt also:

$$\begin{aligned}
\delta_{ij} \cdot \det(A) &\overset{(\dagger)}{=} \det(B) \\
&\overset{\text{Satz 3}}{=} \sum_{k=1}^{n} (-1)^{j+k} \cdot b_{jk} \det(\mathrm{Min}_{\,jk}(B)) \\
&= \sum_{k=1}^{n} a_{ik} \cdot (-1)^{j+k} \cdot \det(\mathrm{Min}_{\,jk}(A)) \\
&= \sum_{k=1}^{n} a_{ik} \cdot \mathrm{adj}_{\,kj}(A) \overset{(*)}{=} \gamma_{ij}
\end{aligned}$$

Also $A \cdot \mathrm{Adj}\,(A) = \det(A) \cdot \mathbf{1}_n$. Die Gleichung $\mathrm{Adj}\,(A) \cdot A = \det(A) \cdot \mathbf{1}_n$ zeigt man entweder analog durch Anwendung derselben Betrachtungen auf Spalten statt Zeilen, oder durch Anwendung von $A^T \mathrm{Adj}\,(A^T) = \det(A^T)\mathbf{1}_n$, denn $\mathrm{Adj}\,(A)^T = \mathrm{Adj}\,(A^T)$ und $\det(A) = \det(A^T)$. Die Formel für A^{-1} folgt unmittelbar. □

Bemerkung: Sei $A \in \mathrm{Mat}\,(m, m, K)$. Sei $1 \leq k \leq \min(m, n)$. Ein $k \times k$-**Minor** von A ist eine $k \times k$-Matrix B, die aus A durch Streichung von $(m - k)$ Zeilen und $(n - k)$ Spalten entsteht. Sei

$$r = \begin{cases} \max\{k \mid \exists\, k \times k\text{-Minor } B \text{ von } A \text{ mit } \det(B) \neq 0\} & \text{für } A \neq 0 \\ 0 & \text{sonst} \end{cases}.$$

Dann stimmt r mit dem Rang von A überein.

4.6 Orientierungen von reellen Vektorräumen

Definition 1: Sei X eine Menge. Eine **Relation** auf X ist eine Untermenge $R \subseteq X \times X$. Wir sagen, daß $x \in X$ in der Relation R zu $y \in X$ steht, und schreiben xRy, falls $(x, y) \in R$. R ist

1. **reflexiv**, falls aus $x \in X$ xRx folgt.
2. **symmetrisch**, falls aus $x, y \in X$ und xRy auch yRx folgt.
3. **transitiv**, falls aus $x, y, z \in X$ und xRy und yRz auch xRz folgt.

R ist eine **Äquivalenzrelation**, falls R reflexiv, transitiv und symmetrisch ist.

Fakt: Sei $\sim$ Äquivalenzrelation auf X. Eine **Äquivalenzklasse** bezüglich $\sim$ ist eine nichtleere Untermenge Y von X mit folgenden Eigenschaften:

1. Aus $x, y \in Y$ folgt $x \sim y$.
2. Aus $y \in Y$ und $x \in X$ und $x \sim y$ folgt $x \in Y$.

Sei $X/\sim$ die Menge aller Äquivalenzklassen bezüglich $\sim$. Dann ist X disjunkte Vereinigung aller Äquivalenzklassen:

$$X = \bigcup_{Y \in X/\sim} Y. \qquad (*)$$

Beweis: Disjunktheit: Seien Y und $\tilde{Y}$ zwei Äquivalenzklassen bezüglich $\sim$. Angenommen, $Y \cap \tilde{Y} \neq \emptyset$, etwa $y \in Y \cap \tilde{Y}$. Für $x \in X$ gilt dann

$$x \in Y \Leftrightarrow x \sim y \Leftrightarrow x \in \tilde{Y},$$

also $Y = \tilde{Y}$ und die Vereinigung $(*)$ ist disjunkt.
Sei weiterhin $x \in X$, $Y = \{y \in X | x \sim y\}$. Dann ist Y Äquivalenzklasse:

- Sei $y, z \in Y$. Dann gilt $x \sim y$ und $x \sim z$ nach Definition von Y , also $y \sim x$ (Symmetrie), also auch $y \sim z$ (Transitivität).
- Sei $y \in Y$ und $y \sim z$. Dann gilt $x \sim y$ (Definition von Y), also auch $x \sim z$ (Transitivität), damit $z \in Y$.

Wegen $x \in Y$ (Reflexivität) gilt $Y \neq \emptyset$, also $Y \in X/\sim$. Also $x \in Y \subseteq \bigcup_{Z \in X/\sim} Z$. □

Beispiel: Sei $X \in \mathbf{Z}$. Für eine natürliche Zahl N betrachten wir die Relation $x \sim y$ genau dann, wenn $x \equiv y \bmod N$, d.h., falls $N \mid x - y$. Man prüft leicht nach, daß es sich um eine Äquivalenzrelation handelt. Die Äquivalenzklassen sind die Restklassen modulo N, sie bilden die Menge $\mathbf{Z}/N\mathbf{Z}$. Wenn $N = P$ eine Primzahl ist, ist $\mathbf{Z}/p\mathbf{Z} = \mathbf{F}_p$.

Definition 2: Sei V ein n-dimensionsionaler reeller Vektorraum. Eine **Halbachse** in V ist eine Äquivalenzklasse bezüglich der folgenden Äquivalenzrelation $\sim$ auf $V - \{0\}$:

$$x \sim y :\Leftrightarrow \exists\, t \in \mathbf{R} : t > 0 \wedge x = t \cdot y$$

Fakt:

1. $\sim$ ist in der Tat eine Äquivalenzrelation.
2. Gilt $n = 1$, so hat die Menge $(V - \{0\})/\sim$ genau zwei Elemente. Für $v \in V - \{0\}$ sind dies die beiden Äquivalenzklassen

$$\begin{aligned} A &= \{\lambda v | \lambda \in \mathbf{R}, \lambda > 0\} \text{ und} \\ B &= \{\lambda v | \lambda \in \mathbf{R}, \lambda < 0\}. \end{aligned}$$

Beweis:

1. - *Reflexivität:* $x = 1 \cdot x$, also $x \sim x$.
   - *Symmetrie:* Aus $x \sim y$ folgt $x = ty$ mit $t > 0$, also $y = t^{-1}x$ mit $t^{-1} > 0$, also $y \sim x$.
   - *Transitivität:* Aus $x \sim y$ und $y \sim z$ folgt $x = ty$ und $y = sz$ mit $s, t > 0$, also $x = (st)z$ mit $st > 0$, also $x \sim z$.

2.
- *A ist Äquivalenzklasse:* Offenbar $A = \{y | y \sim v\}$ Nach dem Beweis des letzten Faktes ist A Äquivalenzklasse.
- *B ist Äquivalenzklasse:* Offenbar $B = \{y | y \sim -v\}$ Nach dem Beweis des letzten Faktes ist B Äquivalenzklasse.
- *A und B sind die einzigen Äquivalenzklassen:* Aus $x \in V$ folgt wegen $\dim(V) = 1$ die Existenz eines $t \in \mathbb{R}$ mit $x = t \cdot v$. Für $x \neq 0$ folgt $t \neq 0$, also $t > 0$ ($x \in A$) oder $t < 0$ ($x \in B$). Also : $V - \{0\} = A \cup B$. Nach dem letzten Fakt folgt, daß A und B die einzigen Äquivalenzklassen sind. □

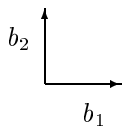
Definition 3: Sei V ein n-dimensionaler reeller Vektorraum. Eine **Orientierung** von V ist eine Halbachse des eindimensionalen Vektorraumes $D(V)$. Sei $\Omega \in D(V)$ eine Orientierung von V und sei $B = (b_1, \ldots, b_n)$ eine Basis von V. Dann hat die reelle Zahl $d(b_1, \ldots, b_n)$ für alle $d \in \Omega$ dasselbe Vorzeichen. Wir sagen, daß B **positiv orientiert** bezüglich Ω ist, falls $d(b_1, \ldots, b_n) > 0$ für alle $d \in \Omega$ ist. Andernfalls sagen wir, daß B negativ orientiert ist.

Bemerkung: Nach Satz 4.4.1 gilt: $d(b_1, \ldots, b_n) \neq 0$ für alle $d \in D(V) \setminus \{0\}$ und alle Basen $B = (b_1, \ldots, b_n)$.

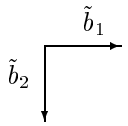
Beispiel: Sei $\dim(V) = 1$. Dann gilt $D(V) = V^*$. Es gibt dann eine Bijektion

$$\begin{array}{rcl} \text{Halbachsen in } D(V) = V^* & \to & \text{Halbachsen in } V \\ (= \text{Orientierung von } V) & & \\ \Omega & \to & \{v \in V \setminus \{0\} \mid v \text{ ist positiv orientiert bzgl } \Omega \\ & & \text{d.h.} l(v) > 0 \text{ für alle } l \in \Omega\} \end{array}$$

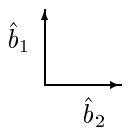
Beispiel: Sei $V = \mathbb{R}^2$. Dann gibt es genau eine Orientierung von V, so daß $((1,0),(0,1))$ positiv orientiert ist. Eine Basis $((a,b),(c,d))$ von V ist dann genau dann positiv orientiert, falls $ad - bc > 0$ ist.



$B = (b_1, b_2) = ((1,0),(0,1))$ +



$\tilde{B} = (\tilde{b}_1, \tilde{b}_2) = ((1,0),(0,-1))$ −



$\hat{B} = (\hat{b}_1, \hat{b}_2) = ((0,1),(1,0))$ −

Beispiel: $\mathbb{R}^n$ wird üblicherweise so definiert, daß die Standardbasis $e_1, \ldots, e_n$ positiv orientiert ist, wobei $e_i = (\delta)_{j=1}^n$. Bei der üblicherweise verwendeten Auswahl der Basis des $\mathbb{R}^3$ in der analytischen Geometrie des Raumes :

- **x-Achse:** In der Zeichenebene (d.h. Tischebene) von links nach rechts.
- **y-Achse:** In der Zeichenebene nach vorne weg vom Betrachter
- **z-Achse:** Senkrecht zur Zeichenebene nach oben

ist also (Einheitsvektor in x, y, z-Richtung) positiv orientiert. Es gilt daher die rechte-Hand-Regel. Das Dreibein (d.h. Tripel von Vektoren) (Daumen, Zeigefinger, Mittelfinger) ist für die rechte Hand positiv orientiert, für die linke Hand negativ. Entsprechend nennt man ein positiv orientiertes Dreibein rechtshändig, ein negativ orientiertes linkshändig.

Bemerkung: Wenn $B = (b_1, \ldots, b_n)$ eine Basis des orientierten Vektorraumes V ist, so hat für $1 \leq i \leq n$ und $\gamma \in \mathbb{R} \setminus \{0\}$ die Basis $(b_1, \ldots, b_{i-1}, \gamma b_i, b_{i+1}, \ldots, b_n)$ dieselbe Orientierung wie B , falls $\gamma > 0$ ist , und die zu B entgegengesezte Orientierung , falls $\gamma < 0$.
Wenn $\pi \in S_n$ ist, so hat die Basis $(b_{\pi(1)}, \ldots, b_{\pi(n)})$ für $\text{sign}(\pi) = 1$ dieselbe Orientierug wie B , für $\text{sign}(\pi) = -1$ die entgegengesetzte.

4.7 Symmetrische n-Linearformen

Sei K ein beliebiger Körper.

Lemma 1: (vgl. Lemma 4.3.1) Sei $f : V \times \ldots \times V \to W$ eine n-lineare Abbildung. Sei $G = \{\pi \in S_n \mid f(x_1, \ldots, x_n) = f(x_{\pi(1)}, \ldots, x_{\pi(n)}) \text{ für alle } (x_1, \ldots, x_n) \in V^n\}$. Dann ist G eine Untergruppe von S_n. Die folgenden Aussagen sind äquivalent :

1. $G = S_n$, d.h. f ändert seinen Funktionswert bei einer beliebigen Permutation der Argumente nicht.
2. G enthält alle Transpositionen (i, j) für $1 \leq i < j \leq n$, d.h., f ändert den Funktionswert bei Vertauschung zweier Argumente nicht.
3. G enthält alle Transpositionen $(i, i+1)$ für $1 \leq i < n$, d.h. f ändert seinen Wert beim Vertauschen benachbarter Argumente nicht.

Beweis: Wie bei Lemma 4.3.1, nur einfacher. □

Definition 1: Eine n-lineare Abbildung, welche diese äquivalenten Bedingungen erfüllt, nennt man **symmetrisch**. Für $W = K$ spricht man von einer **symmetrischen n-Linearform**. In Körpern der Charakteristik 2 sind alle schiefsymmetrischen n-linearen Abbildungen symmetrisch, denn für $\text{char}(K) = 2$ gilt $-1 = +1$ in K.

Beispiel:

1.
$$\begin{aligned} \beta : K^n \times K^n &\to K \\ ((x_1, \ldots, x_n), (y_1, \ldots, y_n)) &\mapsto \sum_{i=1}^{n} x_i y_i \end{aligned}$$
ist eine symmetrische n-Linearform auf K^n.

2.
$$\begin{aligned} \gamma : K \times \ldots \times K &\to K \\ (x_1, \ldots, x_n) &\mapsto x_1 \cdot \ldots \cdot x_n \end{aligned}$$
ist eine symmetrische Bilinearform auf K.

3.

$$\begin{aligned} \delta : (K^n)^m &\to K \\ ((x_1^{(1)}, \ldots, x_n^{(1)}), \ldots, (x_1^{(m)}, \ldots, x_n^{(m)})) &\mapsto \Sigma_{i=1}^n \Pi_{j=1}^m x_i^{(j)} \end{aligned}$$

ist eine symmetrische m-Linearform auf K^n.

5 Quadratische, hermitesche und symplektische Formen

Sei K ein Körper mit von 2 verschiedener Charakteristik.

5.1 Quadratische und symmetrische Bilinearformen

Definition 1: Eine **quadratische Form** auf einem K-Vektorraum ist eine Abbildung $q : V \to K$ mit den folgenden Eigenschaften:

1. $q(\lambda v) = \lambda^2 \cdot q(v)$, $\lambda \in K$, $v \in V$;
2. Die Abbildung

$$\begin{aligned} \beta : V \times V &\to K \\ (x, y) &\mapsto \frac{q(x+y) - q(x) - q(y)}{2} \end{aligned}$$

ist bilinear (und daher eine symmetrische Bilinearform).

Beispiel 1:

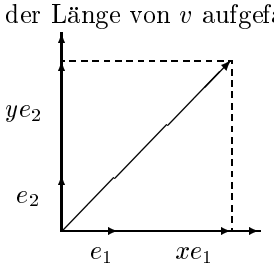
1. Sei

$$\begin{aligned} q : K^n &\to K \\ (x_1, \ldots, x_n) &\mapsto x_1^2 + \ldots + x_n^2. \end{aligned}$$

Dann gilt

$$\begin{aligned} \beta(x, y) &= \frac{1}{2}(q(x+y) - q(x) - q(y)) \\ &= \frac{1}{2} \sum_{i=1}^n ((x_i + y_i)^2 - x_i^2 - y_i^2) \\ &= \sum_{i=1}^n x_i y_i. \end{aligned}$$

Offensichtlich ist β bilinear, Eigenschaft (1) ist klar. Für $K = \mathbf{R}$ kann $q(v)$ als das Quadrat der Länge von v aufgefaßt werden (Satz des Pythagoras).



$v = y \cdot e_2 + x \cdot e_1 = (x, y)$
Länge von $v = |v| = \sqrt{x^2 + y^2} = \sqrt{q(v)}$.

2. Sei $m : \mathbb{R}^4 \to \mathbb{R}$, $m(x_0, x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_0^2$ die MINKOWSKIsche Form. Dann ist $\mu(x, y) = \frac{1}{2}(m(x+y) - m(x) - m(y)) = x_1 y_1 + x_2 y_2 + x_3 y_3 - x_0 y_0$ das MINKOWSKIsche Skalarprodukt. Vektoren $v \in \mathbb{R}^4$ nennt man in der Speziellen Relativitätstheorie

   - **raumartig**, wenn $m(v) > 0$;
   - **zeitartig**, wenn $m(v) < 0$;
   - **lichtartig**, wenn $m(v) = 0, v \neq 0$.

3. Sei V ein K-Vektorraum, $l \in V^*$, $q(x) = l(x)^2$ und

$$\begin{aligned} \beta(x, y) &= \frac{l(x+y)^2 - l(x)^2 - l(y)^2}{2} \\ &= \frac{l(x)^2 + 2l(x)l(y) + l(y)^2 - l(x)^2 - l(y)^2}{2} \\ &= l(x) \cdot l(y). \end{aligned}$$

   β ist bilinear, und es gilt:

$$q(\lambda x) = l(\lambda x)^2 = \lambda^2 l(x)^2 = \lambda^2 q(x).$$

Definition: Offenbar bilden die symmetrischen Bilinearformen einen Untervektorraum des K-Vektorraumes $L(V, V; K)$ aller Bilinearformen auf V (vgl. 4.3). Der Vektorraum aller quadratischen Formen auf V wird wie folgt definiert:

$$(q + \tilde{q})(x) = q(x) + \tilde{q}(x) \qquad (\lambda q)(x) = \lambda q(x),$$

wobei $q, \tilde{q}$ quadratische Formen auf V, $x \in V$, $\lambda \in K$ sind.

Satz 1: Seien $x, y \in V$, q eine quadratische Form und β eine symmetrische Bilinearform. Die Abbildungen

$$\begin{array}{rcl} \{\text{Quadratische Formen auf } V\} & \overset{f,g}{\longleftrightarrow} & \{\text{symmetrische Bilinearformen}\} \\ (f(q))(x, y) = \dfrac{q(x+y) - q(x) - q(y)}{2} & & (\text{d.h. } f(q) = \beta \text{ aus Def.1}) \\ (g(\beta))(x) = \beta(x, x) & & \end{array}$$

sind zueinander inverse Isomorphismen von Vektorräumen.

Beweis:

- $f(q)$ ist nach Definition 1 eine Bilinearform.
- f ist offensichtlich eine lineare Abbildung.
- $q = g(\beta)$ ist eine quadratische Form, denn

$$q(\lambda x) = \beta(\lambda x, \lambda x) = \lambda^2 \beta(x, x) = \lambda^2 q(x).$$

Sei $\tilde{\beta}(x,y) = \frac{q(x+y)-q(x)-q(y)}{2}$. Dann gilt $(*)$:

$$\begin{aligned}
\tilde{\beta}(x,y) &= \frac{1}{2}(\beta(x+y,x+y) - \beta(x,x) - \beta(y,y)) \\
&\overset{\beta \text{ bilinear}}{=} \frac{1}{2}(\beta(x,x) + \beta(x,y) + \beta(y,y) - \beta(x,x) - \beta(y,y)) \\
&= \frac{1}{2}(\beta(x,y) + \beta(y,x)) \\
&\overset{\text{Symmetrie}}{=} \beta(x,y)
\end{aligned}$$

Weil β bilinear ist, ist es auch $\tilde{\beta}$. Es folgt, daß $q = g(\beta)$ eine quadratische Form ist.

- g ist offensichtlich linear.
- $f \circ g = \mathrm{Id}$ folgt aus $(*)$, denn in Bezeichnungsweisen von $(*)$ gilt $f(g(\beta)) = \tilde{\beta}$ nach Definition von f und g.
- $g \circ f = \mathrm{Id}$: Sei q quadratische Form, $\beta = f(q)$, $\tilde{q} = g(\beta)$. Dann gilt

$$\begin{aligned}
\tilde{q} &= \beta(x,x) = \frac{q(x+x) - q(x) - q(x)}{2} \\
&= \frac{2q(x)}{2} = q(x),
\end{aligned}$$

also $(\tilde{q}) = q$. □

Fakt 1: Für quadratische Formen q gilt die **Parallelogrammidentität**:

$$q(x+y) + q(x-y) = 2 \cdot (q(x) + q(y)).$$

Beweis: Sei $\beta = f(q)$. Dann gilt

$$\begin{aligned}
q(x+y) + q(x-y) &= \beta(x+y,x+y) + \beta(x-y,x-y) \\
&= \beta(x,x) + 2\beta(x,y) + \beta(y,y) + \beta(x,x) - 2\beta(x,y) + \beta(y,y) \\
&= 2(q(x) + q(y))
\end{aligned}$$

□

Definition 2:

1. Sei β eine symmetrische Bilinearform auf V und $B = (b_1, \ldots, b_n)$ eine Basis von V. Wir definieren die **Matrixdarstellung von β bezüglich** B als

$$\mathrm{Mat}_B(\beta) = (\beta(b_i, b_j))_{i,j=1}^n.$$

2. Eine $n \times n$-Matrix ist **symmetrisch**, falls $A^T = A$, d.h. für $A = (a_{ij})_{i,j=1}^n$ gilt $a_{ij} = a_{ji}$ für alle $1 \leq i, j \leq n$.

Beispiel:

1. Die symmetrische Bilinearform $\beta((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \sum_{i=1}^n x_i y_i$ hat bezüglich der Standardbasis $B = (e_i)_{i=1}^n$, $e_i = (\delta_{ij})_{j=1}^n$, die Matrixdarstellung

$$\mathrm{Mat}_B(\beta) = (\beta(e_i, e_j))_{i,j=1}^n = \sum_{k=1}^n \delta_{ik}\delta_{jk} = \delta_{ij},$$

also $\text{Mat}_B(\beta) = \mathbf{1}_n$.

2. Die MINKOWSKIsche Form aus Beispiel 1.2 hat bezüglich der Standardbasis die Matrixdarstellung
$$\text{Mat}_B(\mu) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

3. Die symmetrische Bilinearform $\beta((x_0, x_1), (y_0, y_1)) = (x_0 + x_1)(y_0 + y_1)$ hat bezüglich der Standardbasis die Matrixdarstellung
$$\text{Mat}_B(\beta) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Fakt 2: Sei $B = (b_1, \ldots, b_n)$ eine Basis von V, dann ist die Abbildung
$$\text{Mat}_B : \left\{ \begin{array}{c} \text{Symmetrische} \\ \text{Bilinearformen} \\ \text{auf } V \end{array} \right\} \longrightarrow \left\{ \begin{array}{c} \text{Symmetrische } n \times n\text{-} \\ \text{Matrizen mit} \\ \text{Koeffizienten aus } K \end{array} \right\}$$
ein Isomorphismus von K-Vektorräumen.

Beweis: Wir geben eine inverse Abbildung g zu Mat_B an. Zur Erinnerung:
$$i_B : K^n \to V, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^{n} x_i b_i$$
(Elemente von K^n als $n \times 1$-Matrizen) ist ein Isomorphismus.
Für $x, y \in V$ und $A \in \text{Mat}(n, n, K)$ ist
$$(g(A))(x, y) = (i_B^{-1}(x))^T \cdot A \cdot (i_B^{-1}(y))$$
eine 1×1-Matrix, also ein Element von K. Offenbar ist $g(A)$ eine Bilinearform auf V (Linearität von i_B und Distributivität des Matrizenproduktes). Wenn $A^T = A$ gilt, so gilt
$$\begin{aligned} (g(A))(x, y) &= ((g(A))(x, y))^T = (i_B^{-1}(y))^T \cdot A^T \cdot ((i_B^{-1}(x))^T)^T \\ &= (i_B^{-1}(y))^T \cdot A \cdot ((i_B^{-1})(x)) = g(A)(y, x). \end{aligned}$$
Man prüft leicht nach, daß die auf diese Weise definierte Abbildung g zu Mat_B invers ist. □

Folgerung: Sei β eine symmetrische Bilinearform auf K^n und $B = \text{Mat}_B(\beta)$, wobei $B = (e_1, \ldots, e_n)$ die Standardbasis ist. Dann gilt für $x, y \in K^n$
$$\beta(x, y) = x^T \cdot B \cdot y,$$
wobei $x, y \in K^n$ als $n \times 1$-Matrizen und 1×1-Matrizen als Elemente von K betrachtet werden. □

Fakt 3: Sei V ein K-Vektorraum, $B = (b_1, \ldots, b_n)$ eine Basis von V, $B^* = (b_1^*, \ldots, b_n^*)$ die dazu duale Basis des dualen Vektorraumes V^*. Sei $\sigma_\beta : V \to V^*$ durch $(\sigma_\beta(x))(y) = \beta(x, y)$ für $x, y \in V$ definiert. Dann gilt
$$\text{Mat}_B(\beta) = \text{Mat}_{B,B^*}(\sigma_\beta).$$

Beweis: Die Definition der Matrixdarstellung (in 3.1) von $f : V \to \tilde{V}$ bzgl. $B = (b_1, \ldots, b_n)$ und $\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_n)$ lautet

$$\text{Mat}_{B,\tilde{B}} = (a_{ij})_{i,j=1}^n \text{ mit } f(b_j) = \sum_{i=1}^n a_{ij}\tilde{b}_i.$$

Wende dies auf $\tilde{V} = V^*$ und $A = (a_{ij})_{i,j=1}^n = \text{Mat}_{B,B^*}(\sigma_\beta)$ an:

$$\sigma_\beta(b_j) = \sum_{i=1}^n a_{ij} b_i^*. \qquad (*)$$

Gemäß der Definition der dualen Basis

$$b_i^*(b_k) = \delta_{ki} \text{ (allgemein: } b_i^*(\sum_{k=1}^n \lambda_k b_k) = \lambda_i) \qquad (\dagger)$$

folgt dann

$$\begin{aligned} a_{kj} &= \sum_{i=1}^n \delta_{ik} a_{ij} \overset{(\dagger)}{=} \sum_{i=1}^n b_i^*(b_k) \cdot a_{ij} \\ &= \left(\sum_{i=1}^n a_{ij} b_i^*\right)(b_k) \overset{(*)}{=} (\sigma_\beta(b_j))(b_k) \\ &\overset{\text{Def. von } \sigma_\beta}{=} \beta(b_j, b_k). \end{aligned}$$

Aus Definition 2 folgt nun, daß $\text{Mat}_{B,B^*}(\sigma_\beta) = \text{Mat}_B(\beta)$. □

Definition 3: Sei $f : W \to V$ linear, β eine symmetrische Bilinearform und q eine quadratische Form auf V. Wir definieren für $x, y \in W$

$$(f^*(\beta))(x, y) = \beta(f(x), f(y)) \text{ und } (f^*(q))(x) = q(f(x)).$$

$f^*(\beta)$ ist eine symmetrische Bilinearform und $f^*(q)$ eine quadratische Form auf W.

Fakt 4: In den Bezeichnungen der Definition sei $\tilde{B}$ eine Basis von W und B eine Basis von V. Dann gilt

$$\text{Mat}_{\tilde{B}}(f^*(\beta)) = \text{Mat}_{\tilde{B},B}(f)^T \cdot \text{Mat}_B(\beta) \cdot \text{Mat}_{\tilde{B},B}(f).$$

Beweis: Offenbar $\sigma_{f^*(\beta)} = f^* \sigma_\beta f$. Wende Fakt 3.1.4. und Fakt 3.1.3. an. □

Bemerkung: Für $f = \text{Id}_n$ gilt $f^*(\beta) = \beta$, also macht Fakt 4 in diesem Fall eine Aussage über die Änderung von $\text{Mat}_B(\beta)$ bei Änderung der Basis.

5.2 Nicht ausgeartete Bilinearformen. Der Rang.

Wir behalten die Voraussetzung $\text{char}(K) \neq 2$ bei.

Definition 1: Sei β eine symmetrische Bilinearform auf V. Der **Nullraum** von β ist als

$$N(\beta) = \{x \in V \mid \forall\, y \in V : \beta(x, y) = 0\}$$

definiert. Sei V endlichdimensional. Dann nennt man β **nicht ausgeartet**, falls $N(\beta) = \{0\}$ gilt. Der **Rang** von β ist als $\text{Rang}(\beta) = \dim(V) - \dim(N(\beta))$ definiert.

Fakt 1: Offenbar gilt $N(\beta) = \ker(\sigma_\beta)$ (vgl. Definition 2.4.5). Also sind für endlichdimensionale Vektorräume V äquivalent:

1. β ist nicht ausgeartet;
2. σ_β ist Monomorphismus;
3. σ_β ist ein Isomorphismus.

□

Fakt 2: Eine symmetrische Bilinearform hat denselben Rang wie ihre Matrixdarstellung.

Beweis: Nach Satz 2.4.1 und Fakt 3.1.2. gilt

$$\begin{aligned} \text{Rang}\,(\beta) &\overset{\text{Def.1; Fakt 1}}{=} \dim(V) - \dim(\ker(\sigma_\beta)) \\ &\overset{\text{Satz 4.1}}{=} \dim(\text{Bild}\,(\sigma_\beta)) \\ &\overset{\text{Fakt 3.1.2}}{=} \text{Rang}\,(\text{Mat}_{B,B^*}(\sigma_\beta)) \\ &= \text{Rang}\,(\text{Mat}_B(\beta)). \end{aligned}$$

□

Beispiel:

1. Sei β auf $\mathbb{R}^n$ gegeben durch $\beta(x,y) = \sum_{i=1}^n x_i y_i$. Dann ist β nicht ausgeartet, denn $N(\beta) = \{0\}$, weil $\beta(x,x) > 0$ für $x \neq 0$.
2. Sei β auf K^n wie in (1) gegeben. Dann hat β bezüglich der Standardbasis die Matrixdarstellung $\text{Mat}_B(\beta) = \mathbf{1}_n$, also ist $\text{Rang}\,(\beta) = n$.
3. Sei μ auf $\mathbb{R}^4$ das MINKOWSKIsche Skalarprodukt. Dann ist die Matrixdarstellung von μ bezüglich der Standardbasis invertierbar, also ist μ nicht ausgeartet.
4. Sei $l \in V^*$, $\beta(x,y) = l(x) \cdot l(y)$. Dann gilt $N(\beta) = \ker(l)$. Wenn $\dim(V) \geq 2$ ist, gilt $\ker(l) \neq \{0\}$, also ist in diesem Fall β ausgeartet.

Fakt 3: Sei V endlichdimensional und $W \subseteq V$ ein Vektorraum mit $W \oplus N(\beta) = V$. Dann ist die Einschränkung von β auf W nicht ausgeartet.

Beweis: Übungsaufgabe.

Definition 2: Zwei symmmetrische Bilinearformen (V,β) und $(\tilde{V},\tilde{\beta})$ sind **isomorph**, falls ein Isomorphismus $f : V \to \tilde{V}$ mit $f^*(\tilde{\beta}) = \beta$ existiert. Offenbar ist die Isomorphie eine Äquivalenzrelation auf der Klasse aller symmetrischen Bilinearformen. Die Äquivalenzklassen nennt man **Isomorphieklassen**.

Fakt 4:

1. Offenbar ist der Rang eine Isomorphieinvariante, d.h., zwei isomorphe symmetrische Bilinearformen haben denselben Rang.
2. Zwei symmetrische Bilinearformen $\beta, \tilde{\beta}$ auf endlichdimensionalen Vektorräumen $V, \tilde{V}$ sind genau dann isomorph, wenn zwei Basen B und $\tilde{B}$ mit $\text{Mat}_B(\beta) = \text{Mat}_{\tilde{B}}(\tilde{\beta})$ existieren.

Beweis: Der erste Fakt folgt leicht aus dem zweiten. Sei $\text{Mat}_B(\beta) = \text{Mat}_{\tilde{B}}(\tilde{\beta})$ und seien $B = (b_1,\ldots,b_n)$ und $\tilde{B} = (\tilde{b}_1,\ldots,\tilde{b}_n)$, dann existiert ein Isomorphismus $\varphi : V \to \tilde{V}$ mit $\varphi(b_i) = \tilde{b}_i$. Es gilt $\varphi^*(\tilde{\beta}) = \beta$.
Sei umgekehrt $\varphi : V \to \tilde{V}$ ein Isomorphismus mit $\varphi^*(\tilde{\beta}) = \beta$ und sei $B = (b_1,\ldots,b_n)$ eine Basis von V. Dann ist $\tilde{B} = (\varphi(b_1),\ldots,\varphi(b_n))$ eine Basis von $\tilde{V}$ und $\text{Mat}_B(\beta) = \text{Mat}_{\tilde{B}}(\tilde{\beta})$. □

5.3 Klassifikation der quadratischen Formen

Sei $\mathrm{char}\,(K) \neq 2$. Alle Vektorräume werden als endlichdimensional vorausgesetzt.

Lemma 1: Sei V ein Vektorraum und sei β eine symmetrische Bilinearform auf V. Wenn $\beta \neq 0$ gilt, so gibt es ein $x \in V$ mit $\beta(x,x) \neq 0$.

Beweis: Angenommen $\beta(x,x) = 0$ für alle $x \in V$, aber es gäbe $x, y \in V$ mit $\beta(x,y) \neq 0$. Dann gilt $0 = \beta(x+y, x+y) = \beta(x,x) + 2\beta(x,y) + \beta(y,y) = 2\beta(x,y) \neq 0$. Widerspruch. □

Bemerkung 1: Der Beweis hätte auch durch Satz 5.1.1 geführt werden können. Der angegeben Beweis ist konstruktiv, d.h. er gestattet die praktische Berechnunng eines Vektors $v \in V$ mit $\beta(v,v) \neq 0$.

Bemerkung 2: In einem Körper mit Charakteristik 2 ist das Lemma falsch, denn dann ist jede schiefsymmetrische Bilinearform symmetrisch.

Definition 1: Eine Basis $B = (b_1, \ldots, b_n)$ von V nennt man eine **Orthogonalbasis** bezüglich β, falls $\beta(b_i, b_j) = 0$ für $i \neq j$ gilt.

Satz 1: Sei β eine symmetrische Bilinearform auf V. Dann hat V eine Orthogonalbasis bezüglich β.

Beweis: durch Induktion nach $\dim(V)$. Für $\dim(V) = 0$ ist die leere Menge eine Orthogonalbasis. Sei die Behauptung für Vektorräume kleinerer Dimension als V bewiesen. Wenn $\beta = 0$ gilt, so ist jede Basis eine Orthogonalbasis. Sei also $\beta \neq 0$. Nach Lemma 1 gibt es $v \in V$ mit $\beta(v,v) \neq 0$, Sei $W = v^\perp = \{w \in V \mid \beta(v,w) = 0\}$. Offenbar ist W ein Untervektorraum von V. Dann gilt $v \notin W$ nach Vorraussetzung und $W \oplus Kv = V$, denn für $x \in V$ gilt

$$x = \underbrace{\frac{\beta(x,v)}{\beta(v,v)} \cdot v}_{\in K \cdot v} + \underbrace{(x - \frac{\beta(x,v)}{\beta(v,v)} \cdot v)}_{\in W}.$$

Nach Induktionvoraussetzung gibt es eine Orthogonalbasis $(b_1, \ldots, b_n)$ von W bezüglich $\beta|_W$. Dann ist $(v, b_2, \ldots, b_n)$ eine Orthogonalbasis von V. □

Bemerkung 3: Weil der Beweis von Lemma 1 konstruktiv war, ist es auch der Beweis des Satzes.

