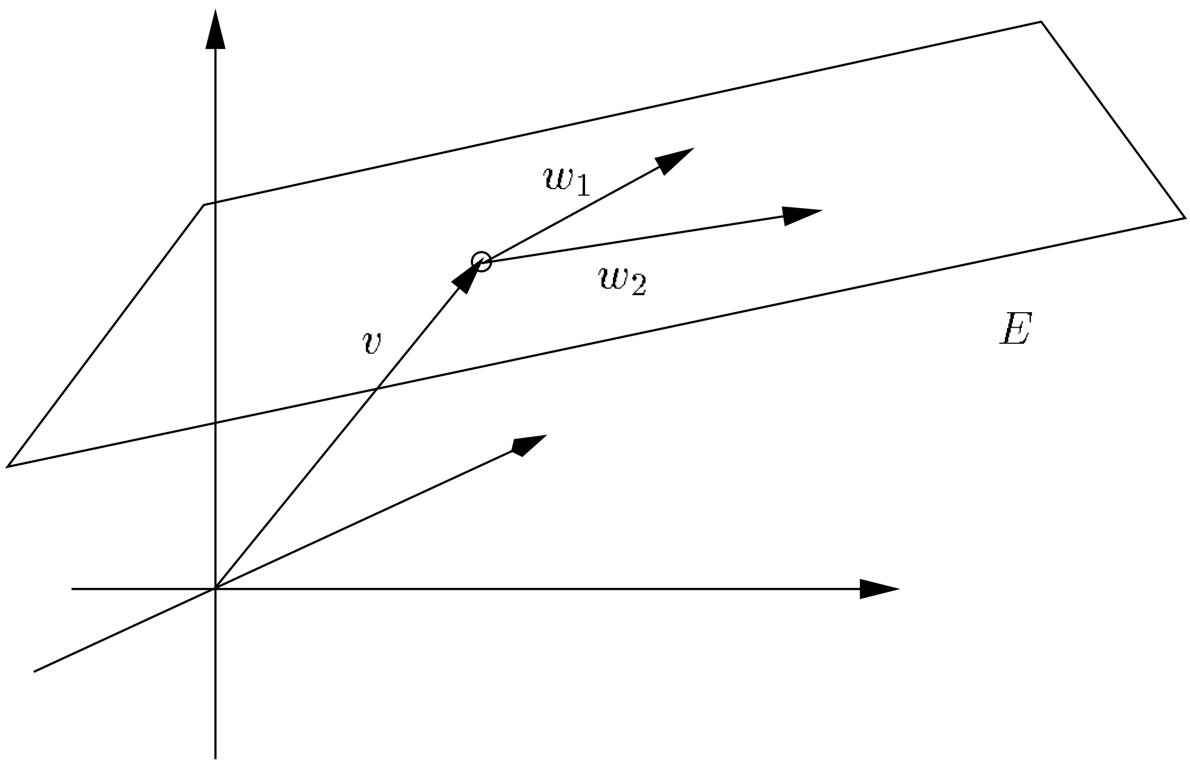


Lineare Algebra I

Eine Vorlesung von Prof. Dr. Klaus Hulek
hulek@math.uni-hannover.de

Institut für Mathematik
Universität Hannover
D – 30060 Hannover
Germany
E-Mail : hulek@math.uni-hannover.de

§ 0 Mengen, Abbildungen

Definition. (G. Cantor) Eine *Menge* ist die Zusammenfassung wohlunterschiedener Objekte unseres Denkens zu einem Ganzen.

Beispiele.
(i) $\mathbb{N} := \{1, 2, 3, 4, 5, \dots\}$ (*natürliche Zahlen*)

(ii) $\mathbb{Z} := \{0, \pm 1, \pm 2, \dots\}$ (*ganze Zahlen*)

(iii) $\mathbb{Q} := \{\frac{p}{q}; p,\ q$ ganze, teilerfremde Zahlen, $q \neq 0\}$ (*rationale Zahlen*)

(iv) $\mathbb{R} :=$ Menge der *reellen Zahlen*

(v) $M :=$ Menge der Studierenden in diesem Hörsaal

Schreibweise.
(1) $a \in A$: a ist *Element* von A $\quad (2 \in \mathbb{N},\ \sqrt{2} \in \mathbb{R},\ \pi \in \mathbb{R})$

$a \notin A$: a ist *kein Element* von A $\quad (-2 \notin \mathbb{N},\ \sqrt{2} \notin \mathbb{Q})$

(2) $A \subset B$: A ist Teilmenge von B. $\quad (a \in A \Rightarrow a \in B)$

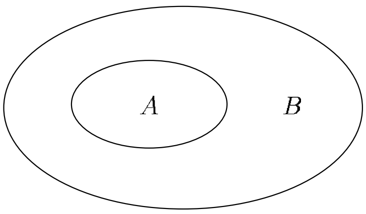


Abb. 1: Teilmenge

Bemerkung. Dies läßt auch $A = B$ zu.

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

Operationen mit Mengen

(1) *Durchschnitt* $A \cap B := \{x;\ x \in A \text{ und } x \in B\}$

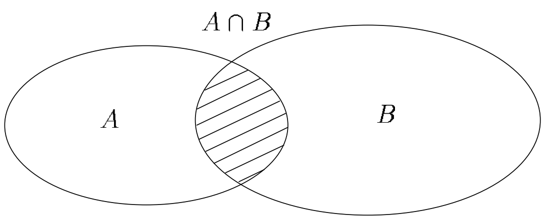


Abb. 2: Durchschnitt

Allgemeiner:

$$\bigcap_{i \in I} A_i = \{x;\ x \in A_i \text{ für alle } i \in I\}$$

(2) Vereinigung

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

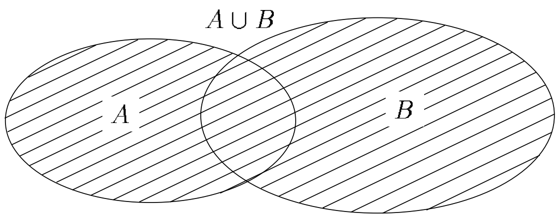


Abb. 3: Vereinigung

(3) Differenz von Mengen

$$A \setminus B := \{x \in A;\ x \notin B\}.$$

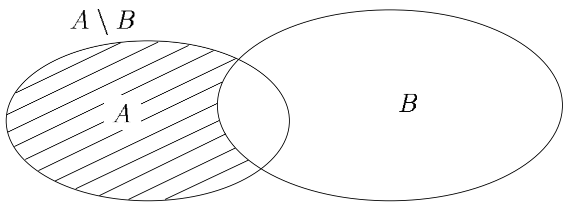


Abb. 4: Differenz

(4) Kartesisches Produkt

$$A \times B := \{(a, b);\ a \in A, b \in B\} \quad \text{(Menge von Paaren)}$$

Beispiel.

$$\mathbb{R}^2 := \{(x_1, x_2);\ x_1, x_2 \in \mathbb{R}\}.$$

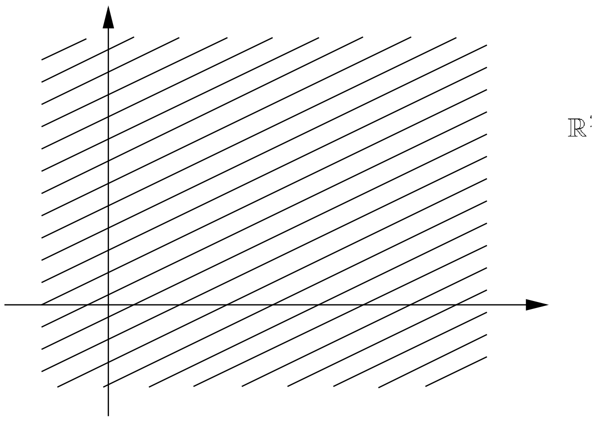


Abb. 5: Zahlenebene

Allgemeiner:

$$A_1 \times \cdots \times A_n := \{(a_1, \ldots, a_n);\ a_i \in A_i\} \qquad \text{(Menge der } n\text{-Tupel)}$$

Abbildungen

X, Y seien Mengen.

Definition. Eine *Abbildung* $f : X \to Y$ ist eine Vorschrift, die jedem Element $x \in X$ genau ein Element $f(x) \in Y$ zuordnet.

X heißt *Definitionsbereich*, Y heißt *Zielbereich* (*Wertebereich*) von f.

Beispiele.

(i)
$$\begin{array}{rrcl} f: & M & \longrightarrow & \mathbb{N} \\ & \text{Student/in} & \longmapsto & \text{Geburtsjahr} \end{array}$$

(ii) Normalparabel:
$$\begin{array}{rrcl} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & f(x) & = & x^2 \end{array}$$

(iii)
$$\begin{array}{rrcl} f: & \mathbb{R} & \longrightarrow & \{0,1\} \\ & f(x) & := & \begin{cases} 0 & \text{falls } x \in \mathbb{Q} \\ 1 & \text{falls } x \notin \mathbb{Q}. \end{cases} \end{array}$$

Für jede Menge X ist die *Identität*

$$\begin{array}{rrcl} \mathrm{id}_X: & X & \longrightarrow & X \\ & x & \longmapsto & x \end{array}$$

definiert.

Definition. Es sei $f : X \to Y$ eine Abbildung und $M \subset X$, $N \subset Y$ Teilmengen.

(i) Das *Bild* von f ist die Menge

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

Ist $M \subset X$ eine Teilmenge, so ist entsprechend

$$f(M) := \{f(x);\ x \in M\}.$$

(ii) Ist $N \subset Y$ eine Teilmenge, so ist das *Urbild* von N unter f definiert als

$$f^{-1}(N) := \{x \in X;\ f(x) \in N\}.$$

Beispiele. Für $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$ gilt:
(i) $f(\mathbb{R}) = \{x \in \mathbb{R};\ x \geq 0\} =: \mathbb{R}_{\geq 0}$,
$f([0,2]) = [0,4]$
(ii) $f^{-1}([0,4]) = [-2,2]$.

Definition. Es sei $f : X \to Y$ eine Abbildung und $M \subset X$ eine Teilmenge. Dann ist die *Einschränkung* von f auf M definiert durch

$$\begin{array}{rccc} f|_M : & M & \longrightarrow & Y \\ & x & \longmapsto & f(x). \end{array}$$

Es seien $f : X \to Y, g : Y \to Z$ Abbildungen.

Definition. Die *Hintereinanderschaltung* (*Komposition*) von f und g ist die Abbildung

$$\begin{array}{rccc} g \circ f : & X & \longrightarrow & Z \\ & x & \longmapsto & g(f(x)). \end{array}$$

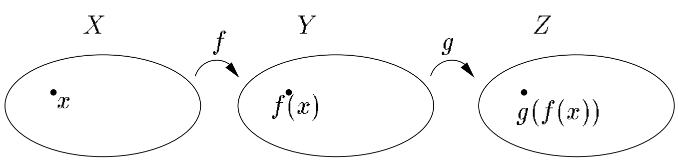


Abb. 6: Komposition

Beispiel.

$$\begin{array}{rlll} f : & \mathbb{R} \longrightarrow \mathbb{R}; & f(x) & = \sqrt{2}x \\ g : & \mathbb{R} \longrightarrow \mathbb{R}; & g(x) & = \dfrac{1}{1+x^2}. \end{array}$$

Dann ist

$$(g \circ f)(x) = g(f(x)) = g(\sqrt{2}x) = \frac{1}{1+(\sqrt{2}x)^2} = \frac{1}{1+2x^2}.$$

Lemma 0.1
Sind $f : X \to Y, g : Y \to Z, h : Z \to W$, so gilt

$$h \circ (g \circ f) = (h \circ g) \circ f \qquad \text{(Assoziativität von Abbildungen)}$$

Beweis. Beides sind Abbildungen von X nach W. Für alle $x \in X$ gilt:

$$\begin{aligned} ((h \circ g) \circ f)(x) &= (h \circ g)(f(x)) = h(g(f(x))) \\ (h \circ (g \circ f))(x) &= h(g \circ f(x)) = h(g(f(x))). \end{aligned}$$

□

Definition. Eine Abbildung $f : X \to Y$ heißt
(i) *injektiv*, falls gilt: Ist $x \neq y$, so ist auch $f(x) \neq f(y)$.
(ii) *surjektiv*, falls $f(X) = Y$, d. h.zu jedem $y \in Y$ gibt es ein $x \in X$ mit $f(x) = y$.
(iii) *bijektiv*, falls f injektiv und surjektiv ist.

Lemma 0.2
Die folgenden beiden Aussagen sind äquivalent:
(i) f ist bijektiv.
(ii) Es gibt eine Abbildung $g : Y \to X$ mit $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$.

Beweis. (i)⇒(ii) Es sei $y \in Y$. Da f bijektiv ist, gibt es genau ein $x \in X$ mit $f(x) = y$. Wir setzen $g(y) := x$. Dann ist nach Konstruktion von g:

$$(f \circ g)(y) = f(x) = y, \quad (g \circ f)(x) = g(f(x)) = x.$$

(ii)⇒(i) Wir zeigen zunächst, daß f injektiv ist:

$$f(x) = f(x') \;\Rightarrow\; g(f(x)) = g(f(x')) \;\Rightarrow\; x = x'.$$

Die Abbildung f ist auch surjektiv, da für $y \in Y$ gilt:

$$f(g(y)) = y.$$

□

Definition. g heißt dann die *Umkehrabbildung* von f, bezeichnet mit f^{-1}.

Beispiel. Die Abbildung $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$ ist weder injektiv $(f(-1) = f(1))$ noch surjektiv. Es ist

$$f(\mathbb{R}) = \mathbb{R}_{\geq 0} = \{x \in \mathbb{R};\ x \geq 0\}.$$

Die Abbildung $f|_{\mathbb{R}_{\geq 0}}$ ist injektiv, aber nicht surjektiv. Dagegen ist die Abbildung

$$\begin{aligned} \overline{f} : \quad \mathbb{R}_{\geq 0} &\longrightarrow \mathbb{R}_{\geq 0} \\ \overline{f}(x) &= x^2 \end{aligned}$$

bijektiv mit Umkehrabbildung

$$\overline{f}^{-1}(x) = \sqrt{x}.$$

§ 1 Elementargeometrie

Man kann die reellen Zahlen als *Zahlengerade* interpretieren.

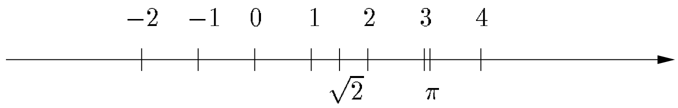


Abb. 7: Zahlengerade

Wir betrachten nun eine *Ebene* E (z. B. die Tafelebene). Legt man einen Ursprung 0 und ein Koordinatenkreuz fest, so kann man die Ebene mit $\mathbb{R}^2$ identifizieren.

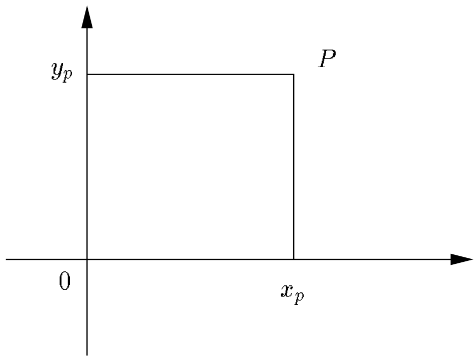


Abb. 8: Ebene

Die Abbildung

$$\begin{array}{rcl} E & \longrightarrow & \mathbb{R}^2 \\ P & \longmapsto & (x_P, y_P) \end{array}$$

ist bijektiv. Man nennt (x_p, y_p) die *kartesischen Koordinaten* des Punktes P. (Descartes, 1596–1650).

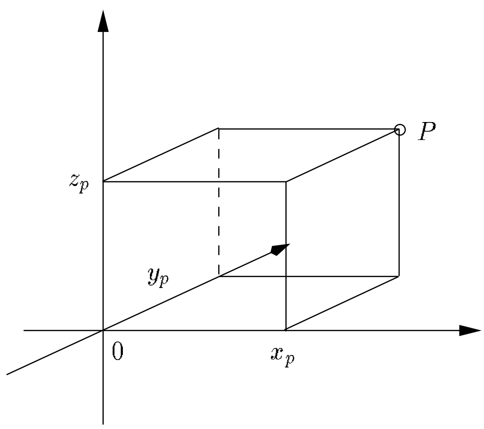


Abb. 9: 3-dimensionaler Raum

Analog sei R der *Anschauungsraum*. Dann hat man nach Wahl eines Koordinatensystems eine Bijektion

$$\begin{array}{rcl} R & \longrightarrow & \mathbb{R}^3 \\ P & \longmapsto & (x_P, y_P, z_P). \end{array}$$

Entsprechend nennt man $\mathbb{R}^n$ den *n-dimensionalen Raum*. Ein Element

$$(x_1, \ldots, x_n) \in \mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{\text{n-mal}}$$

heißt ein *Vektor*, genauer ein *Zeilenvektor*. Wir schreiben

$$x = (x_1, \ldots, x_n).$$

Man kann mit diesen Vektoren *algebraische Operationen* ausführen:

(1) *Vektoraddition:* $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$

$$x + y := (x_1 + y_1, \ldots, x_n + y_n).$$

Geometrische Deutung: $(n = 2)$

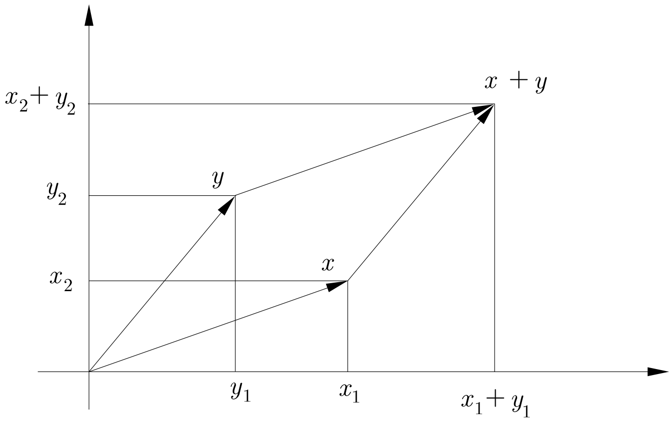
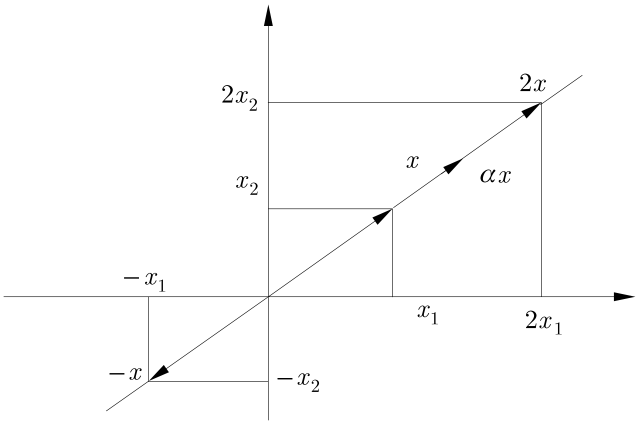


Abb. 10: Vektoraddition

(2) *Skalarmultiplikation:* $x = (x_1, \ldots, x_n) \in \mathbb{R}^n, \alpha \in \mathbb{R}$

$$\alpha \cdot x := (\alpha x_1, \ldots, \alpha x_n).$$



$(-x := -1 \cdot x)$
$(x - y := x + (-y))$

Abb. 11: Skalarmultiplikation

Diese Operationen erfüllen eine Reihe von *Rechenregeln*:

(1) (i) Für alle Vektoren $x, y, z \in \mathbb{R}^n$ gilt

$$(x+y)+z = x+(y+z).$$

(ii) Der *Nullvektor* $0 = (0, \dots, 0)$ hat die Eigenschaft

$$x+0 = 0+x = x \qquad (\text{für alle } x \in \mathbb{R}^n).$$

Zu jedem $x \in \mathbb{R}^n$ gibt es ferner $-x := (-1) \cdot x$ mit

$$x + (-x) = 0.$$

(iii)

$$x+y = y+x \qquad (\text{für alle } x, y \in \mathbb{R}^n).$$

(2) Für $x, y \in \mathbb{R}^n$; $\alpha, \beta \in \mathbb{R}$ gilt

$$\begin{aligned} (\alpha+\beta)\cdot x &= (\alpha \cdot x) + (\beta \cdot x) \\ \alpha \cdot (x+y) &= (\alpha \cdot x) + (\alpha \cdot y) \\ \alpha \cdot (\beta \cdot x) &= (\alpha \cdot \beta) \cdot x \\ 1 \cdot x &= x. \end{aligned}$$

Denn: Wir begründen etwa die zweite Formel von (2)

$$\begin{aligned} \alpha \cdot (x+y) &= \alpha \cdot (x_1+y_1, \dots, x_n+y_n) = (\alpha(x_1+y_1), \dots, \alpha(x_n+y_n)) \\ &\underset{\substack{\text{Distributiv-} \\ \text{gesetz in } \mathbb{R}}}{=} (\alpha x_1 + \alpha y_1, \dots, \alpha x_n + \alpha y_n) \\ &= (\alpha x_1, \dots, \alpha x_n) + (\alpha y_1, \dots, \alpha y_n) = (\alpha \cdot x) + (\alpha \cdot y). \end{aligned}$$

Die obigen Rechenregeln werden grundlegend für die Definition eines Vektorraums werden.