Definition 2: K ist **quadratisch abgeschlossen**, falls für jedes $x \in K$ eine Quadratwurzel (d.h, ein $\gamma \in K$ mit $\gamma^2 = x$) existiert.

Bemerkung 4: Sei K ein Körper, $\mathrm{char}(K) \neq 2$. Dann existieren zu jedem $x \in K - \{0\}$ entweder keine oder zwei Quadratwurzeln.
Sei dazu $x = y^2$. Dann ist auch $-y$ eine Quadratwurzel aus x, diese ist wegen $y \neq 0$ von y verschieden. Die beiden Quadratwurzeln y und $-y$ sind sie einzigen Quadratwurzeln aus x, denn

$$t^2 - x = t^2 - y^2 = (t-y)(t+y).$$

Beispiel 1: $\mathbb{C}$ ist quadratisch abgeschlossen, denn wenn $z = a + ib$ eine komplexe Zahl ist, so ist

$$\sqrt{\frac{\sqrt{(a^2+b^2)} + a}{2}} + \mathrm{sign}(b)\sqrt{\frac{\sqrt{a^2+b^2} - a}{2}} \cdot i$$

eine Quadratwurzel aus z. Dabei ist $\sqrt{t}$ die nichtnegative Quadratwurzel aus der reellen Zahl $t \geq 0$.

Theorem 2:

1. Sei β eine symmetrische Bilinearform auf einem endlichdimensionalen Vektorraum über einem quadratisch abgeschlossenem Körper K, $\mathrm{char}\,(K) \neq 2$. Dann existiert eine Basis $B = (b_1, \ldots, b_n)$ von V , so daß

$$\beta(b_i, b_j) = \begin{cases} \delta_{ij} & \text{für } 1 \leq i \leq r \\ 0 & \text{sonst,} \end{cases}$$

   wobei r der Rang von β ist.

2. Die symmetrischen Bilinearformen (V, β) und $(\tilde{V}, \tilde{\beta})$ auf endlichdimensionalen K-Vektorräumen V und $\tilde{V}$ sind (unter denselben Voraussetztungen an K wie im ersten Punkt) genau dann isomorph, wenn $\dim(V) = \dim(\tilde{V})$ und $\mathrm{Rang}\,(\beta) = \mathrm{Rang}\,(\tilde{\beta})$.

Bemerkung: Wir haben also in diesem Falle eine vollständige Klassifikation der symmetrischen Bilinearformen auf endlichdimensionalen Vektorräumen bekommen: bis auf Isomorphie ist (V, β) eindeutig durch $n = \dim(V)$ und $r = \mathrm{Rang}\,(\beta)$ charakterisiert, wobei alle $(n, r) \in \mathbb{N}^2$ mit $r \leq n$ angenommen werden.

Beweis von Theorem 2: Sei $\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_n)$ eine Orthogonalbasis von V bezüglich β ($\tilde{B}$ existiert nach Satz 1). Durch Umordnen von $\tilde{\beta}$ können wir erreichen, daß eine Zahl r mit $0 \leq r \leq n$ existiert, so daß $s_i = \beta(\tilde{b}_i, \tilde{b}_i)$ für $1 \leq i \leq r$ von 0 verschieden ist und für $r < i \leq n$ verschwindet. Dann stimmt r mit dem Rang von β überein (das folgt aus Fakt 5.2.2).
Sei q_i eine Quadratwurzel aus s_i, dann hat die Basis $B = (b_1, \ldots, b_n)$,

$$b_i = \begin{cases} \frac{\tilde{b}_i}{q_i} & \text{für } 1 \leq i \leq r \\ \tilde{b}_i & \text{für } r < i \leq n \end{cases}$$

die im ersten Punkt des Theorems beschriebene Gestalt. Der zweite Punkt folgt aus dem ersten Punkt und aus Fakt 5.2.4. □

Bemerkung 5: Auch dieser Beweis ist konstruktiv.

Definition 3: Eine **Orthonormalbasis** von V bezüglich β ist eine Basis $B = (b_1, \ldots, b_n)$ von V mit $\mathrm{Mat}_B(\beta) = \mathbf{1}_n$, d.h. $\beta(b_i, b_j) = \delta_{ij}$.

Bemerkung 6: Weil der Rang von $\mathbf{1}_n$ n ist, können Orthonormalbasen nur bezüglich nichtausgearteter symmetrischer Bilinearformen existieren. Umgekehrt garantiert das Theorem die Existenz von Orthonormalbasen für nichtausgeartete symmetrische Bilinearformen auf endlichdimensionalen Vektorräumen über quadratisch abgeschlossenen Körpern mit von 2 verschiedener Charakteristik.

5.4 Quadratische Formen auf reellen Vektorräumen

Theorem 3 (Trägheitssatz von SYLVESTER): Sei β eine symmetrische Bilinearform über einem $\mathbf{R}$-Vektorraum V der endlichen Dimension n. Dann existieren $a, b \in \mathbf{N}$ mit $a+b \leq n$ und eine Orthogonalbasis $B = (b_1, \ldots, b_n)$ von V, so daß

$$\beta(b_i, b_i) = \begin{cases} 1 & \text{für } 1 \leq i \leq a \\ -1 & \text{für } a < i \leq a+b \\ 0 & \text{für } i > a+b \end{cases}$$

gilt. Die Zahlen a und b sind dabei eindeutig durch die Isomorphieklasse von (V, β) bestimmt, umgekehrt bestimmen sie diese Isomorphieklasse gemeinsam mit $n = \dim(V)$ eindeutig. Es gilt $a + b = \text{Rang}\,(\beta)$.

Beweis: Wir führen zunächst den Beweis der Existenz einer Basis B mit der beschriebenen Eigenschaft. Sei $\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_n)$ eine Orthogonalbasis von V, und sei $\hat{B} = (\hat{b}_1, \ldots, \hat{b}_n)$ mit

$$\hat{b}_i = \begin{cases} \frac{\tilde{b}_i}{\sqrt{\beta(\tilde{b}_i, \tilde{b}_i)}} & \text{falls } \beta(\tilde{b}_i, \tilde{b}_i) > 0 \\ \frac{\tilde{b}_i}{\sqrt{-\beta(\tilde{b}_i, \tilde{b}_i)}} & \text{falls } \beta(\tilde{b}_i, \tilde{b}_i) < 0 \\ \tilde{b}_i & \text{falls } \beta(\tilde{b}_i, \tilde{b}_i) = 0 \end{cases}$$

Dann ist $\hat{B}$ eine Orthogonalbasis mit $\beta(\hat{b}_i, \hat{b}_i) \in \{-1, 0, 1\}$, also kann aus $\hat{B}$ durch Umordnen der Basisvektoren eine Basis B erhalten werden, die die im Theorem beschriebene Form hat. Daß a, b und n die Isomorphieklasse eindeutig bestimmen, folgt aus Fakt 5.2.4. Aus Fakt 5.2.2 folgt die letzte Behauptung des Theorems. Zum Beweis der Eindeutigkeit von a und b benötigen wir

Definition 1:

- β ist **positiv definit**, falls $\beta(v, v) > 0$ für $v \in V - \{0\}$;
- β ist **positiv semidefinit**, falls $\beta(v, v) \geq 0$ für alle $v \in V$;
- β ist **negativ definit** (bzw. **negativ semidefinit**), falls $-\beta$ positiv definit (bzw. semidefinit) ist.

Beispiel 1: Offenbar ist das Euklidische Skalarprodukt auf $\mathbf{R}^n$ positiv definit. Daher ist β positiv definit, wenn eine Orthonormalbasis B von V bezüglich β existieren sollte. Wenn eine Basis B mit $\text{Mat}_B(\beta) = -\mathbf{1}_n$ existieren sollte, so ist β negativ definit.

Lemma 1: Sei β eine symmetrische Bilinearform auf einem endlichdimensionalen $\mathbf{R}$-Vektorraum V und seien X_+ und X_- Unterräume von V, so daß die Einschränkung von β auf X_+ positiv und auf X_- negativ definit ist. Dann gilt

$$\dim(X_+) + \dim(X_-) \leq \text{Rang}\,(\beta).$$

Beweis: Es gilt $X_+ \cap X_- = \{0\}$, denn aus $x \in (X_+ \cap X_-) - \{0\}$ würden nach unseren Voraussetzungen aus $\beta|_{X_\pm}$ die widersprüchlichen Ungleichungen $\beta(x, x) > 0$ (wegen $x \in X_+ - \{0\}$) und $\beta(x, x) < 0$ (wegen $x \in X_- - \{0\}$) folgen. Sei $X = X_+ \oplus X_-$, dann gilt $\dim(X) = \dim(X_+) + \dim(X_-)$. $\beta|_X$ ist nicht ausgeartet, denn sei $x = x_+ + x_- \in X - \{0\}$, mit $x_\pm \in X_\pm$. Für $x_+ \neq 0$ sei $y = x_+ \in X$, für $x_+ = 0$ (und damit $x_- \neq 0$) sei $y = x_-$. In beiden Fällen gilt $\beta(x, y) \neq 0$, also ist $\beta|_X$ nicht ausgeartet. Insbesondere folgt $X \cap N(\beta) = \{0\}$,

also $\dim(V) \geq \dim(X) + \dim(N(\beta)) = \dim(X_+) + \dim(X_-) + \dim(N(\beta))$. □

Beweis von Theorem 3, Eindeutigkeit von a und b: Offenbar genügt es zu zeigen, daß a und b durch (V, β) selbst eindeutig bestimmt sind. Sei $\tilde{B}$ eine andere Orthogonalbasis mit $\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_n)$,

$$\beta(\tilde{b}_i, \tilde{b}_i) = \begin{cases} 1 & \text{für } 1 \leq i \leq \tilde{a} \\ 0 & \text{für } i > \tilde{a} + \tilde{b} \\ -1 & \text{für } \tilde{a} < i \leq \tilde{a} + \tilde{b}. \end{cases}$$

Sei X_+ die lineare Hülle von $C_+ = (b_1, \ldots, b_a)$ und X_- die lineare Hülle von $C_- = (\tilde{b}_{\tilde{a}+1}, \ldots, \tilde{b}_{\tilde{a}+\tilde{b}})$ Dann ist C_+ eine Orthogonalbasis von X_+, und $\mathrm{Mat}_{C_-}(\beta|_{X_-}) = -\mathbf{1}_{\tilde{b}}$. Aus Beispiel 1 folgt

$$a + b = \mathrm{Rang}\,(\beta) \geq \dim(X_+) + \dim(X_-) = a + \tilde{b}$$

Also $b \geq \tilde{b}$. Indem wir die Rollen von B und $\tilde{B}$ vertauschen, erhalten wir auch $b \leq \tilde{b}$, also $b = \tilde{b}$. Weiterhin gilt

$$a = \mathrm{Rang}\,(\beta) - b = \mathrm{Rang}\,(\beta) - \tilde{b} = \tilde{a},$$

denn $\mathrm{Rang}\,(\beta) = a + b = \tilde{a} + \tilde{b}$. Also sind a und b eindeutig bestimmt. □

Fakt: β ist genau dann positiv bzw. negativ definit, falls $a = \dim(V)$ bzw. $b = \dim(V)$ in Theorem 3 gilt. β ist genau dann positiv bzw. negativ semidefinit, falls $b = 0$ bzw. $a = 0$ in Theorem 3 gilt. Insbesondere hat V genau dann eine Orthonormalbasis bezüglich β, wenn β positiv definit ist. □

Definition 2: In der Situation von Theorem 3 nennt man $\mathrm{sign}\,(\beta) = a - b$ die **Signatur** von β.

Bemerkung 1: Die Isomorphieklasse (V, β) ist dann eindeutig bestimmt durch $\dim(V)$, $\mathrm{rang}\,(\beta)$ und $\mathrm{sign}\,(\beta)$.

Beispiel 2: $\beta_n((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \sum_{i=1}^n x_i y_i$. β_n ist das Euklidische Skalarprodukt auf dem $\mathbb{R}^n$. β_n ist positiv definit, eine Orthonormalbasis des $\mathbb{R}^n$ ist durch die Standardbasis $(b_1, \ldots, b_n)$ gegeben, $b_i = (\delta_{ij})_{j=1}^n$; $a = n$, $b = 0$, $\mathrm{sign}\,(\beta_n) = n$. Sei $n = a + b$,

$$\beta_{a,b}((x_i)_{i=1}^n, (y_i)_{i=1}^n) = \sum_{i=1}^{a} x_i y_i - \sum_{i=a+1}^{n} x_i y_i.$$

Bezüglich der Standardbasis $(e_1, \ldots, e_n)$ hat $\beta_{a,b}$ die in Theorem 3 beschriebene Gestalt. $\beta_{a,b}$ ist nicht ausgeartet, $\mathrm{sign}\,(\beta_{a,b}) = a - b$. Insbesondere ist $\beta_{3,1}$ das MINKOVSKIsche Skalarprodukt auf dem $\mathbb{R}^4$, es gilt $\mathrm{sign}\,(\beta_{3,1}) = 2$.

Bemerkung 2: Der Beweis der Existenzaussage in Theorem 3 ist konstruktiv. Insbesondere ergibt sich daraus ein praktisches Verfahren zur Bestimmung von $a, b, \mathrm{sign}\,(\beta)$ und zum Test auf positive Definitheit.

Definition 3: Ein bezüglich β positiv definiter Unterraum $W \subseteq V$ ist **maximal positiv definit**, falls für jedes $U \subseteq V$ mit $W \subset U$ die Einschränkung von β auf U nicht positiv definit ist. Analog definiert man maximale negativ definite Unterräume.

Fakt 1:

1. Jeder positiv bzw. negativ definite Unterraum W von V ist in einem maximal positiv bzw. negativ definiten Unterraum enthalten.

2. Alle maximalen positiv bzw. negativ definiten Unterräume haben die Dimension a bzw. b (aus Theorem 3).

3. Sei $V_+ \subseteq V$ ein maximal positiv definiter Unterraum und V_- ein maximal negativ definiter Unterraum. Dann gilt $V = V_+ \oplus V_- \oplus N(\beta)$.

Beweis:

1. Durch absteigende Induktion nach $\dim(W)$. Die Behauptung ist klar für $\dim(W) = \dim(V)$. Sei die Behauptung für alle definiten Unterräume größerer Dimension als W bewiesen. Falls W selbst nicht maximal ist, so gibt es einen Unterraum $U \supset W$, so daß U ebenfalls positiv bzw. negativ definit ist. Dann ist U nach der Induktionsvoraussetzung in einem maximalen positiv bzw. negativ definiten Unterraum enthalten.

2. Sei W maximal positiv definit. Sei $\tilde{a} = \dim(W)$ und sei $v_1, \ldots, v_{\tilde{a}}$ eine Orthonormalbasis von W bezüglich β. Aus dem Beweis von Satz 5.3.1 folgt, daß man $(v_1, \ldots, v_{\tilde{a}})$ zu einer Orthogonalbasis $(v_1, \ldots, v_{\tilde{a}}, v'_{\tilde{a}+1}, \ldots, v'_n)$ ergänzen kann. Sei $\beta(v'_i, v'_i) = 0$ genau dann, wenn $i > \operatorname{rang}(\beta)$. Wir können auch voraussetzen, daß $\beta(v_i, v_i) > 0$ für $1 \leq i \leq a'$ und $\beta(v_i, v_i) < 0$ für $a' < i \leq a' + b'$. Dann $a' \geq \tilde{a}$. Wenn $a' > \tilde{a}$ wäre, so wäre $W + \mathbb{R}v'_{\tilde{a}+1}$ ein positiv definiter Unterraum von V, im Widerspruch zur Eigenschaft von W. Also $a' = \tilde{a}$. Sei $v_i = v_i / \sqrt{|\beta(v'_i, v'_i)|}$ für $i \leq \operatorname{rang}(\beta)$, $v_i = v'_i$ für $i > \operatorname{rang}(\beta)$. Dann

$$\beta(v_i, v_j) = \begin{cases} 1 & 1 \leq i = j \leq a \\ -1 & a' < i = j \leq \operatorname{rang}(\beta) \\ 0 & \text{sonst.} \end{cases}$$

   Wegen der Eindeutigkeit aus Theorem 3 folgt $a' = a$, also $\tilde{a} = a$.

3. - $V_+ \cap V_- = \{0\}$ ist klar.
   - $(V_+ \oplus V_-) \cap N(\beta) = \{0\}$ folgt aus dem Beweis von Lemma 1.
   - $V_+ \oplus V_- \oplus N(\beta) = V$:

$$\begin{aligned} \dim(V_+ \oplus V_- \oplus N(\beta)) &= \dim(V_+) + \dim(V_-) + \dim(N(\beta)) \\ &= a + b + \dim(N(\beta)) = \operatorname{rang}(\beta) + \dim(N(\beta)) \\ &= \dim(V). \end{aligned}$$

□

Fakt 2: Sei β eine nicht ausgeartete symmetrische Bilinearform auf einem endlichdimensionalen reellen Vektorraum V, sei $B = (b_1, \ldots, b_n)$ eine Basis von V und habe b dieselbe Bedeutung wie in unserer Formulierung von Theorem 3. Dann gilt:

$$(-1)^b \det(\operatorname{Mat}_B(\beta)) > 0.$$

Bemerkung: $\operatorname{Mat}_B(\beta) = \operatorname{Mat}_{B,B^*}(\sigma_\beta)$; β nicht ausgeartet $\Leftrightarrow$ σ_β Isomorphismus $\Leftrightarrow$ $\det(\operatorname{Mat}_B(\beta)) \neq 0$. Sei $C = (c_1, \ldots, c_n)$ eine andere Basis von V, dann gilt nach Fakt 5.4.1

$$\operatorname{Mat}_B(\beta) = \operatorname{Mat}_{B,C}(\mathrm{Id})^T \cdot \operatorname{Mat}_C(\beta) \cdot \operatorname{Mat}_{B,C}(\mathrm{Id}). \qquad (*)$$

Beweis von Fakt 2: Sei C eine Basis von V mit

$$(\operatorname{Mat}_C(\beta))_{ij} = \begin{cases} 1 & i = j \leq a \\ -1 & a < i = j \leq a + b \\ 0 & \text{sonst} \end{cases}$$

(vgl. Theorem 3). Nach $(*)$ gilt dann

$$\begin{aligned}(-1)^b \det(\mathrm{Mat}_B(\beta)) &= (-1)^b \det(\mathrm{Mat}_{B,C}(\mathrm{Id})^T) \det(\mathrm{Mat}_C(\beta)) \det(\mathrm{Mat}_{B,C}(\mathrm{Id})) \\ &= (-1)^b \det(\mathrm{Mat}_{B,C}(\mathrm{Id}))(-1)^b \det(\mathrm{Mat}_{B,C}(\mathrm{Id})) \\ &= \det(\mathrm{Mat}_{B,C}(\mathrm{Id}))^2 > 0\end{aligned}$$

□

Satz 1: Sei $B = (b_1, \ldots, b_n)$ eine Basis des reellen Vektorraumes V, sei $V_k = \mathcal{L}(\{b_1, \ldots, b_k\})$, und sei B_k die Basis $(b_1, \ldots, b_k)$ von V_k. Sei β eine symmetrische Bilinearform auf V. Dann ist β genau dann positiv definit, wenn für $1 \leq k \leq n$

$$\det(\mathrm{Mat}_{B_k}(\beta|_{V_k \times V_k})) > 0.$$

Bemerkung: Eine symmetrische Matrix A mit reellen Einträgen ist positiv definit, falls die durch A definierte symmetrische Bilinearform $\beta_A(x, y) = x^T \cdot A \cdot y$ positiv definit ist. Nach Satz 1 ist dies genau dann der Fall, wenn in

$$\left(\begin{array}{c|c|c|c} a_{11} & a_{12} & a_{13} & \ldots \\ \hline a_{21} & a_{22} & a_{23} & \ldots \\ \hline a_{31} & a_{32} & a_{33} & \ldots \\ \hline \vdots & \vdots & \vdots & \ddots \end{array}\right) \quad \begin{array}{l} v_1 = \{(x_1, 0, \ldots)\} \\ v_2 = \{(x_1, x_2, 0, \ldots)\} \\ v_3 = \{(x_1, x_2, x_3, 0, \ldots)\} \end{array}$$

die Determinanten der eingerahmten Matrizen positiv sind. Praktischen Wert hat dieses Kriterium nur für kleine n, etwa für $n = 2$.

Beweis von Satz 1, Notwendigkeit: Sei β positiv definit. Offenbar ist die Einschränkung $\beta|_{V_k \times V_k}$ ebenfalls positiv definit (und daher nicht ausgeartet),also $\det(\mathrm{Mat}_{B_k}(\beta|_{V_k \times V_k})) > 0$.
Hinlänglichkeit, Induktion nach n: Die Behauptung ist klar für $n = 1$. Sei die Behauptung für alle Vektorräume kleinerer Dimension als n bewiesen. Wenn β die Bedingung des Kriteriums erfüllt, so tut dies auch $\beta|_{V_{n-1}}$, also ist $\beta|_{V_{n-1}}$ nach der Induktionsannahme positiv definit. Also ist entweder V_{n-1} oder V_n ein maximaler positiv definiter Unterraum. Im letzten Fall ist β positiv definit, im ersten Fall ergibt Fakt 1 $a = \dim V_{n-1} = n - 1$ und daher $b = 1$, und Fakt 2 würde $\det(\mathrm{Mat}_B(\beta)) < 0$ ergeben. □

Beispiel: $\begin{pmatrix} a & c \\ c & b \end{pmatrix}$ ist positiv definit $\Leftrightarrow a > 0$ und $ab - c^2 > 0$.

Bemerkung: Die Einschränkung einer nichtausgearteten indefiniten Form auf einen Unterraum kann ausgeartet sein.

5.5 Die Gramsche Determinante, die Gruppen $\mathbf{O}(\beta)$ und $\mathbf{SO}(\beta)$

Sei V ein Vektorraum der endlichen Dimension n über einem Körper K, $\mathrm{char}(K) \neq 2$. Sei $D(V)$ der Raum der schiefsymmetrischen n-Linearformen auf V. Wir wissen aus Kapitel 4, daß $\dim(D(V)) = 1$ gilt. Sei $d \in D(V) - \{0\}$. Dann kann jedes $e \in D(V)$ als $e = \lambda d$ geschrieben werden, wobei $\lambda \in K$ eindeutig bestimmt ist.

Satz 1: In der soeben beschriebenen Situation sei β eine symmetrische Bilinearform auf V. Dann gibt es eine eindeutig bestimmte Konstante $c_{\beta,d} \in K$, so daß für $v_1, \ldots, v_n, w_1, \ldots, w_n \in V$ die **Gramsche Determinante**

$$\det(\beta(v_i, w_j))_{i,j=1}^n = c_{\beta,d} d(v_1, \ldots, v_n) d(w_1, \ldots, w_n) \qquad (*)$$

gegeben ist. Es gilt $c_{\lambda\beta,\theta d} = \lambda^n \theta^{-2} c_{\beta,d}$. Es gilt $c_{\beta,d} = 0$ genau dann, wenn β ausgeartet ist.

Beweis: Sei $G(v_1, \ldots, v_n, w_1, \ldots, w_n)$ die linke Seite von $(*)$. Aus den Eigenschaften der Determinante von Matrizen folgt, daß G eine $2n$-Linearform auf V ist, die schiefsymmetrisch in $v_1, \ldots, v_n$ und in $w_1, \ldots, w_n$ ist. Für feste $v_1, \ldots, v_n$ ist also $w_1, \ldots, w_n \to G(v_1, \ldots, v_n, w_1, \ldots, w_n)$ eine schiefsymmetrische Linearform auf V, also gibt es ein eindeutig bestimmtes $c(v_1, \ldots, v_n) \in K$ mit:

$$G(v_1, \ldots, v_n, w_1, \ldots, w_n) = c(v_1, \ldots, v_n) d(w_1, \ldots, w_n).$$

Es gibt eine Basis $(b_1, \ldots, b_n)$ von V mit $d(b_1, \ldots, b_n) = 1$, dann gilt

$$c(v_1, \ldots, v_n) = G(v_1, \ldots, v_n, b_1, \ldots, b_n).$$

Also ist $c(v_1, \ldots, v_n)$ ebenfalls ein Element von $D(V)$ und es gibt $c_{\beta,d} \in K$ mit $c(v_1, \ldots, v_n) = c_{\beta,d} d(v_1, \ldots, v_n)$, und $c_{\beta,d} = G(b_1, \ldots, b_n, b_1, \ldots, b_n)$ eindeutig bestimmt. Die zweite Behauptung ist trivial. Die letzte Aussage folgt aus $G(b_1, \ldots, b_n, b_1, \ldots, b_n) = \det(\mathrm{Mat}_B(\beta)) \overset{\text{Fakt 5.1.3}}{=} \det(\mathrm{Mat}_{B,B^*}(\sigma_\beta))$ und Fakt 5.2.1. □

Bemerkung: $(*)$ gilt auch für nichtsymmetrische Bilinearformen.
Fakt 1: Für $m > n$ und $v_1, \ldots, v_m, w_1, \ldots, w_m \in V$ gilt $\det((\beta(v_i, w_j))_{i,j=1}^m) = 0$.
Der Beweis ist Übungsaufgabe. □
Beispiel 1: Für das Euklidische Skalarprodukt β_n und für $d_n(v_1, \ldots, v_n) = \det(\mathrm{Mat}_Z(v_1, \ldots, v_n)) = \det(\mathrm{Mat}_S(v_1, \ldots, v_n))$, $d_n \in D(\mathbf{R}^n)$, gilt $c_{\beta_n,d_n} = 1$. Die Identität $(*)$ folgt in diesem Fall aus der Multiplikativität der Determinante und aus $\mathrm{Mat}_Z(v_1, \ldots, v_n)\mathrm{Mat}_S(w_1, \ldots, w_n) = (\beta_n(v_i, w_j))_{i,j=1}^n$.

Wir setzen von nun an β als nicht ausgeartet voraus.

Definition 1: $\mathrm{O}(\beta)$ ist die Menge aller Automorphismen u von V mit der Eigenschaft

$$\forall\, v, w \in V : \beta(u(v), u(w)) = \beta(v, w),$$

d.h. $u^*\beta = \beta$, versehen mit der Hintereinanderausführung von Automorphismen als Gruppenoperation. $\mathrm{O}(\beta)$ ist die **orthogonale** Gruppe von β.

Fakt 2: Für alle $u \in \mathrm{O}(\beta)$ gilt $|\det(u)| = 1$.

Beweis: Sei G die linke Seite von $(*)$ wie beim Beweis von Satz 1. Es gilt

$$\begin{aligned}
c_{\beta,d} \det(u)^2 d(v_1, \ldots, v_n) d(w_1, \ldots, w_n) &= \\
c_{\beta,d} d(u(v_1), \ldots, u(v_n)) d(u(w_1), \ldots, u(w_n)) &= \\
G(u(v_1), \ldots, u(v_n), u(w_1), \ldots, u(w_n)) &= \\
\det(\beta(u(v_i), u(w_j))_{i,j=1}^n) &= \\
G(v_1, \ldots, v_n, w_1, \ldots, w_n) &= \\
c_{\beta,d} d(v_1, \ldots, v_n) d(w_1, \ldots, w_n)
\end{aligned}$$

Weil $c_{\beta,d} \neq 0$ (β nicht ausgeartet), folgt daraus $\det(u)^2 = 1$, also $|\det(u)| = 1$. □

Definition 2:

$$\mathrm{SO}(\beta) = \{u \in \mathrm{O}(\beta) \mid \det(u) = 1\} = \ker(\mathrm{O} \stackrel{\det}{\to} \{1, -1\}) \subset O(\beta).$$

$\mathrm{SO}(\beta)$ ist eine Untergruppe von $\mathrm{O}(\beta)$. Erinnere an

$$\beta_{a,b}((x_i)_{i=1}^{a+b}, (y_i)_{i=1}^{a+b}) = \sum_{i=1}^{a} x_i y_i - \sum_{j=a+1}^{a+b} x_i y_i$$

auf $\mathbf{R}^{a+b}$ und $\beta_n = \beta_{n,0}$.

Definition 3:

$$\mathrm{O}(n) = \mathrm{O}(\beta_n),\ \mathrm{SO}(n) = \mathrm{SO}(\beta_n),\ \mathrm{O}(a,b) = \mathrm{O}(\beta_{a,b}),\ \mathrm{SO}(a,b) = \mathrm{SO}(\beta_{a,b}).$$

Beispiel 2: (Wir identifizieren hier 2×2-Matrizen mit Endomorphismen des $\mathbf{R}^2$).

- Sei $u(\varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$. Es gilt $\{u(\varphi) \mid \varphi \in \mathbf{R}\} = \mathrm{SO}(2)$, denn $u(\varphi)$ überführt die Orthonormalbasis $((1,0),(0,1))$ in die Orthonormalbasis $((\cos\varphi, -\sin\varphi), (\sin\varphi, \cos\varphi))$. Es gilt $\det(u(\varphi)) = 1$, also $u(\varphi) \in \mathrm{SO}(2)$. Alle Elemente von $\mathrm{SO}(2)$ können als $u(\varphi)$ dargestellt werden. Dabei gilt $u(\varphi) = u(\varphi')$ genau dann, wenn $\varphi - \varphi' = 2\pi\mathbf{Z}$. Es gilt $u(\varphi)u(\psi) = u(\varphi + \psi)$. $u(\varphi)$ ist eine Drehung um φ.
- Sei $l(t) = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}$. Dann gilt $l(t) \in \mathrm{SO}(1,1)$. Jedes Element von $\mathrm{SO}(1,1)$ kann in dieser Form auf eindeutige Weise dargestellt werden und $l(s)l(t) = l(s+t)$. Die $\mathrm{O}(a,1)$ spielen die Rolle der Lorenzgruppen für die spezielle Relativitätstheorie mit a räumlichen Dimensionen, x_{a+1} spielt dann die Rolle der Zeitkoordinate.

Fakt 3: Sei $q(x) = \beta(x,x)$ die zu β gehörende quadratische Form (vgl. Satz 5.5.1). Dann gilt $u \in \mathrm{O}(\beta)$ genau dann, wenn $q(u(x)) = u(x)$ für alle $x \in V$.

Beweis: Beachte $\beta(x,y) = \frac{1}{2}(\beta(x+y, x+y) - \beta(x,x) - \beta(y,y))$... □

5.6 Euklidische Vektorräume

In diesem Abschnitt ist $K = \mathbf{R}$.

Definition 1: Ein **Euklidischer Vektorraum** ist ein Paar $(V, \langle\cdot,\cdot\rangle)$, wobei V ein reeller Vektorraum und $\langle\cdot,\cdot\rangle : V \times V \to \mathbf{R}$, $(v,w) \to \langle v,w\rangle$ eine positiv definite, symmetrische Bilinearform ist. Dann sei $\|v\| = \sqrt{\langle v,v\rangle}$.

Beispiel: $V = \mathbf{R}^n$, $\langle x,y\rangle = \beta_n(x,y)$. $\|(x_i)_{i=1}^n\| = \sqrt{\sum_{i=1}^n x_i^2}$ ist nach Pythagoras die Länge der Strecke von 0 nach $(x_1, \ldots, x_n)$. Man kann also $\|v\|$ als die Länge von v deuten. Dann sollte die Dreiecksungleichung $\|u+v\| \le \|u\| + \|v\|$ gelten.

Satz 1: (Cauchy-Schwarz). Sei $\langle \cdot, \cdot \rangle$ positiv definit und v beliebig (auch unendlichdimensional). Dann gelten die CAUCHY-SCHWARZsche **Ungleichung**

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

und die **Dreiecksungleichung**

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

Beweis: Sei $\|x\| = \|y\| = 1$. Dann gilt

$$0 \leq \langle x - y, x - y \rangle = \langle x, x \rangle + \langle y, y \rangle - 2\langle x, y \rangle = 2 - 2\langle x, y \rangle. \qquad (*)$$

Es gilt also $\langle x, y \rangle \leq 1$ und auch $\langle x, -y \rangle \leq 1$, also $\langle x, y \rangle \geq -1$. Also gilt $|\langle x, y \rangle| \leq 1$, falls $\|x\| = \|y\| = 1$. Seien nun x und y beliebig. Wenn einer dieser Vektoren verschwindet, ist die Cauchy-Schwarzsche Ungleichung trivial. Andernfalls gilt $\|\frac{x}{\|x\|}\| = 1$, $\|\frac{y}{\|y\|}\| = 1$, also

$$\left| \left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle \right| \leq 1,$$

also $|\langle x, y \rangle| \leq \|x\| \|y\|$.
Beweis der Dreiecksungleichung:

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2. \qquad (**)$$

Bemerkung 1:

- In $(*)$ tritt Gleichheit nur für $x = y$ ein. Deswegen tritt in der Cauchy-Schwarzschen Ungleichung genau dann Gleichheit ein, wenn $\dim(\mathcal{L}(x, y)) \leq 1$. Ebenso wird die Dreiecksungleichung streng, sofern nicht x oder y verschwindet, oder ein positives $\lambda \in \mathbb{R}$ existiert mit $x = \lambda y$.
- $\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\lambda^2 \langle x, y \rangle} = |\lambda| \|x\|$.
- Aus der Dreiecksungleichung folgt, daß (V, d) mit $d(x, y) = \|x - y\|$ ein metrischer Raum ist: $d(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$. V ist ein **reeller Hilbertraum**, falls dieser metrische Raum vollständig ist. Dies ist immer der Fall für $\dim(V) < \infty$.

Definition 2: Sei V wie in Satz 1 definiert. Dann nennen wir zwei Vektoren $x, y \in V$ **orthogonal**, falls $\langle x, y \rangle = 0$; dann schreiben wir $x \perp y$.

Bemerkung 2: (PYTHAGORAS) $x \perp y \Leftrightarrow \|x + y\|^2 = \|x\|^2 + \|y\|^2 \Leftrightarrow \|x - y\|^2 = \|x\|^2 + \|y\|^2$.

Definition 3: Sei V wie in Satz 1 definiert. Dann sei für $x, y \in V - \{0\}$ $w(x, y)$ der durch die Eigenschaften $0 \leq w(x, y) \leq \pi$ und $\langle x, y \rangle = \|x\| \|y\| \cos(w(x, y))$ eindeutig bestimmte Winkel (Er existiert nach der Cauchy-Schwarzschen Ungleichung).

Bemerkung 3: $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x, y \rangle = \|x\|^2 + \|y\|^2 - 2\|x\| \|y\| \cos(w(x, y))$ ist der Cosinussatz der ebenen Trigonometrie.