Wir werden im folgenden oft die Bezeichnungen „Vektoren“ und „Punkte“ gleichbedeutend verwenden.

Geraden im $\mathbb{R}^n$

Es seien $v, w \in \mathbb{R}^n$. Wir setzen

$$v + \mathbb{R} \cdot w := \{u \in \mathbb{R}^n;\ u = v + \alpha w \text{ für ein } \alpha \in \mathbb{R}\}.$$

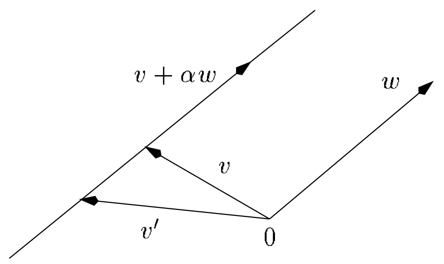


Abb. 12: Gerade

Definition. Eine Teilmenge $A \subset \mathbb{R}^n$ heißt eine *Gerade*, falls es Vektoren $v, w \in \mathbb{R}^n$ mit $w \neq 0$ gibt, so daß

$$A = v + \mathbb{R}w.$$

Bemerkung. Die Abbildung

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & A \\ \alpha & \longmapsto & v + \alpha w \end{array}$$

ist bijektiv.

Die Surjektivität ist klar, die Injektivität folgt aus:

$$\begin{aligned} v + \alpha w &= v + \alpha' w \Rightarrow \alpha w = \alpha' w \Rightarrow \\ \alpha w - \alpha' w &= 0 \Rightarrow (\alpha - \alpha')w = 0 \overset{w \neq 0}{\Rightarrow} \alpha - \alpha' = 0. \end{aligned}$$

Lemma 1.1

(i) Es sei $A = v + \mathbb{R}w$ eine Gerade. Ferner sei $v' \in A$. Dann gilt $A = v' + \mathbb{R}w$.

(ii) Seien $A = v + \mathbb{R}w$ und $A' = v' + \mathbb{R}w'$. Dann gilt

$$A = A' \quad \Leftrightarrow \quad A \cap A' \neq \emptyset \text{ und es gibt ein } \beta \neq 0 \text{ mit } w' = \beta w.$$

Beweis.

(i) $v' \in A \Rightarrow v' = v + \alpha' w$ für ein $\alpha' \in \mathbb{R}$.

$v' + \mathbb{R}w \supset A$: $x \in A$, d. h. $x = v + \alpha w$ für ein $\alpha \in \mathbb{R}$.

$x = v + \alpha w + \alpha' w - \alpha' w = \underbrace{v + \alpha' w}_{=v'} + (\alpha - \alpha')w \in v' + \mathbb{R}w.$

$v' + \mathbb{R}w \subset A$: $x = v' + \alpha w = v + \alpha' w + \alpha w = v + (\alpha' + \alpha)w \in A.$

(ii) „$\Rightarrow$“ Dann ist $A \cap A' \neq \emptyset$ klar.

$$\left.\begin{aligned} v' \in A' = A &\ \Rightarrow v' = v + \alpha' w \\ v' + w' \in A' = A &\ \Rightarrow v' + w' = v + \alpha'' w \end{aligned}\right\} \Rightarrow$$

$$w' = v + \alpha'' w - v' = (v + \alpha'' w) - (v + \alpha' w) = (\underbrace{\alpha'' - \alpha'}_{=:\beta})w.$$

Also gilt $w' = \beta w$. Es ist $\beta \neq 0$, da $w' \neq 0$.

„$\Leftarrow$“ Nach Voraussetzung gibt es $u \in A \cap A'$. Nach Punkt (i) ist dann

$$A = u + \mathbb{R}w, \quad A' = u + \mathbb{R}w'.$$

$A \subset A'$: $\quad A \ni x = u + \alpha w = u + \frac{\alpha}{\beta} \cdot \underbrace{\beta w}_{w'} = u + \frac{\alpha}{\beta} w' \in A'$.

$A' \subset A$: $\quad A' \ni y = u + \alpha w' = u + \alpha(\beta w) = u + (\alpha\beta) w \in A$. □

Lemma 1.2
Durch zwei verschiedene Punkte $v_1 \neq v_2$ in $\mathbb{R}^n$ gibt es stets genau eine Gerade.

Beweis. *Existenz:*

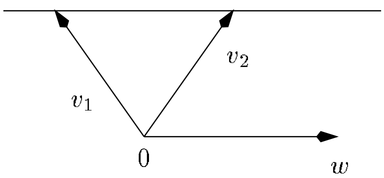


Abb. 13: Gerade durch zwei Punkte

Es sei $w := v_2 - v_1 \neq 0$ und $A := v_1 + \mathbb{R}w$. Dann ist

$$v_1 = v_1 + 0 \cdot w \in A; \; v_2 = v_1 + w = v_1 + 1 \cdot w \in A.$$

Eindeutigkeit: Es sei A' eine weitere Gerade durch v_1 und v_2. Sei etwa

$$A' = v' + \mathbb{R}w'.$$

Nach Lemma (1.1) (i) kann man schreiben

$$A' = v_1 + \mathbb{R}w'.$$

Es gilt, da $v_2 \in A'$, daß $v_2 = v_1 + \beta w'$ für ein $\beta \in \mathbb{R}$. Andererseits ist $v_2 = v_1 + w$, durch Subtraktion erhält man

$$0 = \beta w' - w.$$

D. h.

$$w = \beta w' \text{ mit } \beta \neq 0, \text{ da } w \neq 0.$$

Also

$$w' = \frac{1}{\beta} w.$$

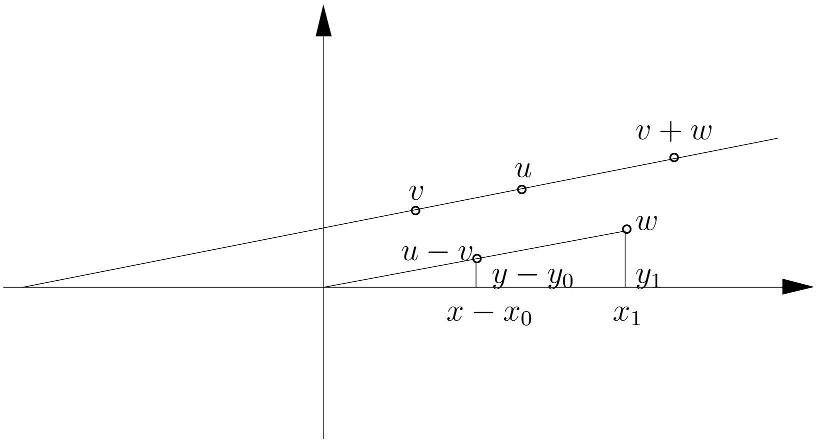
Nach Lemma (1.1) (ii) folgt $A = A'$. □

Der folgende Satz besagt, daß man Geraden in der *Ebene* entweder mit einer Parameterdarstellung oder durch lineare Gleichungen beschreiben kann.

Satz 1.3
Eine Teilmenge $A \subset \mathbb{R}^2$ ist genau dann eine Gerade, wenn es $a, b, c \in \mathbb{R}$ mit $(a, b) \neq (0, 0)$ gibt, so daß

$$A = \{(x, y) \in \mathbb{R}^2;\ ax + by - c = 0\}.$$

Beweis. *1. Schritt:* Sei $A = v + \mathbb{R}w$ mit $w \neq 0$. Wir müssen die *Existenz* einer Gleichung zeigen. Zunächst die *Idee*.



$$\begin{array}{rcl} v & = & (x_0, y_0) \\ w & = & (x_1, y_1) \neq (0, 0) \end{array}$$

Abb. 14: Beweisidee

Wir betrachten $u = (x, y) \in A$. Dann ist

$$u - v = (x - x_0, y - y_0).$$

Elementargeometrische Überlegungen ergeben

$$\frac{x - x_0}{y - y_0} = \frac{x_1}{y_1} \quad \Rightarrow \quad xy_1 - x_0y_1 = yx_1 - y_0x_1$$

also

$$xy_1 - yx_1 + (y_0x_1 - x_0y_1) = 0. \qquad (x, y \text{ als Variable betrachten})$$

Formale Ausführung.

Wir definieren $a := y_1, b := -x_1, c := y_1x_0 - y_0x_1$ und setzen

$$B := \{(x, y) \in \mathbb{R}^2;\ ax + by - c = 0\}.$$

$A \subset B$: $u \in A \Rightarrow u = v + \alpha w = (x_0 + \alpha x_1, y_0 + \alpha y_1)$. Einsetzen in die Gleichung für B ergibt

$$\begin{array}{rcl} (x_0 + \alpha x_1)y_1 - (y_0 + \alpha y_1)x_1 + (y_0x_1 - x_0y_1) & = & \\ x_0y_1 + \alpha x_1y_1 - y_0x_1 - \alpha x_1y_1 + y_0x_1 - x_0y_1 & = & 0 \end{array}$$

d. h. $u \in B$.

$B \subset A$: Nehmen wir zunächst an, daß $a = y_1 \neq 0$. Sei

$$u = (x, y) \in B$$

d. h.

$$ax + by - c = 0.$$

Gesucht: Ein α mit $u = v + \alpha w$.

Dann müßte gelten

$$y = y_0 + \alpha y_1, \text{ d. h. } \alpha = \frac{y - y_0}{y_1} \quad (y_1 \neq 0).$$

Wir setzen also

$$\alpha := \frac{y - y_0}{y_1}.$$

Dann gilt

(1) $$y = y_0 + \alpha y_1$$

Für x gilt

$$x = \frac{c - by}{a} = \frac{y_1 x_0 - y_0 x_1 + x_1 y}{y_1} = x_0 + x_1 \underbrace{\frac{y - y_0}{y_1}}_{=\alpha}$$

also

(2) $$x = x_0 + \alpha x_1.$$

Aus (1) und (2) folgt dann:

$$u = (x, y) = v + \alpha w \in A.$$

Falls $a = y_1 = 0$, so ist $b = -x_1 \neq 0$. Dann argumentiert man analog mit

$$\alpha = \frac{x - x_0}{x_1}.$$

2. Schritt: Es sei nun

$$A = \{(x, y) \in \mathbb{R}^2;\ ax + by - c = 0\} \qquad (a, b) \neq (0, 0).$$

Wir müssen v, w erraten. Sei $a \neq 0$ ($b \neq 0$ verläuft analog). Wir setzen

$$v := (\frac{c}{a}, 0), \quad w := (-b, a) \neq (0, 0).$$
$$B := v + \mathbb{R}w.$$

$A \subset B$: Sei $\alpha := \frac{y}{a}$.

$$(x, y) = \left(\frac{c - by}{a}, y\right) \overset{\text{Def. von } \alpha}{=} \left(\frac{c}{a} - \alpha b, \alpha a\right) = v + \alpha w \in B.$$

$B \subset A$: Sei $\alpha \in \mathbb{R}$. Dann gilt

$$v + \alpha w = \left(\frac{c}{a}, 0\right) + \alpha(-b, a) = \left(\frac{c}{a} - \alpha b, \alpha a\right).$$

Es gilt

$$a\left(\frac{c}{a} - \alpha b\right) + b(\alpha a) - c = c - a\alpha b + b\alpha a - c = 0$$

d. h.

$$v + \alpha w \in A.$$

$\square$

Ebenen im $\mathbb{R}^n$

Idee:

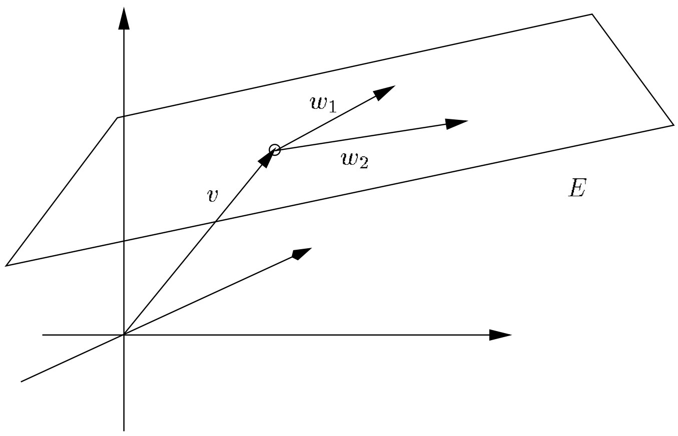


Abb. 15: Ebene im $\mathbb{R}^3$

Wir setzen für w_1, w_2 im $\mathbb{R}^n$:

$$v + \mathbb{R}w_1 + \mathbb{R}w_2 := \{u;\ u = v + \lambda w_1 + \mu w_2 \qquad \text{mit } \lambda, \mu \in \mathbb{R}\}.$$

Wir müssen darauf achten, daß w_1 und w_2 „unabhängig" sind, d. h.nicht in dieselbe Richtung zeigen.

Definition. Zwei Vektoren $w_1, w_2 \in \mathbb{R}^n$ heißen *linear unabhängig*, falls gilt:

$$\lambda_1 w_1 + \lambda_2 w_2 = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = 0.$$

Ansonsten heißen w_1, w_2 *linear abhängig.*

Beispiele.

(i) Gegeben seien die Vektoren $w_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$, $w_2 = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}$. Einsetzen in die obige Definition liefert:

$$\lambda_1 w_1 + \lambda_2 w_2 = \begin{pmatrix} 1\lambda_1 + 0\lambda_2 \\ 0\lambda_1 + 2\lambda_2 \\ 2\lambda_1 + 3\lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 2\lambda_2 \\ 2\lambda_1 + 3\lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \Rightarrow \lambda_1 = \lambda_2 = 0.$$

Also sind die beiden Vektoren linear unabhängig.

(ii) Wählt man hingegen die Vektoren $w_1 = \begin{pmatrix} 2 \\ 6 \\ 4 \end{pmatrix}$ und $w_2 = \begin{pmatrix} 3 \\ 9 \\ 6 \end{pmatrix}$, so hat man die folgende nicht-triviale Linearkombination:

$$3w_1 - 2w_2 = 3 \begin{pmatrix} 2 \\ 6 \\ 4 \end{pmatrix} - 2 \begin{pmatrix} 3 \\ 9 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Daher sind w_1 und w_2 linear abhängig.

(iii) Ist etwa $w_1 = 0$, so sind w_1, w_2 stets linear abhängig, da

$$1w_1 + 0w_2 = 0.$$

Analoges gilt für $w_2 = 0$.

Lemma 1.4
Es seien $w_1, w_2 \neq 0$. Dann sind folgende Aussagen äquivalent:
(i) w_1, w_2 sind linear abhängig.
(ii) Es gibt $\rho \neq 0$ mit $w_1 = \rho w_2$.
(iii) Es gibt $\rho' \neq 0$ mit $w_2 = \rho' w_1$.

Beweis. (i)⇒(ii): Es gibt eine Gleichung

$$\lambda_1 w_1 + \lambda_2 w_2 = 0 \qquad \text{mit } (\lambda_1, \lambda_2) \neq (0,0). \tag{3}$$

Es ist $\lambda_1 \neq 0 \neq \lambda_2$, denn wäre etwa $\lambda_1 = 0$, würde aus $\lambda_2 w_2 = 0$ und $w_2 \neq 0$ folgen, daß $\lambda_2 = 0$ ist. Also können wir (3) auflösen:

$$w_1 = -\frac{\lambda_2}{\lambda_1} w_2 = \rho w_2 \qquad (\rho = -\frac{\lambda_2}{\lambda_1}). \tag{4}$$

(ii)⇒(iii): Es gilt $w_1 = \rho w_2$. Da $w_1 \neq 0$ ist, ist auch $\rho \neq 0$. Also gilt $w_2 = \frac{1}{\rho} w_1$.

(iii)⇒(i): Aus $w_2 = \rho' w_1$ folgt

$$\rho' w_1 + (-1) w_2 = 0,$$

also sind w_1, w_2 linear abhängig. □

Definition. Eine Teilmenge $E \subset \mathbb{R}^n$ heißt eine *Ebene*, falls es Vektoren v, w_1, w_2 gibt, so daß w_1, w_2 linear unabhängig sind, und

$$E = v + \mathbb{R}w_1 + \mathbb{R}w_2$$

gilt. Man nennt eine solche Darstellung die *Parameterdarstellung von E*.

Ist $E = v + \mathbb{R}w_1 + \mathbb{R}w_2$ und $v' \in E$, so zeigt man analog wie in Lemma (1.1) (i), daß

$$E = v' + \mathbb{R}w_1 + \mathbb{R}w_2.$$

Satz 1.5
Es seien v_1, v_2, v_3 Punkte in $\mathbb{R}^n$, die nicht in einer Geraden liegen. Dann gibt es genau eine Ebene durch diese drei Punkte.

Beweis. Der Beweis verläuft ganz analog zum Beweis von Lemma (1.2). Die gesuchte Ebene ist

$$E = v_1 + \mathbb{R}(v_2 - v_1) + \mathbb{R}(v_3 - v_1).$$

$\square$

Für Ebenen im $\mathbb{R}^3$ hat man, ähnlich wie für Geraden im $\mathbb{R}^2$ eine Beschreibung durch Gleichungen.

Satz 1.6
Eine Teilmenge $E \subset \mathbb{R}^3$ ist genau dann eine Ebene, wenn es Zahlen $a, b, c, d \in \mathbb{R}$ mit $(a, b, c) \neq (0, 0, 0)$ gibt, so daß

$$E = \{(x_1, x_2, x_3) \in \mathbb{R}^3;\ ax_1 + bx_2 + cx_3 - d = 0\}.$$

Beweis. Ähnlich wie zu Satz (1.3). $\square$

Längen und Winkelmessung

Die „Länge" eines Vektors $x = (x_1, x_2) \in \mathbb{R}^2$, bzw. der Abstand des Punktes $x = (x_1, x_2)$ vom Ursprung ist gegeben durch

$$\|x\| := d(x, 0) = \sqrt{x_1^2 + x_2^2}.$$

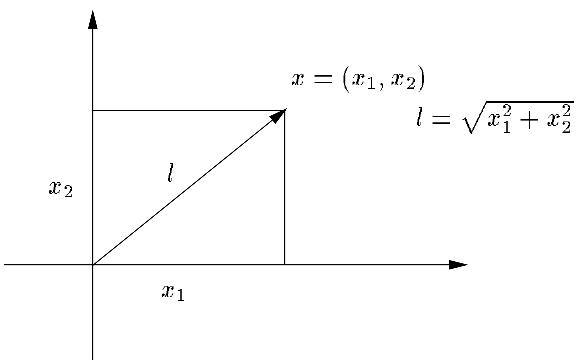


Abb. 16: „Länge" eines Vektors

Analog ist der Abstand zweier Punkte x, y gegeben durch

$$d(x,y) := \|x-y\| = \sqrt{(x_1-y_1)^2+(x_2-y_2)^2}.$$

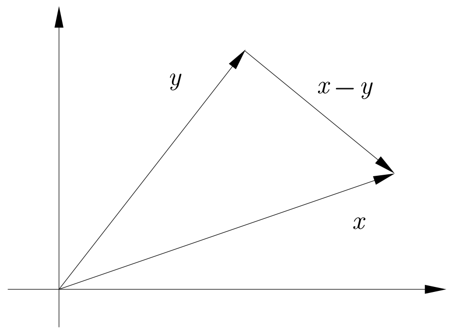


Abb. 17: Abstand zweier Punkte

Für einen Punkt $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ gilt entsprechend

$$\|x\| = d(x,0) = \sqrt{x_1^2+x_2^2+x_3^2}.$$

Dies führt auf den Begriff des *Skalarprodukts*.

Definition. Für zwei Vektoren $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$ ist das *Skalarprodukt* definiert durch

$$\langle x, y\rangle := x_1y_1 + \cdots + x_ny_n = \sum_{i=1}^n x_iy_i.$$

Damit erhalten wir eine Abbildung

$$\begin{array}{rrcl} \langle\ ,\ \rangle : & \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R}, \\ & (x,y) & \longmapsto & \langle x, y\rangle \end{array}$$

die folgende Regeln erfüllt ($x, y, z \in \mathbb{R}^n, \lambda \in \mathbb{R}$):

(1) (i) $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$

(ii) $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$

(2) (i) $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle$

(ii) $\langle x, \lambda y\rangle = \lambda\langle x, y\rangle$

(3) $\langle x, y\rangle = \langle y, x\rangle$

(4) $\langle x, x\rangle \geq 0$ und $\langle x, x\rangle = 0$ genau dann wenn $x = 0$.

Dies läßt sich leicht durch Nachrechnen bestätigen. Etwa bei (2) (i):

$$\langle x, y+z\rangle = \sum_{i=1}^n x_i(y_i+z_i) = \sum_{i=1}^n x_iy_i + \sum_{i=1}^n x_iz_i = \langle x, y\rangle + \langle x, z\rangle.$$

Außerdem ist

$$\langle x, x\rangle = x_1^2 + \cdots + x_n^2 \geq 0$$

und dieser Ausdruck kann nur dann 0 sein, wenn $x_1 = \cdots = x_n = 0$ ist, d. h.wenn $x = 0$.

Definition. Die *Länge* (*Norm*) eines Vektors $x \in \mathbb{R}^n$ ist definiert durch

$$\|x\| := \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Aus der Definition folgen unmittelbar zwei grundlegende Eigenschaften der Norm:

$$\|x\| = 0 \Leftrightarrow x = 0 \qquad \text{und} \qquad \|\lambda x\| = |\lambda|\,\|x\|.$$

Definition. Für $x, y \in \mathbb{R}^n$ definiert man den *Abstand* durch

$$d(x, y) := \|x - y\| = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}.$$

Für den Abstand gilt:

$$\begin{aligned} d(x, y) &\geq 0 \\ d(x, y) &= 0 \Leftrightarrow x = y \\ d(x, y) &= d(y, x). \end{aligned}$$

Lemma 1.7
Es gilt:
(i) $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle$ (Satz von Pythagoras)
(ii) $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$ (Parallelogrammgleichung).

Beweis. Beides zeigt man direkt durch Nachrechnen:
(i)

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle = \langle x, x + y\rangle + \langle y, x + y\rangle = \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 + 2\langle x, y\rangle + \|y\|^2. \end{aligned}$$

(ii) Analog. □

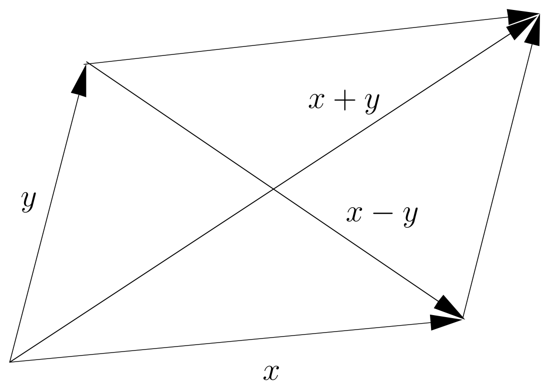


Abb. 18: Zur Parallelogrammgleichung

Satz 1.8 (Cauchy-Schwarzsche Ungleichung)
Für $x, y \in \mathbb{R}^n$ gilt

$$|\langle x, y\rangle| \leq \|x\| \, \|y\|.$$

Gleichheit gilt genau dann, wenn x und y linear abhängig sind.

Beweis. Ist $y = 0$, so sind beide Seiten gleich. Zudem sind x und y linear abhängig. Wir können im folgenden also $y \neq 0$ annehmen. Dann setzen wir

$$\lambda := \langle y, y\rangle > 0, \quad \mu := -\langle x, y\rangle.$$

Dann gilt:

$$\begin{aligned} 0 &\leq \langle \lambda x + \mu y, \lambda x + \mu y\rangle = \lambda^2\langle x, x\rangle + 2\lambda\mu\langle x, y\rangle + \mu^2 \underbrace{\langle y, y\rangle}_{=\lambda} \\ &= \lambda(\langle y, y\rangle\langle x, x\rangle - 2\langle x, y\rangle^2 + \langle x, y\rangle^2) \\ &= \lambda(\langle y, y\rangle\langle x, x\rangle - \langle x, y\rangle^2). \end{aligned}$$

Da $\lambda > 0$ folgt hieraus

$$0 \leq \langle y, y\rangle\langle x, x\rangle - \langle x, y\rangle^2$$

also

$$\langle x, y\rangle^2 \leq \langle x, x\rangle\langle y, y\rangle,$$

und damit durch Wurzelziehen:

$$|\langle x, y\rangle| \leq \|x\| \, \|y\|.$$

Das Gleichheitszeichen gilt genau dann, wenn $\lambda x + \mu y = 0$, also wenn x und y linear abhängig sind. □

Korollar 1.9
Es gilt die Dreiecksungleichung

(i) $\|x + y\| \leq \|x\| + \|y\| \qquad (x, y \in \mathbb{R}^n)$

(ii) $\quad d(x_1, x_3) \leq d(x_1, x_2) + d(x_2, x_3) \qquad (x_1, x_2, x_3 \in \mathbb{R}^n).$

Beweis.
(i) Es gilt

$$\begin{aligned} \langle x+y, x+y\rangle &= \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle \\ &\overset{(1.8)}{\leq} \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2 \end{aligned}$$

Durch Wurzelziehen erhält man

$$\|x+y\| = \sqrt{\langle x+y, x+y\rangle} \leq \|x\| + \|y\|.$$

(ii) Zu zeigen ist die Beziehung

$$\|x_1 - x_3\| \leq \|x_1 - x_2\| + \|x_2 - x_3\|.$$

Dies folgt aus (i) mit

$$x := x_1 - x_2, \quad y := x_2 - x_3.$$

□

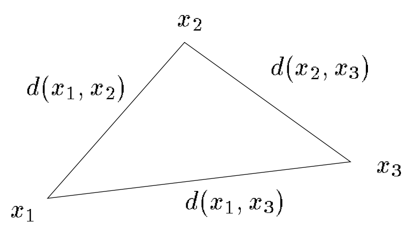


Abb. 19: Dreiecksungleichung

Winkelmessung

Es seien $x, y \in \mathbb{R}^n$, $x, y \neq 0$. Wir wollen den Winkel zwischen x und y definieren.

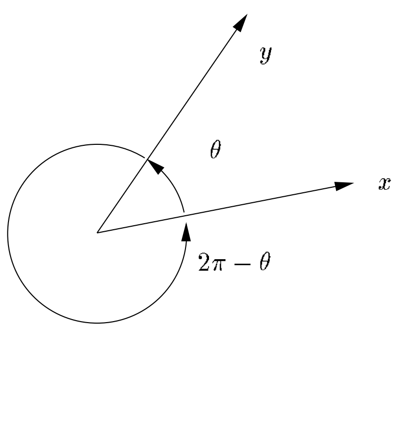


Abb. 20: Zur Winkeldefinition

Nach der Cauchy-Schwarzschen Ungleichung gilt

$$-1 \leq \frac{\langle x, y \rangle}{\|x\| \, \|y\|} \leq 1.$$

Also gibt es genau zwei Winkel θ, θ' mit $\theta' = 2\pi - \theta$, für die gilt

$$\cos\theta = \cos\theta' = \frac{\langle x, y \rangle}{\|x\| \, \|y\|}.$$

Fordert man $\theta \in [0, \pi]$, so ist θ eindeutig bestimmt.

Definition. Die eindeutig bestimmte Zahl $\theta \in [0, \pi]$ mit

$$\cos\theta = \frac{\langle x, y \rangle}{\|x\| \, \|y\|}$$

heißt der *Winkel* zwischen x und y.

Schreibweise. $\sphericalangle(x, y) = \theta$.

Eigenschaften des Winkels

(i) $\langle x, y \rangle = \|x\| \, \|y\| \cos \sphericalangle(x, y)$

(ii) $\sphericalangle(x, y) = \sphericalangle(y, x)$

(iii) $\sphericalangle(\lambda x, \mu y) = \sphericalangle(x, y)$ für $\lambda, \mu > 0$

(iv) $\sphericalangle(x, y) = 0 \Leftrightarrow x = \lambda y$ für ein $\lambda > 0$.

Spezialfall:

(i) Im $\mathbb{R}^2$ ist der oben eingeführte Winkel der übliche. Es seien $x, y \neq 0$. Setzen wir

$$x' = \frac{1}{\|x\|} x, \quad y' = \frac{1}{\|y\|} y$$

so ist $\sphericalangle(x', y') = \sphericalangle(x, y)$. Wir betrachten die folgende Situation (mit $\beta \geq \alpha$):

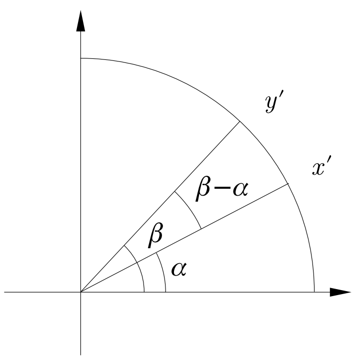


Abb. 21: Winkel in der Ebene

Es ist

$$x' = (\cos\alpha, \sin\alpha), \quad y' = (\cos\beta, \sin\beta).$$

Damit gilt:

$$\langle x', y'\rangle = \cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\beta - \alpha).$$

Also

$$\sphericalangle(x,y) = \begin{cases} \beta - \alpha & \text{falls } \beta - \alpha \leq \pi \\ \alpha - \beta + 2\pi & \text{sonst.} \end{cases}$$

(ii) Im Fall des $\mathbb{R}^2$ kann man auch einen „orientierten" Winkel einführen, der dann Werte in $[0, 2\pi[$ annehmen kann.

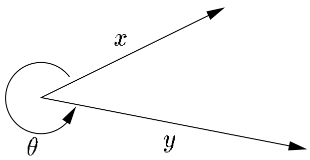


Abb. 22: Orientierter Winkel

Es seien

$$A = u + \mathbb{R}w, \quad B = u + \mathbb{R}w'$$

zwei Geraden, die sich in u schneiden.

Definition. Der *Winkel* zwischen den Geraden A, B ist definiert durch

$$\sphericalangle(A,B) := \min\{\sphericalangle(w,w'), \sphericalangle(w,-w')\}.$$

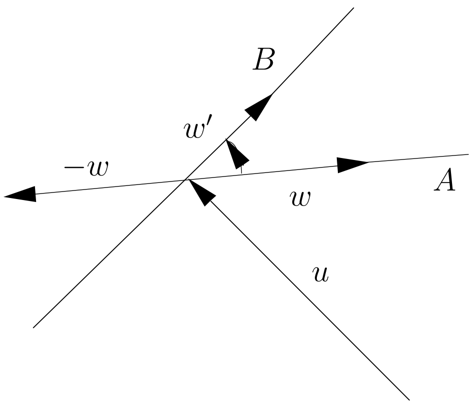


Abb. 23: Winkel zwischen Geraden

Bemerkung. Durch diese Definition hat man erreicht, daß $\sphericalangle(A,B) \in [0, \frac{\pi}{2}]$ ist.

Unsere obigen Überlegungen ergeben noch folgende Eigenschaft des Winkels:

$$\sphericalangle(x, y) = \frac{\pi}{2} \Leftrightarrow \langle x, y\rangle = 0.$$

Definition. Zwei Vektoren $x, y \in \mathbb{R}^n$ heißen zueinander *orthogonal* oder *stehen senkrecht aufeinander*, falls $\langle x, y\rangle = 0$.

Der Nullvektor steht nach dieser Definition auf allen Vektoren senkrecht.

Ist $v \perp w$, so erhält der Satz des Pythagoras seine gewohnte Form

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2$$

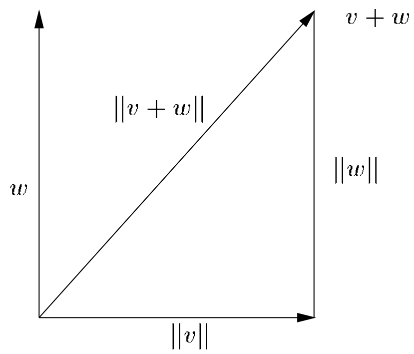


Abb. 24: Satz des Pytagoras in Vektorschreibweise

Definition. Ein Vektor $s \in \mathbb{R}^n$ ist *orthogonal* zu einer Geraden $A \subset \mathbb{R}^n$ falls $\langle s, v_1 - v_2\rangle = 0$ für alle $v_1, v_2 \in A$.

Bemerkung.

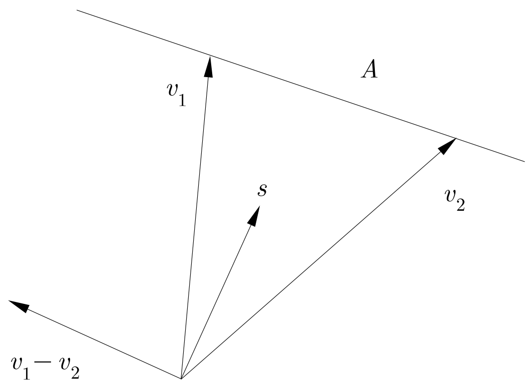


Abb. 25: Normalenvektor zu einer Geraden

(i) Ist $A = v + \mathbb{R}w$, so ist s genau dann orthogonal zu A, wenn $s \perp w$.

(ii) Ist
$$A = \{(x, y) \in \mathbb{R}^2;\ ax + by - c = 0\} \quad ((a, b) \neq (0, 0)),$$

dann ist $s = (a, b)$ orthogonal zu A.

Beweis.

(i) „$\Rightarrow$" Es ist $v, v + w \in A$. Ist s orthogonal zu A, folgt nach Definition $s \perp ((v + w) - v)$, also $s \perp w$.

„$\Leftarrow$" Es seien $v_1, v_2 \in A$, also

$$v_1 = v + \lambda w, \quad v_2 = v + \mu w.$$

Damit gilt:

$$\langle s, v_1 - v_2 \rangle = \langle s, \lambda w - \mu w \rangle = \lambda \langle s, w \rangle - \mu \langle s, w \rangle = 0.$$

(ii) Es seien $v_1 = (x_1, y_1)$, $v_2 = (x_2, y_2) \in A$. Dann gilt

$$\begin{aligned} \langle s, v_1 - v_2 \rangle &= \langle s, v_1 \rangle - \langle s, v_2 \rangle = (ax_1 + by_1) - (ax_2 + by_2) \\ &= c - c = 0. \end{aligned}$$

Satz 1.10 (Normalengleichung)
Es gilt

(i) Ist $0 \neq s \in \mathbb{R}^2$ und $v_1 \in \mathbb{R}^2$, dann ist

$$A = \{u \in \mathbb{R}^2;\ \langle s, u - v_1 \rangle = 0\}$$

eine Gerade durch v_1, die orthogonal zu s ist.

(ii) Ist $A \subset \mathbb{R}^2$ eine Gerade, $v_1 \in A$ und $s \neq 0$ orthogonal zu A, dann ist

$$A = \{u \in \mathbb{R}^2;\ \langle s, u - v_1 \rangle = 0\}.$$

Bemerkung. s heißt *Normalenvektor* von A. Manchmal ersetzt man s durch den *Normaleneinheitsvektor* $n = \frac{s}{\|s\|}$. In diesem Fall erhält man die *Hessesche Normalform*. Diese ist dann bis auf das Vorzeichen eindeutig bestimmt. (Das Vorzeichen kann man etwa noch durch die Bedingung $\langle n, v_1 \rangle > 0$ festlegen, falls $0 \notin A$ ist.)

Beweis.