Bemerkung 4: (GRAM-SCHMIDTsches Orthogonalisierungsverfahren): Sei $B = (b_1, \ldots, b_n)$ eine

Basis des euklidischen Vektorraumes V. Dann erhalten wir durch folgende induktive Vorschrift eine Orthonormalbasis $(v_1, \ldots, v_n)$:

$$\begin{aligned} v_1 &= \frac{b_1}{\|b_1\|} \\ v_k &= \frac{b_k - \sum_{i=1}^{k-1} \langle b_k, v_i \rangle v_i}{\|b_k - \sum_{i=1}^{k-1} \langle b_k, v_i \rangle v_i\|}, \end{aligned}$$

falls $v_1, \ldots, v_{k-1}$ schon berechnet sind. Man prüft induktiv, daß $\mathcal{L}(v_1, \ldots, v_{k-1}) = \mathcal{L}(b_1, \ldots, b_{k-1})$. Also $b_k \neq \sum_{i=1}^{k-1} \langle b_k, v_i \rangle v_i$. Damit ist v_k wohldefiniert. Offenbar $\|v_k\| = 1$. Für $j < k$ gilt $\langle v_k, v_j \rangle = c\langle b_k - \sum_{i=1}^{k-1} \langle b_k, v_i \rangle v_i, v_j \rangle = c\langle b_k, v_j \rangle - c \sum_{i=1}^{k-1} \langle b_k, v_i \rangle \langle v_i, v_j \rangle = 0$, wobei wir im Sinne eines Induktionsbeweises schon vorausgesetzt haben, daß die $(v_1, \ldots, v_{k-1})$ orthogonal sind, d.h. $\langle v_i, v_j \rangle = \delta_{ij}$ für $1 \leq i, j \leq k-1$.

Fakt 1: Sei V ein beliebiger reeller Vektorraum und $\langle \cdot, \cdot \rangle$ positiv definit. Sei $(e_1, \ldots, e_n)$ bezüglich $\langle \cdot, \cdot \rangle$ orthonormal.

1. Für $x \in V$ gilt die PARSEVALsche **Ungleichung**

$$\sum_{i=1}^{n} \langle x, e_i \rangle^2 \leq \|x\|^2.$$

2. Die Gleichheit gilt in dieser Ungleichung genau dann, wenn $x = \sum_{i=1}^{n} \langle x, e_i \rangle e_i$. Dies ist genau dann für alle $x \in V$ der Fall, wenn $(e_1, \ldots, e_n)$ eine Basis von V ist.

3. Sei $(e_1, \ldots, e_n)$ eine Orthonormalbasis, dann gilt

$$\langle x, y \rangle = \sum_{i=1}^{n} \langle x, e_i \rangle \langle y, e_i \rangle$$

für $x, y \in V$.

Beweis: Sei $\tilde{x} = x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i$. Dann gilt

$$\langle \tilde{x}, e_j \rangle = \langle x, e_j \rangle - \sum_{i=1}^{n} \langle x, e_i \rangle \langle e_i, e_j \rangle = 0,$$

also auch $\langle \tilde{x}, \sum_{i=1}^{n} \langle x, e_i \rangle e_i \rangle = 0$. Nach Pythagoras gilt

$$\|x\|^2 = \|\tilde{x}\|^2 + \|\sum_{i=1}^{n} \langle x, e_i \rangle e_i\|^2 = \|\tilde{x}\|^2 + \sum_{i=1}^{n} \langle x, e_i \rangle^2,$$

also $\|x^2\| - \sum_{i=1}^{n} \langle x, e_i \rangle^2 = \|\tilde{x}\|^2$. Daraus folgen (1) und der erste Teil von (2). Wenn $\tilde{x} = 0$ ist für alle $x \in V$, so erzeugen die $b_1, \ldots, b_n$ V und sind eine Basis. Wenn $(e_1, \ldots, e_n)$ eine Basis ist, so folgt aus der Orthogonalität der e_i, daß $x = \sum_{i=1}^{n} \alpha_i e_i$, $\alpha_i = \langle x, e_i \rangle$ und $\tilde{x} = 0$ gilt.
Weiterhin gilt

$$\langle x, y \rangle = \langle \sum_{i=1}^{n} \langle x, e_i \rangle e_i, \sum_{j=1}^{n} \langle y, e_j \rangle e_j \rangle = \sum_{i,j=1}^{n} \langle x, e_i \rangle \langle y, e_j \rangle \langle e_i, e_j \rangle = \sum_{i=1}^{n} \langle x, e_i \rangle \langle y, e_i \rangle.$$

□

Satz 2: Sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer Vektorraum. Sei l ein lineares Funktional auf V. Dann gibt es genau ein $r(l) \in V$ mit der Eigenschaft

$$\forall\, v \in V :\ l(v) = \langle v, r(l)\rangle.$$

Durch $r : V^* \to V$ ist ein Isomorphismus definiert.

Beweis: Erinnere an 5.1 und 5.2. Für eine beliebige symmetrische Bilinearform β auf V war $\sigma_\beta : V \to V^*$ durch $(\sigma_\beta(x))(y) = \beta(x,y)$ für alle $x, y \in V$ definiert. Es gilt $N(\beta) = \ker(\sigma_\beta)$, wegen $\dim(V^*) = \dim(V)$ ist also σ_β genau dann ein Isomorphismus, wenn β nicht ausgeartet ist. Sei $\beta(x,y) = \langle x, y\rangle$. Dann ist β nicht ausgeartet, also ist σ_β ein Isomorphismus. Sei $r = \sigma_\beta^{-1} : V^* \to V$ der zu σ_β inverse Isomorphismus. Dann leistet r das Gewünschte:

$$l(x) = \sigma_\beta(r(l))(x) = \beta(r(l), x) = \langle r(l), x\rangle. \qquad \square$$

Definition 4: Für einen Unterraum $X \subseteq V$ sei

$$X^\perp = \{y \in V \mid \forall\, x \in X : \langle x, y\rangle = 0\}$$

das **orthogonale Komplement** zu X.

Bemerkung: Wenn man den Isomorphismus $r : V^* \to V$ aus Satz 2 benutzt, um V^* mit V zu identifizieren, so geht diese Definition in die in Abschnitt 2 gegebene Definition über.

Satz 3: Für jeden Unterraum X eines Euklidischen Vektorraumes V gilt

$$X \oplus X^\perp = V.$$

Beweis: Sei $(b_1, \ldots, b_d)$ eine Orthonormalbasis von X. Für $v \in V$ sei $p_X(v) = \sum_{i=1}^d \langle v, b_i\rangle b_i \in X$. Man prüft leicht $\langle p_X(v), b_i\rangle = \langle v, b_i\rangle$ für $1 \le i \le d$, also $\langle p_X(v), x\rangle = \langle v, x\rangle$ für alle $x \in X$, also $v - p_X(v) \in X^\perp$. Also $X + X^\perp = V$. Aus $v \in X \cap X^\perp$ folgt nach Definition 4 $\langle v, v\rangle = 0$ und damit $v = 0$, also ist die Summe $X + X^\perp$ eine direkte Summe. $\square$

Definition 5: Sei X ein Unterraum eines Euklidischen Vektorraumes V. Die durch $p_X(v) \in X$ und $v - p_X(v) \in X^\perp$ (nach Satz 3) eindeutig definierte Abbildung $p_X : V \to V$ nennt man den **Orthogonalprojektor** auf X.

Bemerkung: Es gilt $p_X^2 = p_X$, denn $p_X(x) = x$ für $x \in X$.

Bemerkung 5: Satz 2 und Satz 3 gelten für reelle Hilberträume ähnlich, allerdings ist es notwendig, das lineare Funktional aus Satz 2 als stetig vorauszusetzen (Riesscher Satz) und in Satz 3 zu fordern, daß X abgeschlossen ist.

Definition 6: Sei $n = \dim V$. $d \in D(V)$ nennt man **normiert**, wenn die folgenden äquivalenten Bedingungen erfüllt sind:

1. $|d(b_1, \ldots, b_n)| = 1$ für eine geeignete Orthonormalbasis $b_1, \ldots, b_n$ von V.
2. $|d(b_1, \ldots, b_n)| = 1$ für jede Orthonormalbasis $b_1, \ldots, b_n$ von V.
3. Es gilt $c_{\langle \cdot, \cdot \rangle, d} = 1$ in der Formel für die Gramsche Determinante, d.h.

$$\begin{vmatrix} \langle v_1, w_1 \rangle & \ldots & \langle v_1, w_n \rangle \\ \vdots & \ddots & \vdots \\ \langle v_n, w_1 \rangle & \ldots & \langle v_n, w_n \rangle \end{vmatrix} = d(v_1, \ldots, v_n) d(w_1, \ldots, w_n).$$

Bemerkung: (3)⇒(2)⇒(1) ist klar, (1)⇒(3) folgt aus dem Satz über die Gramsche Determinante.

Fakt 2: Mit d ist auch $-d$ normiert. Es gibt genau zwei normierte Elemente von $D(V)$ und zwar in jeder Halbachse von $D(V)$ genau eines.

Beweis: Sei $d \neq 0$ beliebig. Dann gilt $c = c_{\langle \cdot, \cdot \rangle, d} > 0$, denn für $v_i = b_i$ und $v_1, \ldots, v_n$ orthonormal ist die Gramsche Determinante $G(v_1, \ldots, v_n, w_1, \ldots, w_n) = 1$. Dann ist $d/\sqrt{c}$ normiert. Sei d normiert und $\lambda \in \mathbb{R}$. Offenbar ist λd genau dann normiert, wenn $\lambda^2 = 1$, woraus die restlichen Aussagen folgen. □

Lemma 1: Sei $\tilde{V} \subset V$ ein Unterraum der Kodimension eins (d.h. $\dim(V) - \dim(\tilde{V}) = 1 = \dim \tilde{V}^\perp$). Sei $h \in \tilde{V}^\perp$ und $\|h\| = 1$. Sei $d \in D(V)$ normiert. Dann gilt

1. $$\tilde{d}(x_1, \ldots, x_{n-1}) = d(h, x_1, \ldots, x_{n-1}), \; x_1, \ldots, x_{n-1} \in \tilde{V}$$
   definiert ein normiertes Element von $D(\tilde{V})$.
2. Für $v_1 \in V$ und $v_2, \ldots, v_n \in \tilde{V}$ gilt
   $$d(v_1, \ldots, v_n) = \langle v_1, h \rangle \tilde{d}(v_2, \ldots, v_n).$$

Beweis: Die erste Behauptung folgt sofort aus Definition 6, denn wenn $b_1, \ldots, b_{n-1}$ eine Orthonormalbasis für $\tilde{V}$ ist, so ist $(h, b_1, \ldots, b_{n-1})$ eine Orthonormalbasis für V.
Sei $v_1' = v_1 - \langle v_1, h \rangle h$. Dann gilt $v_1' \in (\mathbb{R}h)^\perp = \tilde{V}$, denn aus Satz 3 folgt leicht $X^{\perp\perp} = X$ für alle Unterräume $X \subseteq V$. Also ist $(v_1', v_2, \ldots, v_n)$ linear abhängig, denn alle diese n Vektoren gehören zu dem $(n-1)$-dimensionalen Unterraum $\tilde{V}$. Daraus folgt $d(v_1', v_2, \ldots, v_n) = 0$ wegen der Schiefsymmetrie von d. Also folgt

$$\begin{aligned} d(v_1, \ldots, v_n) &= d(v_1' + \langle v_1, h \rangle h, v_2, \ldots, v_n) = d(v_1', v_2, \ldots, v_n) + \langle v_1, h \rangle d(h, v_2, \ldots, v_n) \\ &= \langle v_1, h \rangle \tilde{d}(v_2, \ldots, v_n). \end{aligned}$$

□

Folgerung 1: Sei

$$S(v_1, \ldots, v_n) = \{ \sum_{i=1}^{n} \lambda_i v_i \mid 0 \leq \lambda \leq 1 \}$$

das von den v_i erzeugte Parallelspat. Möge das n-dimensionale Volumen von Parallelspaten im n-dimensionalen Euklidischen Raum folgende Eigenschaften haben:

1. Das Volumen des 1-Spates ist die Norm des erzeugenden Vektors.

2. Das Volumen eines n-Spates $S(v_1, \ldots, v_n)$ ist gleich dem Produkt aus dem $(n-1)$-dimensionalen Volumens eines Seitenspates (etwa $S(v_2, \ldots, v_n)$) und der Höhe auf dieses Seitenspat.

Dann gilt

$$\text{Volumen}(S(v_1, \ldots, v_n)) = |d(v_1, \ldots, v_n)|, \; d \in D(V) \text{ normiert} .$$

Beispiele: Fläche von

$$\mathbf{R}^2 \supset S((x_1, x_2), (y_1, y_2)) = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}.$$

Volumen von

$$\mathbf{R}^3 \supset S((x_1, x_2, x_3), (y_1, y_2, y_3), (z_1, z_2, z_3)) = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix},$$

denn auf dem $\mathbf{R}^n$ ist $d_n(x_1, \ldots, x_n) = \det(\text{Mat}_z(x_1, \ldots, x_n))$ bezüglich β_n normiert, denn $d_n(e_1, \ldots, e_n) = 1$ für die Standardbasis.

Beweis von Folgerung 1 durch Induktion nach n beginnend mit $n = 1$. Betrachte die $(n-1)$-dimensionale Seite $S(v_2, \ldots, v_n)$ von $S(v_1, \ldots, v_n)$. Sei $\tilde{V}$ der von den $v_2, \ldots, v_n$ erzeugte Unterraum. Wenn $\dim(\tilde{V}) < n-1$, so sind $v_2, \ldots, v_n$ linear abhängig, es folgt durch die Induktionsannahme, daß das Volumen$(S(v_2, \ldots, v_n)) = 0$, also auch Volumen$(S(v_1, \ldots, v_n)) = 0 = d(v_1, \ldots, v_n)$. Sei also $\dim(\tilde{V}) = n-1$ und seien h und $\tilde{d}$ wie in Lemma 1. Dann ist $\langle v_1, h\rangle$ die Höhe von $S(v_1, \ldots, v_n)$ bezüglich der Seite $S(v_2, \ldots, v_n)$, also gilt Volumen$(S(v_1, \ldots, v_n)) = |\langle v_1, h\rangle| \cdot$Volumen$(S(v_2, \ldots, v_n)) = |\langle v_2, h\rangle| |\tilde{d}(v_2, \ldots, v_n)| = |d(v_1, \ldots, v_n)|$. □

5.6.1 Das Vektorprodukt auf einem dreidimensionalen orientierten Euklidischen Raum

Sei $(V, \langle \cdot, \cdot \rangle)$ ein dreidimensionaler Euklidischer Raum und sei $d \in D(V)$ normiert.

Definition 7: Für $x, y \in V$ sei $x \times y \in V$ der durch

$$\forall\, z \in V : \langle x \times y, z\rangle = d(x, y, z) \qquad (*)$$

eindeutig definierte Vektor.

Bemerkung: Die Eindeutigkeit und Existenz von $x \times y$ folgt durch Anwendung von Satz 2 auf das Funktional $z \mapsto d(x, y, z)$.

Definition 8: Eine **Liealgebra** ist ein Paar (V, C), wobei $C : V \times V \to V$ schiefsymmetrisch und bilinear ist, so daß die Jacobi-Identität gilt:

$$C(a, C(b, c)) + C(c, C(a, b)) + C(b, C(c, a)) = 0.$$

Bemerkung: Meistens schreibt man C als Kommutator $C(x, y) = [x, y]$. Für einen Vektorraum W ist $V = L(W, W)$ versehen mit $[x, y] = x \circ y - y \circ x$ eine Liealgebra. Sei V eine Liealgebra und sei $L_x(y)$ der Kommutator von x mit y für $x, y \in V$. Dann ist die Jacobi-Identität zu $L_x L_y - L_y L_x = L_{[x,y]}$ äquivalent.

Fakt 3:

1. $\times$ ist bilinear und schiefsymmetrisch.
2. Sei $u \in O(V, \langle\cdot,\cdot\rangle)$. Dann gilt $u(x \times y) = \det(u)(u(x) \times u(y))$.
3. $x \times y$ ist zu x und y orthogonal.
4. $x \times y = 0$ genau dann, wenn x und y linear abhängig sind.
5. Sei (e_1, e_2, e_3) eine positiv orientierte Orthonormalbasis, dann gilt

$$x \times y = \begin{vmatrix} e_1 & e_2 & e_3 \\ \langle x, e_1\rangle & \langle x, e_2\rangle & \langle x, e_3\rangle \\ \langle y, e_1\rangle & \langle y, e_2\rangle & \langle y, e_3\rangle \end{vmatrix}$$

(zu berechnen nach der Leibnizschen Formel).

Beweis: (1), (2), (3) folgen leicht aus $(*)$, z.B. (2):

$$\begin{aligned} \langle z, \det(u)u(x) \times u(y)\rangle &\overset{(*)}{=} \det(u)d(u(x), u(y), z) = \det(u)d(u(x), u(y), u(u^{-1}(z))) \\ &= \det(u)^2 d(x, y, u^{-1}(z)) \overset{(*)}{=} \langle x \times y, u^{-1}(z)\rangle \\ &= \langle u(x \times y), z\rangle. \end{aligned}$$

Da das für alle z gilt, folgt (2).
Für (3) bemerkt man $\langle x \times y, x\rangle \overset{(*)}{=} d(x, y, x) = 0$.
Wenn x und y linear abhängig sind, so gilt $\langle x \times y, z\rangle = d(x, y, z) = 0$ für alle z wegen der Schiefsymmetrie von d. Also $x \times y = 0$. Wenn x und y linear unabhängig sind, so ist (x, y, z) eine Basis von V und es gilt $0 \neq d(x, y, z) = \langle x \times y, z\rangle$, also $x \times y \neq 0$, was (4) beweist. Weiterhin gilt für alle $c \in V$

$$\begin{aligned} \left\langle \begin{vmatrix} e_1 & e_2 & e_3 \\ \langle x, e_1\rangle & \langle x, e_2\rangle & \langle x, e_3\rangle \\ \langle y, e_1\rangle & \langle y, e_2\rangle & \langle y, e_3\rangle \end{vmatrix}, c \right\rangle &= \begin{vmatrix} \langle c, e_1\rangle & \langle c, e_2\rangle & \langle c, e_3\rangle \\ \langle x, e_1\rangle & \langle x, e_2\rangle & \langle x, e_3\rangle \\ \langle y, e_1\rangle & \langle y, e_2\rangle & \langle y, e_3\rangle \end{vmatrix} \overset{\text{Gram}}{=} d(e_1, e_2, e_3)d(c, x, y) \\ &= d(x, y, c), \end{aligned}$$

denn die Permutation von $d(c, x, y)$ ist gerade und $d(e_1, e_2, e_3) = 1$, also wird $(*)$ erfüllt. □

Beispiel: Im $\mathbf{R}^3$ gilt für das Standardskalarprodukt und die Standardorientierung

$$x \times y = \begin{vmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - y_2x_1)$$

mit $x = (x_i)_{i=1}^3$, $y = (y_i)_{i=1}^3$ und der Standardbasis $(e_i)_{i=1}^3$.

Satz 4: (H. Grassmann) Es gelten die Formeln

1.
$$\langle x \times y, a \times b\rangle = \begin{vmatrix} \langle a, x\rangle & \langle a, y\rangle \\ \langle b, x\rangle & \langle b, y\rangle \end{vmatrix}; \qquad (\dagger)$$

2.
$$a \times (b \times c) = \langle a, c\rangle b - \langle a, b\rangle c. \qquad (\ddagger)$$

Beweis: Die zweite Formel folgt aus der ersten, denn

$$\begin{aligned}\langle a\times(b\times c),z\rangle &\overset{(*)}{=} d(a,b\times c,z)=d(z,a,b\times c)=\langle z\times a,b\times c\rangle\\ &= \begin{vmatrix}\langle z,b\rangle & \langle z,c\rangle\\ \langle a,b\rangle & \langle a,c\rangle\end{vmatrix}=\langle z,\langle a,c\rangle b-\langle a,b\rangle c\rangle.\end{aligned}$$

Damit ist (‡) bewiesen. Es verbleibt der Beweis von (†). Dazu benötigen wir folgendes

Lemma 2: Sei (e_1,e_2,e_3) eine Orthonormalbasis von V, und sei $c,d,g,h\in V$. Dann gilt

$$\sum_{i=1}^{3}\langle e_i,d\rangle\begin{vmatrix}\langle g,c\rangle & \langle g,e_i\rangle\\ \langle h,c\rangle & \langle h,e_i\rangle\end{vmatrix}=\begin{vmatrix}\langle g,c\rangle & \langle g,d\rangle\\ \langle h,c\rangle & \langle h,d\rangle\end{vmatrix},$$

denn die linke Seite ist

$$\sum_{i=1}^{3}\langle g,c\rangle\langle h,e_i\rangle\langle e_i,d\rangle-\sum_{i=1}^{3}\langle h,c\rangle\langle g,e_i\rangle\langle e_i,d\rangle=\langle g,c\rangle\langle h,d\rangle-\langle h,c\rangle\langle d,g\rangle,$$

womit Lemma 2 bewiesen ist. □

Wir können nun die linke Seite von (†) berechnen, sei dazu (e_1,e_2,e_3) eine Orthonormalbasis:

$$\begin{aligned}\langle x\times y,a\times b\rangle &= \sum_{i=1}^{3}\langle x\times y,e_i\rangle\langle a\times b,e_i\rangle=\sum_{i=1}^{3}d(x,y,e_i)d(a,b,e_i)\\ &\overset{\text{Gram}}{=} \sum_{i=1}^{3}\begin{vmatrix}\langle a,x\rangle & \langle a,y\rangle & \langle a,e_i\rangle\\ \langle b,x\rangle & \langle b,y\rangle & \langle b,e_i\rangle\\ \langle e_i,x\rangle & \langle e_i,y\rangle & \langle e_i,e_i\rangle\end{vmatrix} \quad \text{Auflösung nach der letzten Zeile}\\ &= \sum_{i=1}^{3}\left(\langle e_i,x\rangle\begin{vmatrix}\langle a,y\rangle & \langle a,e_i\rangle\\ \langle b,y\rangle & \langle b,e_i\rangle\end{vmatrix}-\langle e_i,y\rangle\begin{vmatrix}\langle a,x\rangle & \langle a,e_i\rangle\\ \langle b,x\rangle & \langle b,e_i\rangle\end{vmatrix}+\begin{vmatrix}\langle a,x\rangle & \langle a,y\rangle\\ \langle b,x\rangle & \langle b,y\rangle\end{vmatrix}\right)\\ &\overset{\text{Lemma 2}}{=} \begin{vmatrix}\langle a,y\rangle & \langle a,x\rangle\\ \langle b,y\rangle & \langle b,x\rangle\end{vmatrix}-\begin{vmatrix}\langle a,x\rangle & \langle a,y\rangle\\ \langle b,x\rangle & \langle b,y\rangle\end{vmatrix}+3\begin{vmatrix}\langle a,x\rangle & \langle a,y\rangle\\ \langle b,x\rangle & \langle b,y\rangle\end{vmatrix}=\begin{vmatrix}\langle a,x\rangle & \langle a,y\rangle\\ \langle b,x\rangle & \langle b,y\rangle\end{vmatrix}.\end{aligned}$$

□

Folgerung 2: Das Vektorprodukt $x\times y$ erfüllt folgende Eigenschaften, die es eindeutig charakterisieren:

1. $x\times y$ ist zu den beiden Argumenten orthogonal.
2. Für seine Länge gilt $\|x\times y\|=\|x\|\|y\|\sin(w(x,y))$.
3. Das Dreibein $(x,y,x\times y)$ ist rechtshändig, d.h. $d(x,y,x\times y)\geq 0$.

Beweis: Daß diese Bedingungen das Vektorprodukt eindeutig bestimmen, ist klar.

1. Fakt 3.3
2. folgt aus (†), angewendet mit $b=y$ und $a=x$:

$$\begin{aligned}\|x\times y\|^2 &= \langle x\times y,x\times y\rangle=\begin{vmatrix}\langle x,x\rangle & \langle x,y\rangle\\ \langle x,y\rangle & \langle y,y\rangle\end{vmatrix}=\|x\|^2\|y\|^2-\langle x,y\rangle^2\\ &\overset{\text{Def. 3}}{=} \|x\|^2\|y\|^2-(\cos(w(x,y))\|x\|\|y\|)^2=\sin^2(w(x,y))\|x\|^2\|y\|^2,\end{aligned}$$

   damit folgt die Behauptung.

3. folgt leicht aus $(*)$, denn $d(x,y,x\times y)=\langle x\times y, x\times y\rangle \geq 0$. $\square$

Folgerung 3: Das Vektorprodukt erfüllt die Jacobi-Identität

$$a\times(b\times c)+c\times(a\times b)+b\times(c\times a)=0.$$

Es ist nicht assoziativ, bilinear und schiefsymmetrisch und daher eine Lie-Algebra.

Beweis: Wende $(\ddagger)$ an:

$$\begin{aligned} a\times(b\times c)+c\times(a\times b)+b\times(c\times a) &= \\ b\langle a,c\rangle - c\langle a,b\rangle + a\langle b,c\rangle - b\langle a,c\rangle + c\langle a,b\rangle - a\langle b,c\rangle &= 0 \end{aligned}$$

5.7 Hermitesche Formen

In diesem Abschnitt arbeiten wir über den komplexen Zahlen. Erinnere an $\overline{x+iy}=x-iy$ und an $\overline{z\cdot z'}=\overline{z}\cdot\overline{z}'$, $\overline{z+z'}=\overline{z}+\overline{z}'$, $z\overline{z}=|z|^2=x^2+y^2>0$ für $z=x+iy$. Sei für $x,y,z\in\mathbb{C}$

$$\gamma:\mathbb{C}\times\mathbb{C}\to\mathbb{C},\quad \gamma(x,y)=x\overline{y}.$$

Dann gilt $\gamma(z,z)=z\overline{z}=|z|^2$. γ ist nicht bilinear, sondern linear im ersten Argument ($\gamma(x+\lambda y,z)=\gamma(x,y)+\lambda\gamma(y,z)$) und antilinear im zweiten Argument ($\gamma(x,y+\lambda z)=\gamma(x,y)+\overline{\lambda}\gamma(x,z)$). Statt der üblichen Symmetrie des Skalarprodukts gilt $\gamma(x,y)=\overline{\gamma(y,x)}$.

Definition 1: Sei V ein komplexer Vektorraum. Eine **hermitesche Form** auf V ist eine Abbildung $\eta: V\times V\to\mathbb{C}$ mit folgenden Eigenschaften:

1. $\eta(x+\lambda\tilde{x},y)=\eta(x,y)+\lambda\eta(\tilde{x},y)$, $x,\tilde{x},y\in V$, $\lambda\in\mathbb{C}$;
2. $\eta(x,y)=\overline{\eta(y,x)}$

Bemerkung: Dann gilt

$$\begin{aligned} \eta(x,y+\lambda\tilde{y}) &= \overline{\eta(y+\lambda\tilde{y},x)}=\overline{\eta(y,x)+\lambda\eta(\tilde{y},x)}=\overline{\eta(y,x)}+\overline{\lambda\eta(\tilde{y},x)} \\ &= \eta(x,y)+\overline{\lambda}\eta(x,\tilde{y}). \end{aligned}$$

Also ist η antilinear im zweiten Argument.

Die Theorie der hermiteschen Formen auf komplexen Vektorräumen verläuft analog zur Theorie der symmetrischen Bilinearformen auf reellen Vektorräumen.

Definition 2: Sei $B=(b_1,\ldots,b_n)$ eine Basis von V (also V endlichdimensional). Wir definieren

$$\mathrm{Mat}_B(\eta)=(\eta(b_i,b_j))_{i,j=1}^n.$$

Fakt 1: Für eine komplexe $n\times n$-Matrix $A=(h_{ij})_{i,j=1}^n$ sei A^* durch $A^*=\overline{A^T}=(\overline{a_{ji}})_{i,j=1}^n$ (koeffizientenweise konjugiert) definiert. A^* ist die **adjungierte Matrix** zu A, es gilt $(A+B)^*=A^*+B^*$, $(AB)^*=B^*A^*$, $(\lambda A)^*=\overline{\lambda}A^*$ für $\lambda\in\mathbb{C}$. Aus der Leibnizschen Determinantenformel folgt leicht $\det(\overline{A})=\overline{\det(A)}$, also folgt $\det A^*=\overline{\det A}$.
Die Menge der hermiteschen Formen auf V bildet mit den üblichen Regeln für Addition und Multiplikation einen reellen Vektorraum: $(\eta+\tilde{\eta})(x,y)=\eta(x,y)+\tilde{\eta}(x,y)$, $(t\eta)(x,y)=t\eta(x,y)$,

$t \in \mathbb{R}$. Ebenso bildet die Menge der selbstadjungierten Matrizen $\{A \in \mathrm{Mat}\,(n,n,\mathbb{C}) \mid A = A^*\}$ einen reellen Vektorraum. Man prüft leicht nach, daß für jede Basis B von V die Abbildung

$$\begin{array}{rcl} \{\text{hermitesche Formen auf } V\} & \to & \{\text{selbstadjungierte } n \times n\text{-Matrizen}\} \\ \eta & \mapsto & \mathrm{Mat}_B(\eta) \end{array}$$

ein Isomorphismus von reellen Vektorräumen ist.
Sei $f : V' \to V$ eine lineare Abbildung zwischen komplexen Vektorräumen. Sei $(f^*\eta)(x,y) = \eta(f(x), f(y))$, dies ist eine hermitesche Form auf V' (wobei η eine hermitesche Form auf V ist). Sei B' eine Basis von V'. Man prüft leicht nach:

$$\mathrm{Mat}_{B'}(f^*\eta) = \mathrm{Mat}_{B',B}(f)^* \mathrm{Mat}_B(\eta) \mathrm{Mat}_{B',B}(f). \qquad \square$$

Bemerkung: Wenn η eine hermitesche Form ist, so gilt $\eta(x,x) = \overline{\eta(x,x)}$, also ist $\eta(x,x)$ reell.

Theorem 4: (Analogon des Sylvesterschen Trägheitssatzes) Sei η eine Hermitesche Form auf einem endlichdimensionalen komplexen Vektorraum V. Dann gibt es eine Basis $B = (b_1, \ldots, b_n)$ von V mit der Eigenschaft, daß a, b existieren mit

$$\eta(b_i, b_j) = \begin{cases} 1 & 1 \le i = j \le a \\ -1 & a < i = j \le a+b \\ 0 & \text{sonst.} \end{cases}$$

Beweis: Ziemlich analog zum Beweis des Sylvesterschen Trägheitssatzes. Die Eindeutigkeit von a und b kann man durch Anwendung von Theorem 3 auf den reellen Vektorraum V, die reelle symmetrische Bilinearform $\mathrm{Re}\,(\eta)$ und die reelle Basis $(b_1, ib_1, \ldots, b_n, ib_n)$ beweisen.

Bemerkung: In der Situation des Theorems nennen wir η nichtentartet, falls die folgenden äquivalenten Bedingungen erfüllt sind:

1. $a + b = n$;
2. $N(\eta) = \{x \in V \mid \forall\, y \in V : \eta(x,y) = 0\} = \{0\}$.

Sei η nicht entartet. Sei $d \in D(V)$. Dann gibt es eine reelle Konstante $c_{\eta,d}$ mit

$$\begin{vmatrix} \eta(v_1,w_1) & \ldots & \eta(v_1,w_n) \\ \vdots & \ddots & \vdots \\ \eta(v_n,w_1) & \ldots & \eta(v_n,w_n) \end{vmatrix} = c_{\eta,d}\, d(v_1,\ldots,v_n)\overline{d(w_1,\ldots,w_n)}.$$

Es gilt $c_{\eta,d} > 0$, falls b gerade und $c_{\eta,d} < 0$ falls b ungerade ($c_{\eta,d} = 0$ falls η entartet sein sollte). Die **unitäre Gruppe** $U(\eta)$ ist als die Gruppe aller Automorphismen u von V mit $\eta(u(x), u(y)) = \eta(x,y)$ für alle $x, y \in V$ definiert. Es gilt $\det(u)\overline{\det(u)} = 1$, also $|\det(u)| = 1$. Die Gruppe $SU(\eta)$ ist als $SU(\eta) = \{u \in U(\eta) \mid \det(u) = 1\}$ definiert.

Definition 3: Die hermitesche Form η auf V ist positiv definit, falls $\eta(x,x) > 0$ für alle $x \in V - \{0\}$ gilt. In diesem Fall schreiben wir η auch als Skalarprodukt $\eta(x,y) = \langle x, y\rangle$ und setzen $\|x\| = \sqrt{\langle x, x\rangle}$. Wenn V endlichdimensional sein sollte, so nennen wir $(V, \langle\cdot,\cdot\rangle)$ einen endlichdimensionalen **Hilbertraum**.

Bemerkung 3: Weil η hermitesch ist, ist $\eta(x,x)$ reell, also ist die Bedingung $\eta(x,x) > 0$

sinnvoll.

Beispiel 1: Das Standardskalarprodukt auf $\mathbb{C}^n$ ist durch $\langle x, y\rangle = \sum_{i=1}^n x_i\overline{y_i}$, $x = (x_i)_{i=1}^n$, $y = (y_i)_{i=1}^n$, definiert. Es gilt $\langle x, x\rangle = \sum_{i=1}^n x_i\overline{x_i} = \sum_{i=1}^n |x_i|^2$.
Sei C der Raum der 2π-periodischen stetigen Funktionen $f : \mathbb{R} \to \mathbb{C}$. Durch

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\,dt$$

wird eine positiv definite hermitesche Form auf C definiert. Es gilt

$$\|f\| = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2\,dt}.$$

Sei l_2 die Menge aller Folgen $(z_i)_{i=0}^\infty$ mit $z_i \in C$ und $\sum_{i=1}^\infty |z|^2 < \infty$. Dann ist l_2 ein komplexer Vektorraum und durch $\langle (x_i)_{i=0}^\infty, (y_i)_{i=0}^\infty\rangle = \sum_{i=0}^\infty x_i\overline{y_i}$ ist eine positiv definite hermitesche Form definiert.
Wir konstruieren analog zur Theorie der Euklidischen Vektorräume die Theorie der positiv definiten Hermiteschen Formen.

Satz 1: (Cauchy-Schwarz) Sei $\langle\cdot,\cdot\rangle$ eine positiv definite Hermitesche Form auf einem komplexen Vektorraum V. Dann gelten

1. $$|\langle v, w\rangle| \le \|v\|\|w\|,$$
   Gleichheit gilt genau dann, wenn v und w linear abhängig sind.
2. $$\|v + w\| \le \|v\| + \|w\| \qquad \text{(Dreiecksungleichung)},$$
   Gleichheit gilt genau dann, wenn $v = 0$ oder wenn eine reelle Zahl $\lambda \ge 0$ mit $w = \lambda v$ existiert.

Beweis:

1. Wie dürfen $v \neq 0$ und $w \neq 0$ voraussetzen, sonst ist die Ungleichung trivial. Indem wir v durch $v/\|v\|$ und w durch $w/\|w\|$ ersetzen, können wir $\|v\| = \|w\| = 1$ annehmen. Falls $\langle v, w\rangle \neq 0$, sei $z = \langle v, w\rangle/|\langle v, w\rangle|$. Sei $\tilde{w} = zw$, dann gilt $\|w\| = \|\tilde{w}\|$ und

   $$\langle v, \tilde{w}\rangle = \langle v, zw\rangle = \overline{z}\langle v, w\rangle = \frac{\overline{\langle v, w\rangle}\langle v, w\rangle}{|\langle v, w\rangle|} = |\langle v, w\rangle|.$$

   Indem wir also w durch $\tilde{w}$ ersetzen, dürfen wir annehmen, daß $\langle v, w\rangle$ reell und ≥ 0 ist. Dann gilt $(*)$:

   $$\begin{aligned} 0 \le \langle v - w, v - w\rangle &= \langle v, v\rangle + \langle w, w\rangle - \langle v, w\rangle - \langle w, v\rangle \\ &= 1 + 1 - \langle v, w\rangle - \overline{\langle v, w\rangle} \\ &= 2 - 2|\langle v, w\rangle| \end{aligned}$$

   Also $|\langle v, w\rangle| \le 1$ für $\|v\| = \|w\| = 1$ und die erste Behauptung in (1) folgt. Offenbar kann Gleichheit in $(*)$ nur für $v = w$ eintreten. Daher kann Gleichheit in der Cauchy-Schwarzschen Ungleichung nur dann eintreten, wenn v und w linear abhängig sind.

2. Betrachte V als reellen Vektorraum mit der positiv definiten symmetrischen Bilinearform $\beta(x,y) = \operatorname{Re}\langle x,y\rangle$. Wende die Dreiecksungleichung aus 5.6 an. □

Bemerkung 4: Sei $d(x,y) = \|x-y\|$. Dann ist (V,d) ein metrischer Raum. $(V,\langle\cdot,\cdot\rangle)$ nennt man einen **Hilbertraum**, falls dieser metrische Raum vollständig ist. Es gilt $\|\lambda v\| = |\lambda|\|v\|$ für $\lambda \in \mathbb{C}$. Insbesondere ändert die Multiplikation mit einer komplexen Zahl vom Betrag 1 die Norm nicht.

Beispiel: Aus der Cauchy-Schwarzschen Ungleichung, angewendet auf $\mathbb{C}^n$ und $\mathbb{C}$ aus Beispiel 1, folgt

$$\begin{aligned}\left|\sum_{i=1}^n x_i\overline{y}_i\right| &\leq \sqrt{\sum_{i=1}^n |x_i|^2}\sqrt{\sum_{i=1}^n |y_i|^2} \qquad (x_i)_{i=1}^n,\ (y_i)_{i=1}^n \in \mathbb{C}^n \\ \left|\int_0^{2\pi} f(x)\overline{g(x)}\,dx\right| &\leq \sqrt{\int_0^{2\pi} |f(x)|^2\,dx}\sqrt{\int_0^{2\pi} |g(x)|^2\,dx}.\end{aligned}$$

Aus der Dreiecksungleichung für $\mathbb{C}^n$ folgt

$$\sqrt{\sum_{i=1}^n |x_i+y_i|^2} \leq \sqrt{\sum_{i=1}^n |x_i|^2}\sqrt{\sum_{i=1}^n |y_i|^2}$$

Also folgt $\sum_{i=1}^\infty |x_i+y_i|^2 < \infty$, falls $\sum_{i=1}^\infty |x_i|^2 < \infty$ und $\sum_{i=1}^\infty |y_i|^2 < \infty$. Damit ist l_2 aus Beispiel 1 tatsächlich ein Vektorraum. Aus der Cauchy-Schwarzschen Ungleichung für $\mathbb{C}^n$ folgt

$$\sum_{i=1}^n |x_iy_i| \leq \sqrt{\sum_{i=1}^n |x_i|^2}\sqrt{\sum_{i=1}^n |y_i|^2} \leq \sqrt{\sum_{i=1}^\infty |x_i|^2}\sqrt{\sum_{i=1}^\infty |y_i|^2}.$$

Also konvergiert $\sum_{i=1}^\infty x_i\overline{y_i}$ absolut, daher ist das Skalarprodukt auf l_2 wohldefiniert. Die Cauchy-Schwarzsche Ungleichung für l_2 besagt

$$\left|\sum_{i=1}^\infty x_i\overline{y_i}\right| \leq \sqrt{\sum_{i=1}^\infty |x_i|^2}\sqrt{\sum_{i=1}^\infty |y_i|^2}.$$

l_2 ist ein Hilbertraum. Der Raum C ist kein Hilbertraum. Man kann ihn zu einem Raum meßbarer Funktionen vervollständigen.

Definition 4: In der Situation von Satz 1 nennen wir ein n-Tupel $(v_1,\ldots,v_n)$ von Vektoren aus V **orthonormal**, falls $\langle v_i,v_j\rangle = \delta_{ij}$.

Fakt 2: Dann gilt

$$\left\|\sum_{i=1}^n \lambda_i v_i\right\| = \sqrt{\sum_{i=1}^n |\lambda_i|^2},$$

denn $\langle\sum_{i=1}^n \lambda_i v_i, \sum_{i=1}^n \lambda_i v_i\rangle = \sum_{i,j=1}^n \lambda_i\overline{\lambda_j}\langle v_i,v_j\rangle = \sum_{i,j=1}^n \lambda_i\overline{\lambda_j}\delta_{ij} = \sum_{i=1}^n |\lambda_i|^2$. □

Bemerkung 5: Das GRAM-SCHMIDTsche Orthonormalisierungsverfahren. Sei $(w_1,\ldots,w_n)$ eine Basis von V. Definiere $(v_1,\ldots,v_n)$ induktiv wie folgt:

$$\begin{aligned} v_1 &= \frac{w_1}{\|w_1\|}; \\ v_k &= \frac{w_k - \sum_{i=1}^{k-1}\langle w_k, v_i\rangle v_i}{\|w_k - \sum_{i=1}^{k-1}\langle w_k, v_i\rangle v_i\|}, \end{aligned}$$

falls $v_1, \ldots, v_{k-1}$ schon berechnet worden sind. Der Beweis erfolgt ähnlich wie im reellen Fall.

Fakt 3:

1. In der Situation von Satz 1 sei $(v_1, \ldots, v_n) \in V^n$ orthonormal. Für $x \in V$ gilt die PARSEVALsche **Ungleichung**
$$\sum_{i=1}^{n} |\langle x, v_i \rangle| \leq \|x\|^2.$$
2. Gleichheit tritt genau dann ein, wenn $x = \sum_{i=1}^{n} \langle x, v_i \rangle v_i$. Dies ist genau dann der Fall, wenn x in der linearen Hülle der v_i liegt.
3. Wenn $(v_1, \ldots, v_n)$ sogar eine Orthonormalbasis von V ist, so gilt
$$\langle x, y \rangle = \sum_{i=1}^{n} \langle x, v_i \rangle \langle v_i, y \rangle.$$

Beweis: Wie im reellen Fall berechnet man

$$\begin{aligned}
\left\| x - \sum_{i=1}^{n} \langle x, v_i \rangle v_i \right\|^2 &= \left\langle x - \sum_{i=1}^{n} \langle x, v_i \rangle v_i, x - \sum_{i=1}^{n} \langle x, v_i \rangle v_i \right\rangle \\
&= \|x\|^2 + \left\| \sum_{i=1}^{n} \langle x, v_i \rangle v_i \right\|^2 - \left\langle x, \sum_{i=1}^{n} \langle x, v_i \rangle v_i \right\rangle - \left\langle \sum_{i=1}^{n} \langle x, v_i \rangle v_i, x \right\rangle \\
&\overset{\text{Fakt 2}}{=} \|x\|^2 + \sum_{i=1}^{n} |\langle x, v_i \rangle|^2 - \sum_{i=1}^{n} \overline{\langle x, v_i \rangle} \langle x, v_i \rangle - \sum_{i=1}^{n} \langle x, v_i \rangle \langle v_i, x \rangle \\
&= \|x\|^2 - \sum_{i=1}^{n} |\langle x, v_i \rangle|^2.
\end{aligned}$$

Damit sind der erste Punkt und die erste Hälfte des zweiten Punktes bewiesen. Möge x zur linearen Hülle der v_i gehören, $x = \sum_{i=1}^{n} \lambda_i v_i$. Es folgt: $\langle x, v_i \rangle = \langle \sum_{j=1}^{n} \lambda_j v_j, v_i \rangle = \sum_{j=1}^{n} \lambda_j \langle v_j, v_i \rangle = \sum_{j=1}^{n} \lambda_j \delta_{ji} = \lambda_i$, was den zweiten Punkt beweist. Der dritte Punkt ist genauso einfach. □

Beispiel 2:

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) \overline{g(t)}\, dt \qquad f_n(t) = e^{int},\ n \in \mathbf{Z}$$

ist ein orthonormales System. $f = \sum_{n=-\infty}^{\infty} \langle f, f_n \rangle f_n$. Auf die analytischen Einzelheiten soll hier nicht eingegangen werden.

Satz 2: Sei V ein endlichdimensionaler Hilbertraum. Dann kann jedes lineare Funktional l auf V auf eindeutige Weise in der Form $l(x) = \langle x, y_l \rangle$ dargestellt werden.

Beweis: Sei $(e_1, \ldots, e_n)$ eine Orthonormalbasis. Dann gilt $l(x) = (\sum_{i=1}^{n} \langle x, e_i \rangle l(e_i)) = \langle x, y_l \rangle$, wobei $y_l = \sum_{i=1}^{n} \overline{l(e_i)} e_i$. Der Beweis der Eindeutigkeit ist einfach. □

Definition 5: Sei V ein komplexer Vektorraum mit einer positiv definiten Hermiteschen Form $\langle \cdot, \cdot \rangle$. Für $x \in V$ sei

$$X^{\perp} = \{y \in V \mid \forall\, x \in X : \langle x, y \rangle = 0\}.$$

Definition 6: Ein **Orthogonalprojektor** π von V auf einen Untervektorraum X ist eine lineare Abbildung $\pi : V \to X$ mit der Eigenschaft $v - \pi(v) \in X^\perp$ für alle $v \in V$.

Bemerkung: Für $x \in X \cap X^\perp$ gilt $\langle x, x\rangle = 0$, also $X \cap X^\perp = \{0\}$. Sei π ein Orthogonalprojektor auf X. Für $x \in X$ gilt dann $x - \pi(x) \in X \cap X^\perp = \{0\}$, also $\pi(x) = x$ für $x \in X$. Also $\pi^2 = \pi$ und π ist tatsächlich ein Projektor auf X.

Satz 3: Sei V ein endlichdimensionaler Hilbertraum und $X \subset V$ ein Untervektorraum. Dann gilt $V = X \oplus X^\perp$. Es existiert ein eindeutig bestimmter Orthogonalprojektor π von V auf X.

Beweis: Die Direktheit der Summe $X \bigoplus X^\perp$ folgt aus der obigen Bemerkung. Um $V = X + X^\perp$ zu zeigen, genügt es, die Existenz eines Orthogonalprojektors auf X zu zeigen. Sei dazu $(e_1, \ldots, e_k)$ eine Orthonormalbasis von X. Setze $\pi(v) = \sum_{i=1}^k \langle v, e_i\rangle e_i \in X$. Dann gilt $\langle \pi(v), e_j\rangle = \langle v, e_j\rangle$ für $j \leq k$. Also gilt für alle $y \in X$, $y = \sum_{j=1}^k \lambda_j e_j$, $\lambda_j \in \mathbb{C}$

$$\langle v - \pi(v), y\rangle = \sum_{j=1}^{k} \overline{\lambda_j}(\langle v, e_j\rangle - \langle \pi(v), e_j\rangle) = 0.$$

π ist also tatsächlich ein Orthogonalprojektor. Eindeutigkeit des Orthogonalprojektors: Sei $\tilde{\pi}$ ein weiterer Orthogonalprojektor auf X. Für $x \in X$ gilt $\tilde{\pi}(x) = x = \pi(x)$. Für $y \in X^\perp$ gilt $\|\tilde{\pi}(y)\|^2 = \langle \tilde{\pi}(y), \tilde{\pi}(y)\rangle = \langle \tilde{\pi}(y), \tilde{\pi}(y) - y\rangle = 0$, denn $\tilde{\pi}(y) \in X$ und $y - \tilde{\pi}(y) \in X^\perp$. Also $\tilde{\pi}(y) = 0 = \pi(y)$. Weil $X \oplus X^\perp = V$ folgt $\tilde{\pi} = \pi$. □

Bemerkung: Diese Sätze gelten in ähnlicher Weise für beliebige Hilberträume. In Satz 2 muß man die Stetigkeit des linearen Funktionales l voraussetzen, Satz 3 gilt für abgeschlossene Unterräume X.

5.8 Symplektische Vektorräume

Sei K ein beliebiger Körper, auch char $K = 2$. Sei V ein endlichdimensionaler K-Vektorraum.

Definition 1: Eine schiefsymmetrische Bilinearform s auf V ist **nichtentartet**, falls zu jedem $x \in V - \{0\}$ ein $y \in V$ mit $s(x, y) \neq 0$ existiert.

Definition 2:

1. Sei $B = (b_1, \ldots, b_n)$ eine Basis von V. Dann sei $\mathrm{Mat}_B(s) = (s(b_i, b_j))_{i,j=1}^n$.
2. Sei $\sigma_s : V \to V^*$ die durch $(\sigma_s(x))(y) = s(x, y)$, $x \in V$ und $y \in V$ definierte lineare Abbildung.

Fakt 1: Durch Mat_B wird ein Isomorphismus des Vektorraumes aller schiefsymmetrischen Bilinearformen auf V auf den Raum aller Matrizen $A \in \mathrm{Mat}(n, n, K)$ mit $A = -A^T$ (d.h. $a_{ij} = -a_{ji}$ für $A = (a_{ij})_{i,j=1}^n$ und $a_{ii} = 0$ für $1 \leq i \leq n$) definiert.

Bemerkung: Für char $K \neq 2$ folgt aus $A = -A^T$ $a_{ii} = -a_{ii}$ und $2a_{ii} = 0$, also $a_{ii} = 0$.

Beweis von Fakt 1: Man prüft leicht nach, daß $\mathrm{Mat}_B(s)$ die behaupteten Eigenschaften hat. Sei umgekehrt $A = -A^T$ und $a_{ii} = 0$, $A = (a_{ij})_{i,j=1}^n$. Sei s_A die durch

$$s_A(\sum_{i=1}^{n} \lambda_i b_i, \sum_{j=1}^{n} \theta_j b_j) = \sum_{i,j=1}^{n} \lambda_i \theta_j a_{ij}$$

eindeutig definierte Bilinearform auf V. Dann gilt

$$s_A(\sum_{i=1}^n \lambda_i b_i, \sum_{j=1}^n \lambda_j b_j) = \sum_{i,j=1}^n a_{ij}\lambda_i\lambda_j = \sum_{1\le i<j\le n} a_{ij}\lambda_i\lambda_j + \underbrace{\sum_{i=1}^n a_{ii}\lambda_i^2}_{=0} + \underbrace{\sum_{1\le j<i\le n} a_{ij}\lambda_i\lambda_j}_{=-\sum_{1\le j<i\le n} a_{ji}\lambda_j\lambda_i} = 0.$$

Man prüft leicht nach, daß $A \mapsto s_A$ zu Mat_B invers ist. □

Fakt 2: Es gilt $\mathrm{Mat}_{B,B^*}(\sigma_s) = -\mathrm{Mat}_B(s)$.

Beweis: Sei $\mathrm{Mat}_B(s) = (a_{ij})_{i,j=1}^n$ und $(b_1, \ldots, b_n)$ eine Basis von V und $(b_1^*, \ldots, b_n^*)$ die dazu duale Basis. Dann gilt

$$\sigma_s(b_j)(\sum_{i=1}^n \lambda_i b_i) = s(b_j, \sum_{i=1}^n \lambda_i b_i) = \sum_{i=1}^n a_{ji}\lambda_i = \sum_{i=1}^n a_{ji} b_i^*(\sum_{k=1}^n \lambda_k b_k) = (\sum_{i=1}^n a_{ji} b_i^*)(\sum_{k=1}^n \lambda_k b_k).$$

Also $\sigma_s(b_j) = \sum_{i=1}^n a_{ji} b_i^*$ und die Behauptung folgt. □

Fakt 3: s ist nicht ausgeartet $\Leftrightarrow$ σ_s ist ein Isomorphismus $\Leftrightarrow$ $\det(\mathrm{Mat}_B(s)) \neq 0$.

Beweis: Wie in 5.2.

Fakt 4: Sei $f : \tilde{V} \to V$ eine lineare Abbildung, und sei $f^*(s)$ die durch $(f^*(s))(x,y) = s(f(x), f(y))$ definierte schiefsymmetrische Bilinearform auf $\tilde{V}$ und sei $\tilde{B}$ eine Basis von $\tilde{V}$. Dann gilt

$$\mathrm{Mat}_{\tilde{B}}(f^*(s)) = \mathrm{Mat}_{\tilde{B},B}(f)^T \mathrm{Mat}_B(s) \mathrm{Mat}_{\tilde{B},B}(f).$$

Beweis: Wie der Beweis von Fakt 5.1.4.

Definition 3: Ein **symplektischer** K-Vektorraum ist ein Paar (V, s), wobei V ein endlichdimensionaler K-Vektorraum und s eine nichtausgeartete schiefsymmetrische Bilinearform auf V ist.

Theorem 5: Sei (V, s) ein symplektischer Vektorraum über einem beliebigen Körper K. Dann gibt es eine **Darboux-Basis** von V, d.h. eine Basis $(v_1, \ldots, v_n, w_1, \ldots, w_n)$ mit $s(v_i, v_j) = s(w_i, w_j) = 0$ für alle $1 \le i, j \le n$ und $s(v_i, w_j) = \delta_{ij}$. Insbesondere ist $\dim V$ gerade.

Bemerkung: Dann gilt $s(w_i, v_j) = -s(v_j, w_i) = \delta_{ij}$. Die Matrixdarstellung von s hat die Form

$$\begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix}$$

Beweis von Theorem 5 durch Induktion nach der Dimension von V beginnend mit dem trivialen Fall $V = \{0\}$. Sei die Behauptung für Vektorräume kleinerer Dimensionen bewiesen. Sei $v_1 \in V - \{0\}$. Nach der Voraussetzung existiert ein $w_1 \in V$ mit $s(v_1, w_1) = 1$. Sei $\tilde{V} = \{x \in V \mid s(x, v_1) = s(x, w_1) = 0\}$. Dann gilt

1. $V = Kv_1 \oplus Kw_1 \oplus \tilde{V}$;
2. $s|_{\tilde{V}\times\tilde{V}}$ ist nicht ausgeartet.

Dann hat nach Induktionsvoraussetzung $\tilde{V}$ eine Darboux-Basis $(v_2, \ldots, v_n, w_2, \ldots, w_m)$ und $(v_1, \ldots, v_n, w_1, \ldots, w_n)$ ist eine Darboux-Basis für V. Beweis von (1): Sei $x \in V$ und sei $y = x - s(v_1, x)w_1 + s(w_1, x)v_1$. Dann gilt

$$s(v_1, y) = s(v_1, x) - s(v_1, x)\underbrace{s(v_1, w_1)}_{=1} + s(w_1, x)\underbrace{s(v_1, v_1)}_{=0} = 0$$

und

$$s(w_1,y) = s(w_1,x) - s(v_1,x)\underbrace{s(w_1,w_1)}_{=0} + s(w_1,x)\underbrace{s(w_1,v_1)}_{=-s(v_1,w_1)=-1} = 0.$$

Also gilt $y \in \tilde{V}$ und daher $x = y + s(v_1,x)w_1 - s(w_1,x)v_1 \in \tilde{V} + Kv_1 + Kw_1$. Es verbleibt, die Direktheit der Summe zu zeigen. Angenommen, $\alpha v_1 + \beta w_1 \in \tilde{V}$. Dann gilt $0 = s(v_1, \alpha v_1 + \beta w_1) = \beta$ und $0 = s(w_1, \alpha v_1 + \beta w_1) = -\alpha$. Also $\alpha = \beta = 0$. Daher ist die Summe $Kv_1 + Kw_1 + \tilde{V}$ direkt. Beweis von (2): Sei $z \in \tilde{V} - \{0\}$. Nach Voraussetzung gibt es ein $x \in V$ mit $s(x,z) \neq 0$. Nach Behauptung 1 kann man x als $x = \alpha v_1 + \beta w_1 + y$ mit $\alpha, \beta \in K$ und $y \in \tilde{V}$ darstellen. Dann gilt

$$s(y,z) = s(x,z) - \alpha\underbrace{s(v_1,z)}_{=0} - \beta\underbrace{s(w_1,z)}_{=0} = s(x,z) \neq 0.$$

Weil $y \in \tilde{V}$ beweist dies, daß $s|_{\tilde{V}\times\tilde{V}}$ nicht ausgeartet ist. □

Definition 4: Für einen symplektischen Vektorraum (V,s) definieren wir die **symplektische Gruppe** $\mathrm{Sp}(V,s)$ als

$$\mathrm{Sp}(V,s) = \{f \in L(V,V) \mid s(f(x),f(y)) = s(x,y)\}.$$

Fakt 5: Elemente der symplektischen Gruppe haben die Determinante 1.
Beweis: Man zeigt:

$$\begin{vmatrix} s(v_1,w_1) & \dots & s(v_1,w_{2n}) \\ \vdots & \ddots & \vdots \\ s(v_{2n},w_1) & \dots & s(v_{2n},w_{2n}) \end{vmatrix} = c_{s,d}\, d(v_1,\dots,v_{2n})\, d(w_1,\dots,w_{2n}),\ 2n = \dim V,\ d \in D(V),$$

ebenso wie die Formel für die Gramsche Determinante. Daraus folgt schon $\det(f)^2 = 1$ für $f \in \mathrm{Sp}(V,s)$, also $\det(f) \in \{\pm 1\}$. Für $\mathrm{char}\,(K) = 2$ ist dies schon der ganze Beweis. Für $\mathrm{char}\,(K) \neq 2$ brauchen wir ein schärferes Argument. Beweisskizze: Sei M die Menge aller Zerlegungen von $\{1,\dots,2n\}$ in n zweielementige Teilmengen. Sei $z_0 \in M$ die Zerlegung $\{1,\dots,2n\}$ in $\{1,n+1\} \cup \{2,n+2\} \cup \dots \cup \{n,2n\}$. Für $z : \{1,\dots,2n\} = x_1 \cup \dots \cup x_n$ und $\pi \in S_{2n}$ sei $\pi(z)$ die Zerlegung $\{1,\dots,2n\} = \pi(x_1) \cup \dots \cup \pi(x_n)$. Diese Regel definiert eine **Wirkung** von S_{2n} auf der Menge M, d.h., $(\pi\rho)(z) = \pi(\rho(z))$, $\mathrm{Id}(z) = z$, also einen Gruppenhomomorphismus $S_{2n} \to S(M)$. Man kann leicht zeigen, daß die Wirkung von S_{2n} auf M **transitiv** ist, d.h. für $x,y \in M$ existiert $\pi \in S_{2n}$ mit $\pi x = y$. Sei $\varepsilon = (\varepsilon_1,\dots,\varepsilon_n) \in (\mathbb{Z}/2\mathbb{Z})^n$ und sei $p_\varepsilon \in S_{2n}$ die durch $p_\varepsilon(i) = i, p_\varepsilon(n+i) = n+i$ falls $\varepsilon_i = 0$ und $p_\varepsilon(i) = n+i, p_\varepsilon(n+i) = i$, falls $\varepsilon_i = 1$ für $1 \leq i \leq n$ definierte Permutation. Für $\rho \in S_n$ sei $q_\rho \in S_{2n}$ die durch $q_\rho(i) = \rho(i)$, $q_\rho(n+i) = n+\rho(i)$, $1 \leq i \leq n$ definierte Permutation. (q_ρ permutiert die untere und die obere Schicht entsprechend ρ. p_ε vertauscht die untere und die obere Schicht an den Stellen i, für die $\varepsilon_i = 1$ gilt.) Es gilt $p_\varepsilon(z_0) = z_0$, $q_\rho(z_0) = z_o$. Sei H_n die Menge aller $\pi \in S_{2n}$, für die $\rho \in S_n$ und $\varepsilon \in (\mathbb{Z}/2\mathbb{Z})^n$ mit $\pi = p_\varepsilon q_\rho$ existieren. Man kann nun zeigen, daß H_n eine Untergruppe in S_{2n} ist und daß $H_n = \{\pi \in S_{2n} \mid \pi z_0 = z_0\}$. Es gilt $\mathrm{sign}\,(p_\varepsilon q_\rho) = (-1)^{|\{i|e_i=1, 1\leq i\leq n\}|}$. Für $\pi \in S_{2n}$ sei

$$d_\pi(x_1,\dots,x_{2n}) = \mathrm{sign}\,(\pi) \prod_{k=1}^{n} s(x_{\pi(k)}, x_{\pi(n+k)}), \quad x_1,\dots,x_{2n} \in V. \tag{$\dagger$}$$

Man prüft leicht

$$d_{\pi p_\varepsilon q_\rho}(x_1,\dots,x_n) = d_\pi(x_1,\dots,x_{2n}), \tag{$*$}$$

denn q_ρ vertauscht nur Faktoren in (†) und p_ε ändert nur die Reihenfolge der Argumente in Faktoren von (†), mithin $d_{\pi\tau} = d_\pi$ für $\tau \in H_n$. Weiterhin gilt

$$d_{\rho\pi}(x_1, \ldots, x_{2n}) = \operatorname{sign}(\rho) d_\pi(x_{\rho(1)}, \ldots, x_{\rho(2n)}). \qquad (\ddagger)$$

Für $z \in M$ wähle $\pi \in S_{2n}$ mit $\pi z_0 = z$. Setze $d_z(x_1, \ldots, x_{2n}) = d_\pi(x_1, \ldots, x_{2n})$. Wenn $\tilde{\pi} \in S_{2n}$ mit $\tilde{\pi} z_0 = z$, so gilt für $\rho = \pi^{-1}\tilde{\pi}$ gilt $\rho z_0 = z_0$, also $\rho \in H_n$, also $d_{\tilde{\pi}} = d_{\pi\rho} \overset{(*)}{=} d_\pi$. Aus (‡) folgt

$$d_z(x_{\pi(1)}, \ldots, x_{\pi(2n)}) = d_{\pi(z)}(x_1, \ldots, x_n) \operatorname{sign}(\pi).$$

Offenbar sind alle d_z $2n$-Linearformen auf V. Sei

$$d = \sum_{z \in M} d_z.$$

Aus (‡) folgt leicht $d(x_1, \ldots, x_{2n}) = \operatorname{sign}(\pi) d(x_{\pi(1)}, \ldots, x_{\pi(2n)})$ für $\pi \in S_{2n}$. Weil $\operatorname{char}(K) \neq 2$ gilt, folgt daraus $d \in D(V)$ (auch für $\operatorname{char}(K) = 2$ gilt $d \in D(V)$ aus anderen Gründen). Sei $(v_1, \ldots, v_n, w_1, \ldots, w_n)$ wie in Theorem 5. Dann gilt

$$d_z(v_1, \ldots, v_n, w_1, \ldots, w_n) = \begin{cases} 1 & z = z_0 \\ 0 & \text{sonst} \end{cases}.$$

Also $d(v_1, \ldots, v_n, w_1, \ldots, w_n) = 1$, damit $d \neq 0$. Aber aus der Definition von d folgt $d(f(x_1), \ldots, f(x_{2n})) = d(x_1, \ldots, x_{2n})$, denn dasselbe gilt für die d_t, falls $f \in \operatorname{Sp}(V, s)$. Also $\det(f) = 1$. □

Bemerkung: In der Hamiltonschen Mechanik betrachtet man

$$V = \{(\underbrace{q_1, \ldots, q_n}_{\text{Ortskoordinaten}}, \underbrace{p_1, \ldots, p_n}_{\text{Impulskoordinaten}})\}$$

mit

$$s((q_1, \ldots, q_n, p_1, \ldots, p_n), (\tilde{q}_1, \ldots, \tilde{q}_n, \tilde{p}_1, \ldots, \tilde{p}_n)) = \sum_{i=1}^{n} (q_i \tilde{p}_i - \tilde{q}_i p_i).$$

Dann sind die Elemente der symplektischen Gruppe gerade die linearen kanonischen Transformationen. Aus Fakt 5 folgt, daß solche Transformationen das Phasenvolumen erhalten.

6 Eigenwerttheorie

Sei K ein beliebiger Körper, V ein K-Vektorraum, A ein Endomorphismus von V.

Definition 1:

1. Das **Spektrum** von A ist die Menge aller $\lambda \in K$, für die $A - \lambda \mathrm{Id}_V : V \to V$ kein Automorphismus von V (d.h. nicht bijektiv) ist.

2. Für $\lambda \in K$ sei
$$V_\lambda(A) = \{v \in V \mid Av = \lambda v\}.$$
Wir nennen $V_\lambda(A)$ den zu λ gehörigen **Eigenraum** von A.

3. Wir nennen $\lambda \in K$ einen **Eigenwert** von A, falls der Eigenraum $V_\lambda(A) \neq \{0\}$. In diesem Fall nennen wir die von 0 verschiedenen Elemente von $V_\lambda(A)$ die **Eigenvektoren** zum Eigenwert λ.

4. Ein Unterraum $W \subseteq V$ ist A-**invariant**, falls $A(W) \subseteq W$.

Fakt 1: Eigenwerte gehören zum Spektrum.

Beweis: Es gilt: $V_\lambda(A) = \ker(A - \lambda\mathrm{Id})$. Also gilt: λ ist Eigenwert $\Leftrightarrow \ker(A - \lambda\mathrm{Id}) \neq \{0\} \Leftrightarrow A - \lambda\mathrm{Id}_V$ ist nicht injektiv $\Rightarrow A - \lambda\mathrm{Id}_V$ ist nicht bijektiv $\Leftrightarrow \lambda$ liegt im Spektrum von A. □

Fakt 2: Sei V endlichdimensional. Dann sind folgende Aussagen äquivalent:

1. λ ist ein Eigenwert von A.
2. λ gehört zum Spektrum von A.
3. $\det(A - \lambda\mathrm{Id}) = 0$.

Beweis: (1) $\Rightarrow$ (2) ist Fakt 1. (2) $\Rightarrow$ (1) folgt wegen der Endlichdimensionalität von V aus der Tatsache, daß ein Endomorphismus eines endlichdimensionalen Vektorraumes genau dann bijektiv ist, wenn er injektiv ist. (2) $\Leftrightarrow$ (3) folgt aus der Determinantentheorie. □

Bemerkung: Die Menge aller Eigenwerte nennt man auch das **Eigenwertspektrum**.

Beispiele:

1. Sei $A : K^n \to K^n$ der Operator der Multiplikation mit der Matrix

$$\begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$$

   Dann ist $A - \lambda\mathrm{Id}$ der Operator der Multiplikation mit der Matrix

$$\begin{pmatrix} \lambda_1 - \lambda & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n - \lambda \end{pmatrix},$$

   deren Determinante $(\lambda_1 - \lambda) \cdot \ldots \cdot (\lambda_n - \lambda)$ ist. Daher ist λ genau dann ein Eigenwert von A, wenn $\lambda = \lambda_i$, $1 \leq i \leq n$, ist. In diesem Fall ist der Eigenraum $(K^n)_\lambda(A) = \{(x_i)_{i=1}^n \in K^n \mid x_i = 0 \text{ falls } \lambda_i \neq \lambda\}$. Seine Dimension ist gleich der Anzahl der Indizes i, $1 \leq i \leq n$, für die $\lambda_i = \lambda$ gilt. Man kann das Problem der Bestimmung der Eigenwerte und Eigenvektoren von A auch auffassen als das Problem der Auffindung einer Basis B von V, für die die Matrixdarstellung von A bezüglich B eine Diagonalmatrix ist (was nicht immer gelingt). Offenbar ist $\mathrm{Mat}_{B,B}(A)$ genau dann eine Diagonalmatrix, falls B aus Eigenvektoren besteht.

2. Als unendlichdimensionales Beispiel betrachten wir $C([0,1]) = \{\text{stetige Funktionen } f : [0,1] \to \mathbb{C}\}$, $(Af)(t) = tf(t)$. Dann gilt: $Af = \lambda f \Leftrightarrow$ für alle $t \in [0,1]$ gilt $(t-\lambda)f(t) = 0 \Leftrightarrow$ für alle $t \in [0,1] - \{\lambda\}$ gilt $f(t) = 0 \overset{\text{Stetigkeit}}{\Leftrightarrow} f = 0$. Also hat A keine Eigenwerte. Andererseits prüft man leicht, daß das Spektrum von A mit $[0,1]$ übereinstimmt: eine Funktion g mit $g(\lambda) \neq 0$ kann nicht zum Bild von $A - \lambda\mathrm{Id}$ gehören, aber für $\lambda \notin [0,1]$ ist $f \mapsto g(t) = \frac{f(t)}{t-\lambda}$ ein Inverses zu $A - \lambda\mathrm{Id}$.

Fakt 3: Für jedes $\lambda \in K$ ist $V_\lambda(A)$ A-invariant.

Beweis: Aus $x \in V_\lambda(A)$ folgt $Ax = \lambda x \in V_\lambda(A)$. □

6.1 Eigenwerttheorie für normale und selbstadjungierte Endomorphismen eines endlichdimensionalen Hilbertraums

Sei H ein endlichdimensionaler Hilbertraum, und A ein Endomorphismus von H. Für alle $y \in H$ ist die Abbildung $H \ni x \mapsto \langle Ax, y\rangle$ ein lineares Funktional auf H. Nach Satz 5.7.2 gibt es also einen Vektor $A^*y \in H$ mit der Eigenschaft $\langle Ax, y\rangle = \langle x, A^*y\rangle$ für alle $x \in H$. A^*y ist durch diese Eigenschaft eindeutig bestimmt, und die Abbildung $y \mapsto A^*y$ ist linear.

Definition 1: Wir nennen A^* den zu A **adjungierten** Endomorphismus. Wir nennen A **selbstadjungiert**, falls $A = A^*$ ist. Wir nennen A **normal**, falls A und A^* kommutieren, d.h., falls $AA^* = A^*A$ gilt.

Fakt 1: Es gelten folgende Regeln:

1. $(A + B)^* = A^* + B^*$;
2. $(\lambda B)^* = \overline{\lambda} B^*$ für $\lambda \in \mathbb{C}$;
3. $(AB)^* = B^*A^*$;
4. $A^{**} = A$.