(i) Es sei $s = (a, b)$, $u = \langle x, y \rangle$, $c = \langle s, v_1 \rangle$. Dann gilt

$$\langle s, u - v_1 \rangle = \langle s, u \rangle - \langle s, v_1 \rangle = ax + by - c.$$

Aus Satz (1.3) folgt zunächst, daß A eine Gerade ist. Da $\langle s, v_1 - v_1 \rangle = \langle s, 0 \rangle$ ist, ist $v_1 \in A$. Liegt $v \in A$ mit $v_1 \neq v$, so ist $w := v - v_1$ ein Richtungsvektor von A. Es gilt

$$\langle s, w \rangle = \langle s, v - v_1 \rangle = 0$$

und damit ist nach obiger Bemerkung s orthogonal zu A.

(ii) Es sei $v \in A$ mit $v \neq v_1$. Mit $w := v_1 - v$ gilt

$$A = v + \mathbb{R}w.$$

Da s orthogonal zu A ist, gilt $\langle s, w \rangle = 0$. Wir setzen

$$B := \{u \in \mathbb{R}^2;\ \langle s, u - v_1 \rangle = 0\}.$$

Nach (i) ist B eine Gerade. Offensichtlich ist $v_1 \in B$. Es ist auch $v \in B$, da $\langle s, v - v_1 \rangle = \langle s, -w \rangle = 0$. Da eine Gerade durch zwei Punkte eindeutig bestimmt ist, gilt $A = B$. □

Wir wollen nun den *Abstand* eines Punktes von einer Geraden bestimmen.

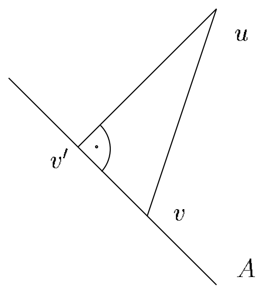


Abb. 26: Abstand eines Punktes von einer Geraden

Wir definieren diesen Abstand als

$$d(u, A) := \min\{d(u, v);\ v \in A\}.$$

(Daß dieses Minimum existiert, werden wir gleich sehen.)

Lemma 1.11
Es sei $A \subset \mathbb{R}^n$ eine Gerade und $u \in \mathbb{R}^n$. Dann gibt es genau ein $v' \in A$, so daß $u - v'$ orthogonal zu A ist. Ferner gilt

$$d(u, A) = d(u, v')$$

(d. h.der senkrechte Abstand ist der kürzeste).

Beweis. Es sei $A = v + \mathbb{R}w$. Für $v' = v + \lambda w$ ist $u - v'$ genau dann orthogonal zu A, wenn

$$0 = \langle u - v', w \rangle = \langle u - v - \lambda w, w \rangle = \langle u - v, w \rangle - \lambda \langle w, w \rangle.$$

Dies ist genau dann der Fall, wenn

$$\lambda = \frac{\langle u - v, w \rangle}{\|w\|^2}.$$

Damit ist λ und daher auch v' eindeutig bestimmt. Für beliebiges $v_1 \in A$ ist also $(v_1 - v') \perp (u - v')$ und nach dem Satz des Pythagoras gilt

$$\|v_1 - u\|^2 = \|v_1 - v'\|^2 + \|u - v'\|^2 \geq \|u - v'\|^2.$$

Damit wird das Minimum genau für $v' = v_1$ angenommen. □

Bemerkung. Wir haben damit eine Formel für $d(u, A)$ hergeleitet. Da

$$v' = v + \lambda w = v + \frac{\langle u - v, w\rangle}{\|w\|^2} w$$

ist, gilt

$$d(u, A) = d(u, v') = \|v' - u\|$$

also

$$\boxed{d(u, A) = \|v + \frac{\langle u - v, w\rangle}{\|w\|^2} w - u\|}$$

Ist $A \subset \mathbb{R}^2$ eine Gerade in der Ebene, so vereinfacht sich dies noch. Es sei $0 \neq s$ orthogonal zu A, $v_1 \in A$. Dann ist eine Normalengleichung von A gegeben durch

$$A = \{v \in \mathbb{R}^2;\ \langle s, v - v_1\rangle = 0\}.$$

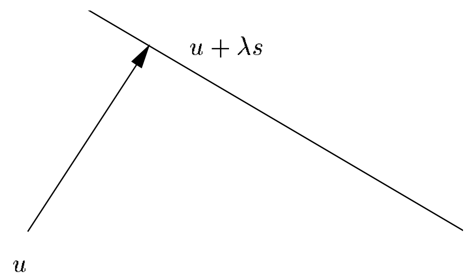


Abb. 27: Zur Bestimmung des Abstandes $d(u, A)$.

Wir suchen ein $\lambda \in \mathbb{R}$, so daß $u + \lambda s \in A$ ist. Dann gilt nämlich

$$d(u, A) = d(u, u + \lambda s) = \|\lambda s\| = |\lambda|\,\|s\|.$$

Die Bedingung $u + \lambda s \in A$ ist äquivalent zu

$$\langle s, u + \lambda s - v_1\rangle = 0 \Leftrightarrow \langle s, u - v_1\rangle + \lambda\langle s, s\rangle = 0.$$

Dies ist äquivalent zu

$$\lambda = \frac{\langle s, v_1 - u\rangle}{\|s\|^2}.$$

Damit gilt

$$\boxed{d(u, A) = \frac{|\langle s, u - v_1\rangle|}{\|s\|}}$$

Die Normalengleichung für A hat die Form

$$A = \{(x, y) \in \mathbb{R}^2;\ ax + by - c = 0\} \qquad (a, b) \neq (0, 0)$$

wobei $s = (a, b)$, $c = \langle s, v_1 \rangle$. Also gilt für $u = (x_0, y_0)$, daß

$$\boxed{d(u, A) = \frac{|ax_0 + by_0 - c|}{\sqrt{a^2 + b^2}}}$$

Ist s ein Normaleneinheitsvektor, so vereinfacht sich diese Gleichung wegen $\|s\| = \sqrt{a^2 + b^2} = 1$ zu

$$\boxed{d(u, A) = |ax_0 + by_0 - c|}$$

D. h.man kann den Abstand eines Punktes von einer Geraden durch Einsetzen in die Hesse-Normalform berechnen.

Es gibt auch eine Normalenform für Ebenen im $\mathbb{R}^3$.

Definition. Ein Vektor $s \in \mathbb{R}^n$ ist *orthogonal* zu einer Ebene $E \subset \mathbb{R}^n$, falls $\langle s, v_1 - v_2 \rangle = 0$ ist für alle $v_1, v_2 \in E$.

Ist $E = v + \mathbb{R}w_1 + \mathbb{R}w_2$ eine Parameterdarstellung von E, so ist s genau dann senkrecht zu E, wenn $s \perp w_1$ und $s \perp w_2$.

Analog wie bei Geraden hat man nun die folgenden Aussagen:

Satz 1.12

(i) Es sei $s \in \mathbb{R}^3$, $s \neq 0$ und $v_1 \in \mathbb{R}^3$. Dann ist

$$E := \{u \in \mathbb{R}^3;\ \langle s, u - v_1 \rangle = 0\}$$

eine Ebene durch v_1, die orthogonal zu s ist.

(ii) Ist umgekehrt E eine Ebene, die durch v_1 geht und orthogonal zu s ist, so ist

$$E = \{u \in \mathbb{R}^3;\ \langle s, u - v_1 \rangle = 0\}.$$

Für einen Punkt $u \in \mathbb{R}^n$ definiert man wieder den *Abstand* zu einer Ebene $E \subset \mathbb{R}^n$ durch

$$d(u, E) = \min\{d(u, v);\ v \in E\}.$$

Satz 1.13

Ist $E \subset \mathbb{R}^3$ eine Ebene durch v_1, die senkrecht zu $s \neq 0$ ist, so ist

$$d(u, E) = \frac{|\langle s, u - v_1 \rangle|}{\|s\|}.$$

Ist speziell

$$E = \{a_1x_1 + a_2x_2 + a_3x_3 - b = 0\}$$

und $u = (u_1, u_2, u_3)$, so ist

$$d(u, E) = \frac{|a_1u_1 + a_2u_2 + a_3u_3 - b|}{\sqrt{a_1^2 + a_2^2 + a_3^2}}.$$

§ 2 Algebraische Grundstrukturen

Definition. Eine *Verknüpfung* auf einer Menge G ist eine Abbildung

$$\begin{array}{rrcl} \circ : & G \times G & \longrightarrow & G \\ & (a,b) & \longmapsto & a \circ b. \end{array}$$

Schreibweise. $a \circ b$, $a \cdot b$, ab, $a + b$.

Beispiele.

(i) $G = \mathbb{N}$

$$\begin{array}{rrcl} \circ : & \mathbb{N} \times \mathbb{N} & \longrightarrow & \mathbb{N} \\ & (a,b) & \longmapsto & a + b. \end{array}$$

(ii) $G = \mathbb{R}$

$$\begin{array}{rrcl} \circ : & \mathbb{R} \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ & (a,b) & \longmapsto & a \cdot b. \end{array}$$

(iii) M sei eine beliebige Menge.

$$G := \mathrm{Abb}(M, M) = \{f;\ f : M \to M \text{ ist eine Abbildung}\}.$$

Als Verknüpfung wählen wir die Komposition von Abbildungen:

$$\begin{array}{rrcl} \circ : & G \times G & \longrightarrow & G \\ & (f,g) & \longmapsto & f \circ g. \end{array}$$

Definition. Eine *Halbgruppe* ist ein Paar $(G, \circ)$, so daß die Verknüpfung $\circ$ das Assoziativgesetz erfüllt, d. h.

(G1) $$(a \circ b) \circ c = a \circ (b \circ c) \qquad (\text{für alle } a, b, c \in G).$$

Alle oben angeführten Beispiele erfüllen das Assoziativgesetz.

Definition. Eine Halbgruppe $(G, \circ)$ heißt eine *Gruppe*, wenn gilt:

(G2) Es gibt ein Element $e \in G$ mit: $e \circ a = a$ für alle $a \in G$.

(G3) Zu jedem $a \in G$ gibt es ein $a' \in G$ mit $a' \circ a = e$.

Definition.

(i) e heißt *neutrales Element* (*linksneutrales Element*, *Einselement*).

(ii) a' heißt zu a *inverses Element (linksinverses Element)*.

Definition. Eine Gruppe $G, \circ$ heißt *kommutativ* (*abelsch*), falls

(G4) $$a \circ b = b \circ a \quad (a, b \in G).$$

Schreibweise. Bei abelschen Gruppen schreibt man in der Regel

$$a + b := a \circ b, \qquad 0 := e.$$

Beispiele.

(i) $(\mathbb{N}, +)$ ist *keine* Gruppe. Man hat zwar

$$0 + n = n,$$

für $n > 0$ gibt es aber kein n' mit

$$n' + n = 0.$$

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

(ii) $(\mathbb{R}, \cdot)$ ist *keine* Gruppe, da 0 kein inverses Element besitzt. Es ist zwar stets

$$1 \cdot x = x,$$

aber 0 besitzt kein multiplikatives inverses Element.

Es sei

$$\mathbb{R}^* := \{x \in \mathbb{R},\ x \neq 0\}.$$

Dann ist $(\mathbb{R}^*, \cdot)$ eine kommutative Gruppe.

(iii) $\mathrm{Abb}(M, M)$ ist keine Gruppe. Zwar gilt stets

$$\mathrm{id}_M \circ f = f,$$

im allgemeinen existiert aber kein Inverses.

(iv) Es sei

$$\mathrm{Bij}(M, M) = \{f;\ f : M \longrightarrow M \text{ ist bijektiv}\} \subset \mathrm{Abb}(M, M).$$

Dies ist eine Gruppe, denn nach Lemma (0.2) gibt es zu jedem f ein g mit

$$g \circ f = \mathrm{id}_M.$$

(v) Für den Spezialfall $M = \{1, \ldots, n\}$ definiert man

$$S_n := \mathrm{Bij}(M, M).$$

Definition. S_n heißt die *symmetrische Gruppe* in n *Variablen*. Die Elemente von S_n heißen *Permutationen* von n Elementen.

Bemerkung. Für $n \geq 3$ ist S_n nicht abelsch.

Um ein Gegenbeispiel zur Kommutativität anzugeben betrachten wir

$$\begin{aligned} f: & \quad \{1,\ldots,n\} \longrightarrow \{1,\ldots,n\} \\ f(1) &:= 1,\ f(2) := 3,\ f(3) := 2,\ f(n) = n \text{ für } n \geq 4 \\ g(1) &:= 3,\ g(2) := 2,\ g(3) := 1,\ g(n) = n \text{ für } n \geq 4. \end{aligned}$$

Dann ist

$$\begin{aligned} (f \circ g)(1) &= f(g(1)) = f(3) = 2 \\ (g \circ f)(1) &= g(f(1)) = g(1) = 3 \end{aligned}$$

also ist $f \circ g \neq g \circ f$.

Einfache Folgerungen aus den Gruppenaxiomen

Es sei $(G, \circ)$ eine Gruppe.

Lemma 2.1

(i) $a \circ e = a$ für alle $a \in G$.

(ii) Es gibt genau ein neutrales Element e.

(iii) Ist a' zu a invers, d. h. $a' \circ a = e$, so gilt auch $a \circ a' = e$.

(iv) Zu jedem a gibt es genau ein inverses Element.

Bemerkung.

(i) Man kann also von *dem* neutralen Element sprechen.

(ii) Ebenso kann man von *dem* inversen Element a' zu a sprechen. Man setzt $a^{-1} := a'$ (bzw. $-a := a'$ im abelschen Fall).

Schreibweise. Wir lassen im folgenden meist das Zeichen $\circ$ weg.

Beweis. (iii) Nach (G3) gibt es a'' mit $a''a' = e$.

$$\begin{aligned} aa' &\overset{\text{(G2)}}{=} e(aa') \overset{\text{Def. von } a''}{=} (a''a')(aa') \overset{\text{(G1)}}{=} a''(a'(aa')) \\ &\overset{\text{(G1)}}{=} a''((\underbrace{a'a}_{=e})a') = a''(ea') \overset{\text{(G2)}}{=} a''a' \\ &\overset{\text{Def.} a''}{=} e. \end{aligned}$$

(i) $ae \overset{\text{(G3)}}{=} a(a'a) \overset{\text{(G1)}}{=} (aa')a \overset{\text{(iii)}}{=} ea \overset{\text{(G2)}}{=} a.$

(ii) e' sei ein weiteres neutrales Element

$$e' \overset{\text{(i)}}{=} e'e \overset{\text{Def. von } e'}{=} e.$$

(iv) Es sei $a'a = e$, $a''a = e$.

$$a'' \overset{\text{(i)}}{=} a''e \overset{\text{(iii)}}{=} a''(aa') \overset{\text{(G1)}}{=} (a''a)a' \overset{\text{(Def. von } a'')}{=} ea' \overset{\text{(G2)}}{=} a'.$$

□

Lemma 2.2
$(G, \circ)$ sei eine Halbgruppe. Dann sind äquivalent:
(i) $(G, \circ)$ ist eine Gruppe.
(ii) Zu je zwei Elementen $a, b \in G$ gibt es genau ein $x \in G$ und ein $y \in G$ mit $xa = b, ay = b$.

Bemerkung. Die Aussage (ii) ist eine Aussage über die Lösbarkeit von Gleichungen in der Halbgruppe G.

Beweis. (i)⇒(ii): Die Idee ergibt sich aus folgender Rechnung. Falls $xa = b$, dann gilt:

$$b = xa \;\Rightarrow\; ba^{-1} = (xa)a^{-1} \overset{\text{(G1)}}{=} x(aa^{-1}) \overset{\text{(2.1)(iii)}}{=} xe \overset{\text{(2.1)(i)}}{=} x.$$

Dies zeigt, daß x, falls es existiert, eindeutig ist. Analog zeigt man die Eindeutigkeit von y.

Wir sehen also $x := ba^{-1}$. Dann gilt

$$xa = (ba^{-1})a \overset{\text{(G1)}}{=} b(a^{-1}a) \overset{\text{(G3)}}{=} be \overset{\text{(2.1)(i)}}{=} b.$$

Analog für $y := a^{-1}b$:

$$ay = a(a^{-1}b) \overset{\text{(G1)}}{=} (aa^{-1})b \overset{\text{(2.1)(iii)}}{=} eb \overset{\text{(G2)}}{=} b.$$

(ii)⇒(i): Es sind (G2) und (G3) zu prüfen.

(G2): Es sei $a \in G$ fest gewählt. Dann gibt es ein $e \in G$ mit $ea = a$. Wir haben also ein neutrales Element für a gefunden. Es sei nun $b \in G$ beliebig. Nach Voraussetzung gibt es ein y mit $ay = b$. Für dieses y gilt dann:

$$eb = e(ay) \overset{\text{(G1)}}{=} (ea)y \overset{\text{Konstr.von } e}{=} ay = b.$$

(G3): Zu $a \in G$ gibt es nach Voraussetzung ein $a' \in G$ mit $a'a = e$. □

Lemma 2.3
(i) $(a^{-1})^{-1} = a$

(ii) $(ab)^{-1} = b^{-1}a^{-1}$.

Beweis. (i) Nach (2.1) (iii) ist $aa^{-1} = e$. Wegen der Eindeutigkeit des Inversen (Lemma (2.1) (iv)) ist $a = (a^{-1})^{-1}$.

(ii) Es gilt $(b^{-1}a^{-1})(ab) \overset{(G1)}{=} b^{-1}(\underbrace{(a^{-1}a)}_{=e} b) \overset{(G2)}{=} b^{-1}b = e$.

Wiederum wegen der Eindeutigkeit des Inversen folgt $b^{-1}a^{-1} = (ab)^{-1}$. □

Untergruppen

Es sei $(G, \circ)$ eine Gruppe und $U \subset G$ eine nicht-leere Teilmenge.

Definition. $U \subset G$ heißt eine *Untergruppe* von G, falls gilt:
(i) U ist multiplikativ abgeschlossen, d. h.für $a, b \in U$ gilt $a \circ b \in U$.

(ii) U ist zusammen mit der von G induzierten Verknüpfung

$$\begin{array}{rcl} U \times U & \longrightarrow & U \\ (a, b) & \longmapsto & a \circ b, \end{array}$$

eine Gruppe.

Beispiele.
(i) $(\mathbb{Z}, +) \subset (\mathbb{Q}, +) \subset (\mathbb{R}, +)$

(ii) $(\mathbb{Q}^*, \cdot) \subset (\mathbb{R}^*, \cdot)$ (wobei $\mathbb{Q}^* := \{x \in \mathbb{Q};\ x \neq 0\}$).

Satz 2.4
Es sind äquivalent:
(i) $U \subset G$ ist Untergruppe.

(ii) Mit $u, v \in U$ ist stets auch $uv^{-1} \in U$.

Ist U eine Untergruppe von G, dann stimmen das neutrale Element von U und G überein, und für jedes $u \in U$ stimmen das Inverse von u in U und G überein.