Beweis: (1) und (2) sind einfach, z.B. $\langle \lambda Bx, y\rangle = \lambda\langle Bx, y\rangle = \lambda\langle x, B^*y\rangle = \langle x, \overline{\lambda} B^*y\rangle$ für alle $x, y \in H$, also $(\lambda B)^* = \overline{\lambda} B^*$. Weiterhin gilt $\langle (AB)(x), y\rangle = \langle A(B(x)), y\rangle = \langle Bx, A^*y\rangle = \langle x, B^*A^*y\rangle$ für alle $x, y \in H$, also folgt (3). (4) wird weiter unten bewiesen. □

Folgerungen:

1. Die selbstadjungierten Operatoren bilden einen reellen Unterraum im Raum aller Endomorphismen von H. Die normalen Endomorphismen bilden keinen Unterraum im Raum aller Endomorphismen von H, aber die Multiplikation eines normalen Endomorphismus mit einer komplexen Zahl ist normal.
2. Wenn A und B selbstadjungiert sind, so gilt $(AB)^* = B^*A^* = BA$, also ist AB genau dann selbstadjungiert, wenn A und B kommutieren, d.h. wenn $AB = BA$ gilt. Die Summe zweier kommutierender normaler Endomorphismen ist normal.
3. Es gilt: $(A^*A)^* = A^*A^{**} = A^*A$, denn für alle A gilt $A^{**} = A$: $\langle A^*x, y\rangle = \overline{\langle y, A^*x\rangle} = \overline{\langle Ay, x\rangle} = \langle x, Ay\rangle$. Also ist A^*A selbstadjungiert, ebenso sind es AA^*, $A + A^*$ und $i(A - A^*)$.

Fakt 2: Sei B eine Orthonormalbasis von H. Dann gilt:

$$\mathrm{Mat}_{B,B}(A^*) = \mathrm{Mat}_{B,B}(A)^*.$$

Insbesondere ist A genau dann selbstadjungiert bzw. normal, wenn es $\mathrm{Mat}_{B,B}(A)$ ist.

Bemerkung: In Fakt 5.7.1 hatten wir $A^* = \overline{A^T}$ definiert, wobei $\overline{(a_{ij})_{i,j=1}^n} = (\overline{a_{ij}})_{i,j=1}^n$. Für $A = (a_{ij})_{i,j=1}^n$ gilt $A^* = (\overline{a_{ji}})_{i,j=1}^n$. Also ist A genau dann selbstadjungiert, falls $a_{ij} = \overline{a_{ji}}$ für alle i, j gilt, was für $i = j$ impliziert, daß a_{ij} reell ist.
Beweis: Einfaches Nachrechnen... □

Fakt 3: Sei A selbstadjungiert. Dann sind alle Eigenwerte von A reell, und die Eigenräume zu verschiedenen Eigenwerten sind zueinander orthogonal.

Beweis: Sei $\lambda \in \mathbb{C}$ ein Eigenwert. Dann gibt es ein $x \in H - \{0\}$ mit $Ax = \lambda x$. Wir können $\|x\| = 1$ annehmen, sonst $x \mapsto x/\|x\|$. Dann gilt: $\lambda = \langle \lambda x, x\rangle = \langle Ax, x\rangle = \langle x, Ax\rangle = \langle x, \lambda x\rangle = \overline{\lambda}\langle x, x\rangle = \overline{\lambda}$. Also ist λ reell. Seien $\lambda \neq \vartheta$ reell, $x \in H_\lambda(A)$, $y \in H_\vartheta(A)$. Dann gilt $\lambda\langle x, y\rangle =$

$\langle Ax, y\rangle = \langle x, Ay\rangle = \langle x, \vartheta y\rangle = \vartheta\langle x, y\rangle$, da $\lambda \neq \vartheta$ folgt $\langle x, y\rangle = 0$. □

Es liegt daher nahe, zu versuchen, eine Orthonormalbasis von H zu konstruieren, die aus Eigenvektoren von A besteht. Wenn jedes selbstadjungierte A Eigenwerte hat, kann man das wie folgt tun: sei λ ein Eigenwert, sei $(b_1, \ldots, b_k)$ eine Orthonormalbasis des entsprechenden Eigenraumes $H_\lambda(A)$ und sei H' das orthogonale Komplement dieses Eigenraumes. Man zeigt dann leicht, daß $A|_{H'}$ ein selbstadjungierter Endomorphismus von H' ist, und die Induktionsannahme impliziert die Existenz einer Orthonormalbasis $(b_{k+1}, \ldots, b_n)$ von H', die aus Eigenvektoren besteht. Dann leistet $(b_1, \ldots, b_n)$ das Gewünschte.
Motivation für die nachfolgende Konstruktion von Eigenwerten: Angenommen, es existiert eine Orthonormalbasis $(b_1, \ldots, b_n)$ von H, so daß (b_i) Eigenvektor zum Eigenwert λ ist, wobei $\lambda_1 \geq \ldots \geq \lambda_n$. Für $x = \sum_{i=1}^n x_i b_i$ mit $\|x\|^2 = \sum_{i=1}^n |x_i|^2 = 1$ gilt

$$\langle Ax, x\rangle = \langle \sum_{i=1}^n \lambda_i x_i b_i, \sum_{j=1}^n x_j b_j\rangle = \sum_{i=1}^n \lambda_i |x_i|^2 = \lambda_1 \sum_{i=1}^n |x_i|^2 - \sum_{i=1}^n (\lambda_1 - \lambda_i)|x_i|^2 \leq \lambda_1,$$

und die Schranke λ_1 wird tatsächlich angenommen, nämlich von allen x, für die $x_i = 0$ für alle i mit $\lambda_i < \lambda_1$ gilt. Solche Vektoren sind Eigenvektoren zu λ_1. Es liegt also nahe, zu zeigen, daß $\max\{\langle Ax, x\rangle \mid \|x\| = 1\}$ ein Eigenwert ist (natürlich ohne die Existenz einer Orthonormalbasis aus Eigenvektoren vorauszusetzen).

Lemma 1: Ein Endomorphismus eines endlichdimensionalen Hilbertraumes H ist genau dann selbstadjungiert, wenn für alle $x \in H$ $\langle Ax, x\rangle$ reell ist.

Beweis: Sei A selbstadjungiert. Dann gilt $\langle Ax, x\rangle = \langle x, A^*x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle}$. Also ist $\langle Ax, x\rangle$ reell. Die (etwas schwierigere) Rückrichtung ist Übungsaufgabe.

Da die Funktion $\langle Ax, x\rangle$ auf der kompakten Menge $\{x \in H \mid \|x\| = 1\}$ stetig ist und reelle Werte annimmt, existiert das Maximum auf dieser Menge.

Satz 1: Sei A selbstadjungiert, und seien

$$\lambda_+ = \max\{\langle Ax, x\rangle \mid x \in H \wedge \|x\| = 1\} \text{ und } \lambda_- = \min\{\langle Ax, x\rangle \mid x \in H \wedge \|x\| = 1\}.$$

Dann sind λ_+ und λ_- Eigenwerte, und jedes x mit $\|x\| = 1$, an dem diese Extrema angenommen werden, ist ein Eigenvektor.

Beweis: Es genügt, die Aussage über λ_+ zu beweisen, die Aussage über λ_- zeigt man genauso. Sei $\lambda = \lambda_+$. Wir wissen, daß es Vektoren $x \in K$ mit $\|x\| = 1$ und $\langle Ax, x\rangle = \lambda$ gibt. Es genügt, zu zeigen, daß jedes x mit dieser Eigenschaft ein Eigenvektor zu λ ist. Wenn x ein Eigenvektor ist, so ist λ der dazugehörige Eigenwert. Angenommen, x wäre kein Eigenvektor. Sei H' das orthogonale Komplement von $\mathbb{C}x$ in H und sei y' die orthogonale Projektion von Ax auf H'. Dann gilt $Ax - y' \in \mathbb{C}x$, weil x kein Eigenvektor ist, gilt also $y' \neq 0$. Es gilt $\langle Ax, y'\rangle = \langle y', y'\rangle > 0$. Indem wir $y = y'/\|y'\|$ setzen, erhalten wir ein zu x orthogonales normiertes y, so daß $\langle Ax, y\rangle$ reell und von Null verschieden ist. Sei $x(t) = \frac{x+ty}{\sqrt{1+t^2}}$. Dann gilt $\|x(t)\| = 1$ für alle $t \in \mathbb{R}$. Also nimmt die Funktion $f(t) = \langle Ax(t), x(t)\rangle$ an der Stelle $t = 0$ ihr globales Maximum λ an, insbesondere ist $t = 0$ auch ein lokales Maximum für diese Funktion. Es gilt $f(t) = \frac{1}{1+t^2}(\langle Ax, x\rangle + t^2\langle Ay, y\rangle + t\langle Ax, y\rangle + t\langle x, Ay\rangle)$, also ist f in Abhängigkeit von t stetig differenzierbar. Die Ableitung an der Stelle 0 ist gerade $f'(0) = \langle Ax, y\rangle + \langle x, Ay\rangle = 2\langle Ax, y\rangle \neq 0$. Also hat f an der Stelle $t = 0$ kein lokales Extremum. Also kann $f(0)$ nicht das Maximum aller $f(t)$ sein. Widerspruch. □

Bemerkungen:

1. Wir haben damit gezeigt, daß jeder selbstadjungierte Endomorphismus eines von 0 verschiedenen Hilbertraumes H einen Eigenwert hat (falls $H \neq \{0\}$).

2. Offenbar gilt $\langle Ax, x\rangle = \lambda_\pm$, falls x ein normierter Eigenvektor zu dem Eigenwert $\lambda_\pm$ ist.

Lemma 2: Sei A selbstadjungiert. Dann ist das orthogonale Komplement A eines A-invarianten Unterraumes A-invariant.

Beweis: Sei $X \subseteq H$ A-invariant. Es gilt $X^\perp = \{y \in H \mid \forall\, x \in X : \langle y, x\rangle = 0\}$. Sei nun $y \in X^\perp$. Dann gilt für alle $x \in X$

$$\langle Ay, x\rangle = \langle y, Ax\rangle = 0,$$

denn $Ax \in X$. Also ist $Ay \in X^\perp$. □

Theorem 1: Für jeden selbstadjungierten Endomorphismus A eines endlichdimensionalen Hilbertraumes H gilt die orthogonale direkte Summenzerlegung in die Eigenräume von A

$$A = \bigoplus_{\lambda \in \mathbf{R}} H_\lambda(A).$$

Beweis: Induktion nach $\dim(H)$ beginnend mit dem trivialen Fall $\dim(H) = 0$. Sei die Behauptung für Hilberträume kleinerer Dimension als H bewiesen. Sei $H \neq \{0\}$. Laut Bemerkung 1 folgt aus Satz 1, daß A einen Eigenwert λ hat. Sei $H' = H_{\lambda_+}(A)^\perp$. Nach Fakt 6.3 aus der Einleitung ist $H_\lambda(A)$ A-invariant, also ist es nach Lemma 2 auch H'. Also ist $A|_{H'}$ ein selbstadjungierter Endomorphismus von H', auf den die Induktionsannahme angewendet werden kann. Es gilt:

$$H'_\lambda(A|_{H'}) = H_\lambda(A) \cap H' = \begin{cases} \{0\} & \text{für } \lambda = \lambda_+ \\ H_\lambda(A) & \text{für } \lambda \neq \lambda_+. \end{cases}$$

Die Induktionsannahme zeigt also

$$H' = \bigoplus_{\lambda \in \mathbf{R}, \lambda \neq \lambda_+} H_\lambda(A).$$

Nach Satz 5.7.3 gilt

$$H = H_{\lambda_+}(A) \oplus H_{\lambda_+}(A)^\perp = H_{\lambda_+}(A) \oplus H' = H_{\lambda_+}(A) \oplus \bigoplus_{\lambda \in \mathbf{R}, \lambda \neq \lambda_+} H_\lambda(A) = \bigoplus_{\lambda \in \mathbf{R}} H_\lambda(A).$$

□

Bemerkung: Insbesondere hat H eine Orthonormalbasis, welche aus Eigenvektoren zu A besteht (Wähle eine Orthonormalbasis für jedes $H_\lambda(A)$).
Es gilt $\dim(H) = \sum_{\lambda \in \mathbf{R}} \dim(H_\lambda(A))$. Man kann $\dim(H_\lambda(A))$ als die **Vielfachheit** des Eigenwertes λ betrachten. Wir haben dann bewiesen, daß es, gezählt nach Vielfachheit, genau $\dim(H)$ Eigenwerte gibt.

Folgerung 1: Sei η eine Hermitesche Form auf H. Dann gibt es eine Orthonormalbasis von H, in der die Matrixdarstellung von η eine Diagonalmatrix ist (Hauptachsenbasis, Hauptachsentransformation).

Beweis: Nach Satz 5.7.2 gibt es für alle $y \in H$ genau ein $Ay \in H$ mit $\langle x, Ay\rangle = \eta(x, y)$. Weil η eine Hermitesche Form ist, ist A linear und selbstadjungiert. Wende die Bemerkung nach Theorem 1 auf A an. □

Verallgemeinerung: Sei V ein beliebiger endlichdimensionaler Vektorraum über einem beliebigen Körper K und seien $A_1, \ldots, A_n$ Endomorphismen von V. Gibt es eine Basis B von V, deren Elemente für alle A_i Eigenvektoren sind, d.h. für die $\mathrm{Mat}_B(A_i)$ eine Diagonalmatrix ist. Weil die Multiplikation von Diagonalmatrizen kommutativ ist, impliziert die Existenz einer solchen Basis $\mathrm{Mat}_B(A_iA_j) = \mathrm{Mat}_B(A_i)\mathrm{Mat}_B(A_j) = \mathrm{Mat}_B(A_j)\mathrm{Mat}_B(A_i) = \mathrm{Mat}_B(A_jA_i)$, also $A_iA_j = A_jA_i$. Mit anderen Worten, A_i und A_j kommutieren, die $A_1, \ldots, A_n$ kommutieren also paarweise. Wenn die A_i selbstadjungierte Endomorphismen eines endlichdimensionalen Hilbertraumes sind, ist diese Bedingung auch hinreichend.

Definition 2: Seien $A_1, \ldots, A_n$ paarweise kommutierende Endomorphismen eines beliebigen Vektorraumes V. Dann sei

$$V_{\lambda_1,\ldots,\lambda_n}(A_1,\ldots,A_n) = \bigcap_{i=1}^{n} V_{\lambda_i}(A_i) = \{v \in V \mid A_i(v) = \lambda_i v, 1 \le i \le n\} \qquad \lambda_1,\ldots,\lambda_n \in K$$

der gemeinsame Eigenraum von $(A_1, \ldots, A_n)$ zu $(\lambda_1, \ldots, \lambda_n)$.

Bemerkung: $V_{\lambda_1,\ldots,\lambda_n}(A_1,\ldots,A_n)$ kann man auch für nicht paarweise kommutierende A_i in dieser Form definieren (allerdings wenig sinnvoll).

Lemma 3: Wenn B mit $A_1, \ldots, A_n$ kommutiert, so ist $V_{\lambda_1,\ldots,\lambda_n}(A_1,\ldots,A_n)$ B-invariant.

Beweis: Sei $v \in V_{\lambda_i}(A_i)$. Dann gilt $A_iBv = BA_iv = B(\lambda_i v) = \lambda_i Bv$, also $Bv \in V_{\lambda_i}(A_i)$. Also ist $V_{\lambda_i}(A_i)$ B-invariant. Der Durchschnitt B-invarianter Unterräume ist B-invariant, also ist $V_{\lambda_1,\ldots,\lambda_n}(A_i,\ldots,A_n) = \bigcap_{i=1}^{n} V_{\lambda_i}(A_i)$ B-invariant. □

Die Definition und das Lemma beziehen sich auf *beliebige* Vektorräume über beliebigen Körpern. Sei nun H ein endlichdimensionaler Hilbertraum, $K = \mathbb{C}$.

Theorem 2: Seien $A_1, \ldots, A_n$ paarweise kommutierende selbstadjungierte Endomorphismen von H. Dann gilt die orthogonale direkte Summenzerlegung

$$H = \bigoplus_{(\lambda_1,\ldots,\lambda_n)\in\mathbf{R}^n} H_{\lambda_1,\ldots,\lambda_n}(A_1,\ldots,A_n).$$

Beweis durch Induktion nach n. Der Fall $n = 1$ ist gerade Theorem 1. Sei die Aussage für $n-1$ Operatoren bewiesen. Dann gilt

$$H = \bigoplus_{(\lambda_1,\ldots,\lambda_{n-1})\in\mathbf{R}^{n-1}} H_{\lambda_1,\ldots,\lambda_{n-1}}(A_1,\ldots,A_{n-1}).$$

Nach Lemma 3 ist $H' = H_{\lambda_1,\ldots,\lambda_{n-1}}(A_1,\ldots,A_{n-1})$ A_n-invariant. Die Einschränkung von A_n auf den A_n-invarianten Unterraum H' ist ein selbstadjungierter Endomorphismus von H', Theorem 1 ergibt folglich

$$\begin{aligned} H' &= \bigoplus_{\lambda\in\mathbf{R}} H'_\lambda(A_n|_{H'}) = \bigoplus_{\lambda\in\mathbf{R}} (H' \cap H_\lambda(A_n)) = \bigoplus_{\lambda\in\mathbf{R}} (H_{\lambda_1,\ldots,\lambda_{n-1}}(A_1,\ldots,A_{n-1}) \cap H_\lambda(A_n)) \\ &= \bigoplus_{\lambda\in\mathbf{R}} H_{\lambda_1,\ldots,\lambda_{n-1},\lambda}(A_1,\ldots,A_n). \end{aligned}$$

Also

$$H = \bigoplus_{(\lambda_1,\ldots,\lambda_{n-1})\in\mathbf{R}^{n-1}} H_{\lambda_1,\ldots,\lambda_{n-1}}(A_1,\ldots,A_{n-1})$$

$$= \bigoplus_{(\lambda_1,\ldots,\lambda_{n-1})\in\mathbf{R}^{n-1}} \left(\bigoplus_{\lambda_n\in\mathbf{R}} H_{\lambda_1,\ldots,\lambda_n}(A_1,\ldots,A_n)\right) = \bigoplus_{(\lambda_1,\ldots,\lambda_n)\in\mathbf{R}^n} H_{\lambda_1,\ldots,\lambda_n}(A_1,\ldots,A_n).$$

□

Folgerung 2: Sie A ein normaler Endomorphismus von H. Dann gilt die orthogonale direkte Summenzerlegung

$$H = \bigoplus_{\lambda\in\mathbb{C}} H_\lambda(A).$$

Beweis: A ist genau dann normal, wenn A und A^* miteinander kommutieren. Zum Beweis benötigen wir

Fakt: Wenn A und B kommutieren, so kommutieren auch $\alpha A+\beta B$ und $\gamma A+\delta B^*$ für $\alpha,\beta,\gamma,\delta\in K=\mathbb{C}$. □

Für einen normalen Operator A kommutieren also die beiden selbstadjungierten Endomorphismen von H $A_1=\frac{1}{2}(A+A^*)$ und $A_2=\frac{1}{2i}(A-A^*)$. Nach Theorem 2 gilt

$$H = \bigoplus_{(\lambda_1,\lambda_2)\in\mathbf{R}^2} H_{\lambda_1,\lambda_2}(A_1,A_2). \qquad (*)$$

Behauptung: $H_{\lambda_1,\lambda_2}(A_1,A_2)=H_{\lambda_1+i\lambda_2}(A)$. Wenn diese Behauptung bewiesen ist, so ergibt $(*)$

$$H = \bigoplus_{\lambda=\lambda_1+i\lambda_2\in\mathbb{C}} H_\lambda(A).$$

Beweis der Behauptung: Sei $v\in H_{\lambda_1,\lambda_2}(A_1,A_2)$. Dann gilt

$$Av = \left(\frac{A+A^*}{2}\right)v + i\left(\frac{A-A^*}{2i}\right)v = \lambda_1 v + i\lambda_2 v = (\lambda_1+i\lambda_2)v.$$

Damit $H_{\lambda_1,\lambda_2}(A_1,A_2)\subseteq H_{\lambda_1+i\lambda_2}(A)$. Wir müssen zeigen, daß Gleichheit eintritt. Sei $H'=H_{\lambda_1+i\lambda_2}(A)$. Weil A_1 und A_2 mit A kommutieren, ist H' invariant unter A_1 und A_2 (Lemma 3), also

$$H' = \bigoplus_{\alpha,\beta\in\mathbf{R}^2} H_{\alpha,\beta}(A_1|_{H'},A_2|_{H'}).$$

Aber für $(\alpha,\beta)\neq(\lambda_1,\lambda_2)$ gilt

$$H'_{\alpha,\beta}(A_1|_{H'},A_2|_{H'}) = H'\cap H_{\alpha,\beta}(A_1,A_2)\subseteq H'\cap H_{\alpha+i\beta}(A) = H_{\lambda_1+i\lambda_2}(A)\cap H_{\alpha+i\beta}(A)=\{0\},$$

denn es gilt folgender

Fakt: Wenn A ein beliebiger Endomorphismus eines beliebigen K-Vektorraumes V ist, so gilt für $\lambda\neq\vartheta$ $V_\lambda(A)\cap V_\vartheta(A)=\{0\}$.
In der Tat, wenn x zu dem Durchschnitt gehört, so gilt $\lambda x=Ax=\vartheta x$, also $(\lambda-\theta)x=0$, also $x=0$. □

Damit ist Folgerung 2 bewiesen. □Wir wollen Folgerung 2 auf unitäre Endomorphismen von H anwenden. U ist unitär genau dann, wenn $\langle Ux,Uy\rangle=\langle x,y\rangle$ für alle $x,y\in H$ gilt.

Fakt 4: Für einen Endomorphismus U eines endlichdimensionalen Hilbertraumes H sind folgende Aussagen äquivalent:

1. U ist unitär;

2. $\|Ux\| = \|x\|$ für alle $x \in H$;

3. $U^*U = \mathrm{Id}$;

4. $UU^* = \mathrm{Id}$.

Insbesondere ist jeder unitäre Endomorphismus normal.
Beweis: (1) $\Rightarrow$ (2) ist klar. (2) $\Rightarrow$ (1): Sei $\|Ux\| = \|x\|$ für alle $x \in H$. Dann gilt

$$2\mathrm{Re}\,\langle x, y\rangle = \|x+y\|^2 - \|x\|^2 - \|y\|^2 = \|Ux+Uy\|^2 - \|Ux\|^2 - \|Uy\|^2 = 2\mathrm{Re}\,\langle Ux, Uy\rangle.$$

Es folgt

$$\langle x, y\rangle = \mathrm{Re}\,\langle x, y\rangle + i\mathrm{Re}\,\langle x, iy\rangle = \mathrm{Re}\,\langle Ux, Uy\rangle + i\mathrm{Re}\,\langle Ux, iUy\rangle = \langle Ux, Uy\rangle.$$

(1) $\Leftrightarrow$ (3):

$$\begin{aligned}(1) \quad &\Leftrightarrow \quad \forall\, x, y \in H : \langle x, y\rangle = \langle Ux, Uy\rangle \Leftrightarrow \forall\, x, y \in H : \langle x, y\rangle = \langle x, U^*Uy\rangle \\ &\Leftrightarrow \quad \forall\, y \in H : U^*Uy = y \Leftrightarrow (3).\end{aligned}$$

(3) $\Leftrightarrow$ (4): Weil H endlichdimensional ist, ist U^* genau dann ein Rechtsinverses zu U, wenn es ein Linksinverses zu U ist.
Warnung: Für unendlichdimensionale Hilberträume sind (3) *und* (4) *nicht äquivalent. Man verlangt dann von unitären Operatoren explizit, daß sie invertierbar sind.*
Wegen der Äquivalenz von (1), (3) und (4) ist jeder unitäre Endomorphismus normal. □

Wir können also Folgerung 2 auf den unitären Endomorphismus U anwenden. Sei $H_\lambda(U) \neq \{0\}$, $v \in H_\lambda(U) - \{0\}$. Dann gilt $\|v\| = \|Uv\| = \|\lambda v\| = |\lambda|\|v\|$, also $|\lambda| = 1$. Wir erhalten
Folgerung 3: Für unitäres U gilt die direkte orthogonale Summenzerlegung

$$H = \bigoplus_{\lambda \in \mathbb{C}, |\lambda|=1} H_\lambda(U). \qquad \square$$

Folgerung 4: Seien $A_1, \ldots, A_n$ paarweise kommutierende normale Endomorphismen von H. Dann gilt die orthogonale direkte Summenzerlegung

$$H = \bigoplus_{(\lambda_1,\ldots,\lambda_n) \in \mathbb{C}^n} H_{\lambda_1,\ldots,\lambda_n}(A_1, \ldots, A_n).$$

Wenn dabei A_i unitär ist, so verschwinden die Summanden zu $(\lambda_1, \ldots, \lambda_n)$ mit $|\lambda_i| \neq 1$.

Beweis: Dies folgt aus Lemma 3 und Folgerung 2 in derselben Weise wie Theorem 2 aus Theorem 1 und Lemma 3 folgt.

Fakt 5: Sei p ein Endomorphismus eines endlichdimensionalen Hilbertraumes H. Dann ist p genau dann ein Orthogonalprojektor auf einem Unterraum von H, wenn $p^2 = p$ und $p = p^*$ gilt, d.h., wenn p ein **selbstadjungiertes Idempotent** ist.

Beweis: Sei p der Orthogonalprojektor auf den Unterraum $X \subseteq H$. Es gilt $H = X \oplus X^\perp$. Für $a, b \in X$, $\tilde{a}, \tilde{b} \in X^\perp$ gilt

$$\langle p(a+\tilde{a}), b+\tilde{b}\rangle \overset{p(\tilde{a})=0}{=} \langle a, b+\tilde{b}\rangle \overset{\tilde{b}\in X^\perp}{=} \langle a, b\rangle = \langle a+\tilde{a}, p(b+\tilde{b})\rangle$$

und

$$p^2(a+\tilde{a}) = p(p(a+\tilde{a})) = p(a) = a.$$

Weil alle $x, y \in H$ als $x = a + \tilde{a}$ und $y = b + \tilde{b}$ mit $a, b \in X$ und $\tilde{a}, \tilde{b} \in X^\perp$ dargestellt werden können, zeigt dies, daß $p = p^*$ und $p^2 = p$ gelten.
Sei umgekehrt $p^2 = p$ und $p = p^*$. Sei $X = \text{Bild}(p)$ und $Y = \ker(p)$. Dann $X^\perp = Y$, denn $y \in X^\perp \Leftrightarrow \forall\, x \in X : \langle y, x\rangle = 0 \Leftrightarrow \forall\, z \in H : \langle y, p(z)\rangle = 0 \Leftrightarrow \forall\, z \in H : \langle p(y), z\rangle = 0 \Leftrightarrow p(y) = 0$. Es gilt also $H = X + Y = X + X^\perp$. Sei $x \in X$, dann gilt $x = p(z)$ für ein $z \in H$. Also gilt $p(x) = p(p(z)) = p(z) = x$. Für $y \in X^\perp = Y = \ker(p)$ gilt $p(y) = 0$. Also ist p Orthogonalprojektor auf X. □

Bemerkung 2:

1. Sei V ein endlichdimensionaler $\mathbb{R}$- oder $\mathbb{C}$-Vektorraum und sei $A \in \text{End}(V)$. Wir betrachten die V-wertige Differentialgleichung

$$\frac{dv}{dt} = Av, \quad v : \mathbb{R} \to V. \tag{$*$}$$

   Angenommen, V hat eine Basis $(e_1, \ldots, e_n)$, die aus Eigenvektoren für A besteht, $Ae_i = \lambda_i e_i$. Dann ist die allgemeine Lösung von $(*)$ durch $v(t) = \sum_{i=1}^n C_i e^{\lambda_i t} e_i$ gegeben, wobei $C_1, \ldots, C_n$ frei zu wählen sind.

2. Eigentlich wird die Spektraltheorie erst im unendlichdimensionalen Fall *richtig* interessant...

6.2 Eigenwerttheorie für normale und selbstadjungierte Endomorphismen eines Euklidischen Vektorraumes

Sei V ein Euklidischer Vektorraum, $K = \mathbb{R}$, $\langle \cdot, \cdot \rangle$ das Skalarprodukt auf V.

Definition 1: Für einen Endomorphismus $A : V \to V$ sei A^* der durch die Gleichung $\langle Ax, y\rangle = \langle x, A^* y\rangle$ eindeutig bestimmte Endomorphismus von V. A ist **selbstadjungiert**, falls $A = A^*$ gilt. A ist **normal**, wenn $AA^* = A^*A$ gilt, d.h., falls A und A^* kommutieren.

Fakt 1: p ist genau dann Orthogonalprojektor auf einem Unterraum $U \subseteq V$, wenn $p = p^*$ und $p^2 = p$ gilt.

Beweis: Analog zu Fakt 6.1.5.

Fakt 2: $(A + B)^* = A^* + B^*$, $(\lambda A)^* = \lambda A^*$, $(AB)^* = B^* A^*$, $A^{**} = A$.

Beweis: Analog zu Fakt 6.1.1.

Folgerung: Die selbstadjungierten Endomorphismen von V bilden also einen Unterraum im Raum aller Endomorphismen von V. Das Produkt zweier selbstadjungierter Endomorphismen ist genau dann selbstadjungiert, wenn die beiden Faktoren kommutieren. Für jeden Endomorphismus A sind $A + A^*$, AA^* und A^*A selbstadjungiert.

Beweis: Analog zum Beweis derselben Folgerungen im vorherigen Kapitel.

Fakt 3: Sei B eine Orthonormalbasis von V. Dann gilt $\text{Mat}_B(A^*) = \text{Mat}_B(A)^T$, insbesondere ist A genau dann selbstadjungiert, wenn $\text{Mat}_B(A)$ symmetrisch entlang der Hauptdiagonale ist.

Beweis: Analog zum Beweis im vorherigen Kapitel.

Sei von nun an A selbstadjungiert.

Fakt 4: Dann sind die Eigenräume von A zu verschiedenen Eigenwerten zueinander orthogonal.

Beweis: Wie Fakt 6.1.3: Sei $x \in V_\lambda(A) = \{x \in V \mid Ax = \lambda x\}$ und $y \in V_\vartheta(A)$ für $\lambda \neq \vartheta$, dann $\lambda\langle x, y\rangle = \langle \lambda x, y\rangle = \langle Ax, y\rangle = \langle x, Ay\rangle = \langle x, \vartheta y\rangle = \vartheta\langle x, y\rangle$. Also $(\lambda - \vartheta)\langle x, y\rangle = 0$, also $\langle x, y\rangle = 0$.

□

Wir möchten zeigen, daß eine Orthonormalbasis von V existiert, die aus Eigenvektoren von A besteht. Wir gehen wie im komplexen Fall vor.

Satz 1: Sei $S = \{x \in V \mid \|x\| = 1\}$. Die Funktion $x \mapsto \langle Ax, x\rangle$ ist auf der kompakten Menge S stetig und nimmt daher auf S ein Maximum und ein Minimum an. Diese Extremwerte sind Eigenwerte von A, die Argumente, an denen sie angenommen werden, sind Eigenvektoren von A.

Beweis: Wie Satz 6.1.1.

Bemerkung 1: Offenbar sind auch alle normierten Eigenvektoren von A zu den Eigenwerten $\lambda_+ = \max\{\langle Ax, x\rangle \mid x \in V \wedge \|x\| = 1\}$, $\lambda_- = \min\{\langle Ax, x\rangle \mid x \in V \wedge \|x\| = 1\}$ Extrema der in Satz 1 betrachteten Funktion.

Lemma 1: Sei A selbstadjungiert. Dann ist das orthogonale Komplement eines A-invarianten Unterraumes A-invariant.

Beweis: Wie Lemma 6.1.2.

Theorem 3: Sei A ein selbstadjungierter Endomorphismus eines Euklidischen Vektorraumes V. Dann gilt die orthogonale Summenzerlegung

$$V = \bigoplus_{\lambda \in \mathbf{R}} V_\lambda(A).$$

Insbesondere hat V eine Orthonormalbasis, die aus Eigenvektoren von A besteht.

Beweis: Wie Theorem 1 im letzten Kapitel.

Folgerung 1: (Hauptachsentransformation für symmetrische Bilinearformen) Sei β eine symmetrische Bilinearform auf einem Euklidischen Vektorraum V. Dann hat V eine Orthonormalbasis B, so daß $\mathrm{Mat}_B(\beta)$ eine Diagonalmatrix ist.

Beweis: Es gibt einen selbstadjungierten Endomorphismus A von V mit $\beta(x, y) = \langle Ax, y\rangle$, das folgt aus Satz 5.6.2. Anwendung von Theorem 3 auf A ergibt den Beweis. □

Bemerkung: Sei $\beta(x, y)$ eine symmetrische Bilinearform auf $\mathbf{R}^2$, die in der Standardbasis Diagonalgestalt hat. Dann $\beta((x_1, x_2), (y_1, y_2)) = ax_1y_1 + bx_2y_2$. Wir nehmen weiterhin an, daß β positiv definit ist, also $a, b > 0$, und $a < b$. Dann ist $E = \{x \in \mathbf{R}^2 \mid \beta(x, x) = 1\}$ eine Ellipse mit der x-Achse als Hauptachse und der y-Achse als Nebenachse. Für positiv definites β auf V ist $\{x \in V \mid \beta(x, x) = 1\}$ ein Ellipsoid, eine Basis mit der in Folgerung 1 beschriebenen Eigenschaft liefert die Haupt- und Nebenachsen für dieses Ellipsoid. Falls $\dim(V) = 2$ und β nicht ausgeartet, aber indefinit ist, so ist $\{x \in V \mid \beta(x, x) = 1\}$ eine Hyperbel.

Theorem 4: Seien $A_1, \ldots, A_n$ paarweise kommutierende selbstadjungierte Endomorphismen eines Euklidischen Vektorraumes V. Dann gilt die orthogonale Summenzerlegung

$$V = \bigoplus_{(\lambda_1, \ldots, \lambda_n) \in \mathbf{R}^n} V_{\lambda_1, \ldots, \lambda_n}(A_1, \ldots, A_n).$$

Insbesondere hat V eine Orthonormalbasis, in der alle A_i Diagonalgestalt haben.

Beweis: Dies leitet man mit Hilfe von Lemma 6.2.1 aus Theorem 3 in derselben Weise her, wie Theorem 2 aus Theorem 1.

Definition 2: Sei A ein normaler Endomorphismus von V. Für $r > 0$ und $0 < \varphi < \pi$ setzen wir

$$V_{r,\varphi}(A) = \{x \in V \mid AA^*x = r^2x \wedge (A + A^*)x = 2r\cos\varphi \cdot x\}.$$

Theorem 5: Sei A ein normaler Endomorphismus eines Euklidischen Vektorraumes V.

1. Es gilt die orthogonale direkte Summenzerlegung

$$V = \bigoplus_{\lambda \in \mathbf{R}} V_\lambda(A) \oplus \bigoplus_{r>0, 0<\varphi<\pi} V_{r,\varphi}(A).$$

2. Für $r > 0$ und $0 < \varphi < \pi$ hat $V_{r,\varphi}(A)$ eine Orthonormalbasis $B = (e_1, \ldots, e_k, f_1, \ldots, f_k)$ mit $Ae_i = r\cos\varphi e_i + r\sin\varphi f_i$, $Af_i = r\cos\varphi f_i - r\sin\varphi e_i$ für $1 \leq i \leq k$. Mit anderen Worten,

$$\mathrm{Mat}_B(A|_{V_{r,\varphi}}) = \begin{pmatrix} r\cos\varphi \mathbb{1}_k & -r\sin\varphi \mathbb{1}_k \\ r\sin\varphi \mathbb{1}_k & r\cos\varphi \mathbb{1}_k \end{pmatrix}.$$

Insbesondere ist die Dimension von $V_{r,\varphi}(A)$ gerade.