Beweis. (i)⇒(ii) Es sei e^* das neutrale Element in U. Damit gilt, da e das neutrale Element in G ist

$$e^*u = u = eu.$$

Multiplizieren wir diese Gleichung mit dem inversen Element, u^{-1} von u in U folgt

$$(e^*u)u^{-1} = (eu)(u^{-1}) \ \Rightarrow\ e^*(uu^{-1}) = e(uu^{-1}) \ \Rightarrow\ e^* = e$$

wobei wir Lemma (2.1) (iii) und Lemma (2.1) (i) verwendet haben. Es sei nun v^{-1} das inverse Element von v in G und v' das inverse Element von v in U. Wir wollen zeigen, daß $v^{-1} = v'$ gilt. Dann haben wir auch $uv^{-1} \in U$ gezeigt. Es gilt in G:

$$v'v = e^* = e = v^{-1}v.$$

Wegen der Eindeutigkeit des Inversen nach Lemma (2.1) (iv) folgt $v' = v^{-1}$.

(ii)⇒(i) Es sei $u \in U$. Nach Voraussetzung gilt mit $v = u$, daß $e = uu^{-1} \in U$. Damit besitzt U ein neutrales Element. Ferner gilt wieder nach Voraussetzung, daß $u^{-1} = eu^{-1} \in U$, also besitzt u ein inverses Element in U. Es bleibt zu zeigen, daß U multiplikativ abgeschlossen ist. Seien hierzu $u, v \in U$. Dann gilt nach dem, was wir gerade gezeigt haben, auch $v^{-1} \in U$. Nach Voraussetzung gilt dann $uv = u(v^{-1})^{-1} \in U$. □

Homomorphismen

G, G' seien Gruppen. (Wir geben die Verknüpfungen nicht explizit an.)

Definition. Eine Abbildung $f : G \to G'$ heißt ein *(Gruppen-)homomorphismus*, falls gilt

(H) $$f(ab) = f(a)f(b) \qquad \text{(für alle } a, b \in G).$$

Beispiele.
(i) $G = G' = (\mathbb{Z}, +)$

$$\begin{array}{rccc} f: & \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ & n & \longmapsto & 3n. \end{array}$$

Es gilt: $f(n+m) = 3(n+m) = 3n + 3m = f(n) + f(m)$.

(ii) *Gegenbeispiel:*

$$\begin{array}{rccc} f: & \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ & n & \longmapsto & n+1. \end{array}$$

ist kein Homomorphismus, da gilt $f(n+m) = n+m+1$, aber $f(n) + f(m) = (n+1) + (m+1) = n+m+2$.

Definition. Ein Homomorphismus $f : G \to G'$ heißt
(i) *Monomorphismus*, falls f injektiv ist,

(ii) *Epimorphismus*, falls f surjektiv ist,

(iii) *Isomorphismus*, falls es einen Homomorphismus $g : G' \to G$ gibt mit $f \circ g = \mathrm{id}_{G'}$ und $g \circ f = \mathrm{id}_G$.

Lemma 2.5
Für einen Gruppenhomomorphismus $f : G \to G'$ sind äquivalent:
(i) f ist bijektiv.

(ii) f ist ein Gruppenisomorphismus.

Beweis. (ii)⇒(i) folgt aus Lemma (0.2).

(i)⇒(ii) Nach Lemma (0.2) gibt es eine Abbildung $f^{-1}: G' \to G$ mit $f^{-1} \circ f = \mathrm{id}_G, f \circ f^{-1} = \mathrm{id}_{G'}$. Sei $g := f^{-1}$. Zu zeigen ist, daß g ein Gruppenhomomorphismus ist, d. h.daß gilt:

$$f^{-1}(ab) = f^{-1}(a)f^{-1}(b).$$

Dies folgt aus der Injektivität von f, da gilt

$$f(f^{-1}(a)f^{-1}(b)) \overset{(\mathrm{H})}{=} f(f^{-1}(a))\, f(f^{-1}(b)) = ab, \qquad f(f^{-1}(ab)) = ab.$$

□

Es sei $f: G \to G'$ ein Gruppenhomomorphismus; e, e' seien die neutralen Elemente von G bzw. G'.

Definition.

(i) Der *Kern* von f ist definiert als

$$\mathrm{Ker}\, f = \{g \in G;\ f(g) = e'\}.$$

(ii) Das *Bild* von f ist

$$\mathrm{Im}\, f = \{f(g) \in G';\ g \in G\}.$$

(D. h.Im f ist das mengentheoretische Bild von f.)

Beispiel. Es sei

$$G' := C[a,b] := \{f;\ f: [a,b] \to \mathbb{R} \text{ ist eine stetige Funktion}\}.$$

Auf G' wird $f + g$ definiert durch

$$(f+g)(x) := f(x) + g(x).$$

Dann ist $(C[a,b], +)$ eine Gruppe. Ferner sei

$$G := C^1[a,b] := \{f;\ f: [a,b] \longrightarrow \mathbb{R} \text{ ist stetig differenzierbar}\}.$$

Dann ist G eine Untergruppe von G'. Die Abbildung

$$\begin{aligned} \frac{d}{dx}: \quad G &\longrightarrow G' \\ f &\longmapsto f' = \frac{df}{dx}. \end{aligned}$$

ist ein Homomorphismus, da

$$\frac{d}{dx}(f+g) = \frac{d}{dx}f + \frac{d}{dx}g.$$

Es ist

$$\operatorname{Ker}\frac{d}{dx} = \{f;\ f \equiv c\},$$

d. h.der Kern besteht genau aus der Menge der konstanten Funktionen.

Lemma 2.6
Es sei $f : G \to G'$ ein Gruppenhomomorphismus. Dann gilt:
(i) $f(e) = e'$.
(ii) $(f(a))^{-1} = f(a^{-1})$.
(iii) $\operatorname{Ker} f \subset G$ ist Untergruppe.
(iv) $\operatorname{Im} f \subset G'$ ist Untergruppe.
(v) f ist genau dann injektiv, wenn $\operatorname{Ker} f = \{e\}$.

Beweis. (i) $f(e)f(e) = f(ee) = f(e)$. Durch Multiplikation mit $f(e)^{-1}$ folgt $f(e) = e'$.

(ii) $f(a^{-1})f(a) = f(a^{-1}a) = f(e) \overset{\text{(i)}}{=} e'$. Also folgt $f(a^{-1}) = (f(a))^{-1}$.

(iii) $a, b \in \operatorname{Ker} f$. Nach Satz (2.4) genügt es zu zeigen, daß $ab^{-1} \in \operatorname{Ker} f$. Dies gilt, da

$$f(ab^{-1}) = f(a)f(b^{-1}) \overset{\text{(ii)}}{=} f(a)f(b)^{-1} = e'(e')^{-1} = e'e' = e'.$$

(iv) $u, v \in \operatorname{Im} f$. Zu zeigen ist, daß $uv^{-1} \in \operatorname{Im} f$. Da $u, v \in \operatorname{Im} f$ gibt es $a, b \in G$ mit $f(a) = u$, $f(b) = v$.

$$uv^{-1} = f(a)(f(b))^{-1} \overset{\text{(ii)}}{=} f(a)f(b^{-1}) = f(ab^{-1}) \in \operatorname{Im} f.$$

(v) Es ist $e \in \operatorname{Ker} f$ nach (i). $\Rightarrow$ Ist f injektiv, so folgt aus $e \in \operatorname{Ker} f$, daß $\operatorname{Ker} f = \{e\}$ gilt.

Umgekehrt sei f nicht injektiv. D. h.es gibt $a \neq b$ mit $f(a) = f(b)$. Wegen $a \neq b$ ist $ab^{-1} \neq e$. Aber $ab^{-1} \in \operatorname{Ker} f$, da

$$f(ab^{-1}) = f(a)f(b^{-1}) \overset{\text{(ii)}}{=} f(a)f(b)^{-1} = f(a)f(a)^{-1} = e'.$$

□

Ringe

R sei eine Menge. Ferner seien $+, \cdot$ Verknüpfungen auf R.

Definition. Das Tripel $(R, +, \cdot)$ heißt ein *Ring*, falls

(R1) $(R, +)$ ist eine abelsche Gruppe.
(R2) $(R, \cdot)$ ist eine Halbgruppe.

Für $r, s, t \in R$ gilt:

(R3)
$$\begin{aligned} (r+s)\cdot t &= r\cdot t + s\cdot t \\ r\cdot(s+t) &= r\cdot s + r\cdot t. \end{aligned}$$

Beispiele.
(i) $(\mathbb{Z}, +, \cdot)$
(ii) $(\mathbb{Q}, +, \cdot)$
(iii) $(\mathbb{R}, +, \cdot)$

Definition.
(i) Ein Ring heißt *kommutativ*, falls

$$r\cdot s = s\cdot r \qquad \text{(für alle } r, s \in R).$$

(ii) Man sagt R ist ein *kommutativer Ring* mit *Einselement*, falls es zusätzlich ein Element $1 \neq 0$ gibt (0 ist das neutrale Element von $(R, +)$) mit

$$1\cdot r = r \qquad \text{(für alle } r \in R).$$

Beispiele. Wie oben.

Schreibweise. Meist schreibt man rs statt $r\cdot s$.

Definition. R_1, R_2 seien Ringe. Eine Abbildung $f : R_1 \to R_2$ heißt ein *Ringhomomorphismus*, falls für alle $r, s \in R$ gilt:
(i) $f(r+s) = f(r) + f(s)$
(ii) $f(rs) = f(r)f(s)$.

Körper

Es sei K eine Menge und $+, \cdot$ seien Verknüpfungen auf K.

Definition. Das Tripel $(K, +, \cdot)$ heißt ein *Körper*, falls gilt:

(K1) $(K, +)$ ist abelsche Gruppe.

(K2) Ist 0 das neutrale Element von K, so setzt man

$$K^* := K \backslash \{0\}.$$

Für $a, b \in K^*$ gilt $a\cdot b \in K^*$ und $(K^*, \cdot)$ ist eine abelsche Gruppe.

(K3) Es gelten die *Distributivgesetze*:

$$\begin{aligned} a\cdot(b+c) &= (a\cdot b) + (a\cdot c) \\ (a+b)\cdot c &= (a\cdot c) + (b\cdot c). \end{aligned}$$

Beispiele.

(i) $(\mathbb{Q}, +, \cdot)$

(ii) $(\mathbb{R}, +, \cdot)$

(iii) $(\mathbb{F}_2, +, \cdot)$ Hierbei ist $\mathbb{F}_2 = \{0, 1\}$ und Addition und Multiplikation sind durch die folgenden Tafeln erklärt:

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

Schreibweise.

$$\begin{aligned} 0 &:= \text{neutrales Element von } (K, +) \\ 1 &:= \text{neutrales Element von } (K^*, \cdot) \quad (\text{Es gilt } 1 \neq 0) \\ -a &:= \text{inverses Element zu } a \text{ in } (K, +) \\ \frac{1}{a} &:= a^{-1} := \text{inverses Element von } (K^*, \cdot)(a \neq 0) \\ a - b &:= a + (-b) \\ \frac{a}{b} &:= ab^{-1} \qquad (b \neq 0). \end{aligned}$$

Lemma 2.7

(i) $0 \cdot a = a \cdot 0 = 0 \quad (a \in K)$.

(ii) Ist $a \cdot b = 0$, so folgt $a = 0$ oder $b = 0$ (Nullteilerfreiheit der Multiplikation.)

(iii) Ist $a \neq 0$, so gibt es genau eine Lösung von $a \cdot x = b$, nämlich $x = \frac{b}{a}$.

(iv) $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$, sowie $(-a) \cdot (-b) = a \cdot b$.

Beweis. (i) $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$. Also folgt $0 \cdot a = 0$. Die Aussage $a \cdot 0 = 0$ zeigt man analog.

(ii) Dies folgt sofort aus (K2).

(iii) Ist $b \neq 0$, so folgt dies aus Lemma (2.2) und (i). Ist $b = 0$, so folgt nach (K2) aus $a \cdot x = 0$, $a \neq 0$, daß $x = 0 \overset{\text{(i)}}{=} 0 \cdot \frac{1}{a} = \frac{0}{a}$.

(iv) $a \cdot b + a \cdot (-b) = a \cdot (b + (-b)) = a \cdot 0 \overset{\text{(i)}}{=} 0$. Also folgt $a \cdot (-b) = -(a \cdot b)$. Ferner gilt $(-a) \cdot b + a \cdot b = ((-a) + a) \cdot b = 0 \cdot b \overset{\text{(i)}}{=} 0$. Damit folgt auch $(-a) \cdot b = -(a \cdot b)$.

Schließlich folgt
$$(-a) \cdot (-b) = -(a \cdot (-b)) = -(-(a \cdot b)) = a \cdot b,$$
wobei wir beim letzten Gleichheitszeichen Lemma (2.3) (i) benutzt haben. □

Es seien K und K' Körper.

Definition.

(i) Eine Abbildung $f : K \to K'$ heißt ein *Körperhomomorphismus*, falls gilt

$$\begin{aligned} f(a+b) &= f(a)+f(b) \\ f(a \cdot b) &= f(a) \cdot f(b) \end{aligned}$$

(ii) f heißt *Körpermonomorphismus* (*-epimorphismus*), falls zusätzlich f injektiv (surjektiv) ist.

(iii) f heißt ein *Körperisomorphismus*, falls es einen Körperhomomorphismus $g : K' \to K$ gibt mit $f \circ g = \mathrm{id}_{K'}$ und $g \circ f = \mathrm{id}_K$.

Bemerkung. Wir verwenden hier mißbräuchlich dasselbe Zeichen für die Operationen in K und K'.

Bemerkung. Ein Körperhomomorphismus, der nicht identisch 0 ist, ist stets injektiv.

Lemma 2.8
Ein Körperhomomorphismus $f : K \to K'$ ist genau dann ein Körperisomorphismus, wenn f bijektiv ist.

Beweis. Genau wie in Lemma (2.5). □

Schreibweise. Man schreibt auch hier meist ab statt $a \cdot b$.

Der Körper $\mathbb{C}$ der komplexen Zahlen

Wir definieren

$$\mathbb{C} := \mathbb{R} \times \mathbb{R}$$

Hierauf definieren wir eine Addition und eine Multiplikation wie folgt:

Addition:

$$(a,b) + (a',b') := (a+a', b+b')$$

Mit dieser Definition gilt:

(K1) $(\mathbb{C}, +)$ ist eine abelsche Gruppe und es gilt

$$e = (0,0); \qquad -(a,b) = (-a,-b).$$

Multiplikation:

$$(a,b) \cdot (a',b') := (aa' - bb', ab' + a'b).$$

(K2): Zu zeigen: Ist $(a,b) \neq (0,0) \neq (a',b')$, so folgt $(a,b) \cdot (a',b') \neq (0,0)$.

Wir nehmen hierzu an, daß:

$$(a,b) \cdot (a',b') = (aa' - bb', ab' + a'b) = (0,0).$$

1. Fall: $a \neq 0$. Dann gilt:

$$aa' - bb' = 0 \Rightarrow a' = \frac{bb'}{a}. \tag{1}$$

Einsetzen ergibt:

$$ab' + a'b = 0 \overset{(1)}{\Rightarrow} ab' + \frac{b^2 b'}{a} = 0 \Rightarrow \underbrace{(a^2 + b^2)}_{\neq 0} b' = 0 \Rightarrow b' = 0.$$

Mit (1) folgt auch $a' = 0$, und damit $(a', b') = (0, 0)$.

2. Fall: Der Fall $b \neq 0$ kann analog behandelt werden.

Neutrales Element: Das neutrale Element ist $(1, 0)$, es gilt nämlich

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

Inverses Element: Das zu $(a, b) \neq (0, 0)$ inverse Element ist

$$(a, b)^{-1} = (\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}).$$

Es gilt nämlich

$$(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}) \cdot (a, b) = (\frac{a^2}{a^2 + b^2} + \frac{b^2}{a^2 + b^2}, \frac{ab}{a^2 + b^2} - \frac{ba}{a^2 + b^2})$$

Das Assoziativitätsgesetz für die Multiplikation und die Kommutativität rechnet man leicht direkt nach.

(K3) Distributivgesetze: Man rechnet sofort nach:

$$\begin{aligned} (a, b) \cdot ((a', b') + (a'', b'')) &= (a, b) \cdot (a' + a'', b' + b'') \\ &= (aa' + aa'' - bb' - bb'', ab' + ab'' + a'b + a''b). \end{aligned}$$

Andererseits ist

$$\begin{aligned} (a, b) \cdot (a', b') + (a, b) \cdot (a'', b'') &= (aa' - bb', ab' + a'b) + (aa'' - bb'', ab'' + a''b) \\ &= (aa' - bb' + aa'' - bb'', ab' + a'b + ab'' + a''b). \end{aligned}$$

Das zweite Distributivgesetz folgt analog.

Definition. $\mathbb{C}$ heißt der Körper der *komplexen Zahlen*.

Wir betrachten die Abbildung

$$\begin{aligned} \varphi : \mathbb{R} &\longrightarrow \mathbb{C} \\ a &\longmapsto (a, 0). \end{aligned}$$

Dies ist ein *Körpermonomorphismus*: Die Injektivität ist klar. Es gilt außerdem:

$$\begin{aligned}\varphi(a+b) &= (a+b,0) = (a,0)+(b,0) = \varphi(a)+\varphi(b)\\ \varphi(a\cdot b) &= (a\cdot b,0) = (a,0)\cdot(b,0) = \varphi(a)\cdot\varphi(b).\end{aligned}$$

Mittels φ wird der Körper der reellen Zahlen in den Körper der komplexen Zahlen eingebettet.

Das Element

$$i := (0,1)$$

heißt *imaginäre Einheit*. Es gilt

$$i^2 = (0,1)\cdot(0,1) = (0-1,0) = (-1,0) = -(1,0) = \varphi(-1).$$

Identifiziert man $\mathbb{R}$ mit $\varphi(\mathbb{R})$, so schreibt sich die letzte Formel in der Form

$$\boxed{i^2 = -1}.$$

Gaußsche Zahlenebene

Man kann $\mathbb{C}$ mit der Ebene $\mathbb{R}^2$ identifizieren. Üblicherweise verwendet man die Schreibweise

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

Die reellen Zahlen sind durch $b=0$ charakterisiert.

Geometrische Darstellung

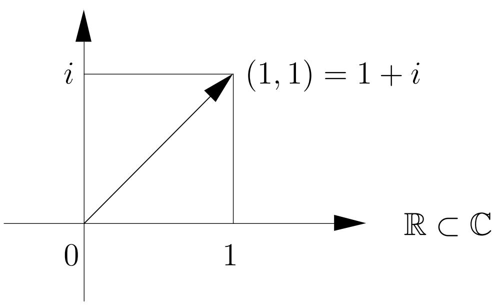


Abb. 28: Die komplexe Zahlenebene

Addition und Multiplikation drücken sich mit obiger Schreibweise wie folgt aus:

$$(2)\qquad (a+bi)+(c+di) = (a+c)+(b+d)i$$

$$(3)\qquad (a+bi)\cdot(c+di) = ac+adi+bci-bd = (ac-bd)+(ad+bc)i.$$

Definition. Es sei $z = a + bi$ mit $a, b \in \mathbb{R}$. Dann heißt a der *Realteil* von z und b der *Imaginärteil* von z.

Schreibweise. $a = \operatorname{Re} z$, $b = \operatorname{Im} z$.

Definition. Ist $z = a + bi$ eine komplexe Zahl, so heißt

$$\overline{z} := a - bi$$

die zu z konjugiert komplexe Zahl.

Bemerkung. Es gilt $\overline{z + w} = \overline{z} + \overline{w}$, $\overline{zw} = \overline{z}\,\overline{w}$, $\overline{\overline{z}} = z$, $z = \overline{z} \Leftrightarrow z \in \mathbb{R}$.

Definition. Ist $z = a + bi$, so heißt

$$|z| := \sqrt{z\overline{z}} = \sqrt{(a + bi)(a - bi)} = \sqrt{a^2 + b^2}$$

der *Betrag* von z. Der Betrag von $z = a + bi$ ist also die Länge des Vektors (a, b).

In vielen Fällen ist es auch nützlich, komplexe Zahlen durch Polarkoordianten darzustellen. Eine Zahl $0 \neq z \in \mathbb{C} = \mathbb{R}^2$ wird durch ihren Betrag $r = |z|$ und den Winkel φ zur x-Achse eindeutig bestimmt ($\varphi \in [0, 2\pi[$):

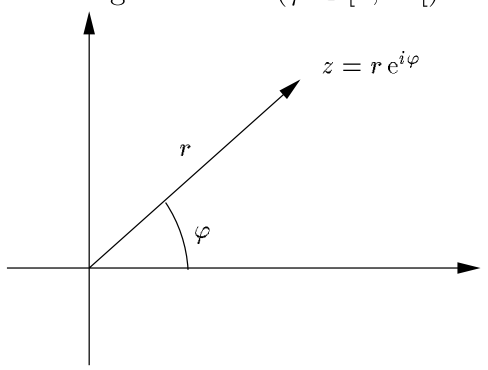


Abb. 29: Polarkoordinaten zur Beschreibung der komplexen Zahlenebene

Setzt man

$$\mathrm{e}^{i\varphi} := \cos\varphi + i \sin\varphi,$$

so ist

$$z = r\,\mathrm{e}^{i\varphi}\,.$$

Polarkoordinaten eignen sich besonders gut zur Beschreibung der Multiplikation. Mit

$$z = r\,\mathrm{e}^{i\varphi}, \quad w = r'\,\mathrm{e}^{i\psi}$$

gilt

$$\begin{aligned}
zz' &= (r\,\mathrm{e}^{i\varphi})(r'\,\mathrm{e}^{i\psi}) \\
&= rr'(\cos\varphi + i\sin\varphi)(\cos\psi + i\sin\psi) \\
&= rr'[(\cos\varphi\cos\psi - \sin\varphi\sin\psi) + i(\sin\varphi\cos\psi + \cos\varphi\sin\psi)] \\
&= rr'(\cos(\varphi + \psi) + i\sin(\varphi + \psi)) \\
&= rr'\,\mathrm{e}^{i(\varphi+\psi)}\,.
\end{aligned}$$

Bemerkung. In der Analysis wird oft zuerst die komplexe Exponentialfunktion e^z eingeführt und Sinus und Cosinus werden dann durch

$$e^{i\varphi} = \cos\varphi + i\sin\varphi,$$

bzw. durch ihre Reihenentwicklungen erklärt.

§ 3 Der Begriff des Vektorraums

Im folgenden sei stets K ein Körper und V eine Menge.

Gegeben seien Verknüpfungen

$$+: \begin{array}{ccc} V \times V & \longrightarrow & V \\ (x,y) & \longmapsto & x+y \end{array} \quad \text{und} \quad \cdot: \begin{array}{ccc} K \times V & \longrightarrow & V \\ (\alpha,x) & \longmapsto & \alpha \cdot x =: \alpha x. \end{array}$$

Definition. $(V,+,\cdot)$ heißt ein *Vektorraum über* K (K*-Vektorraum*), falls gilt:

(V1) $(V,+)$ ist eine abelsche Gruppe

Für $\alpha,\beta \in K$, $x,y \in V$ gilt:

(V2)
$$\begin{array}{rrcl} \text{(i)} & (\alpha+\beta)\cdot x & = & \alpha \cdot x + \beta \cdot x \\ \text{(ii)} & \alpha \cdot (x+y) & = & \alpha \cdot x + \alpha \cdot y \\ \text{(iii)} & (\alpha \cdot \beta) \cdot x & = & \alpha \cdot (\beta \cdot x) \\ \text{(iv)} & 1 \cdot x & = & x \qquad (1 = 1_K). \end{array}$$

Bezeichnungen.

(i) Die Elemente $\alpha, \beta, \ldots \in K$ heißen *Skalare*, K heißt *Skalarenkörper*.

(ii) Die Elemente $x, y, \ldots \in V$ heißen *Vektoren*.

(iii) Wir bezeichnen mit $0 = 0_K$ $(1 = 1_K)$ das additive (multiplikative) neutrale Element in K. Das neutrale Element in V wird mit $0 = 0_V$ bezeichnet.

(iv) Man schreibt oft αx statt $\alpha \cdot x$.

Lemma 3.1
V sei ein K-Vektorraum. Für $x \in V, \alpha \in K$ gilt:

(i) $0_K x = 0_V$.

(ii) $\alpha\, 0_V = 0_V$.

(iii) $\alpha x = 0_V \;\Rightarrow\; \alpha = 0_K$ oder $x = 0_V$.

(iv) $(-\alpha)x = \alpha(-x) = -(\alpha x)$.

Beweis.

(i) $0_K x = (0_K + 0_K)x = 0_K x + 0_K x$, also $0_K x = 0_V$.

(ii) $\alpha\, 0_V = \alpha(0_V + 0_V) \overset{\text{(V2)(ii)}}{=} \alpha\, 0_V + \alpha\, 0_V$, also $\alpha\, 0_V = 0_V$.

(iii) Es sei $\alpha \neq 0$. Dann gilt:

$$\alpha x = 0_V \Rightarrow \alpha^{-1}(\alpha x) = \alpha^{-1} 0_V \overset{\text{(ii)}}{=} 0_V \Rightarrow 0_V = \alpha^{-1} 0_V = \alpha^{-1}(\alpha x) = 1x \overset{\text{(V2)(iv)}}{=} x.$$

(iv) $(-\alpha)x + \alpha x \overset{\text{(V2)(i)}}{=} ((-\alpha) + \alpha))x = 0_K x \overset{\text{(i)}}{=} 0_V$.

Also folgt $(-\alpha)x = -(\alpha x)$.

Ferner gilt $\alpha(-x) + \alpha(x) \overset{\text{(V2)(ii)}}{=} \alpha((-x) + x) = \alpha\, 0_V \overset{\text{(ii)}}{=} 0_V$, und damit $\alpha(-x) = -(\alpha x)$. □

Schreibweise. Man schreibt meist $0_K = 0$ und $0_V = 0$.

Beispiele.

(1) Standardbeispiel: Für einen Körper K setzen wir

$$\begin{aligned} K^n &:= \underbrace{K \times \cdots \times K}_{n\text{-mal}} \qquad \text{mit} \\ x &= (x_1, \ldots, x_n) \in K^n \quad (x_i \in K). \end{aligned}$$

Im Fall $K = \mathbb{R}$ erhalten wir das früher beschriebene Beispiel $\mathbb{R}^n$.

Addition: $x + y = (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$.

Skalarmultiplikation: $\alpha x = \alpha(x_1, \ldots, x_n) := (\alpha x_1, \ldots, \alpha x_n)$.

(V1): $(K^n, +)$ ist abelsche Gruppe mit

$$0 = (0, \ldots, 0), \quad -x = (-x_1, \ldots, -x_n).$$

Von den Vektorraumaxiomen rechnen wir z. B. (V2) (i) nach:

$$\begin{aligned} (\alpha + \beta)x &= (\alpha + \beta)(x_1, \ldots, x_n) = ((\alpha + \beta)x_1, \ldots, (\alpha + \beta)x_n)) \\ &= (\alpha x_1 + \beta x_1, \ldots, \alpha x_n + \beta x_n) \\ &= (\alpha x_1, \ldots, \alpha x_n) + (\beta x_1, \ldots, \beta x_n) \\ &= \alpha x + \beta x. \end{aligned}$$

Die anderen Regeln verifiziert man analog.

Man nennt K^n den n-dimensionalen Raum der *Zeilenvektoren* über dem Körper K. Man hat auch die Möglichkeit, mit Spaltenvektoren zu arbeiten.

$${}^nK := \{x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix};\ x_i \in K\}.$$

Man bezeichnet dann ${}^nK$ als den Raum der *Spaltenvektoren*.

Addition: $x + y = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$.

Skalarmultiplikation: $\alpha x = \alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}$.

Mittels der Identifikation $(x_1, \ldots, x_n) \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ werden wir meist K^n und ${}^nK$ identifizieren.

(2) Abbildungsräume:

(a) Man kann Vektorräume von Abbildungen betrachten. Für einen Körper K und eine Menge M definieren wir:

$$\mathrm{Abb}(M, K) := \{f;\ M \longrightarrow K \text{ ist eine Abbildung}\}$$

Auf $\mathrm{Abb}(M, K)$ führen wir wie folgt eine Addition und eine Skalarmultiplikation ein:

Addition: $f + g$ wird definiert durch

$$(f + g)(x) := f(x) + g(x) \quad (x \in M).$$

Skalarmultiplikation: Für $\alpha \in K$ definieren wir

$$(\alpha f)(x) := \alpha f(x).$$

(V1): $(\mathrm{Abb}(M, K), +)$ ist eine abelsche Gruppe:

$$\begin{array}{rlrcl} 0 & \text{wird definiert durch} & 0(x) & := & 0 \text{ für alle } x \in M. \\ -f & \text{wird definiert durch} & (-f)(x) & := & -(f(x)). \end{array}$$

(V2): Wir zeigen beispielsweise: $\alpha(f + g) = \alpha f + \alpha g$.

Dies ist leicht nachzurechnen:

$$\begin{aligned} (\alpha(f + g))(x) &= \alpha((f + g)(x)) = \alpha(f(x) + g(x)) \\ &= \alpha f(x) + \alpha g(x) = (\alpha f)(x) + (\alpha g)(x) \\ &= (\alpha f + \alpha g)(x). \end{aligned}$$

Die anderen Gesetze verifiziert man analog.

(b) Spezialfall

$$K = \mathbb{R}, \quad M := [a, b] \subset \mathbb{R}.$$
$$\mathrm{Abb}(M, K) = \mathrm{Abb}([a, b], \mathbb{R}) = \{f;\ f : [a, b] \longrightarrow \mathbb{R}\}.$$

Dies kann man spezialisieren:

$$\begin{aligned} C[a, b] &:= \{f;\ f : [a, b] \longrightarrow \mathbb{R} \text{ stetig}\} \subset \mathrm{Abb}(M, \mathbb{R}) \\ C^1[a, b] &:= \{f;\ f : [a, b] \longrightarrow \mathbb{R} \text{ stetig diff.bar}\} \\ C^\infty[a, b] &:= \{f;\ f : [a, b] \longrightarrow \mathbb{R} \text{ beliebig oft diff.bar}\} \end{aligned}$$

Die Verknüpfungen sind immer wie oben definiert.

(3) Der Raum $\mathrm{Abb}(\mathbb{N},\mathbb{R})$ ist der Vektorraum der reellen Folgen.

(4) $K=\mathbb{R}$, $V=(\mathbb{C},+)$

$$\begin{array}{rcl} K\times V=\mathbb{R}\times\mathbb{C} & \longrightarrow & \mathbb{C} \\ (\alpha,x) & \longmapsto & \alpha x. \end{array}$$

(D. h.man faßt $\mathbb{C}$ als $\mathbb{R}$-Vektorraum auf.)

(5) $K=\mathbb{Q}$, $V=(\mathbb{R},+)$

$$\begin{array}{rcl} K\times V=\mathbb{Q}\times\mathbb{R} & \longrightarrow & \mathbb{R} \\ (\alpha,x) & \longmapsto & \alpha x. \end{array}$$

(D. h.man faßt $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum auf.)

(6) $K=\mathbb{R}$, $V=\mathbb{R}[x]=\{a_0+\cdots+a_nx^n;\ a_i\in\mathbb{R}\}$.

V heißt der *Raum der Polynome* in der Variablen x. Für $\alpha\in\mathbb{R}$ und

$$\begin{array}{c} P=P(x)=a_0+\cdots+a_nx^n \\ Q=Q(x)=b_0+\cdots+b_mx^m \end{array}$$

definieren wir (falls $n\leq m$, ansonsten analog):

$$\begin{array}{rcl} P+Q & := & (a_0+b_0)+\cdots+(a_n+b_n)x^n+\cdots+b_mx^m, \\ \alpha P & := & (\alpha a_0)+\cdots+(\alpha a_n)x^n. \end{array}$$

V sei ein K-Vektorraum.

Definition. Eine Teilmenge $U\subset V$ heißt ein *Untervektorraum*, falls:

(U0) $$U\neq\emptyset$$

(U1) $$u,v\in U\Rightarrow u+v\in U$$

(U2) $$\alpha\in K,u\in U\Rightarrow \alpha u\in U.$$

Lemma 3.2
U ist zusammen mit den eingeschränkten Verknüpfungen $U\times U\to U$; $(u,v)\mapsto u+v$ bzw. $K\times U\to U$; $(\alpha,u)\mapsto\alpha u$ selbst ein Vektorraum.

Beweis. (V1): $0\in U$, denn da $U\neq\emptyset$ gibt es ein $u\in U$. Also ist $0\,u\in U$ nach (U2). Nach Lemma (3.1) (i) ist aber $0\,u=0\in V$. Mit $u\in U$ ist auch $(-1)u\in U$ nach (U2). Nach Lemma (3.1) (iv) ist $(-1)u=-u$. Damit ist $(U,+)$ eine abelsche Gruppe.

(V2): Die Rechenregeln (i)–(iv) gelten in U, da sie bereits in V gelten. □

Lemma 3.3
V sei ein K-Vektorraum.

(i) Sind U_1 und U_2 Unterräume, so ist auch $U_1 \cap U_2$ ein Unterraum.

(ii) Sind $(U_i)_{i \in I}$ Unterräume, so ist auch $\bigcap_{i \in I} U_i$ ein Unterraum.

Beweis. Es genügt (ii) zu zeigen.

(U0): Da $0 \in U_i$ für alle $i \in I$, folgt $0 \in \bigcap_{i \in I} U_i$, also ist $\bigcap_{i \in I} U_i \neq \emptyset$.

(U1): $u, v \in \bigcap_{i \in I} U_i \Rightarrow u, v \in U_i$ für alle $i \in I \Rightarrow u+v \in U_i$ für alle $i \in I \Rightarrow u+v \in \bigcap_{i \in I} U_i$.

(U2): $\alpha \in K, u \in \bigcap_{i \in I} U_i \Rightarrow u \in U_i$ für alle $i \in I \Rightarrow \alpha u \in U_i$ für alle $i \in I \Rightarrow \alpha u \in \bigcap_{i \in I} U_i$.

□

Bemerkung. Sind U_1, U_2 Unterräume, so ist im allgemeinen $U_1 \cup U_2$ kein Unterraum von V.

Beispiel. $V = \mathbb{R}^2$

$$\begin{aligned} U_1 &= \{(x, 0);\ x \in \mathbb{R}\} \subset \mathbb{R}^2 \\ U_2 &= \{(0, y);\ y \in \mathbb{R}\} \subset \mathbb{R}^2 \end{aligned}$$

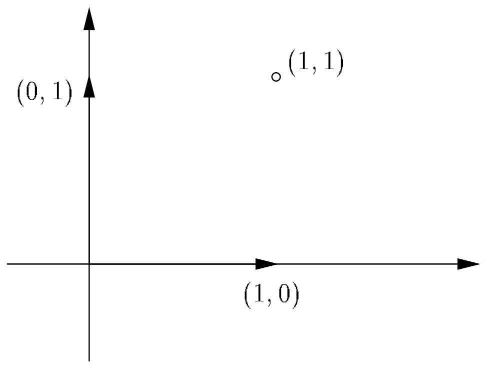


Abb. 30: Die Vereinigung von Unterräumen ist i. a. kein Unterraum

Offensichtlich sind U_1 und U_2 Unterräume mit $(1, 0) \in U_1$ und $(0, 1) \in U_2$. Es ist aber $(1, 0) + (0, 1) = (1, 1) \notin U_1 \cup U_2$.

Definition. Die *Summe* zweier Unterräume U_1 und U_2 ist definiert als

$$U_1 + U_2 := \{u;\ u = u_1 + u_2 \text{ mit } u_1 \in U_1, u_2 \in U_2\} \subset V.$$

Lemma 3.4
Sind U_1, U_2 Unterräume von V, so ist auch $U_1 + U_2$ ein Unterraum von V.

Beweis. (U0): $0 \in U_1, 0 \in U_2 \Rightarrow 0 + 0 = 0 \in U_1 + U_2$.

(U1): $\left.\begin{array}{l} u \in U_1 + U_2 \\ v \in U_1 + U_2 \end{array}\right\} \Rightarrow \begin{array}{l} u = u_1 + u_2;\ u_i \in U_i \\ v = v_1 + v_2;\ v_i \in U_i \end{array}$

Also ist

$$u + v = \underbrace{(u_1 + v_1)}_{\in U_1} + \underbrace{(u_2 + v_2)}_{\in U_2} \in U_1 + U_2.$$

(U2): Es sei $u \in U_1 + U_2$, d. h.$u = u_1 + u_2$ mit $u_i \in U_i$.

Dann gilt:

$$\alpha u = \alpha(u_1 + u_2) = (\alpha u_1) + (\alpha u_2) \in U_1 + U_2.$$

$\square$

Bemerkung. $U_1 + U_2$ ist der kleinste Unterraum von V, der $U_1 \cup U_2$ enthält.

Beispiel. Im obigen Beispiel ist $U_1 + U_2 = \mathbb{R}^2$.

§ 4 Elementare Vektorraumtheorie

Im folgenden sei K stets ein Körper.

Definition.
(i) Eine *homogene Gleichung* in den Unbekannten $\xi_1, \dots, \xi_n$ ist ein Ausdruck der Gestalt

$$\alpha_1\xi_1 + \cdots + \alpha_n\xi_n = 0. \qquad (\alpha_i \in K) \tag{1}$$

(ii) Ein *homogenes Gleichungssystem* von m Gleichungen in den Unbekannten $\xi_1, \dots, \xi_n$ besteht aus m homogenen Gleichungen

$$\begin{array}{c} \alpha_{11}\xi_1 + \cdots + \alpha_{1n}\xi_n = 0 \\ \alpha_{21}\xi_1 + \cdots + \alpha_{2n}\xi_n = 0 \\ \vdots \qquad\qquad \vdots \\ \alpha_{m1}\xi_1 + \cdots + \alpha_{mn}\xi_n = 0. \end{array} \qquad (\alpha_{ij} \in K; 1 \leq i \leq m, 1 \leq j \leq n) \tag{2}$$

(iii) Eine *Lösung* von (2) ist ein Vektor $x = (x_1, \dots, x_n) \in K^n$ mit

$$\begin{array}{c} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n = 0 \\ \alpha_{21}x_1 + \cdots + \alpha_{2n}x_n = 0 \\ \vdots \qquad\qquad \vdots \\ \alpha_{m1}x_1 + \cdots + \alpha_{mn}x_n = 0. \end{array}$$

Bemerkung. Es gibt stets die *triviale* Lösung $0 = (0, \dots, 0)$.