Beweis: Beim Beweis von Folgerung 6.1.2 hatten wir Theorem 2 auf $\frac{1}{2}(A + A^*)$ und $\frac{1}{2i}(A - A^*)$ angewendet. Eine analoge Anwendung von Theorem 4 ist hier nicht möglich, denn der zweite dieser Ausdrücke macht keinen Sinn mehr. Statt dessen wenden wir Theorem 4 auf die kommutierenden selbstadjungierten Endomorphismen $A + A^*$ und AA^* an. Wir erhalten

$$V = \bigoplus_{(r,s) \in \mathbf{R}^2} V_{r,s}(AA^*, A + A^*) \qquad \text{(orthogonal)}\ . \tag{$*$}$$

Lemma 2:

1. Wenn $V_{t,s}(AA^*, A + A^*) \neq \{0\}$ gilt, so gilt $0 \leq (\frac{s}{2})^2 \leq t$. Mit anderen Worten, $t \geq 0$ und $|s| \leq 2\sqrt{t}$.
2. $V_{(\frac{s}{2})^2,s}(AA^*, A + A^*) = V_{\frac{s}{2}}(A)$.
3. Sei $|s| < 2\sqrt{t}$, also $s = 2\sqrt{t}\cos\varphi$ für ein eindeutig bestimmtes $0 < \varphi < \pi$. Dann gilt $V_{t,s}(AA^*, A + A^*) = V_{\sqrt{t},\varphi}(A)$.

Offenbar beweist das Lemma den ersten Punkt des Theorems, denn es identifiziert $(*)$ mit der zu beweisenden Orthogonalzerlegung.

Beweis von Lemma 2: Wir benutzen eine Hälfte von

Lemma 3: Ein Endomorphismus eines Euklidischen Vektorraumes V ist genau dann normal, wenn für alle $x \in V$ gilt $\|Ax\| = \|A^*x\|$.

Bemerkung: Dies gilt auch für endlichdimensionale Hilberträume.
Beweis von Lemma 3: Sei A normal. Dann ist $\|Ax\| = \sqrt{\langle Ax, Ax\rangle} = \sqrt{\langle A^*Ax, x\rangle} = \sqrt{\langle A^*x, A^*x\rangle} = \|A^*x\|$. Die andere Richtung ist Übungsaufgabe. □

Beweis von Lemma 2: Zu (1): Sei $x \in V_{t,s}(AA^*, A + A^*) - \{0\}$. Wir setzen o.B.d.A. $\|x\| = 1$ voraus. Es gilt: $\|A^*x\|^2 = \langle A^*x, A^*x\rangle = \langle x, AA^*x\rangle = \langle x, tx\rangle = t\langle x, x\rangle = t$. Also gilt $t \geq 0$ und $\|A^*x\| = \sqrt{t}$. Nach Lemma 3 gilt $\|Ax\| = \|A^*x\|$. Es gilt $|s| = \|(A + A^*)x\| \leq \|Ax\| + \|A^*x\| = 2\sqrt{t}$. Damit ist (1) bewiesen. Gleichheit tritt in der letzten Zeile nur dann ein, wenn Gleichheit in der Dreiecksungleichung eintritt, also wenn einer der Vektoren Ax und A^*x ein nicht negatives

Vielfaches des anderen ist. Wegen $\|Ax\| = \|A^*x\|$ ist das der Fall genau dann, wenn $Ax = A^*x$. Dann gilt $2Ax = (A + A^*)x = sx$, und $x \in V_{\frac{s}{2}}(A)$, womit "$\subseteq$" in (2) bewiesen ist. (3) läuft auf eine einfache Anwendung von Definition 2 hinaus: $V_{r,\varphi}(A) = \{x \in V \mid AA^*x = r^2x, (A + A^*)x = 2r\cos\varphi x\} = V_{r^2,2r\cos\varphi}(AA^*, A + A^*)$. "$\supseteq$" von (2): Es gilt nach (1), (3) und $(*)$

$$V = \bigoplus_{s\in\mathbf{R}} V_{(\frac{s}{2})^2,s}(AA^*, A + A^*) \oplus \bigoplus_{\substack{r>0\\ 0<\varphi<\pi}} V_{r,\varphi}(A).$$

Weil A und A^* kommutieren, ist diese Zerlegung A-invariant. Daraus folgt für $\lambda \in \mathbf{R}$ leicht:

$$\begin{aligned} V_\lambda(A) &= \underbrace{\bigoplus_{s\in\mathbf{R}} V_\lambda(A) \cap \underbrace{V_{(\frac{s}{2})^2,s}(AA^*, A + A^*)}_{\subseteq V_{\frac{s}{2}}(A)}}_{\neq\{0\} \text{ nur für } s=2\lambda} \oplus \bigoplus_{\substack{r>0\\ 0<\varphi<\pi}} V_\lambda(A) \cap V_{r,\varphi}(A) \\ &= V_{\lambda^2,2\lambda}(AA^*, A + A^*) \oplus \bigoplus_{\substack{r>0\\ 0<\varphi<\pi}} \underbrace{V_\lambda(A) \cap V_{r,\varphi}(A)}_{=\{0\}} \\ &= V_{\lambda^2,2\lambda}(AA^*, A + A^*) \end{aligned}$$

Es verbleibt der Beweis der versprochenen Behauptung $V_\lambda(A) \cap V_{r,\varphi}(A) = 0$ für $0 < \varphi < \pi$ und $r > 0$. Andernfalls sei x ein von 0 verschiedener Vektor dieses Durchschnitts, wobei wir $\|x\| = 1$ voraussetzen dürfen. Dann gilt $|\lambda| = \|\lambda x\| = \|Ax\| = \sqrt{\langle Ax, Ax\rangle} = \sqrt{\langle A^*Ax, x\rangle} = \sqrt{\langle AA^*x, x\rangle} = \sqrt{r^2\langle x, x\rangle} = r$. Also $|\lambda| = r$. Es gilt $\lambda = \langle Ax, x\rangle = \langle x, A^*x\rangle = \langle A^*x, x\rangle$. Also gilt $2\lambda = \langle (A + A^*)x, x\rangle = 2r\cos\varphi$. Dies kann für $0 < \varphi < \pi$ und $r = |\lambda| \neq 0$ nicht der Fall sein. Dieser Widerspruch beweist die versprochene Behauptung $V_\lambda(A) \cap V_{r,\varphi}(A) = \{0\}$ und der Beweis von (2) ist vollständig. □

Beweis von Theorem 5.2: Der Beweis erfolgt durch eine Kette von Lemmata.

Lemma 4: Für einen Endomorphismus U eines Euklidischen Vektorraumes V sind folgende Aussagen äquivalent:

1. U ist orthogonal;
2. $\|Ux\| = \|x\|$ für alle $x \in V$;
3. $UU^* = \mathrm{Id}_V$;
4. $U^*U = \mathrm{Id}_V$.

Insbesondere sind alle $U \in O(V, \langle\cdot,\cdot\rangle)$ normal.

Beweis: wie Fakt 6.1.4.

Lemma 5: Sei V ein Euklidischer Vektorraum mit $\dim V \neq 0$ und A ein normaler Endomorphismus von V, so daß $V = V_{r,\varphi}(A)$ für ein $r > 0$ und $0 < \varphi < \pi$ gilt.

1. Dann ist jedes Element von v in einem zweidimensionalen A-invarianten Unterraum enthalten.
2. $U = A/r \in O(V)$. Insbesondere gilt $\langle Ax, Ay\rangle = 0$ falls $\langle x, y\rangle = 0$.
3. Das orthogonale Komplement eines A-invarianten Unterraumes ist A-invariant.
4. V hat eine orthogonale Zerlegung in zweidimensionale A-invariante Unterräume.

Beweis:

1. Es gilt $A^2 = A(A + A^*) - AA^* = 2r\cos\varphi A - r^2\mathrm{Id}_V$. Also ist für alle $x \in V$ der Unterraum $\mathcal{L}(\{x, Ax\})$ A-invariant. Für $x \neq 0$ ist dieser Unterraum zweidimensional, denn wenn er eindimensional wäre, wäre x ein Eigenvektor von A, im Widerspruch zu der schon bewiesenen Summenzerlegung aus Theorem 5.

2. Es gilt $\|Ax\|^2 = \langle Ax, Ax\rangle = \langle x, A^*Ax\rangle = r^2\langle x, x\rangle$, also $\|Ax\| = r\|x\|$, also $\|Ux\| = \|x\|$. Nach Lemma 4 ist U orthogonal.

3. Sei $W \subseteq V$ A-invariant, dann ist W auch U-invariant, also ist $U|_W$ ein orthogonaler Automorphismus von W. Sei $x \in W^\perp$, $y \in W$, dann gibt es ein $\tilde{y} \in W$ mit $y = U\tilde{y}$, und $\langle Ux, y\rangle = \langle Ux, U\tilde{y}\rangle = \langle x, \tilde{y}\rangle = 0$. Also ist $W^\perp$ U-invariant und damit auch A-invariant.

4. Beweis durch Induktion nach $\dim(V)$, beginnend mit dem trivialen Fall $\dim(V) = 0$. Sei die Behauptung für Vektorräume kleinerer Dimension als V bewiesen. Sei $x \in V - \{0\}, W \subseteq V$ ein zweidimensionaler A-invarianter Unterraum mit $x \in W$ (vgl. (1)). Dann gilt $V = W \oplus W^\perp$. Nach (3) ist $W^\perp$ A-invariant und man sieht leicht $W^\perp = W^\perp_{r,\varphi}(A|_{W^\perp})$. Nach der Induktionsvoraussetzung ist $W^\perp$ die orthogonale direkte Summe von zweidimensionalen A-invarianten Unterräumen. □

Lemma 6: Unter den Voraussetzungen von Lemma 5 sei zusätzlich V zweidimensional. Wenn B eine Orthonormalbasis von V ist, so gilt

$$\mathrm{Mat}_B(A) = \begin{pmatrix} r\cos\varphi & -r\sin\varphi \\ r\sin\varphi & r\cos\varphi \end{pmatrix} \text{ oder } \mathrm{Mat}_B(A) = \begin{pmatrix} r\cos\varphi & r\sin\varphi \\ -r\sin\varphi & r\cos\varphi \end{pmatrix}.$$

Dabei kann B stets so gewählt werden, daß der erste dieser Fälle eintritt.

Beweis: Sei $B = (e_1, e_2)$ und sei $\mathrm{Mat}_B(A) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann $Ae_1 = ae_1 + ce_2$. Es gilt $\langle Ae_1, e_1\rangle = \langle e_1, A^*e_1\rangle = \langle A^*e_1, e_1\rangle$, also $a = \langle Ae_1, e_1\rangle = \langle \frac{1}{2}(A + A^*)e_1, e_1\rangle = \langle r\cos\varphi\ e_1, e_1\rangle = r\cos\varphi$. Ebenso zeigt man $d = r\cos\varphi$. Es gilt $r^2 = \|Ae_1\|^2 = a^2 + c^2 = r^2\cos^2\varphi + c^2$, also $c = \pm r\sin\varphi$. Ebenso zeigt man $b = \pm r\sin\varphi$. Wenn $b = c$ wäre, so wäre $\mathrm{Mat}_B(A)$ symmetrisch, also wäre A selbstadjungiert. Aber A kann in V keine Eigenvektoren haben. Also $c = -d$. Damit tritt einer der beiden behaupteten Fälle ein. Falls der zweite Fall eintreten sollte, so vertauschen wir e_1 und e_2 und erhalten auf diese Weise eine Orthonormalbasis von V, auf die der erste Fall zutrifft. □

Beweis von Theorem 5, (2): Nach Lemma 5 zerfällt $V_{r,\varphi}(A)$ in eine orthogonale direkte Summe $V_{r,\varphi}(A) = \bigoplus_{i=1}^k W_i$ zweidimensionaler A-invarianter Unterräume. Nach Lemma 6 hat W_i eine Orthonormalbasis (e_i, f_i), in der $A|_{W_i}$ die Matrixdarstellung $\begin{pmatrix} r\cos\varphi & -r\sin\varphi \\ r\sin\varphi & r\cos\varphi \end{pmatrix}$ hat. Dann hat die Orthonormalbasis $(e_1, \ldots, e_k, f_1, \ldots, f_k)$ von $V_{r,\varphi}$ die geforderte Eigenschaft. □

Bemerkung: Auf $(V_{r,\varphi}(A))_{\mathbb{C}}$ ist die Komplexifizierung der Einschränkung von A ein normaler Endomorphismus mit den Eigenwerten $re^{\pm i\varphi}$.

Folgerung 2: Sei V ein Euklidischer Vektorraum, U ein orthogonaler Endomorphismus von V.

1. Es gilt
$$V = V_1(U) \oplus V_{-1}(U) \oplus \bigoplus_{0<\varphi<\pi} V_{1,\varphi}(U).$$

2. $V_{1,\varphi}(U)$ hat eine Basis $B = (e_1, \ldots, e_k, f_1, \ldots, f_k)$ mit

$$\mathrm{Mat}_B(U|_{V_{1,\varphi}}) = \begin{pmatrix} \cos\varphi \mathbb{1}_k & -\sin\varphi \mathbb{1}_k \\ \sin\varphi \mathbb{1}_k & \cos\varphi \mathbb{1}_k \end{pmatrix}.$$

3. $V_{-1}(U)$ hat gerade Dimension, falls $U \in \mathrm{SO}(V, \langle\cdot,\cdot\rangle)$ (d.h., $\det(U) = 1$). Falls $\det(U) = -1$, hat $V_{-1}(U)$ ungerade Dimension.

4. Wenn die Dimension von V ungerade ist, hat U einen Eigenwert. Falls zusätzlich $\det(U) = 1$ gilt, so ist 1 ein Eigenwert.

Beweis:

1. Es gilt $UU^* = \mathrm{Id}_V$, also gilt $V_{r,\varphi}(U) = \{0\}$ für $r \neq 1$. Wegen $\|Ux\| = \|x\|$, sind ± 1 die einzigen Eigenwerte von U. Die Behauptung folgt nun aus Theorem 5.

2. folgt ebenfalls aus Theorem 5.

3. **Fakt:** Sei A ein Endomorphismus eines endlichdimensionalen Vektorraumes X, und sei $X = Y \oplus Z$ eine direkte Zerlegung von X in A-invariante Unterräume. Dann gilt $\det(A) = \det(A|_Y)\det(A|_Z)$, d.h.

$$\begin{vmatrix} A & 0 \\ 0 & B \end{vmatrix} = \det(A)\det(B),\ A \in \mathrm{Mat}\,(k,k,K),\ B \in \mathrm{Mat}\,(l,l,K).$$

Der Beweis erfolgt durch Induktion nach k, Auflösung nach der ersten Zeile. Es folgt

$$\begin{aligned} \det(U) &= \det(U|_{V_1(U)})\det(U|_{V_{-1}(U)}) \prod_{0<\varphi<\pi} \underbrace{\det(U|_{V_{1,\varphi}(U)})}_{=1,\ \text{falls } V_{1,\varphi}(U)=\{0\}} \\ &= 1(-1)^{\dim(V_{-1}(U))} \prod_{0<\varphi<\pi} \begin{vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{vmatrix}^{\dim(V_{1,\varphi}(U))/2} = (-1)^{\dim(V_{-1}(U))} \end{aligned}$$

4. folgt, weil $V_{1,\varphi}(V)$ für $0 < \varphi < \pi$ gerade Dimension hat. Wenn zusätzlich $\det(U) = 1$ gilt, so hat auch $V_{-1}(U)$ gerade Dimension (nach (3)), also ist $\dim(V_1(U))$ ungerade und daher ungleich Null. □

Folgerung 3: Jedes Element von $\mathrm{SO}(2) - \{\mathrm{Id}\}$ hat bezüglich der Standardbasis des $\mathbb{R}^2$ die Matrixdarstellung

$$\begin{pmatrix} r\cos\varphi & r\sin\varphi \\ -r\sin\varphi & r\cos\varphi \end{pmatrix}$$

mit $0 < \varphi < 2\pi$ und φ eindeutig bestimmt.

Beweis: folgt aus Lemma 6.

Folgerung 4: Sei V ein dreidimensionaler Euklidischer Vektorraum und sei $U \in \mathrm{SO}(V, \langle\cdot,\cdot\rangle) - \{\mathrm{Id}_V\}$. Dann gibt es ein eindeutig bestimmtes φ mit $0 < \varphi < \pi$, den Drehwinkel von U, so daß $V = V_1(U) \oplus V_{1,\varphi}(U)$.

Beweis: folgt aus Folgerung 2.

6.3 Das charakteristische Polynom

Sei K ein Körper.

Definition 1: Ein Polynom mit Koeffizienten aus K ist eine Folge $p = (p_k)_{k=0}^{\infty}$, so daß $p_i \in K$ und ein i_0 mit $p_i = 0$ für $i > i_0$ existiert. Wir schreiben

$$p(T) = \sum_{k=0}^{i_0} p_k T^k.$$

Sei $K[T]$ die Menge aller Polynome mit Koeffizienten aus K. Diese Menge bildet einen K-Vektorraum mit den Operationen

$$\begin{aligned} (p_i)_{i=0}^{\infty} + (q_i)_{i=0}^{\infty} &= (p_i + q_i)_{i=0}^{\infty}, \\ \lambda(p_i)_{i=0}^{\infty} &= (\lambda p_i)_{i=0}^{\infty}. \end{aligned}$$

Das konstante Polynom λ ist durch die Folge $(\lambda, 0, \ldots)$ definiert. Die Multiplikation zweier Polynome $(p_k)_{k=0}^{\infty}$ und $(q_k)_{k=0}^{\infty}$ ist durch die Folge $(v_k)_{k=0}^{\infty}$ gegeben, wobei $v_m = \sum_{k=0}^{m} p_k q_{m-k}$. Für $x \in K$ und $p = (p_k)_{k=0}^{\infty} \in K[T]$ sei der **Funktionswert** von p an der Stelle x durch $p(x) = \sum_{k=0}^{\infty} p_k x^k$ definiert (diese Summe ist tatsächlich endlich). x ist eine **Nullstelle**, falls $p(x) = 0$ gilt. Der **Grad** eines von Null verschiedenen Polynoms p ist das größte i mit $p_i \neq 0$. Der Grad des Nullpolynoms sei $-\infty$. Wir bezeichnen den Grad von p mit $\deg p$.

Bemerkung 1: Wenn K endlich ist, gibt es von Null verschiedene Polynome $p \in K[T]$ mit $p(x) = 0$ für alle $x \in K$. Deshalb konnten wir Polynome nicht einfach als Funktionen von K nach K definieren. Zum Beispiel gilt $x^2 + x = 0$ für alle $x \in \mathbb{F}_2$. In $p(T) = \sum_{k=0}^{\deg p} p_k T^k$ muß man T als eine Variable auffassen, in die man auch Elemente von Erweiterungskörpern von K einsetzen kann.

Fakt 1: Seien $P, Q \in K[T]$. Dann gilt für $x \in K$

1. $(PQ)(x) = P(x)Q(x)$, $(P+Q)(x) = P(x) + Q(x)$;
2. $\deg(PQ) = \deg P + \deg Q$,$\deg(P+Q) \leq \max\{\deg P, \deg Q\}$ $(-\infty + -\infty = -\infty, -\infty + c = -\infty \ \forall c \in \mathbb{N})$. Falls $\deg Q < \deg P$, so gilt insbesondere $\deg(P+Q) = \deg P$.

Beweis: Sei $PQ = R = (r_k)_{k=0}^{\infty}$. Dann gilt

$$\begin{aligned} R(x) &= \sum_{k=0}^{\infty} r_k x^k = \sum_{k=0}^{\infty} \sum_{i=0}^{k} p_i q_{k-i} x^k \\ &= \sum_{i,j=0}^{\infty} p_i q_j x^{i+j} = \left(\sum_{i=0}^{\infty} p_i x^i \right) \left(\sum_{j=0}^{\infty} q_j x^j \right) = P(x)Q(x) \end{aligned}$$

Die zweite Formel ist noch einfacher.
Die dritte Formel ist offenbar richtig, falls P oder Q verschwindet, denn dann sind beide Seiten $-\infty$. Sei $P \neq 0$ und $Q \neq 0$, $a = \deg P$, $b = \deg Q$. Sei $R = PQ = (r_k)_{k=0}^{\infty}$. Dann gilt

$$r_{a+b} = \sum_{i=0}^{a+b} p_i q_{a+b-i}.$$

Für $i > a$ verschwindet in $p_i q_{a+b-i}$ der erste Faktor, für $i < a$ gilt $a+b-i > b$ und es verschwindet der zweite Faktor. Also bleibt in der letzten Summe nur ein Summand übrig. $v_{a+b} = p_a q_b \neq 0$. Für $k > a+b$ gilt

$$r_k = \sum_{i=0}^{k} p_i q_{k-i},$$

und in $p_i q_{k-i}$ verschwindet für $i > a$ der erste Faktor, für $i \leq a$ der zweite Faktor, es gilt nämlich $k-i > a+b-i \geq b$. Also gilt $r_{a+b} \neq 0$ und $r_i = 0$ für $i > a+b$, also $\deg R = a+b = \deg P + \deg Q$. Sei $S = P + Q = (s_k)_{k=0}^{\infty}$. Für $k > \max\{\deg P, \deg Q\}$ gilt $s_k = p_k + q_k = 0$. Also $\deg S \leq \max\{\deg P, \deg Q\}$. Falls $\deg Q < \deg P$ gilt $s_{\deg p} = p_{\deg p} + q_{\deg p} = p_{\deg p} \neq 0$, also $\deg S = \deg P$. □

Bemerkung:

1. Insbesondere folgt aus $P \neq 0$ und $Q \neq 0$ $\deg(PQ) = \deg P + \deg Q \geq 0$, also $PQ \neq 0$.
2. P ist genau dann ein konstantes Polynom, wenn $\deg P = 0$ oder $\deg P = -\infty$.
3. Sei K ein endlicher Körper, seien $x_1, \ldots, x_q$ die Elemente von K und sei $P(T) = \prod_{k=1}^{q}(T - x_k)$. Dann gilt $\deg P = q$, $P \neq 0$, aber $P(x) = \prod_{k=1}^{q}(x - x_k) = 0$ für $x \in K$.
4. Polynomaddition und Polynommultiplikation sind kommutativ und assoziativ. Es gilt auch das Distributivgesetz $P(Q+R) = PQ+PR$. Für das konstante Polynom 1 gilt $1P = P$, für das konstante Polynom 0 gilt $P + 0 = P$ und $0P = 0$. $(K[T], +)$ ist eine abelsche Gruppe. Man sagt dazu, $(K[T], +, \cdot)$ ein **kommutativer Ring mit Einselement** ist (ebenso wie $(\mathbb{Z}, +, \cdot)$).

Satz 1: (Division mit Rest) Sei $P, Q \in K[T]$ und $Q \neq 0$. Dann kann P mit Rest durch Q geteilt werden; $P = XQ + R$ mit $X \in K[T]$, $R \in K[T]$ und $\deg R < \deg Q$. Dabei sind X und R durch diese Eigenschaften eindeutig bestimmt.

Beweis: Existenz der Division mit Rest, Induktion nach $\deg P$. Für $\deg P < \deg Q$ ist $P = 0Q + P$ eine Division mit Rest. Sei die Existenz der Division mit Rest durch Q für Polynome kleineren Grades als P bewiesen. Sei weiterhin $\deg P \geq \deg Q$. Sei

$$\tilde{P}(T) = P(T) - \frac{p_{\deg P}}{q_{\deg Q}} T^{\deg P - \deg Q} Q(T).$$

Der Grad des Subtrahenden ist gleich $\deg P$, sein führender Koeffizient ist

$$\frac{p_{\deg P}}{q_{deg Q}} q_{\deg Q} = p_{\deg P}.$$

Also gilt $\tilde{p}_{\deg P} = 0$. Also $\deg \tilde{P} < \deg P$. Nach Induktionsvoraussetzung kann $\tilde{P}$ als $\tilde{P} = \tilde{X}Q + R$ dargestellt werden mit $\deg R < \deg Q$. Dann gilt aber $P = XQ + R$ mit

$$X(T) = \tilde{X}(T) + \frac{p_{\deg P}}{q_{\deg Q}} T^{\deg P - \deg Q}.$$

Damit ist die Existenz der Division mit Rest bewiesen.
Beweis der Eindeutigkeit: Angenommen, $XQ+R = P = \tilde{X}Q+\tilde{R}$ mit $\deg R < \deg Q$ und $\deg \tilde{R} < \deg Q$. Falls $X = \tilde{X}$ folgt sofort $R = \tilde{R}$, denn $(K[T], +)$ ist eine abelsche Gruppe. Wenn $X \neq \tilde{X}$, so wäre in $Q(X - \tilde{X}) = \tilde{R} - R$ nach Fakt 1 der Grad der linken Seite $\deg Q + \deg(X - \tilde{X}) \geq \deg Q$, der Grad der rechten Seite dagegen $< \deg Q$. Widerspruch. □

Bemerkung: Der führende Koeffizient eines von 0 verschiedenen Polynoms $P = (p_k)_{k=0}^{\infty}$ ist $p_{\deg P}$. Bei der Multiplikation multiplizieren sich die führenden Koeffizienten. Der Grad einer Summe zweier Polynome gleichen Grades ist genau dann kleiner als der Grad der Summanden, wenn die Summe der führenden Koeffizienten Null ergibt.

Fakt 2: Jedes Polynom $P \neq 0$ kann auf eindeutige Weise als

$$P(T) = \prod_{i=1}^{k} (T - x_i)^{v_i} Q(T)$$

dargestellt werden, wobei $Q \in K[T]$ mit $Q(x) \neq 0$ für alle $x \in K$ und $x_1, \ldots, x_k$ paarweise verschieden sind; $k, v_i \in \mathbf{N}$ eindeutig bestimmt. Man nennt die v_i die **Vielfachheit** der Nullstelle x_i.

Beweis: Existenz durch Induktion nach $\deg P$. Für $\deg P = 0$ sei $k = 0$ und $Q = P$. Sei die Behauptung für Polynome kleineren Grades bewiesen. Wenn P keine Nullstelle hat, sei $k = 0$ und $Q = P$. Wenn x eine Nullstelle von P ist, dann dividieren wir P mit Rest durch $(T - x)$: $P(T) = (T - x)\tilde{P}(T) + R$. Wegen $\deg R < \deg(T - x) = 1$ ist R konstant. Wegen $P(x) = (x-x)\tilde{P}(x)+R(x) = R(x)$ gilt $R = 0$. Also $P(T) = (T-x)\tilde{P}(x)$. Sei $\tilde{P}(T) = \prod_{i=1}^{\tilde{k}} (T-x_i)^{\tilde{v}_i} Q(T)$ eine Zerlegung der beschriebenen Art für $\tilde{P}$. Sie existiert nach der Induktionsvoraussetzung, denn $\deg P = \deg(T - x) + \deg \tilde{P} = 1 + \deg \tilde{P}$. Falls kein i mit $1 \leq i \leq \tilde{k}$ und $x = x_i$ existiert, dann sei $k = \tilde{k} + 1$, $x_i = \tilde{x}_i$, $v_i = \tilde{v}_i$ für $1 \leq i \leq \tilde{k}$ und $x_k = x$, $v_k = 1$. Falls $x = x_i$ für ein i, $1 \leq i \leq \tilde{k}$, gilt, sei $k = \tilde{k}$, $x_j = \tilde{x}_j$ und $v_j = \tilde{v}_j$ für $i \neq j$ und $v_i = \tilde{v}_i + 1$. In beiden Fällen gilt $P(T) = \prod_{i=1}^{k} (T - x_i)^{v_i} Q(T)$. Damit ist der Beweis der Existenz abgeschlossen.
Beweis der Eindeutigkeit der Zerlegung: Angenommen, x ist eine Nullstelle von P. Dann muß einer der Faktoren $Q(x)$ oder $(x - x_i)^{v_i}$, $1 \leq i \leq k$, verschwinden. Weil Q keine Nullstelle hat, muß $x \in \{x_1, \ldots, x_k\}$ sein. Auf der anderen Seite sind alle x_i Nullstellen von P. Also ist $\{x_1, \ldots, x_k\} = \{x \in K \mid P(x) = 0\}$ durch P eindeutig bestimmt. Eine andere Zerlegung der beschriebenen Art für P hat also die Form $P(T) = \prod_{i=1}^{k} (T - x_i)^{\tilde{v}_i} \tilde{Q}(T)$. Angenommen, es gäbe ein i mit $1 \leq i \leq k$ und $\tilde{v}_i \neq v_i$. Wir können o.B.d.A. $i = 1$ und $\tilde{v}_1 < v_1$ annehmen. Dann gilt

$$0 = \tilde{P}(T) - P(T) = (T - x_1)^{\tilde{v}_1} \left[\prod_{i=2}^{k} (T - x_i)^{\tilde{v}_i} \tilde{Q}(T) - (T - x_1)^{v_1 - \tilde{v}_1} \prod_{i=2}^{k} (T - x_i)^{v_i} Q(T) \right].$$

Aber $K[T]$ ist **nullteilerfrei**, d.h., das Produkt zweier von 0 verschiedener Polynome ist von 0 verschieden. Wegen $(T-x_1)^{\tilde{v}_1} \neq 0$ muß in der letzten Gleichung der Faktor in eckigen Klammern verschwinden:

$$\prod_{i=2}^{k} (T - x_i)^{\tilde{v}_i} \tilde{Q}(T) = (T - x_1)^{v_1 - \tilde{v}_1} \prod_{i=2}^{k} (T - x_i)^{v_i} Q(t).$$

Dies ist nicht möglich, denn x_1 ist eine Nullstelle der rechten Seite, nicht aber der linken. Also $v_i = \tilde{v}_i$ für $1 \leq i \leq k$. Dann gilt $0 = \prod_{i=1}^{k} (T - x_i)^{v_i} (Q(T) - \tilde{Q}(T))$, also $Q = \tilde{Q}$. □

Folgerung 1: Ein Polynom $P \neq 0$ n-ten Grades hat höchstens n verschiedene Nullstellen.

Beweis: Sei $P(T) = \prod_{i=1}^{k} (T - x_i)^{v_i} Q(T)$ wie in Fakt 2, dann

$$\deg P = \sum_{i=1}^{k} \deg(T - x_i)^{v_i} + \deg Q = \sum_{i=1}^{k} v_i + \deg Q \geq k = \text{Anzahl Nullstellen von } P \text{ in } K.$$

□

Bemerkung: Die Zahl k aus Fakt 2 ist die Anzahl der verschiedenen Nullstellen von P in K.

v_i nennt man dabei die **Vielfachheit** der Nullstelle x_i.

Folgerung 2: Wenn K unendlich viele Elemente hat, ist ein Polynom $P \in K[T]$ durch die Funktion $K \to K$, $x \mapsto P(x)$, eindeutig bestimmt. Insbesondere ist dies für Körper mit $\mathrm{char}(K) = 0$ der Fall.

Beweis: Folgt aus Folgerung 1. □

Bemerkung: Folgerung 1 hätte auch mit der Formel für die Vandermondsche Determinante bewiesen werden können.

Definition 2: Ein Körper K ist **algebraisch abgeschlossen**, wenn jedes nichtkonstante Polynom eine Nullstelle in K hat.

Fakt 3: Dann zerfällt jedes Polynom $P \neq 0$ in **Linearfaktoren**, d.h. die Zerlegung aus Fakt 2 nimmt die Form

$$P(T) = q \prod_{i=1}^{L} (T - x_i)^{v_i}$$

an. Faktoren $(T - x_i)$ nennt man Linearfaktoren.

Beweis: Nach Definition 2 muß das Polynom Q aus Fakt 2 konstant sein. □

Bemerkung: Ein endlicher Körper ist niemals algebraisch abgeschlossen.

Beweis: Sei $K = \{x_1, \ldots, x_n\}$, dann hat das Polynom $P(T) = 1 + \prod_{i=1}^{n} (T - x_i)$ keine Nullstelle in K. □

Fakt 4: $\mathbb{R}$ ist nicht algebraisch abgeschlossen. Jedes Polynom ungeraden Grades hat eine Nullstelle.

Beweis: $x^2 + 1 = 0$ hat keine reelle Lösung. Sei $\deg P$ ungerade und der führende Koeffizient 1. Dann $\lim_{T \to \infty} P(T) = \infty$ und $\lim_{T \to -\infty} P(T) = -\infty$, da P stetig ist, hat P nach dem Zwischenwertsatz eine Nullstelle. □

Theorem 6: (C.F. Gauss, Fundamentalsatz der Algebra) $\mathbb{C}$ ist algebraisch abgeschlossen.

Bemerkung 2: Eigentlich betrachtet man Theorem 6 als einen Satz der Analysis. Der erste weitgehend algebraische Beweis wurde von E. Artin und Schreier angegeben (1922).

Lemma 1: Für jedes $n \in \mathbb{N}$ kann man aus jeder komplexen Zahl z eine n-te Wurzel ziehen.

Beweis: Sei $z = re^{i\varphi} = r\cos\varphi + ir\sin\varphi$, $r \geq 0$, $0 \leq \varphi < 2\pi$. Dann ist

$$\sqrt[n]{r} e^{i\frac{\varphi}{n}} = \sqrt[n]{r}\cos\frac{\varphi}{n} + i\sqrt[n]{r}\sin\frac{\varphi}{n}$$

eine n-te Wurzel aus z. □

Bemerkung: Beim Beweis von Theorem 6 können wir uns auf Polynome mit führendem Koeffizienten 1 beschränken. Sei also

$$P(T) = T^n + \sum_{k=0}^{n-1} p_k T^k. \qquad (*)$$

Lemma 2: Sei x_0 ein lokales Minimum der Funktion $z \mapsto |P(z)|$, $\mathbb{C} \to \mathbb{R}$. Dann ist x_0 eine Nullstelle von P.

Beweis: Angenommen, $p = P(x_0) \neq 0$. Sei $P(T) = p + (T - x_0)^k Q(T)$, mit $Q(x_0) = q \neq 0$ (vgl.

Satz 1, Fakt 2), $k > 0$. Sei z eine k-te Wurzel aus $-\frac{p}{q}$. Für $r > 0$ gilt

$$\begin{aligned} |P(x_0 + rz)| &= |p + (rz)^k Q(x_0 + rz)| = |p - \frac{p}{q} r^k (q + \underbrace{a_0 rz + \ldots + a_k (rz)^k}_{O(r)})| \\ &= |p - pr^k + O(r^{k+1})| \leq |p|(1 - r^k) + cr^{k+1} \end{aligned}$$

für eine Konstante c. Wenn also $0 < r < \min(1, \frac{p}{c})$ ist, so gilt $|P(x_0 + rz)| < |P(x_0)|$. Widerspruch zu unserer Annahme $p = P(x_0) \neq 0$. □

Lemma 3: Die (stetige) Funktion $f(z) = |P(z)|$ nimmt auf $\mathbb{C}$ ein Minimum an.