Definition. Die Menge $L := \{x \in K^n;\ x \text{ ist Lösung von (2)}\}$ heißt der *Lösungsraum* des Gleichungssystems (2).

Lemma 4.1
$L \subset K^n$ ist Unterraum.

Beweis. (U0): $L \neq \emptyset$, da $0 \in L$.

(U1): $x, y \in L$. Dann gilt:

$$\begin{aligned} \sum_{j=1}^{n} \alpha_{ij}x_j &= 0 \qquad (i = 1, \dots, m) \\ \sum_{j=1}^{n} \alpha_{ij}y_j &= 0 \qquad (i = 1, \dots, m). \end{aligned}$$

Damit

$$0 = \sum_{j=1}^{n} \alpha_{ij}x_j + \sum_{j=1}^{n} \alpha_{ij}y_j = \sum_{j=1}^{n} \alpha_{ij}(x_j + y_j) \qquad (i = 1, \dots, m),$$

also gilt $(x+y) \in L$.

(U2):

$$\begin{aligned} x \in L \Rightarrow \sum_{j=1}^{n} \alpha_{ij} x_j = 0 \quad &\Rightarrow \quad \alpha(\sum_{j=1}^{n} \alpha_{ij} x_j = 0) \\ &\Rightarrow \quad \sum_{j=1}^{n} \alpha_{ij} (\alpha x_j) = 0 \Rightarrow \alpha x \in L. \end{aligned}$$

□

Lemma 4.2 (Fundamentallemma)
Es sei $m < n$. Dann besitzt jedes homogene lineare Gleichungssystem von m Gleichungen in n Unbekannten eine nichttriviale Lösung.

Beweis. Wir betrachten

$$\begin{aligned} \alpha_{11}\xi_1 + \cdots + \alpha_{1n}\xi_n &= 0 \\ \vdots \qquad\quad & \vdots \\ \alpha_{m1}\xi_1 + \cdots + \alpha_{mn}\xi_n &= 0. \end{aligned}$$

1. Schritt: Sind alle $\alpha_{ij} = 0$, so ist jedes $x \in K^n$ Lösung, und man ist fertig. Es sei also ein $\alpha_{ij} \neq 0$. Nach Umnumerieren der Gleichungen bzw. Unbekannten kann man annehmen, daß $\alpha_{11} \neq 0$ ist.

2. Schritt: Man kann annehmen, daß $m = n-1$ ist. (Gilt die Aussage für $m = n-1$, so auch für $m < n-1$.)

3. Schritt: Wir können $\alpha_{21} = \cdots = \alpha_{n-1,1} = 0$ annehmen. Dazu betrachten wir

$$\begin{array}{ccccccccl} \alpha_{11}\xi_1 & + & \alpha_{12}\xi_2 & + & \cdots & + & \alpha_{1n}\xi_n & = 0 & (\text{mit } \alpha_{11} \neq 0) \\ \alpha_{21}\xi_1 & + & \alpha_{22}\xi_2 & + & \cdots & + & \alpha_{2n}\xi_n & = 0 & \\ \vdots & & & & \vdots & & & & \\ \alpha_{n-1,1}\xi_1 & + & \alpha_{n-1,2}\xi_2 & + & \cdots & + & \alpha_{n-1,n}\xi_n & = 0. & \end{array}$$

Dieses Gleichungssystem hat dieselbe Lösungsmenge wie

$$\begin{array}{ccccccccc} \alpha_{11}\xi_1 & + & \alpha_{12}\xi_2 & + & \cdots & + & \alpha_{1n}\xi_n & = 0 \\ \underbrace{(\alpha_{21} - \alpha_{11}\frac{\alpha_{21}}{\alpha_{11}})}_{=0}\xi_1 & + & (\alpha_{22} - \alpha_{12}\frac{\alpha_{21}}{\alpha_{11}})\xi_2 & + & \cdots & + & (\alpha_{2n} - \alpha_{1n}\frac{\alpha_{21}}{\alpha_{11}})\xi_n & = 0 \\ \vdots & & & & \vdots & & & \\ \alpha_{n-1,1}\xi_1 & + & \cdots & & \cdots & + & \alpha_{n-1,n}\xi_n & = 0. \end{array}$$

Analog verfährt man mit den anderen Gleichungen. Dies führt auf ein homogenes Gleichungssystem der Form:

$$(*)\qquad \begin{array}{ccccccc} \alpha_{11}\xi_1 & + & \alpha_{12}\xi_2 & + \cdots + & \alpha_{1n}\xi_n & = 0 \\ & & \alpha'_{22}\xi_2 & + \cdots + & \alpha'_{2n}\xi_n & = 0 \\ & & \vdots & \vdots & & \\ & & \alpha'_{n-1,2}\xi_2 & + \cdots + & \alpha'_{n-1,n}\xi_n & = 0. \end{array}$$

Die letzten $n-2$ Gleichungen sind Gleichungen in den $n-1$ Unbekannten $\xi_2, \dots, \xi_n$.

4. Schritt: Wir beweisen nun die Lösbarkeit mittels vollständiger Induktion nach n. Für $n = 2$ ist die Aussage klar, wir haben die nicht-triviale Lösung $(-\alpha_{12}, \alpha_{11})$.

Induktionsschritt: Es genügt ein Gleichungssystem der obigen Gestalt anzusehen. Nach Induktionsvoraussetzung besitzt das aus $(n-2)$ Gleichungen bestehende System $(*)$ eine Lösung $(x_2, \dots, x_n) \neq 0$. Wir setzen

$$x_1 := -\frac{1}{\alpha_{11}}(\alpha_{12}x_2 + \cdots + \alpha_{1n}x_n).$$

Dann ist

$$\alpha_{11}x_1 + \cdots + \alpha_{1n}x_n = 0$$

und $0 \neq (x_1, \dots, x_n)$ ist eine nicht-triviale Lösung des ursprünglichen Gleichungssystems.

□

Erzeugung von Unterräumen

Es sei V ein K-Vektorraum und $v_1, \dots, v_n \in V$ seien Vektoren.

Definition. Man nennt

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n \qquad (\alpha_i \in K;\ i = 1, \dots, n)$$

eine *Linearkombination* von $v_1, \dots, v_n$. Die Linearkombination heißt *trivial* falls $\alpha_1 = \cdots = \alpha_n = 0$ ist.

Es sei $A \subset V$ eine Teilmenge.

Definition. Man nennt

$$\operatorname{Span} A := \{v \in V;\ v = \alpha_1 v_1 + \cdots + \alpha_n v_n,\ v_i \in A, \alpha_i \in K\}$$

den *Spann* oder das *Erzeugnis* von A.

Bemerkung.

(i) $\operatorname{Span} A$ ist ein *Unterraum* von V.

(ii) $\operatorname{Span} A$ ist der *kleinste* Unterraum, der A enthält.

(iii) Ist $A = \{v_1, \dots, v_n\}$, so schreibt man auch

$$\operatorname{Span} A = \operatorname{Span}(v_1, \dots, v_n).$$

Regeln.

(i) $\operatorname{Span}\{v\} = Kv$.

(ii) $A \subset B \Rightarrow \operatorname{Span} A \subset \operatorname{Span} B$.

(iii) $A = \operatorname{Span} A \Leftrightarrow A$ ist Unterraum von V.

(iv) $\operatorname{Span}(\operatorname{Span} A) = \operatorname{Span} A$.

Konvention. $\operatorname{Span} \emptyset := \{0\}$.

Definition. Eine Teilmenge $E \subset V$ heißt *Erzeugendensystem* von V, falls $V = \operatorname{Span} E$ ist.

Bemerkung. Jeder Vektorraum V besitzt ein Erzeugendensystem, etwa $E = V$.

Definition. Ein Vektorraum V heißt *endlich erzeugt*, falls es eine endliche Menge $E \subset V$ mit $V = \operatorname{Span} E$ gibt.

Beispiele.

(1) $V = K^n$. Dann ist

$$V = \operatorname{Span}(e_1, \dots, e_n) \quad \text{mit } e_i = (0, \dots, \overset{\substack{i\text{-te Stelle} \\ \downarrow}}{1}, 0, \dots, 0),$$

da für $x = (x_1, \dots, x_n) \in V$ gilt

$$x = x_1 e_1 + \dots + x_n e_n.$$

(2) Der Raum $V = \mathbb{R}[x]$ der Polymone in einer Variablen ist *nicht* endlich erzeugt, da es Polynome von beliebig hohem Grad gibt.

Lineare Abhängigkeit/Unabhängigkeit

Definition. Die Vektoren $v_1, \dots, v_n \in V$ heißen *linear abhängig*, falls es $\alpha_1, \dots, \alpha_n \in K$ gibt, so daß nicht alle $\alpha_i = 0$ sind und

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0.$$

Ansonsten heißen $v_1, \dots, v_n$ *linear unabhängig*.

Bemerkung.

(i) Linear unabhängig bedeutet also:

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0 \Rightarrow \alpha_1 = \dots = \alpha_n = 0.$$

(ii) Ein Vektor v_1 ist linear abhängig genau dann wenn $v_1 = 0$.

(iii) Zwei Vektoren v_1, v_2 sind linear abhängig, genau wenn es $(\alpha_1, \alpha_2) \neq (0,0)$ gibt mit:

$$\alpha_1 v_1 + \alpha_2 v_2 = 0.$$

Ist $\alpha_1 \neq 0$, so folgt $v_1 = -\frac{\alpha_2}{\alpha_1} v_2$. Ist $\alpha_1 = 0$, so ist $\alpha_2 \neq 0$ und damit $v_2 = 0$.

Definition. Eine Teilmenge $A \subset V$ heißt *linear unabhängig*, falls je endlich viele Vektoren $v_1, \dots, v_n \in A$ linear unabhängig sind.

Beispiele.

(i) Im K^n sind $e_1, \dots, e_n$ linear unabhängig.

(ii) Im $\mathbb{R}^3$ sind je vier Vektoren der Form $v = (\alpha, \beta, \gamma), e_1, e_2, e_3$ linear abhängig, da

$$1v - \alpha e_1 - \beta e_2 - \gamma e_3 = 0.$$

(iii) In $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ ist die Menge der Monome

$$A = \{1, x, x^2, \dots, x^n, \dots\}$$

linear unabhängig.

Lemma 4.3 (Abhängigkeitslemma)
Es seien $v_1, \dots, v_n \in V$ linear unabhängig. Sind $x, v_1, \dots, v_n$ linear abhängig, so ist x eine Linearkombination von $v_1, \dots, v_n$.

Beweis. Da $x, v_1, \dots, v_n$ linear abhängig sind, gibt es eine Linearkombination

$$\alpha_0 x + \alpha_1 v_1 + \cdots + \alpha_n v_n = 0 \qquad (\text{nicht alle } \alpha_i = 0).$$

Wäre $\alpha_0 = 0$, so hätte man $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$, und damit einen Widerspruch zur Annahme, daß $v_1, \dots, v_n$ linear unabhängig sind. Also ist $\alpha_0 \neq 0$, und damit

$$x = -\frac{\alpha_1}{\alpha_0} v_1 - \cdots - \frac{\alpha_n}{\alpha_0} v_n.$$

□

Lemma 4.4 (Schrankenlemma)
V besitze ein Erzeugendensystem bestehend aus n Elementen. Dann sind je $n+1$ Vektoren in V linear abhängig.

Beweis.

(1) $V = \{0\}$. Dann ist $V = \mathrm{Span}\,\emptyset$ und man hat $n = 0$. Je 1 Element ist linear abhängig, da 0 dies ist.

(2) Wir erklären zunächst die Beweis*methode* am Beispiel $n = 1$. Dann ist $V = \operatorname{Span} v$. Es seien $x, y \in V$.

Dann gilt
$$x = \alpha v;\ y = \beta v \qquad (\alpha, \beta \in K).$$
Wir suchen $\gamma, \delta \in K$ mit $(\gamma, \delta) \neq (0, 0)$, so daß
$$0 = \gamma x + \delta y = (\gamma\alpha + \delta\beta)v.$$
Dies ist möglich, da das homogene Gleichungssystem
$$\alpha\xi_1 + \beta\xi_2 = 0$$
nach dem Fundamentallemma nichttriviale Lösungen besitzt.

(3) Es sei nun $V \neq \{0\}$, und $V = \operatorname{Span}(v_1, \ldots, v_n)$. Ferner seien $w_1, \ldots, w_{n+1} \in V$. Dann gibt es Darstellungen
$$w_i = \sum_{j=1}^{n} \alpha_{ji} v_j \qquad (i = 1, \ldots, n + 1;\ \alpha_{ji} \in K).$$
Wir suchen Elemente $\beta_1, \ldots, \beta_{n+1} \in K$, nicht alle zugleich 0, so daß
$$\sum_{i=1}^{n+1} \beta_i w_i = 0.$$
Nun ist
$$\sum_{i=1}^{n+1} \beta_i w_i = \sum_{i=1}^{n+1} \beta_i (\sum_{j=1}^{n} \alpha_{ji} v_j) = \sum_{j=1}^{n} (\sum_{i=1}^{n+1} \beta_i \alpha_{ji}) v_j.$$
Man nun betrachte das Gleichungssystem
$$\sum_{i=1}^{n+1} \alpha_{ji} \xi_i = 0 \qquad (j = 1, \ldots, n).$$
Dies hat $n+1$ Unbekannte und n Gleichungen, besitzt also nach dem Fundamentallemma eine nichttriviale Lösung $(\beta_1, \ldots, \beta_{n+1})$. Diese tut es. □

Der Begriff der Basis

Wie stets sei V ein K-Vektorraum.

Definition. Es sei $V \neq \{0\}$. Eine *(ungeordnete) Basis* von V ist eine Teilmenge $B \subset V$ mit den folgenden Eigenschaften:
(B1) B ist Erzeugendensystem.

(B2) B ist linear unabhängig, (d. h.je endlich viele Elemente von B sind linear unabhängig).

Konvention. $\emptyset$ ist Basis von $V = \{0\}$.

Lemma 4.5
$B = \{v_i\}_{i\in I}$ sei Basis von $V \neq \{0\}$. Dann besitzt jedes Element $v \in V$ eine eindeutige Darstellung

$$v = \sum_{i\in I} \alpha_i v_i \qquad \text{(nur endlich viele } \alpha_i \neq 0\text{)}.$$

Beweis. Die *Existenz* einer solchen Darstellung folgt sofort aus (B1). Die *Eindeutigkeit* sieht man wie folgt:

Es sei

$$\begin{aligned} v &= \alpha_1 v_1 + \cdots + \alpha_n v_n, \\ v &= \beta_1 w_1 + \cdots + \beta_m w_m. \end{aligned}$$

Nach einer eventuellen Umnumerierung kann man annehmen, daß

$$v_1 = w_1, \ldots, v_r = w_r$$

ist und daß die anderen Vektoren $v_{r+1}, \ldots, v_n, w_{r+1}, \ldots, w_m$ paarweise verschieden sind. Aus

$$\begin{aligned} v &= \alpha_1 v_1 + \cdots + \alpha_r v_r + \alpha_{r+1} v_{r+1} + \cdots + \alpha_n v_n, \\ v &= \beta_1 w_1 + \cdots + \beta_r w_r + \beta_{r+1} w_{r+1} + \cdots + \beta_m w_m \end{aligned}$$

folgt durch Subtraktion, daß

$$0 = (\alpha_1 - \beta_1)v_1 + \cdots + (\alpha_r - \beta_r)v_r + \alpha_{r+1}v_{r+1} + \cdots + \alpha_n v_n - \beta_{r+1}w_{r+1} - \cdots - \beta_m w_m.$$

Da B eine linear unabhängige Menge ist, folgt hieraus

$$\alpha_1 = \beta_1, \ldots, \alpha_r = \beta_r \quad \text{und} \quad \alpha_i = \beta_i = 0 \text{ für } i, j \geq r+1,$$

und damit die gesuchte Eindeutigkeit. $\square$

Beispiele.
(1) Der Raum $V = K^n$ hat die Basis

$$e_1 = (1, 0, \ldots, 0), \ \ldots, \ e_i = (0, \ldots, 1, \ldots, 0), \ \ldots, \ e_n = (0, \ldots, 0, 1).$$

Man nennt dies die *Standardbasis* des K^n.

(2) Der Raum $\mathbb{R}[x]$ hat die (unendliche) Basis $\{1, x, x^2, \ldots\}$.

(3) Wir betrachten den Raum der reellen Folgen, d. h.

$$F := \mathrm{Abb}(\mathbb{N}, \mathbb{R})$$

und darin den Unterraum

$$F' := \mathrm{Abb}[\mathbb{N}, \mathbb{R}] := \{a \in \mathrm{Abb}(\mathbb{N}, \mathbb{R});\ a(n) \neq 0 \text{ für nur endlich viele } n\}.$$

Für jedes $n \in \mathbb{N}$ definieren wir

$$\delta_n : \mathbb{N} \longrightarrow \mathbb{R}; \quad \delta_n(m) = \begin{cases} 1 & \text{für } m = n \\ 0 & \text{für } m \neq n. \end{cases}$$

Dann ist $(\delta_n)_{n\in\mathbb{N}}$ eine Basis von F', aber kein Erzeugendensystem von F. Legt man das Zornsche Lemma zu Grunde, so hat jeder Vektorraum, und damit auch F eine Basis.

Satz 4.6 (Basissatz für endlich erzeugte Vektorräume)
Es sei $V \neq \{0\}$ ein endlich erzeugter K-Vektorraum. Dann gilt:

(i) V besitzt eine endliche Basis.

(ii) Je zwei Basen von V haben gleichviele Elemente.

(iii) Sind $v_1, \ldots, v_r \in V$ linear unabhängig, dann ist entweder $\{v_1, \ldots, v_r\}$ Basis von V, oder es gibt $v_{r+1}, \ldots, v_n$, so daß $\{v_1, \ldots, v_n\}$ Basis von V ist.

Bemerkung. Man nennt die Anzahl der Elemente einer Basis auch die *Länge* der Basis.

Beweis. (iii) Ist $\{v_1, \ldots, v_r\}$ ein Erzeugendensystem, so ist man fertig. Ansonsten ist

$$U = \mathrm{Span}(v_1, \ldots, v_r) \subsetneqq V.$$

Es sei $v_{r+1} \notin U$, $v_{r+1} \in V$. Dann sind $\{v_1, \ldots, v_{r+1}\}$ linear unabhängig, da sonst nach dem Abhängigkeitslemma $v_{r+1} \in \mathrm{Span}(v_1, \ldots, v_r) = U$ gilt. Damit ist $\{v_1, \ldots, v_{r+1}\}$ linear unabhängig. Falls dies eine Basis ist, so ist man fertig. Sonst setze man das Verfahren fort, bis es abbricht, was durch das Schrankenlemma garantiert wird. Man kommt so in endlich vielen Schritten zu einer Basis, die $v_1, \ldots, v_r$ enthält.