Beweis: Sei n der Grad des Polynoms (vgl. $(*)$). Sei $R = \max\{|p_i| \mid 0 \leq i < n\}$. Für $|x| > S = \max(1, |p_0| + nR)$, $x \in \mathbb{C}$, gilt

$$f(x) = |P(x)| \geq |x|^n - |\sum_{i=0}^{n-1} p_i x^i| \overset{|x| \geq 1}{\geq} |x|^n - nR|x|^{n-1} = |x|^{n-1}(|x| - nR) > p_0 = |P(0)| = f(0).$$

Nach Sätzen aus der Analysis nimmt die stetige Funktion $f(z)$ auf der kompakten Menge $\{z \mid |z| \leq S\}$ ein Minimum an. Nach der vorhergehenden Überlegung ist dieses Minimum auch ein Minimum von f auf ganz $\mathbb{C}$. □

Theorem 6 folgt nun aus Lemma 2 und Lemma 3. □

Bemerkung 3: Aus Gründen der Einfachheit nehmen wir an, daß K unendlich viele Elemente hat. Dann ist $P \in K[T]$ eindeutig durch die Funktion $K \to K$, $x \mapsto P(x)$, definiert.

Satz 2:

1. Sei V ein endlichdimensionaler K-Vektorraum und A ein Endomorphismus von V. Dann gibt es genau ein Polynom $P_A \in K[T]$ mit $\det(\lambda \mathrm{Id}_V - A) = P(\lambda)$ für alle $\lambda \in K$. $\deg P_A = \dim V$.
2. Sei $A \in \mathrm{Mat}\,(n, n, K)$. Dann existiert genau ein Polynom $P_A \in K[T]$ mit $\det(\lambda \mathbb{1}_n - A) = P_A(\lambda)$. Es gilt $\deg P_A = n$.

Beweis: Die Eindeutigkeit von P_A folgt in beiden Fällen aus Folgerung 2. Es genügt, (2) zu beweisen ($P_A = P_{\mathrm{Mat}_B(A)}$). Sei also $A = (a_{ij})_{i,j=1}^n$. Nach der Leibnizschen Determinantenformel gilt

$$\begin{aligned} \det(\lambda \mathbb{1}_n - A) &= \det(\lambda \delta_{ij} - a_{ij})_{i,j=1}^n = \sum_{\pi \in S_n} \mathrm{sign}(\pi) \prod_{i=1}^n (\lambda \delta_{i,\pi(i)} - a_{i,\pi(i)}) \\ &= \sum_{\pi \in S_n} P_{A,\pi}(\lambda), \end{aligned}$$

wobei $P_{A,\pi} \in K[T]$ durch

$$\begin{aligned} P_{A,\pi}(T) &= \mathrm{sign}(\pi) \prod_{i=1}^n (T\delta_{i,\pi(i)} - a_{i,\pi(i)}) \\ &= \mathrm{sign}(\pi) \prod_{\substack{1 \leq i \leq n \\ \pi(i) \neq i}} (-a_{i,\pi(i)}) \prod_{\substack{1 \leq i \leq n \\ \pi(i) = i}} (T - a_{ii}), \end{aligned}$$

definiert ist. Also $\det(\lambda \mathbb{1}_n - A) = P_A(\lambda)$, wobei $P_A = \sum_{\pi \in S_n} P_{A,\pi}$. Es verbleibt die Bestimmung des Grades von P_A.

$$\deg(P_{A,\pi}) \overset{(*)}{=} |\{i \mid 1 \leq i \leq n \wedge \pi(i) = i\}| = \begin{cases} n & \pi = \mathrm{Id} \\ \leq n-2 & \pi \neq \mathrm{Id} \end{cases}, \qquad (*)$$

denn jede Permutation von $\{1, \ldots, n\}$, die $n-1$ dieser Zahlen festläßt, läßt wegen ihrer Bijektivität auch die verbleibende Zahl fest. Aus $(*)$ und Fakt 1 folgt $\deg(P_A) = n$, denn

$$P_A = \underbrace{P_{A,\mathrm{Id}}}_{\deg = n} + \underbrace{\sum_{\pi \neq \mathrm{Id}} P_{A,\pi}}_{\deg < n \text{ nach Fakt 1}} . \qquad \square$$

Bemerkung: Die Definition von $P_{A,\pi}$ und damit auch die Definition von P_A funktioniert auch für endliche Körper K. In der Situation von Satz 2, (1) setzt man dann $P_A = P_{\mathrm{Mat}_B(A)}$ und hat dann noch die Unabhängigkeit dieses Polynomes von B zu zeigen.

Definition 3: P_A nennt man das **charakteristische Polynom** von A.

Bemerkung: Es gibt auch die Konvention $\tilde{P}_A(\lambda) = \det(A - \lambda \mathrm{Id})$ für das charakteristische Polynom.

Definition 4: Die **Spur** $tr(A)$ von $A \in \mathrm{Mat}(n, n, K)$ ist die Summe ihrer Diagonalelemente. Die Spur eines Endomorphismus eines endlichdimensionalen Vektorraumes ist die Spur seiner Matrixdarstellung bezüglich einer beliebigen Basis.

Fakt 6: Seien $A, B \in \mathrm{Mat}(n, n, K)$. Es gilt:

1. $tr(AB) = tr(BA)$
2. $tr(A + B) = tr(A) + tr(B)$
3. Der zweithöchste Koeffizient des charakteristischen Polynoms von A ist $-tr(A)$.

Insbesondere ist die Spur eines Endomorphismus von V unabhängig von der Wahl der Basis von V.

Beweis:

1. Sei $A = (a_{ij})_{i,j=1}^n$, $B = (b_{ij})_{i,j=1}^n$, $AB = (c_{ij})_{i,j=1}^n$, $BA = (d_{ij})_{i,j=1}^n$. Die Formel des Matrixproduktes liefert $c_{ik} = \sum_{j=1}^n a_{ij} b_{jk}$, $d_{ik} = \sum_{j=1}^n b_{ij} a_{jk}$, also

$$tr(AB) = \sum_{i=1}^n c_{ii} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ji} = \sum_{j=1}^n \sum_{i=1}^n b_{ji} a_{ij} = \sum_{j=1}^n d_{jj} = tr(BA).$$

2. ist klar. Allgemeiner: $tr : \mathrm{Mat}(n, n, K) \to K$ ist ein lineares Funktional.

3. Erinnere an den Beweis von Satz 2: $\det(\lambda \mathrm{Id} - A) = P_A(\lambda) = \sum_{\pi \in S_n} P_{A,\pi}(\lambda)$.

$$P_{A,\pi}(\lambda) = \mathrm{sign}(\pi) \prod_{\substack{1 \leq i \leq n \\ \pi(i) \neq i}} (-a_{i,\pi(i)}) \prod_{\substack{1 \leq i \leq n \\ \pi(i) = i}} (\lambda - a_{ii}),$$

$\deg(P_{A,\pi}) \leq n-2$, falls $\pi \neq \mathrm{Id}$. $\deg(P_{A,\mathrm{Id}}) = n$. Also stimmt der zweithöchste Koeffizient von P_A mit dem zweithöchsten Koeffizienten von $P_{A,\mathrm{Id}}$ überein.

$$\begin{aligned} P_{A,\mathrm{Id}}(\lambda) &= \prod_{1\leq i\leq n}(\lambda - a_{ii}) = \sum_{S\subset\{1,\ldots,n\}}(-1)^{n-|S|}\prod_{i\notin S} a_{ii}\lambda^{|S|} \\ &= \lambda^n + \sum_{i=1}^{n} -a_{ii}\lambda^{n-1} + \text{Terme vom Grad} \leq n-2 \\ &= \lambda^n - tr(A)\lambda^{n-1} + \text{Terme vom Grad} \leq n-2. \end{aligned}$$

Sei nun A ein Endomorphismus von V. Wir möchten zeigen, daß $tr(\mathrm{Mat}_{B,B}(A))$ von der Auswahl dieser Basis unabhängig ist. Sei $\tilde{B}$ eine andere Basis von V. Dann gilt

$$\begin{aligned} \mathrm{Mat}_{\tilde{B},\tilde{B}}(A) &= \mathrm{Mat}_{B,\tilde{B}}(\mathrm{Id})\mathrm{Mat}_{\tilde{B},B}(A) \\ &= \mathrm{Mat}_{B,\tilde{B}}(\mathrm{Id})\mathrm{Mat}_{B,B}(A)\mathrm{Mat}_{\tilde{B},B}(\mathrm{Id}) = M\mathrm{Mat}_{B,B}M^{-1} \end{aligned}$$

$tr(\mathrm{Mat}_{\tilde{B},\tilde{B}}(A)) = tr(M\mathrm{Mat}_{B,B}M^{-1}) \overset{(1)}{=} tr(M^{-1}M\mathrm{Mat}_{B,B}(A)) = tr(\mathrm{Mat}_{B,B}(A))$. □

Bemerkung:

1. (1)-(3) von Fakt 6 können nun auch auf Endomorphismen von endlichdimensionalen Vektorräumen angewendet werden.
2. (3) macht nach dem, was wir über das charakteristische Polynom bewiesen haben, nur über unendlichen Körpern Sinn (obwohl es allgemein gilt). Alle anderen Teile von Fakt 6 haben wir über beliebigen Körpern bewiesen.

Fakt 7: Sei $\dim(V) = n$, sei A ein Endomorphismus von V.

1. Das charakteristische Polynom von A hat die Form
$$P_A(\lambda) = \lambda^n - tr(A)\lambda^{n-1} + \ldots + (-1)^n \det(A).$$
2. λ ist Eigenwert von A genau dann, wenn λ Nullstelle des charakteristischen Polynoms P_A ist.
3. Wenn K algebraisch abgeschlossen ist (z.B. $\mathbb{C}$), dann hat A einen Eigenwert.

Beweis: ist klar.

Definition 5: Sei V ein Vektorraum.

1. Ein **Projektor** ist ein Endomorphismus p von V mit $p^2 = p$.
2. Eine **Involution** ist ein Endomorphismus i von V mit $i^2 = \mathrm{id}$.
3. Sei $X \subseteq V$ ein Unterraum. Ein **Komplementärraum** zu X ist ein Unterraum $Y \subseteq V$ mit $V = X \oplus Y$.

Beispiel: Wenn V ein Euklidischer Vektorraum oder ein Euklidischer Hilbertraum ist, so ist $X^\perp$ ein Komplementärraum zu X, und der Orthogonalprojektor ist ein Projektor.

Bemerkung 5: Sei p ein Projektor. Dann ist $Y = \ker(p)$ ein Komplementärraum zu $X = \mathrm{Bild}\,(p)$. Wenn umgekehrt $X \subseteq V$ und ein Komplementärraum Y zu X vorgegeben sind, dann

existiert genau ein Projektor p mit $\ker(p) = Y$ und $\text{Bild}\,(p) = X$.

Beweis: Sei p vorgegeben. Aus $v \in \ker(p) \cap \text{Bild}\,(p)$, etwa $v = p(w)$, folgt

$$0 = p(v) = p(p(w)) = p^2(w) = p(w) = v.$$

Also $\ker(p) \cap \text{Bild}\,(p) = \{0\}$. Wenn $v \in V$ vorgegeben ist, so gilt:

$$v = \underbrace{p(V)}_{\in \text{Bild}(p)} + \underbrace{(v - p(V))}_{\in \ker(p)}.$$

Also $V = \ker(p) \oplus \text{Bild}\,(p)$. Wenn umgekehrt $V = X \oplus Y$ gilt, so sei $p : V \to V$ durch $p(v) = x$ definiert, wobei $v = x + y$, $x \in X$, $y \in Y$ die eindeutige Zerlegung von V in der direkten Summe $X \oplus Y$ ist. Dann ist $p^2 = p$, $\ker(p) = Y$, $\text{Bild}\,(p) = X$. □

Bemerkung 6: Sei $\text{char}\,(K) \neq 2$.

1. Durch

$$\begin{array}{rcl} \text{Involution auf } V & & \text{Projektion in } V \\ i & \to & p = \frac{1}{2}(\text{id}_V + i) \\ i = 2p - \text{id}_V & \leftarrow & p \end{array}$$

   sind zwei zueinander inverse Bijektionen definiert.

2. Wenn i eine Involution auf V ist, so gilt $V = V_1(i) \oplus V_{-1}(i)$.

Beweis: ist klar.

Bemerkung 7: Insbesondere nehmen Projektoren und, falls $\text{char}\,(K) \neq 2$, auch Involutionen eines endlichdimensionalen Vektorraumes in einer geeigneten Basis stets Diagonalgestalt an. Die Spur eines Projektors p ist $\dim(\text{Bild}\,(p))$, das charakteristische Polynom von p ist $P_p(T) = (T-1)^{\dim(\text{Bild}(p))} T^{\dim(\ker(p))}$. Wenn i eine Involution und $\text{char}\,(K) \neq 2$ ist, so gilt $P_i(T) = (T-1)^{\dim(V_1(i))}(T+1)^{\dim(V_{-1}(i))}$.

$$\det(\lambda \text{Id} - p) = \left|\begin{array}{ccc:ccc} \lambda-1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \lambda-1 & 0 & \dots & 0 \\ \hdashline 0 & \dots & 0 & \lambda & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & \lambda \end{array}\right| \left.\begin{array}{l} \\ \\ \\ \\ \\ \\ \end{array}\right\} \begin{array}{l} \dim(\text{Bild}\,(p)) \\ \\ \\ \\ \dim(\ker(p)) \end{array}$$

$$= (\lambda - 1)^{\dim(\text{Bild}(p))} \lambda^{\dim(\ker(p))}.$$

6.4 Die Jordansche Normalform

Definition 1: Ein Endomorphismus N von V ist **nilpotent**, falls ein $k \in \mathbb{N}$ mit $N^k = 0$ existiert. Ein Endomorphismus U von V ist **unipotent**, falls $U - \text{id}_V$ nilpotent ist.

Bemerkung: Für einen Körper K mit $\text{char}\,(K) = 0$ ist für ein unipotentes U nicht gesichert, daß ein $k > 0$ existiert mit $U^k = \text{id}$.

Fakt 1: Das Spektrum eines nilpotenten Endomorpismus besteht nur aus 0.

Beweis, falls $\dim(V) < \infty$. Wir zeigen, daß für $\lambda \neq 0$ $V_\lambda(N) = \{0\}$. In der Tat aus $\lambda \neq 0$ und $v \in V_\lambda(N) - \{0\}$ würde $N^k v = \lambda^k v \neq 0$ für alle k folgen. □

Bemerkung 1: In Körpern mit Charakteristik 0 gilt die NEUMANN**sche Reihe**:

$$(\lambda \mathrm{id} - N)^{-1} = \sum_{k=0}^{\infty} \lambda^{-k-1} N^k.$$

Beispiel:

$$N = \begin{pmatrix} 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 0 \end{pmatrix} \in \mathrm{Mat}(k,k)$$

erfüllt $N^k = 0$, aber $N^{k-1} \neq 0$, denn

$$N^i = \begin{pmatrix} 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \dots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \dots & \vdots & \dots & 0 & 1 \\ \underbrace{0 \quad \dots \quad 0}_{i} & \dots & 0 & 0 \end{pmatrix}$$

Satz 1: Sei N ein nilpotenter Endomorphismus eines endlichdimensionalen Vektorraumes V. Dann gibt es $p_1 \geq p_2 \geq \dots \geq p_L \geq 1$ mit $\sum_{i=1}^{L} p_i = \dim(V)$ und eine Basis $(e_{ij})_{1 \leq i \leq L, 1 \leq j \leq p_i}$ mit

$$Ne_{ij} = \begin{cases} e_{i,j+1} & \text{falls } j < p_i \\ 0 & \text{falls } j = p_i \end{cases}.$$

Dabei sind L und $p_1, \dots, p_L$ eindeutig bestimmt.

Bemerkung:

$$N = \begin{pmatrix} \boxed{\begin{matrix} 0 & \dots & \dots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 1 & 0 \end{matrix}} & \dots & 0 & \dots & \dots & 0 \\ \vdots & \dots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \dots & \vdots & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & \dots & \boxed{\begin{matrix} 0 & \dots & \dots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 1 & 0 \end{matrix}} \end{pmatrix}$$

Beweis von Satz 1 durch Induktion nach $\dim(V)$, beginnend mit $\dim(V) = 0$. Sei die Existenz für Vektorräume kleinerer Dimension als V bewiesen. Sei p_1 das kleinste p mit $N^p = 0$. Sei $e_{1,1} \in V$ mit $N^{p_i - 1}(e_{1,1}) \neq 0$. Für $1 \leq j \leq p_1$ sei $e_{1,j} = N^{j-1} e_{1,1}$. *Behauptung:* Die $e_{1,j}$, $1 \leq j \leq p_1$, sind linear unabhängig. Angenommen, $v = \sum_{j=1}^{p_1} \lambda_j e_{1,j} = 0$ mit $\lambda_j \in K$, die nicht alle verschwinden. Sei k die kleinste Zahl mit $\lambda_k \neq 0$. Dann gilt

$$N^{p_1 - k} v \quad = \quad \lambda_k N^{p_1 - k} e_{1,k} + \sum_{j=k+1}^{p_1} \lambda_j N^{p_1 - k} e_{1,j}$$

$$
\begin{aligned}
&= \lambda_k N^{p_1-k} N^{k-1} e_{1,1} + \sum_{j=k+1}^{p_1} \lambda_j \underbrace{N^{p_1-k} N^{j-1} e_{1,1}}_{=N^{p_1+j-k-1} e_{1,1}=0 \text{ ,denn } p_1+j-k-1 \geq p_1} \\
&= \lambda_k e_{1,p_1} \neq 0,
\end{aligned}
$$

im Widerspruch zu unserer Annahme, daß $v = 0$ ist. Also ist die lineare Unabhängigkeit der $e_{1,j}$ bewiesen.
Die $e_{1,j}$ bilden also eine Basis eines Teilraumes X von V. Wir werden einen komplementären N-invarianten Unterraum Y zu X finden. Sei $(e_{1,1}, \ldots, e_{1,p_1}, f_1, \ldots, f_r)$ eine Basis von V. Durch $l(\sum_{j=1}^{p_1} \lambda_j e_{1,j} + \sum_{k=1}^{r} \vartheta_k f_k) = \lambda_{p_1}$ ist ein Funktional $l \in V^*$ mit $l(e_{1,j}) = \delta_{p_1,j}$ definiert. Für $x \in V$ sei

$$
q(x) = \sum_{j=1}^{p_1} l(N^{p_1-j}(x)) e_{1,j}.
$$

Dann ist q ein Projektor auf X, denn Bild $(q) \subseteq X$ und

$$
q(e_{1,k}) = \sum_{j=1}^{p_1} l(\underbrace{N^{p_1-j}(e_{1,k})}_{= \begin{cases} 0 & j < k \\ e_{1,p_1+k-j} & j \geq k \end{cases}}) e_{1,j} = e_{1,k}.
$$

Weiterhin ist q mit N vertauschbar:

$$
\begin{aligned}
q(Nx) &= \sum_{j=1}^{p_1} l(N^{p_1+1-j} x) e_{1,j} \overset{N^{p_1}=0}{=} \sum_{j=2}^{p_1} l(N^{p_1+1-j} x) N(e_{1,j-1}) \\
&= N \left(\sum_{\bar{j}=1}^{p_1-1} l(N^{p_1-\bar{j}} x) e_{1,\bar{j}} \right) = N \left(\sum_{\bar{j}=1}^{p_1} l(N^{p_1-\bar{j}} x) e_{1,\bar{j}} \right)
\end{aligned}
$$

denn $N e_{1,p_1} = 0$. Also $qN = Nq$, und $Y = \ker(q)$ ist ein N-invarianter komplementärer Unterraum zu X. Nach der Induktionsannahme gibt es eine Basis $(e_{i,j})_{i=2,j=1}^{L,p_i}$ von Y, so daß $N e_{i,j} = \begin{cases} e_{i,j+1} & j < p_i \\ 0 & j = p_i \end{cases}$. Dann ist $(e_{i,j})_{i=1,j=1}^{L,p_i}$ eine Basis von X mit der geforderten Eigenschaft.
Beweis der Eindeutigkeit: Sei $(\tilde{e}_{i,j})_{i=1,j=1}^{\bar{L},\bar{p}_1}$ eine andere Basis von V mit der beschriebenen Eigenschaft. Wir können $\tilde{L} \leq L$ voraussetzen. Angenommen, $\tilde{L} < L$ und für $1 \leq i \leq \tilde{L}$ gilt $p_i = \tilde{p}_i$. Dann gilt

$$
\dim(V) = \sum_{i=1}^{\bar{L}} \tilde{p}_i = \sum_{i=1}^{\bar{L}} p_i < \sum_{i=1}^{L} p_i = \dim(V).
$$

Widerspruch. Angenommen, es gäbe ein i mit $1 \leq i \leq \tilde{L}$ und $p_i \neq \tilde{p}_i$ und $p_j = \tilde{p}_j$ für $1 \leq j \leq i$. Falls $\tilde{p}_i < p_i$ gilt, führt folgende Betrachtung zum Widerspruch: Eine Basis des Bildes von N^K ist gegeben durch $\{e_{i,j} \mid 1 \leq i \leq L \wedge K+1 \leq j \leq p_i\}$. Also

$$
\dim(\text{Bild}\,(N^K)) = \sum_{i=1}^{L} \max(p_i - K, 0),
$$

denn $e_{i,j} = N^K e_{i,j-K}$ für $j > K$ oder $N^K e_{i,j} = e_{i,j+K}$ oder 0. Ebenso

$$
\dim(\text{Bild}\,(N^K)) = \sum_{j=1}^{\bar{L}} \max(0, \tilde{p}_i - K).
$$

Für $\tilde{p}_i < p_i$ gilt einmal

$$\begin{aligned}\dim(\text{Bild}\,(N^{\tilde{p}_i})) &= \sum_{k=1}^{\tilde{L}} \max(0, \tilde{p}_k - \tilde{p}_i) \\ &\overset{\tilde{p}_k \le \tilde{p}_i \text{ für } k \ge i}{=} \sum_{k=1}^{i-1} (\tilde{p}_k - \tilde{p}_i) \\ &= \sum_{k=1}^{L} \max(p_k - \tilde{p}_i, 0) = \dim(\text{Bild}\,(N^{\tilde{p}_i}))\end{aligned}$$

Widerspruch. Da die Annahme $\tilde{L} \le L$ bei diesem Argument nicht benutzt wurde, sondern nur die Annahme $1 \le \min(L, \tilde{L})$, kann die Annahme $p_i < \tilde{p}_i$ ebenso (durch Betrachtung von Bild (N^{p_i})) zum Widerspruch geführt werden. □

Definition 2:

$$V_\lambda^-(A) = \{v \in V \mid \exists\, k : (A - \lambda \mathrm{id}\,)^k v = 0\}$$

ist der **verallgemeinerte Eigenraum** oder **Hauptraum** zum Eigenwert λ.

Bemerkung: $V_\lambda^-(A)$ ist A-invariant, $\dim(V_\lambda^-(A)) < \infty$ (insbesondere ist $A|_{V_\lambda^-(A)} - \lambda\mathrm{id}$ nilpotent, falls $\dim(V) < \infty$).

Beweis: Aus $(A - \lambda\mathrm{id}\,)^k v =: (A-\lambda)^k v = 0$ folgt $(A-\lambda)^k Av = A(A-\lambda)^k v = 0$, also ist $V_\lambda^-(A)$ A-invariant. Falls $\dim(V_\lambda^-(A)) =: D < \infty$, sei $e_1, \ldots, e_D$ eine Basis dieses Unterraumes. Für $1 \le i \le D$ gibt es k_i mit $(A-\lambda)^{k_i} e_i = 0$. Dann ist $(A_{V_\lambda^-(A)} - \lambda)^k = 0$ mit $k = \max(k_1, \ldots, k_D)$.

Theorem 7: Sei A ein beliebiger Endomorphismus eines endlichdimensionalen Vektorraumes V über einem algebraisch abgeschlossenen Körper K.

1. Es gilt die direkte Summenzerlegung:

$$V = \bigoplus_{\lambda \in K} V_\lambda^-(A), \tag{1}$$

   die in Wahrheit endlich ist.

2. $V_\lambda^-(A)$ hat eine Basis $(e_{\lambda,i,j})_{i=1,j=1}^{L_\lambda, P_{\lambda,i}}$ mit

$$Ae_{\lambda,i,j} = \begin{cases} \lambda e_{\lambda,i,j} & \text{falls } j = P_{\lambda,i} \\ \lambda e_{\lambda,i,j} + e_{\lambda,i,j+1} & \text{falls } j < P_{\lambda,i} \end{cases}.$$

   L_λ und $P_{\lambda,i}$ sind dabei eindeutig bestimmt.

3. $P_A(T) = \prod_{\lambda \in K} (T-\lambda)^{\dim(V_\lambda^-(A))}$.

Beweis: Da uns die Existenz von Eigenwerten (für von $\{0\}$ verschiedene V) garantiert ist (K ist algebraisch abgeschlossen), besteht das Problem bei (1) darin, zu $V_\lambda^-(A)$ einen Komplementärraum zu finden. (2) folgt aus Satz 1, (3) ist einfach.

Lemma 1: Sei $\dim(V) < \infty$, K, A beliebig. Dann gilt die direkte Summenzerlegung

$$V = V_\lambda^-(A) \oplus \bigcap_{k=0}^{\infty} \text{Bild}\,(A-\lambda)^k.$$

Beweis: Wie wir schon bemerkt haben, gibt es ein k_0 mit $(A-\lambda)^{k_0}|_{V_\lambda^-(A)} = 0$. Dann gilt

$$\ker(A-\lambda)^k = V_\lambda^-(A) \qquad (*)$$

für $k \geq k_0$. Die absteigende Folge natürlicher Zahlen

$$\dim(\text{Bild}\,(A-\lambda)) \geq \ldots \geq \dim(\text{Bild}\,(A-\lambda)^i) \geq \dim(\text{Bild}\,(A-\lambda)^{i+1} \geq \ldots$$

stabilisiert sich (die Anzahl der echten Ungleichheiten ist $\leq \dim(V)$). Also gibt es ein k_1 mit

$$\text{Bild}\,(A-\lambda)^k = \text{Bild}\,(A-\lambda)^{k_1} \qquad (\dagger)$$

für $k \geq k_1$. $x \in \text{Bild}\,(A-\lambda)^k \Leftarrow \exists\, z : x = (A-\lambda)^k z \Leftarrow \exists\, y : x = (A-\lambda)^{k+1} y \Leftarrow x \in \text{Bild}\,(A-\lambda)^{k+1}$. Sei $i = \max(k_0, k_1)$. Dann gilt (α) $\text{Bild}\,(A-\lambda)^i \cap \ker(A-\lambda)^i = \{0\}$ und (β) $V = \text{Bild}\,(A-\lambda)^i + \ker(A-\lambda)^i$. Wegen $(*)$ und $(\dagger)$ genügt es, das zu zeigen. Zu (α): Sei $x \in \ker(A-\lambda)^i \cap \text{Bild}\,(A-\lambda)^i$. Also gibt es ein $y \in V$ mit $x = (A-\lambda)^i y$ und es gilt $(A-\lambda)^{2i} y = 0$. Wegen $(*)$ folgt $y \in \ker(A-\lambda)^i$ und $x = 0$. Zu (β): Sei $x \in V$ vorgegeben. Wegen $(\dagger)$ gilt $(A-\lambda)^i x \in \text{Bild}\,(A-\lambda)^i = \text{Bild}\,(A-\lambda)^{2i}$. Also gibt es $y \in V$ mit $(A-\lambda)^i x = (A-\lambda)^{2i} y$. Dann gilt

$$x = \underbrace{(A-\lambda)^i y}_{\in \text{Bild}(A-\lambda)^i} + \underbrace{(x - (A-\lambda)^i y)}_{\in \ker(A_\lambda)^i}. \qquad (*)$$

□

Beweis von Theorem 7 durch Induktion nach $\dim(V)$. Für $\dim(V) = 0$ ist die Behauptung klar. Sei $\dim(V) > 0$ und die Aussage (1) für Räume kleinerer Dimension als V bewiesen. Weil K algebraisch abgeschlossen ist, ist die Existenz eines Eigenwertes $\vartheta \in K$ von A garantiert. Dann $\{0\} \neq V_\vartheta(A) \subseteq V_\vartheta^-(A)$. Sei $W = \bigcap_{k\geq 0} \text{Bild}\,((A-\vartheta)^k)$. Nach Lemma 1 ist W ein A-invarianter (leicht zu zeigen) Komplementärraum zu $V_\vartheta^-(A)$. Also $\dim(W) < \dim(V)$. Nach Induktionsvoraussetzung gilt $W = \bigoplus_{\lambda\in K} W_\lambda^-(A|_W)$. Behauptung: (α) $W_\vartheta^-(A|_W) = \{0\}$ und (β): Für $\lambda \neq \vartheta$ gilt $W_\lambda^-(A|_W) = V_\lambda^-(A)$. Dann

$$V = V_\vartheta^-(A) \oplus W = V_\vartheta^-(A) \oplus \bigoplus_{\lambda\in K} W_\lambda^-(A|_W) = V_\vartheta^-(A) \oplus \bigoplus_{\lambda\in K-\{\vartheta\}} V_\lambda^-(A) = \bigoplus_{\lambda\in K} V_\lambda^-(A).$$

Daraus folgt (1). Zu (α): $W_\vartheta^-(A) \subseteq W \cap V_\vartheta^-(A) = \{0\}$. Zu (β): Offenbar gilt $W_\lambda^-(A) \subseteq V_\lambda^-(A)$. Für $\lambda \neq \vartheta$ ist nun $A|_{V_\lambda^-(A)} - \vartheta = (A|_{V_\lambda^-(A)} - \lambda) - (\lambda - \vartheta)$ invertierbar ($A|_{V_\lambda^-(A)} - \lambda$ ist nilpotent, also Spektrum $= 0$.), also $V_\lambda^-(A) = (A-\vartheta)^k V_\lambda^-(A) \subseteq \text{Bild}\,(A-\vartheta)^k$ für alle k, also $V_\lambda^-(A) \subseteq W$ für $\lambda \neq \vartheta$, und $V_\lambda^-(A) \subseteq W_\lambda^-(A) = W \cap V_\lambda^-(A)$.
(2) folgt durch Anwendung von Satz 1 auf den nilpotenten Endomorphismus $(A|_{V_\lambda^-(A)} - \lambda\text{id}\,)$. Sei weiterhin

$$J_{\lambda,k} = \begin{pmatrix} \lambda & 0 & \ldots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 1 & \lambda \end{pmatrix} \in \text{Mat}\,(k,k,K).$$

Seien $\lambda_1, \ldots, \lambda_N$ die Eigenwerte von A, $L_k = L_{\lambda_k, p_{k,i}} = p_{\lambda_k,i}$. Sei $(e_{\lambda,i,j})$ eine Basis wie in (2) des Theoremes und sei $(f_1, \ldots, f_M)$ die Basis $(e_{\lambda_1,1,1}, e_{\lambda,1,2}, \ldots, e_{\lambda_1,1,p_{1,1}}, e_{\lambda_1,2,1}, \ldots, e_{\lambda_1,2,p_{1,2}}, \ldots, e_{\lambda_1,L_1,1}, \ldots, e_{\lambda_N,$
Dann gilt

$$\text{Mat}_B(A) = \text{Diag}(J_{\lambda_k,p_{k,i}})_{k=1,i=1}^{N,L_k};$$

$\det(\vartheta\text{id} - A) = \det(\vartheta\mathbb{1}_M - \text{Mat}_B(A)) = \det(\text{Diag}(\vartheta\mathbb{1}_{p_{k,i}} - J_{\lambda_k p_{k,i}}))_{k=1,i=1}^{N,L_k}$. Die letzte Matrix ist aber eine untere Dreiecksmatrix, ihre Determinante ist das Produkt ihrer Diagonalelemente: Die

Diagonaleinträge dieser Matrix sind $\vartheta - \lambda_k$ von der Vielfachheit $p_{k,1} + \ldots + p_{k,L_k} = \dim(V_{\lambda_k}^{\cdot}(A))$. Damit folgt die Behauptung. □

Folgerung 1: Unter den Voraussetzungen von Theorem 7 ist A genau dann diagonalisierbar, falls für alle $\lambda \in K$ $V_\lambda(A)$ und $V_\lambda^{\cdot}(A)$ übereinstimmen. Dies ist genau dann der Fall, wenn für alle Eigenwerte λ $\dim(V_\lambda(A))$ mit der Vielfachheit der Nullstelle des charakteristischen Polynoms übereinstimmt, denn sie ist gleich $\dim(V_\lambda^{\cdot}(A))$.

Beweis: Die letzte Aussage ist wegen $V_\lambda(A) \subseteq V_\lambda^{\cdot}(A)$ klar. Sei A diagonalisierbar. Dann $V = \sum_{\lambda \in K} V_\lambda(A)$. Wegen $V = \bigoplus_{\lambda \in K} V_\lambda^{\cdot}(A)$ kann dies nur für $V_\lambda(A) = V_\lambda^{\cdot}(A)$ gelten. Sei $V_\lambda(A) = V_\lambda^{\cdot}(A)$. Dann $V = \bigoplus_{\lambda \in K} V_\lambda(A)$, und V hat eine Basis aus Eigenvektoren. □

Bemerkung 2: Theorem 7 und Folgerung 1 gelten auch für nicht algebraisch abgeschlossene Körper, wenn $P_A(T)$ über K in Linearfaktoren zerfällt.

Bemerkung 3: Eine andere Formulierung für den Satz über die Jordansche Normalform ist: Jeder Endomorphismus eines endlichdimensionalen Vektorraumes V über einem algebraisch abgeschlossenen Körper K kann auf eindeutige Weise als $A = D + N$ geschrieben werden, wobei D diagonalisierbar und N nilpotent ist und N und D kommutieren.

Beweis: Existenz der Zerlegung: Sei D durch $Dx = \lambda x$ für $x \in V_\lambda^{\cdot}(A)$ und N durch $A - D$ definiert. Eindeutigkeit: Übungsaufgabe. Zeige $V_\lambda(D) = V_\lambda^{\cdot}(D) \stackrel{!}{=} V_\lambda^{\cdot}(A)$.

Bemerkung 4: Wenn N ein nilpotenter Endomorphismus von V ist, so gilt $N^{\dim(V)} = 0$.

Beweis: folgt aus Satz 1.

Bemerkung 5: N ist nilpotent genau dann, wenn $P_A(T) = T^{\dim(V)}$. Für $\dim(V) = 2$ ist dies genau dann der Fall, wenn $\mathrm{tr}(N) = \det(N) = 0$ gilt.

Beweis: folgt aus Theorem 7.

Bemerkung 6: Wenn $P_A(T)$ $\dim(V)$ verschiedene Nullstellen hat, so ist A diagonalisierbar.

Beweis: Es gilt $\dim(V_\lambda^{\cdot}(A))$ =Vielfachheit der Nullstelle $\lambda = 1$, falls λ ein Eigenwert ist, denn $P_A(T)$ kann keine mehrfache Nullstelle haben. Weil für einen Eigenwert λ $V_\lambda(A) \neq \{0\}$ gilt, folgt $V_\lambda(A) = V_\lambda^{\cdot}(A)$. Mit Anwendung von Bemerkung 2 folgt die Behauptung. □

Theorem 8 (Cayley-Hamilton): Jeder Endomorphismus A eines endlichdimensionalen Vektorraumes über einem beliebigen Körper erfüllt seine eigene charakteristische Gleichung

$$P_A(A) := \sum_{k=0}^{\dim V} p_{A,k} A^k = 0,$$

wobei $P_A(T) = \sum_{k=0}^{\dim V} p_{A,k} T^k$.

Beweis: (Nur für K algebraisch abgeschlossen.) Es genügt nach Theorem 7.1, $P_A(A)|_{V_\lambda^{\cdot}(A)} = 0$ zu zeigen für jeden Eigenwert λ. Es gilt $P_A(T) = (T - \lambda)^{\dim(V_\lambda^{\cdot}(A))} Q(T)$ nach Theorem 7.3. Der Endomorphismus $N = A|_{V_\lambda^{\cdot}(A)} - \lambda \mathrm{id}_{V_\lambda^{\cdot}(A)}$ von $V_\lambda^{\cdot}(A)$ ist nilpotent. Also

$$P_A(A)|_{V_\lambda^{\cdot}(A)} \stackrel{6.3.1}{=} (A|_{V_\lambda^{\cdot}(A)} - \lambda)^{\dim(V_\lambda^{\cdot}(A))} Q(A|_{V_\lambda^{\cdot}(A)}) = N^{\dim(V_\lambda^{\cdot}(A))} Q(A|_{V_\lambda^{\cdot}(A)}) = 0$$

nach Bemerkung 5. □

7 Affine Räume und Faktorräume

7.1 Affine Räume

Definition 1: Ein **affiner Raum** ist ein Tripel $\mathcal{X} = (X, TX, \delta)$, wobei X eine Menge, die Menge der Punkte des affinen Raumes, TX ein K-Vektorraum, der **Translationsvektorraum** von $\mathcal{X}$ und $\delta : X \times X \to TX$ eine Abbildung, die $x, y \in X$ den Vektor zuordnet, der von x nach y zeigt, sind. Folgende Bedingungen müssen gelten:

1. $\delta(x, y) + \delta(y, z) = \delta(x, z)$
2. Für jedes $x \in X$ ist die Abbildung $X \to TX$, $y \mapsto \delta(x, y)$ bijektiv.

Bemerkung: Wenn (2) für x gilt, so ist für alle $x' \in X$ die Abbildung $y \mapsto \delta(x', y) = \delta(x', x) + \delta(x, y)$ bijektiv. Daher reicht es auch, (2) für nur ein $x \in X$ zu fordern. Aus (1) folgt $\delta(x, x) + \delta(x, x) = \delta(x, x)$, also $\delta(x, x) = 0$. Außerdem $\delta(x, y) + \delta(y, x) = \delta(x, x) = 0$, also $\delta(x, y) = -\delta(y, x)$.

Bemerkung 1: Für $x \in X$ und $v \in TX$ sei $x + v \in X$ das durch die Bedingung $\delta(x, x + v) = v$ eindeutig bestimmte Element. Dann $x + 0 = x$, denn $\delta(x, x) = 0$ und $(x + v) + w = x + (v + w)$, denn $\delta(x, (x + v) + w) = \delta(x, x + v) + \delta(x + v, (x + v) + w) = v + w$. Für alle $x \in X$ ist die Abbildung $V \to X$, $v \mapsto x + v$ bijektiv.

Beispiel 1:

1. Für jeden Vektorraum V ist

   $$\mathcal{V} = (V, V, \delta_{\mathcal{V}}), \ \delta_{\mathcal{V}}(x, y) = y - x$$

   ein affiner Raum. $+$ in $\mathcal{V}$ (Bemerkung 1) ist das $+$ in V.

2. Sei V ein K-Vektorraum, $W \subseteq V$ ein Unterraum und $a \in V$.

   $$W + a = \{w + a \mid w \in W\}.$$

   Dann ist

   $$(W + a, W, \delta_{\mathcal{V}}|_{(W+a)\times(W+a)})$$

   ein affiner Raum.

Definition 2:

1. Ein **Morphismus** affiner Räume von $\mathcal{X} = (X, TX, \delta_{\mathcal{X}})$ nach $\mathcal{Y} = (Y, TY, \delta_{\mathcal{Y}})$ ist ein Paar (f, g), wobei $f : X \to Y$ eine Abbildung und $g : TX \to TY$ eine lineare Abbildung ist und gilt
$$\delta_{\mathcal{Y}}(f(x), f(y)) = g(\delta_{\mathcal{X}}(x, y)).$$

2. Ein affiner Unterraum von (X, TX, δ) ist ein affiner Raum $(Y, TY, \delta_{Y\times Y})$, wobei $Y \subseteq X$ eine Teilmenge und $TY \subseteq TX$ ein Unterraum ist. Dann definieren die Einbettungsabbildungen $Y \stackrel{f}{\to} X$ und $TY \stackrel{g}{\to} TX$ einen Morphismus affiner Räume.

3. Die Verknüpfung zweier Morphismen affiner Räume
$$(X, TX, \delta_{\mathcal{X}}) \stackrel{(f,g)}{\to} (Y, TY, \delta_{\mathcal{Y}}) \stackrel{(a,b)}{\to} (Z, TZ, \delta_{\mathcal{Z}})$$
ist durch das Paar (af, bg) gegeben. Der identische Morphismus $\mathcal{X} \stackrel{\mathrm{Id}}{\to} \mathcal{X}$ ist durch $(\mathrm{id}_X, \mathrm{id}_{TX})$ gegeben. Ein **Isomorphismus** affiner Räume ist ein Morphismus $F : \mathcal{X} \to \mathcal{Y}$, für den ein $F^{-1} : \mathcal{Y} \to \mathcal{X}$ mit $FF^{-1} = \mathrm{id}_{\mathcal{Y}}$ und $F^{-1}F = \mathrm{id}_{\mathcal{X}}$ existiert. Dann ist F^{-1} eindeutig bestimmt.

Bemerkung:

1. Für jeden affinen Raum $\mathcal{X} = (X, V, \delta_{\mathcal{X}})$ und für jedes $x \in X$ ist durch $\mathcal{X} \stackrel{(f,g)}{\to} \mathcal{V}$ ($\mathcal{V}$ aus Beispiel 1), $f(y) = \delta(x, y)$, $g = \mathrm{id}_V$, ein Isomorphismus affiner Räume gegeben.

2. Eine **Kategorie** $\mathcal{A}$ besteht aus

   (a) einer Menge $\mathrm{Ob}\,\mathcal{A}$ der Objekte von $\mathcal{A}$ (z.B. Menge aller Gruppen, K-Vektorräume, K-affine Räume);

   (b) für Objekte $X, Y \in \mathrm{Ob}\,\mathcal{A}$ ist eine Menge $\mathrm{Hom}_{\mathcal{A}}(X, Y)$ der Morphismen in $\mathcal{A}$ von X nach Y definiert (z.B. Gruppenhomomorphismen, K-lineare Abbildungen, Morphismen affiner Räume von X nach Y);

   (c) für alle $X \in \mathrm{Ob}\,\mathcal{A}$ ist $\mathrm{id}_X \in \mathrm{Hom}_{\mathcal{A}}(X, Y)$ gegeben;

   (d) für $X, Y, Z \in \mathrm{Ob}\,\mathcal{A}$ ist die Verknüpfungsabbildung $\mathrm{Hom}_{\mathcal{A}}(X, Y) \times \mathrm{Hom}_{\mathcal{A}}(Y, Z) \to \mathrm{Hom}_{\mathcal{A}}(X, Z)$, $(f, g) \mapsto gf$ gegeben.

   Es müssen dann die Regeln $\mathrm{id}_D f = f = f\mathrm{id}_C$ und $f(gh) = (fg)h$ für $A \stackrel{h}{\to} B \stackrel{g}{\to} C \stackrel{f}{\to} D$ gelten.

   Dann haben wir die Kategorien aller Gruppen, K-Vektorräume oder K-affinen Räume kennengelernt.

3. In Definition 2.1 ist die lineare Abbildung $TX \stackrel{f}{\to} TY$ durch die Abbildung $f : X \to Y$ eindeutig bestimmt.

Satz 1: Sei $\mathcal{V} = (V, V, \delta_{\mathcal{V}})$ wie in Beispiel 1 definiert. Dann sind alle affinen Unterräume von $\mathcal{V}$ von der Form $(W + a, W, \delta_{\mathcal{V}}|_{(W+a)\times(W+a)})$, wobei $W \subseteq V$ ein Untervektorraum und $a \in V$ ist.

Beweis: Nach Definition 2.2 ist ein affiner Unterraum von $\mathcal{V}$ ein Tripel $(X, W, \delta_{\mathcal{V}}|_{X\times X})$, wobei $W \subseteq V$ ein Untervektorraum ist und für alle $a, x \in X$ $\delta_{\mathcal{V}}(a, x) = x - a \in W$ und für alle $a \in X$

die Abbildung $X \to W$, $x \mapsto x - a$ bijektiv ist. Sei $a \in X$. Dann gilt $X = W + a$. □

Bemerkung: Eine Untermenge X von V nennt man einen **affinen Unterraum** von V, wenn ein Untervektorraum W von V existiert, so daß $(X, W, \delta_{\mathcal{V}}|_{X\times X})$ ein affiner Unterraum von $\mathcal{V}$ ist. Wenn dies der Fall ist, so ist $W = \{x - y \mid x, y \in X\}$ eindeutig bestimmt.

Definition 3: Sei $\mathcal{X} = (X, TX, \delta)$ ein affiner Raum. Seien $x_1, \ldots, x_n \in X$ und seien $\lambda_1, \ldots, \lambda_n \in K$ und $\sum_{i=1}^n \lambda_i = 1$. Dann sei $\sum_{i=1}^n \lambda_i x_i \in X$ das durch

$$\forall\, y \in X: \qquad \delta(y, \sum_{i=1}^n \lambda_i x_i) = \sum_{i=1}^n \lambda_i \delta(y, x_i)$$

eindeutig bestimmte Element.

Bemerkung:

1. Sei y fest. Dann gibt es nach Definition 1 genau ein p mit $\delta(y,p) = \sum_{i=1}^n \lambda_i \delta(y, x_i)$. Für $z \in X$ gilt dann $\delta(z,p) = \delta(z,y) + \delta(y,p) = \sum_{i=1}^n \lambda_i \delta(z,y) + \sum_{i=1}^n \lambda_i \delta(y, x_i) = \sum_{i=1}^n \lambda_i(\delta(z,y) + \delta(y,x_i)) = \sum_{i=1}^n \lambda_i \delta(z, x_i)$. Also gilt für $\sum_{i=1}^n \lambda_i x_i = p$ die Bedingung in der Definition für alle Punkte $y \in X$.

2. Für $1 \le i \le n$, $1 \le j \le m_i$ seien $\lambda_i \in K$, $\vartheta_{ij} \in K$ und $x_{ij} \in X$ mit $\sum_{i=1}^n \lambda_i = 1$ und $\sum_{j=1}^{m_i} \vartheta_{ij} = 1$. Dann
$$\sum_{i=1}^n \lambda_i (\sum_{j=1}^{m_i} \vartheta_{ij} x_{ij}) = \sum_{i=1}^n \sum_{j=1}^{m_i} (\lambda_i \vartheta_{ij}) x_{ij}.$$

3. Man kann für $x, y \in X$ $\{\lambda x + (1-\lambda)y \mid \lambda \in K\} \subseteq X$ für $x \neq y$ als die Gerade durch x und y auffassen.

4. $\lambda x + (1-\lambda)x = x$.

Satz 2: Sei $\operatorname{char} K \neq 2$. Eine nichtleere Teilmenge $X \subseteq V$ ist genau dann ein affiner Unterraum eines K-Vektorraumes V, wenn für alle $x, y \in X$ und $\lambda \in K$ auch $\lambda x + (1 - \lambda y) \in X$ gilt.

Beweis: Die Notwendigkeit dieser Bedingung ist klar. Zum Beweis der Hinlänglichkeit sei $a \in X$. Wir können X durch $X - a$ ersetzen und erreichen $0 \in X$. Dann ist X ein Untervektorraum von V: für $x \in X$ und $\lambda \in K$ gilt $\lambda x = \lambda x + (1-\lambda)0 \in X$. Für $x, y \in X$ gilt $\frac{x+y}{2} = \frac{1}{2}x + \frac{1}{2}x \in X$, also $x + y = 2(\frac{x+y}{2}) \in X$. □

Bemerkung:

1. Der Satz gilt auch für $\operatorname{char} K = 2$, dann muß man die Abgeschlossenheit von X unter $\sum_{i=1}^3 \lambda_i x_i$, $\sum_{i=1}^3 \lambda_i = 1$ fordern.

2. Ein analoger Sachverhalt gilt auch für Untermengen beliebiger affiner Räume (die zu einem affinen Raum der Form $\mathcal{V}$ aus Beispiel 1.1 isomorph sind).

Bemerkung: Sei $K = \mathbb{R}$. Dann kann man $\{\lambda x + (1-\lambda)y \mid 0 \le \lambda \le 1\}$ als die **Strecke** zwischen x und y auffassen.

Satz 3: Sei $\mathcal{X} = (X, TX, \delta)$ ein affiner Raum und $K = \mathbb{R}$. Dann sind folgende Forderungen an eine Untermenge $M \subseteq X$ äquivalent:

1. Für $x, y \in M$ und $0 \leq \lambda \leq 1$ gilt $\lambda x + (1-\lambda)y \in M$.
2. Für $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit $\sum_{i=1}^n \lambda_i = 1$, $\lambda_i \geq 0$, und $x_1, \ldots, x_i \in M$ gilt $\sum_{i=1}^n \lambda_i x_i \in M$.

Beweis: (2)⇒(1) ist klar. (1)⇒(2) folgt mit Induktion nach n: der Fall $n = 2$ ist klar. Falls $\lambda_n = 1$ folgt $\lambda_1 = \ldots = \lambda_{n-1} = 0$ und $\sum \lambda_i x_i \in M$. Für $\lambda_n \neq 1$ gilt

$$\sum_{i=1}^{n} \lambda_i x_i = (1-\lambda_n)(\sum_{i=1}^{n-1} \frac{\lambda_i}{1-\lambda_n} x_i) + \lambda_n x_n \in M$$

nach (1). □

7.2 Der Faktorraum V/W nach einem Unterraum

Definition 1: Elemente von V/W sind Untermengen von V, die die Form $W + a$ haben, wobei $a \in V$. Die Addition und die Multiplikation sind wie folgt definiert:

- $(W + a) + (W + b) = W + (a + b)$
- $\lambda(W + a) = W + \lambda a$

für $a, b \in V$ und $\lambda \in K$. Man nennt V/W den **Faktor-** oder **Quotientenraum** von V nach W.

Bemerkung: Elemente von V/W sind gerade die affinen Unterräume von V mit Translationsraum W.

Bemerkung 1: V/W ist der Quotient von V nach der Äquivalenzrelation $a \sim b \Leftrightarrow a - b \in W$. Die Äquivalenzklassen sind die Untermengen der Form $W + a$.

Satz 1: Sei $p : V \to V/W$ die durch $p(a) = a + W$ definierte Abbildung. Auf V/W ist die in Definition 1 beschriebene Vektorraumstruktur die einzige, für die p eine lineare Abbildung ist.

Beweis: Offenbar ist p linear. Wegen der Surjektivität von p ist die Vektorraumstruktur auf V/W durch die Vektorraumstruktur auf V eindeutig bestimmt, wenn p linear ist: $(W + a) + (W + b) = p(a) + p(b) = p(a + b) = W + (a + b)$. □

Theorem 1: (Homomorphiesatz) Seien A, V K-Vektorräume. Sei W ein Untervektorraum von V. Dann ist

$$\begin{array}{rcl} L(V/W, A) & \to & \{f \in L(V, A) \mid \ker(f) \supseteq W\} \\ g & \mapsto & gp \end{array}$$

eine Bijektion (also ein Isomorphismus von Vektorräumen).

Beweis: Wegen $\ker(p) = W$ gilt $\ker(gp) \supseteq \ker(p) = W$, daher ist die Abbildung (1) wohldefiniert. Wegen der Surjektivität von p folgt aus $gp = 0$, daß $g = 0$, also ist die lineare Abbildung injektiv. Sei f ein Element der rechten Seite. Definiere eine lineare Abbildung $g : V/W \to A$ durch $g(W + a) = f(a)$. Dies ist eine korrekte Definition, denn aus $W + a = W + \tilde{a}$ folgt $a - \tilde{a} \in W$

und $f(a) = f(\tilde{a}) + f(a - \tilde{a}) = f(\tilde{a})$. Man prüft nun leicht nach, daß g linear ist und daß $gp = f$ gilt. Also ist die Abbildung auch surjektiv. □

Bemerkung: $\ker(p) = W$. Wir bezeichnen p als $p_{V,W}$, falls weitere Faktorräume auftreten.

Theorem 2: (Isomorphiesätze) Sei X ein Vektorraum, $Y, Z \subseteq X$ Unterräume.

1. Wenn $Z \subseteq Y$, so gibt es genau einen Isomorphismus $\alpha : X/Y \to (X/Z)/(Y/Z)$ mit $p_{X/Z,Y/Z} p_{X,Z} = \alpha p_{X,Y}$.

$$\begin{array}{rcl} X & \stackrel{p_{X,Z}}{\longrightarrow} & X/Z \\ {\scriptstyle p_{X,Y}}\downarrow & & \downarrow{\scriptstyle p_{X/Z,Y/Z}} \\ X/Y & \stackrel{\alpha}{\longrightarrow} & (X/Z)/(Y/Z) \end{array}$$

2. Es gibt genau einen Isomorphismus $Y/Y \cap Z \stackrel{\beta}{\to} Y + Z/Z$ mit $\beta p_{Y,Y\cap Z} = p_{Y+Z,Z} i$, wobei $i : Y \to Y + Z$ die Inklusion ist.

Beweis:

1. Offenbar gilt $Y/Z = \{y + Z \mid y \in Y\} \subseteq \{x + Z \mid x \in X\} = X/Z$. Also ist Y/Z ein Unterraum von X/Z. Sei $y \in Y$, dann gilt:

$$p_{X/Z,Y/Z}(p_{X,Z}(y)) = p_{X/Z,Y/Z} \underbrace{(y + Z)}_{\in Y/Z} = 0.$$

   Nach Theorem 1 gibt es genau eine lineare Abbildung α, so daß das Quadrat kommutiert. Diese Abbildung ist surjektiv, denn für $a \in (X/Z)/(Y/Z)$ gibt es $b \in X/Z$ mit $p_{X/Z,Y/Z}(b) = a$ und $c \in X$ mit $p_{X,Z}(c) = b$. Dann $\alpha(p_{X,Y}(c)) = p_{X/Z,Y/Z}(p_{X,Z}(c)) = a$. Sei andererseits $a \in \ker(\alpha)$. Dann gibt es $b \in X$ mit $a = p_{X,Y}(b)$. Dann $p_{X/Z,Y/Z}(p_{X,Z}(b)) = \alpha(p_{X,Y}(b)) = \alpha(a) = 0$. Also $p_{X,Z}(b) \in \ker(p_{X/Z,Y/Z}) = Y/Z$. Also $Z + b \subseteq Y$, also $b \in Y$, also gilt $a = p_{X,Y}(b) = 0$.

2. Der Beweis verläuft analog zu (1): Für $z \in Y \cap Z$ gilt $p_{Y+Z,Z}(i(z)) = p_{Y+Z,Z}(z) = 0$ wegen $z \in Z$. Nach Theorem 1 gibt es genau eine lineare Abbildung $\beta : Y/Y \cap Z \to Y + Z/Z$ mit $\beta p_{Y,Y\cap Z} = p_{Y+Z,Z} i$. β ist surjektiv, denn sei $a \in Y + Z/Z$, dann gibt es $b \in Y + Z$ mit $p_{Y+Z,Z}(b) = a$, und es gibt $y \in Y$, $z \in Z$ mit $b = y + z$. Dann $a = p_{Y+Z,Z}(y + z) = p_{Y+Z,Z}(y) = \beta(p_{Y,Y\cap Z}(y))$. Also ist β surjektiv. Sei andererseits $c \in \ker(\beta)$. Dann gibt es $d \in Y$ mit $p_{Y,Y\cap Z}(b) = c$. Dann gilt $0 = \beta(p_{Y,Y\cap Z}(d)) = p_{Y+Z,Z}(d)$. Also $d \in Z$, also $d \in Y \cap Z$, damit $c = 0$ und β ist injektiv. □

Satz 2: Sei $W \subseteq V$ ein Unterraum und sei $X \subseteq V$ ein Komplementärraum zu W. Dann ist $X \stackrel{p_{V,W}}{\to} V/W$ ein Isomorphismus.

Beweis: Der Kern der Abbildung ist gerade $X \cap \ker(p_{V,W}) = X \cap W = \{0\}$. Also ist die Abbildung injektiv. Sei $a \in V/W$ gegeben. Dann gibt es $b \in V$ mit $p_{V,W}(b) = a$, und $x \in X$, $w \in W$ mit $b = x + w$. Dann $a = p_{V,W}(b) = p_{V,W}(x) + p_{V,W}(w) = p_{V,W}(x)$. Damit ist die Abbildung auch surjektiv. □

Satz 3: Sei $W \subseteq V$ ein Unterraum, $\dim(V) < \infty$. Sei f ein Endomorphismus von V mit $f(W) \subseteq W$. Dann gibt es genau einen Endomorphismus $\hat{f}$ von V/W mit $p_{V,W} f = \hat{f} p_{V,W}$, und es gilt

$$\begin{aligned} \det(f) &= \det(f|_W)\det(\hat{f}) \\ \mathrm{tr}\,(f) &= tr(f|_W) + \mathrm{tr}\,(\hat{f}) \\ P_f(T) &= P_{f|_W}(T) P_{\hat{f}} \end{aligned}$$

Beweis: Für $w \in W$ gilt $p_{V,W}(f(w)) = 0$. Nach Theorem 1 erhalten wir die Existenz und Eindeutigkeit von $\hat{f}$. Sei $B' = (b_1, \ldots, b_n)$ eine Basis von W. Ergänze zu einer Basis $B = (b_1, \ldots, b_m)$ von V. Sei $\hat{b}_i = p_{V,W}(b_i)$. Dann ist $\hat{B} = (\hat{b}_{n+1}, \ldots, \hat{b}_m)$ eine Basis des Faktorraumes nach Satz 2, denn $b_{n+1}, \ldots, b_m$ ist eine Basis des Komplementärraums. Es gilt

$$\mathrm{Mat}_B(f) = \begin{pmatrix} \mathrm{Mat}_{B'}(f|_W) & \mathcal{X} \\ 0 & \mathrm{Mat}_{\hat{B}}(\hat{f}) \end{pmatrix}.$$

Die zweite Aussage (über die Spuren) ist klar. Die Aussage über die Determinanten folgt aus folgendem

Lemma: Für L, N quadratisch gilt $\det \begin{pmatrix} L & 0 \\ M & N \end{pmatrix} = \det(L)\det(N)$.

Die Aussage über die charakterischen Polynome folgt (für unendliche Körper) aus der Aussage über die Determinanten. □

Beweis des Lemmas: Induktion nach a, der Größe von L. Falls $a = 0$ ist die Behauptung klar, sonst Auflösung nach der ersten Zeile und Induktionsvoraussetzung. □

8 Tensorprodukte

Definition 1: Ein **Tensorprodukt** zweier K-Vektorräume X, Y ist ein Paar $(X \otimes Y, \otimes)$ bestehend aus einem K-Vektorraum $X \otimes Y$ und einer bilinearen Abbildung

$$\otimes : X \times Y \to X \otimes Y,\ (x, y) \mapsto x \otimes y$$

mit folgender Eigenschaft: wenn Z ein K-Vektorraum und $B : X \times Y \to Z$ eine bilineare Abbildung ist, so gibt es genau eine lineare Abbildung $f_B : X \otimes Y \to Z$, so daß $B(x, y) = f_B(x \otimes y)$ für alle $x \in X, y \in Y$.

Bemerkung 1: Das Tensorprodukt ist eindeutig bis auf eindeutigen Isomorphismus bestimmt, falls es existiert.
Seien dazu $(X \otimes Y, \otimes)$ und $(X \tilde{\otimes} Y, \tilde{\otimes})$ zwei Tensorprodukte von X und Y. Wenn wir die Bedingung aus der Definition für $X \otimes Y$ auf die bilineare Abbildung $\tilde{\otimes}$ anwenden, so bekommen wir eine eindeutige lineare Abbildung $f_{\tilde{\otimes}} : X \otimes Y \to X \tilde{\otimes} Y$ mit $f_{\tilde{\otimes}}(x \otimes y) = x \tilde{\otimes} y$. Dasselbe Argument mit Vertauschung der Rollen von $\otimes$ und $\tilde{\otimes}$ ergibt ein eindeutig bestimmtes $f_{\otimes} : X \tilde{\otimes} Y \to X \otimes Y$ mit $f_{\otimes}(x \tilde{\otimes} y) = x \otimes y$. Dann sind $f_{\otimes}$ und $f_{\tilde{\otimes}}$ zueinander inverse Isomorphismen. In der Tat, $g = f_{\otimes} f_{\tilde{\otimes}}$ erfüllt $g(x \otimes y) = f_{\otimes}(f_{\tilde{\otimes}}(x \otimes y)) = f_{\otimes}(x \tilde{\otimes} y) = x \otimes y$. Dieselbe Eigenschaft hat aber auch $\mathrm{id}_{X \otimes Y}$. Anwendung der Eindeutigkeitsaussage aus der Definition auf $Z = X \otimes Y$ und $B = \otimes$ ergibt $g = \mathrm{id}_{X \otimes Y}$. Ebenso zeigt man $f_{\tilde{\otimes}} f_{\otimes} = \mathrm{id}_{X \tilde{\otimes} Y}$. Also sind $f_{\otimes}$ und $f_{\tilde{\otimes}}$ zueinander inverse Isomorphismen.

Theorem 1:

1. $X \otimes Y$ existiert.
2. Wenn $(e_i)_{i=1}^m$ und $(f_j)_{j=1}^n$ Basen von X und Y sind, so ist $(e_i \otimes f_j)_{i,j=1}^{m,n}$ eine Basis von $X \otimes Y$. Insbesondere ist $X \otimes Y$ endlichdimensional, wenn die beiden Faktoren endlichdimensional sind, und es gilt $\dim(X \otimes Y) = \dim(X)\dim(Y)$.

Beweis: Zu (1): Sei V die Menge aller Abbildungen $f : X \times Y \to K$ mit endlichem Träger, $\mathrm{supp}(f) = \{(x,y) \mid f(x,y) \neq 0\}$ endlich. Sei $\delta_{(x,y)}(x',y') = \begin{cases} 1 & (x,y) = (x',y') \\ 0 & (x,y) \neq (x',y') \end{cases}$, $\delta_{(x,y)} \in V$. Wenn Z ein K-Vektorraum und $G : X \times Y \to Z$ eine Abbildung ist, so gibt es genau eine lineare Abbildung $\varphi_G : V \to Z$ mit $G(x,y) = \varphi_G(\delta_{(x,y)})$. In der Tat,

$$\varphi_G(f) = \sum_{(x,y) \in X \times Y} f(x,y)G(x,y)$$

(die Summe ist endlich, denn f hat endlichen Träger) leistet das Gewünschte, und wenn $\tilde{\varphi}_G(\delta_{(x,y)}) = G(x,y)$, $\tilde{\varphi}_G$ linear, so gilt

$$\begin{aligned} \tilde{\varphi}_G(f) &= \tilde{\varphi}_G\Big(\sum_{(x,y) \in X \times Y} f(x,y)\delta_{(x,y)}\Big) \\ &= \sum_{(x,y) \in X \times Y} f(x,y)\tilde{\varphi}_G(\delta_{(x,y)}) = \sum_{(x,y) \in X \times Y} f(x,y)G(x,y) = \varphi_G(f). \end{aligned}$$

Sei $W \subseteq V$ der Unterraum, welcher von allen Vektoren, die die Abweichung der Abbildung $X \times Y \to V$, $(x,y) \mapsto \delta_{(x,y)}$ von der Bilinearität messen, erzeugt wird: W wird erzeugt von $(*)$:

- $\delta_{(\lambda x,y)} - \lambda\delta_{(x,y)}$
- $\delta_{(x,\lambda y)} - \lambda\delta_{(x,y)}$
- $\delta_{(x+x',y)} - \delta_{(x,y)} - \delta_{(x',y)}$
- $\delta_{(x,y+y')} - \delta_{(x,y)} - \delta_{(x,y')}$

Sei $X \otimes Y = V/W$, und sei $p : V \to X \otimes Y$ die Projektion auf den Faktorraum, und sei $x \otimes y = p(\delta_{(x,y)})$. Wegen $(*)$ ist $\otimes$ bilinar. Sei $B : X \times Y \to Z$ bilinear und sei $\varphi_B : V \to Z$ die soeben beschriebene Abbildung ($\varphi_B(\delta_{(x,y)}) = B(x,y)$). Weil B bilinear ist, annuliert φ_B alle Vektoren aus $(*)$, z.B. $\varphi_B(\delta_{(\lambda x,y)} - \lambda\delta_{(x,y)}) = \varphi_B(\delta_{(\lambda x,y)}) - \lambda\varphi_B(\delta_{(x,y)}) = B(\lambda x,y) - \lambda B(x,y) = 0$. Also $\varphi_B|_W = 0$. Nach Theorem 7.1 über Faktorräume gibt es genau eine lineare Abbildung $f_B : X \otimes Y \to Z$ mit $f_B p = \varphi_B$. Dann $f_B(x \otimes y) = f_B(p(\delta_{(x,y)})) = \varphi_B(\delta_{(x,y)}) = B(x,y)$. Wenn $\tilde{f}_B : X \otimes Y \to Z$ linear ist mit $\tilde{f}_B(x \otimes y) = B(x,y)$, so $\tilde{f}_B p(\delta_{(x,y)}) = B(x,y)$, also ist $\tilde{f}_B p = \varphi_B$ nach der schon bewiesenen Eindeutigkeit von φ_B, also $\tilde{f}_B p = \varphi_B = f_B p$. Wegen der Surjektivität von p folgt $f_B = \tilde{f}_B$.
Zu (2): Wir behaupten, daß für beliebige X, Y $X \otimes Y$ durch $\{x \otimes y \mid x \in X, y \in Y\}$ erzeugt wird. Für den in (1) konstruierten Raum ist das klar, und für beliebige Tensorprodukte von X und Y folgt dann derselbe Sachverhalt nach Bemerkung 1. Sei $W \subseteq X \otimes Y$ die lineare Hülle der Vektoren $e_i \otimes f_j \in X \otimes Y$, $1 \leq i \leq n$, $1 \leq j \leq n$. Für $x \in X$, $y \in Y$, gilt $x = \sum_{i=1}^m \lambda_i e_i$ und $y = \sum_{i=1}^n \vartheta_j f_j$, also

$$x \otimes y = \left(\sum_{i=1}^m \lambda_i e_i\right) \otimes \left(\sum_{j=1}^n \vartheta_j f_j\right) = \sum_{i=1}^m \sum_{j=1}^n \lambda_i \vartheta_j e_i \otimes f_j \in W.$$

Weil die $x \otimes y$ $X \otimes Y$ erzeugen, gilt $W = X \otimes Y$. Beweis der linearen Unabhängigkeit der $(e_i \otimes f_j)_{i,j=1}^{m,n}$: Sei $(e_i^*)_{i=1}^m$ die zu $(e_i)_{i=1}^m$ duale Basis von X^* und $(f_j)_{j=1}^n$ die zu $(f_j)_{j=1}^n$ duale Basis von Y^*. Sei $B_{i,j}(x,y) = e_i^*(x) f_j^*(y)$, $x \in X, y \in Y$. Dann gibt es $l_{i,j} : X \otimes Y \to K$ mit $B_{i,j}(x,y) = l_{i,j}(x \otimes y)$. Dann:

$$l_{i,j}(e_k \otimes f_l) = B_{i,j}(e_k, f_l) = e_i^*(e_k) f_j^*(f_l) = \delta_{ik}\delta_{jl} = \delta_{(i,j),(k,l)}.$$

Daraus folgt die lineare Unabhängigkeit. □

Fakt 1: Es gibt eindeutig bestimmte Isomorphismen mit folgenden Eigenschaften:

1. $X \otimes Y \to Y \otimes X$, $x \otimes y \mapsto y \otimes x$.
2. $X \otimes (Y \otimes Z) \to (X \otimes Y) \otimes Z$, $x \otimes (y \otimes z) \mapsto (x \otimes y) \otimes z$.
3. $(X \oplus X') \otimes Y \to (X \otimes Y) \oplus (X' \otimes Y)$, $(x, x') \otimes y \mapsto (x \otimes y, x' \otimes y)$.

Fakt 2: Wenn V und W Vektorräume sind, so gibt es genau eine lineare Abbildung

$$V^* \otimes W \to L(V,W),\ l \otimes w \mapsto f_{l,w} : V \to W,\ f_{l,w}(v) = l(v)(w).$$

Diese Abbildung ist injektiv, und ihr Bild besteht aus allen $f \in L(V,W)$ mit $\dim(\text{Bild}\,(f)) < \infty$. Für $\dim(B) < \infty$ ist sie ein Isomorphismus.

Beispiel 1: Für eine Raum-Zeit mit Tangentialraum T an einem Punkt sei $(e^i)_{i=1}^4$ eine Basis von T und (e_i^*) die dazu duale Basis. $f_{i,j} \to \sum_{i,j=1}^{4,4} f_{i,j} e^i \otimes e^j \in T \otimes T$, $f^{i,j} = f_{i,j} e^i \otimes e^j$ (Einsteinsche Summenkonvention). $f^{i,j} \to f^{i,j} e_i^* \otimes e_j^* \in T^* \otimes T^*$, $R^i_{jkl} \to R^i_{jkl} e_i^* \otimes e^j \otimes e^k \otimes e^l \in T^* \otimes T \otimes T \otimes T$.
Beispiel 2: Sei $K \subset L$ eine Körpererweiterung, sei V ein K-Vektorraum. Dann ist $V \otimes_K L$ ein L-Vektorraum mit Multiplikation $l(V \otimes l') = V \otimes (ll')$. Wenn $(b_i)_{i=1}^n$ eine Basis von V (über K) ist, so ist $(b_i \otimes 1)_{i=1}^n$ eine Basis von $V \otimes_K L$ über L. Für $K = \mathbf{R}$, $L = \mathbf{C}$ erhält man $(V)_{\mathbb{C}}$ aus den Übungsaufgaben.