(i) Da $V \neq \{0\}$ ist, gibt es ein $v \neq 0$. Dann wende man (iii) auf $v_1 = v$ an.

(ii) Es sei B eine Basis bestehend aus n Elementen. Nach dem Schrankenlemma sind je $n + 1$ Elemente in V linear abhängig. Sei also C eine weitere Basis. Dann ergibt dieses Argument, daß

$$\#C \leq n = \#B.$$

Durch Vertauschung von C und B erhält man

$$\#B \leq \#C,$$

also $\#B = \#C$. $\square$

Satz 4.7 (Basisauswahlsatz von Steinitz)
Es sei $\{v_1, \dots, v_r\}$ ein Erzeugendensystem eines Vektorraums $V \neq \{0\}$. Dann gibt es $i_1, \dots, i_n \in \{1, \dots, r\}$, so daß $\{v_{i_1}, \dots, v_{i_n}\}$ eine Basis von V ist.

Beweis. Es sei

$$n := \max\{k;\ \text{es gibt } k \text{ linear unabhängige Elemente in } \{v_1, \dots, v_r\}\}.$$

Es seien $v_{i_1}, \dots, v_{i_n}$ linear unabhängige Vektoren. Wir betrachten

$$U := \operatorname{Span}(v_{i_1}, \dots, v_{i_n}) \subset V.$$

Ist $U = V$, so ist man fertig. Ansonsten gibt es v_k mit

$$v_k \notin U.$$

Dann sind $(v_{i_1}, \dots, v_{i_n}, v_k)$ linear unabhängig (Abhängigkeitslemma). Dies ist ein Widerspruch zur Wahl von n. □

Der Dimensionsbegriff

Es sei V ein K-Vektorraum.

Definition. $\mathbb{N}(V) := \{m \in \mathbb{N} \cup \{0\};$ Je $m+1$ Elemente in V sind linear abhängig.$\}$

Bemerkung. $m \in \mathbb{N}(V), n \geq m \Rightarrow n \in \mathbb{N}(V)$.

Beispiele.

(1) $V = K^n$. Dann ist

$$\mathbb{N}(V) = \{n, n+1, \dots\}.$$

(2) $V = \operatorname{Abb}[\mathbb{N}, \mathbb{R}]$. Dann ist $\mathbb{N}(V) = \emptyset$.

(3) $V = \operatorname{Abb}(\mathbb{N}, \mathbb{R})$. Auch hier ist $\mathbb{N}(V) = \emptyset$.

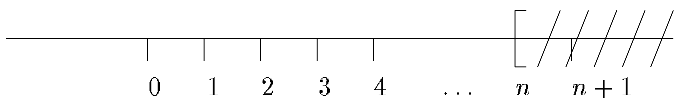


Abb. 31: Die Menge $\mathbb{N}(V)$

Definition. Man nennt V *endlich-dimensional*, falls $\mathbb{N}(V) \neq \emptyset$ ist. Dann heißt

$$\dim V := \operatorname{Min} \mathbb{N}(V)$$

die *Dimension* von V. Ansonsten heißt V unendlich-dimensional und man setzt

$$\dim V = \infty.$$

Bemerkung. $\dim V = 0 \Leftrightarrow V = \{0\}$.

Beispiele.

(1) $\dim K^n = n$.

(2) $\dim \mathrm{Abb}[\mathbb{N}, \mathbb{R}] = \infty$; $\dim \mathrm{Abb}(\mathbb{N}, \mathbb{R}) = \infty$.

(3) $\dim \mathbb{R}[x] = \infty$.

Satz 4.8 (Äquivalenzsatz)
Es sind äquivalent:

(i) V ist endlich-dimensional.

(ii) V ist endlich erzeugt.

Beweis. (ii)$\Rightarrow$(i): Man kann $V \neq \{0\}$ annehmen. V ist endlich erzeugt, etwa

$$V = \mathrm{Span}(v_1, \ldots, v_n).$$

Nach dem Schrankenlemma ist dann $n \in \mathbb{N}(V)$. Also ist $\mathbb{N}(V) \neq \emptyset$ und damit ist V endlich-dimensional.

(i)$\Rightarrow$(ii) Es sei $0 < n = \dim V = \mathrm{Min}\, \mathbb{N}(V) < \infty$. Also gibt es Elemente $v_1, \ldots, v_n$, die linear unabhängig sind. Für alle $x \in V$ sind $x, v_1, \ldots, v_n$ linear abhängig (nach Definition von n). Also gilt nach dem Abhängigkeitslemma

$$x \in \mathrm{Span}(v_1, \ldots, v_n).$$

Damit ist

$$V = \mathrm{Span}(v_1, \ldots, v_n)$$

d. h. V ist endlich erzeugt. □

Korollar 4.9
Es sei V ein endlich erzeugter Vektorraum und $U \subset V$ ein Unterraum. Dann gilt

(i) U ist endlich erzeugt und $\dim U \leq \dim V$.

(ii) $\dim U = \dim V$ gilt genau dann wenn $U = V$.

Beweis. (i) Es sei $U \subset V$. Dann gilt

$$\emptyset \neq \mathbb{N}(V) \subset \mathbb{N}(U),$$

also ist $\mathbb{N}(U) \neq \emptyset$, d. h. U ist endlich-dimensional und nach dem Äquivalenzsatz auch endlich erzeugt. Aus $\mathbb{N}(V) \subset \mathbb{N}(U)$ folgt $\mathrm{Min}\, \mathbb{N}(V) \geq \mathrm{Min}\, \mathbb{N}(U)$ und damit $\dim V \geq \dim U$.

(ii) Es sei $n = \dim U = \dim V$. Wie beim Äquivalenzsatz zeigt man, daß es linear unabhängige Vektoren $u_1, \dots, u_n$ gibt mit

$$U = \mathrm{Span}(u_1, \dots, u_n).$$

Sei $x \in V$. Dann sind $x, u_1, \dots, u_n$ linear abhängig. Nach dem Abhängigkeitslemma folgt $x \in \mathrm{Span}(u_1, \dots, u_n)$, d. h.

$$V = U.$$

□

Satz 4.10 (Basis-Kriterium)
Es sei $V \neq \{0\}$, $n = \dim V < \infty$. Dann implizieren je zwei der folgenden Aussagen, daß $\{v_1, \dots, v_m\}$ eine Basis ist.

(i) $m = n$.

(ii) $v_1, \dots, v_m$ sind linear unabhängig.

(iii) $\{v_1, \dots, v_m\}$ ist Erzeugendensystem.

Beweis. (ii) und (iii): Dies ist die Definition einer Basis.

(i) und (ii): Es sei $x \in V$. Dann sind wegen $n = m = \dim V$ die Vektoren $x, v_1, \dots, v_m$ linear abhängig. Aus dem Abhängigkeitslemma folgt $x \in \mathrm{Span}(v_1, \dots, v_m)$ und damit $V = \mathrm{Span}(v_1, \dots, v_m)$.

(i) und (iii): Wären $v_1, \dots, v_m$ linear abhängig, könnte man ein v_i aus den anderen linear kombinieren, etwa v_1. Also wäre $V = \mathrm{Span}(v_2, \dots, v_n)$. Nach dem Schrankenlemma wäre dann $n - 1 \in \mathbb{N}(V)$. Dies ist ein Widerspruch zu $n = \dim V = \mathrm{Min}\, \mathbb{N}(V)$. □

Satz 4.11 (Dimensionssatz)
Es sei V ein K-Vektorraum. Dann tritt genau einer der folgenden Fälle ein:

(i) $\dim V = 0$ und $V = \{0\}$.

(ii) $0 < \dim V = n < \infty$. Dann gilt:

 (1) Es gibt eine Basis mit n Elementen und jede Basis hat n Elemente.

 (2) Je $n + 1$ Vektoren sind linear abhängig.

(iii) $\dim V = \infty$. Zu jeder Zahl $n \in \mathbb{N}$ gibt es n linear unabhängige Vektoren.

Beweis. Klar aus dem bisherigen. □

Beispiele.

(1) $\dim_{\mathbb{R}} \mathbb{C} = 2$. Eine $\mathbb{R}$-Basis von $\mathbb{C}$ über $\mathbb{R}$ wird gegeben durch $1, i$.

(2) $\dim_{\mathbb{Q}} \mathbb{R} = \infty$. Um eine unendliche linear unabhängige Menge zu finden, betrachten wir die Menge der Primzahlen

$$\mathbb{P} = \{p \in \mathbb{N};\ p \text{ ist ein Primzahl}\}.$$

Zunächst ist $\mathbb{P}$ eine unendliche Menge. Gäbe es nämlich nur endlich viele Primzahlen $p_1, \ldots, p_k$, so betrachte man

$$p = p_1 \cdots p_k + 1.$$

Offensichtlich ist $p \neq p_i$ für alle $p_1, \ldots, p_k$. Wir behaupten, daß p, im Gegensatz zur Annahme, wieder eine Primzahl ist. Ansonsten gäbe es eine Zahl p_i, die p teilt. Aber dann gilt

$$1 = p - p_1 \cdots p_k = p_i p' - p_1 \cdots p_k = p_i(p' - \prod_{\substack{j=1 \\ j \neq i}}^{k} p_j).$$

Das heißt, daß p_i ein Teiler von 1 wäre, ein Widerspruch.

Wir behaupten nun, daß $\{\log p;\ p \in \mathbb{P}\}$ eine linear unabhängige Menge ist. Falls dies nicht der Fall ist, gibt es verschiedene Primzahlen $p_1, \ldots, p_n$ und rationale Zahlen $\alpha_i = \frac{k_i}{m_i} \neq 0$ mit

$$\alpha_1 \log p_1 + \cdots + \alpha_n \log p_n = 0.$$

Durch Potenzieren erhalten wir

$$p_1^{\alpha_1} \cdots p_n^{\alpha_n} = 1.$$

Wir können nun annehmen, daß $k_1, \ldots, k_r > 0$ und $k_{r+1}, \ldots, k_n < 0$, sowie alle $m_i > 0$ sind. Dann haben wir die Gleichung

$$p_1^{\frac{k_1}{m_1}} \cdots p_r^{\frac{k_r}{m_r}} = p_{r+1}^{\frac{-k_{r+1}}{m_{r+1}}} \cdots p_n^{\frac{-k_n}{m_n}}.$$

Man erhält einen Widerspruch zur eindeutigen Zerlegbarkeit einer natürlichen Zahl in Primfaktoren, wenn man diese Gleichung mit dem kleinsten gemeinsamen Vielfachen N von $m_1, \ldots, m_n$ potenziert.

Alternativ kann man damit argumentieren, daß $\mathbb{Q}$ abzählbar ist, $\mathbb{R}$ aber nicht.

Bisher haben wir die Existenz von Basen für endlich erzeugte Vektorräume gezeigt. Allgemeiner gilt:

Satz 4.12
Es sei V ein Vektorraum und $\{v_i\}_{i \in I}$ eine Familie von linear unabhängigen Vektoren. Dann gibt es eine Familie $\{v_j\}_{j \in J}$ mit $I \subset J$, so daß $\{v_j\}_{j \in J}$ eine Basis von V ist.

Beweis. Hierzu benötigt man das Zornsche Lemma, oder das dazu äquivalente Auswahlaxiom der Mengenlehre. Hier verzichten wir auf die Details. $\square$

Korollar 4.13
Jeder Vektorraum besitzt eine Basis.

Homomorphismen von Vektorräumen

V, V' seien Vektorräume über demselben Grundkörper K.

Definition. Eine Abbildung : $V \to V'$ heißt ein *Homomorphismus* (*Vektorraumhomomorphismus*, *lineare Abbildung*, *lineare Transformation*, *linearer Operator*), falls gilt

(H1) $$f(x+y) = f(x) + f(y) \qquad (x, y \in V)$$

(H2) $$f(\alpha x) = \alpha f(x) \qquad (x \in V, \alpha \in K).$$

Lemma 4.14
$f : V \to V'$ sei ein Homomorphismus. Dann gilt
(i) $f(\alpha_1 x_1 + \cdots + \alpha_n x_n) = \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n)$.

(ii) $f(0) = 0$, $f(-x) = -f(x)$.

(iii) Sind $f : V \to V'$ und $g : V' \to V''$ Homomorphismen, so ist auch $g \circ f : V \to V''$ ein Homomorphismus.

(iv) Ist f bijektiv, so ist auch $f^{-1} : V' \to V$ ein Homomorphismus.

Beweis.
(i) Sofort durch vollständige Induktion.

(ii) $f : V \to V'$ ist ein Homomorphismus von abelschen Gruppen. Hierfür hatten wir die Aussagen bereits gesehen (Lemma (2.6)).

(iii) Die Eigenschaft (H1) zeigt man wie bei Gruppenhomomorphismen.

Die Eigenschaft (H2) folgt aus:

$$(f \circ g)(\alpha x) = f(g(\alpha x)) = f(\alpha(g(x))) = \alpha(f(g(x))) = \alpha((f \circ g)(x)).$$

(iv) Wir hatten bereits in Lemma (2.5) gesehen, daß $f^{-1} : V' \to V$ ein Homomorphismus abelscher Gruppen ist. Um (H2) für f^{-1} zu zeigen, betrachten wir die Gleichung

$$f(\alpha f^{-1}(x)) \overset{\text{(H2)}}{=} \alpha f(f^{-1}(x)) = \alpha x.$$

Da f bijektiv ist, folgt hieraus

$$\alpha f^{-1}(x) = f^{-1}(\alpha x).$$

□

Definition. Ein Homomorphismus $f : V \to V'$ heißt ein
(i) *Monomorphismus*, falls f injektiv ist,

(ii) *Epimorphismus*, falls f surjektiv ist,

(iii) *Isomorphismus*, falls es einen Homomorphismus $g : V' \to V$ mit $g \circ f = \mathrm{id}_V$ und $f \circ g = \mathrm{id}_{V'}$ gibt,

(iv) *Endomorphismus*, falls $V' = V$ ist,

(v) *Automorphismus*, falls f ein Isomorphismus und $V' = V$ ist.

Wie bereits früher bei Gruppenhomomorphismen gilt auch hier:

Lemma 4.15
Ein Homomorphismus $f : V \to V'$ ist genau dann ein Isomorphismus, wenn f bijektiv ist.

Beweis. Sofort aus Lemma (4.14) (iv). □

Definition. Zwei Vektorräume V und V' heißen *isomorph* ($V \cong V'$) falls es einen Isomorphismus $f : V \to V'$ gibt.

Bemerkung. Es gilt:

(1) $V \cong V$

(2) $V \cong V' \Leftrightarrow V' \cong V$

(3) $V \cong V', V' \cong V'' \Rightarrow V \cong V''$.

Das heißt, daß der Begriff Isomorphismus eine *Äquivalenzrelation* auf der Menge aller Vektorräume definiert. Meist identifiziert man isomorphe Vektorräume.

Beispiele.

(1) Die Abbildung

$$\begin{array}{rrcl} f: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & f(x_1, x_2, x_3) & = & (x_1 + x_2 + x_3, x_2 + x_3, x_3) \end{array}$$

ist ein Automorphismus mit Umkehrabbildung

$$f^{-1}(y_1, y_2, y_3) = (y_1 - y_2, y_2 - y_3, y_3).$$

(2) Die Abbildung

$$\begin{array}{rrcl} g: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ & (x_1, x_2) & \longmapsto & (x_1 + x_2, 7x_2, x_1^2) \end{array}$$

ist *kein* Homomorphismus.

(3) Die Abbildung

$$\begin{array}{rrcl} \varphi: & K^n & \longrightarrow & {}^nK \\ & (x_1, \dots, x_n) & \longmapsto & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{array}$$

ist ein Isomorphismus.

Satz 4.16
V, V' seien Vektorräume über K. Es sei $v_1, \dots, v_n$ eine Basis von V, sowie $v'_1, \dots, v'_n \in V'$. Dann gibt es genau einen Homomorphismus $f : V \to V'$ mit $f(v_i) = v'_i, i = 1, \dots, n$.

Beweis. *Existenz:* Es sei $x \in V$. Dann besitzt x nach Lemma (4.5) eine eindeutige Darstellung:

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n. \qquad (\alpha_i \in K)$$

Wir setzen

$$f(x) := \alpha_1 v_1' + \cdots + \alpha_n v_n'.$$

Nach Konstruktion ist $f(v_i) = v_i'$. f ist ein Homomorphismus; denn für

$$\begin{aligned} x &= \alpha_1 v_1 + \cdots + \alpha_n v_n, \\ y &= \beta_1 v_1 + \cdots + \beta_n v_n \end{aligned}$$

gilt:

$$\begin{aligned} f(x+y) &= (\alpha_1 + \beta_1) v_1' + \cdots + (\alpha_n + \beta_n) v_n' \\ &= (\alpha_1 v_1' + \cdots + \alpha_n v_n') + (\beta_1 v_1' + \cdots + \beta_n v_n') \\ &= f(x) + f(y). \end{aligned}$$

Die Eigenschaft (H2) weist man analog nach.

Eindeutigkeit: Es sei $g : V \to V'$ ein weiterer Homomorphismus mit $g(v_i) = v_i'$. Dann gilt

$$\begin{aligned} g(x) &= g(\alpha_1 v_1 + \cdots + \alpha_n v_n) \overset{\text{(H1),(H2)}}{=} \alpha_1 g(v_1) + \cdots + \alpha_n g(v_n) \\ &= \alpha_1 v_1' + \cdots + \alpha_n v_n' = f(x). \end{aligned}$$

□

Kern und Bild

Es sei $f : V \to V'$ ein Homomorphismus von Vektorräumen.

Definition.

(i) Für $U \subset V$ setzt man

$$f(U) := \{f(u);\ u \in U\}.$$

(ii) Das *Bild* von f ist

$$\operatorname{Im} f := f(V) = \{f(v);\ v \in V\}.$$

(iii) Der *Kern* von f ist

$$\operatorname{Ker} f := \{v \in V;\ f(v) = 0\} = f^{-1}\{0\}.$$

Lemma 4.17

(i) $\operatorname{Ker} f \subset V$ ist ein Unterraum von V.

(ii) $\operatorname{Im} f \subset V'$ ist ein Unterraum von V'.

(iii) f ist injektiv $\Leftrightarrow \operatorname{Ker} f = \{0\}$.

(iv) $f(\operatorname{Span}(v_1, \dots, v_n)) = \operatorname{Span}(f(v_1), \dots, f(v_n))$.

(v) Ist E ein Erzeugendensystem von V, so ist $f(E)$ ein Erzeugendensystem von $\operatorname{Im} f$.

(vi) Sind $v_1, \dots, v_n \in V$ linear abhängig, so auch $f(v_1), \dots, f(v_n)$.

(vii) Sind $f(v_1), \dots, f(v_n)$ linear unabhängig, so auch $v_1, \dots, v_n$.

(viii) Ist V endlich erzeugt, so ist auch $f(V)$ endlich erzeugt, und es gilt $\dim \operatorname{Im} f \leq \dim V$.

Beweis. Die Beweise von (i)–(iii) verlaufen analog denen bei Gruppenhomomorphismen.

(v) Es sei $E = \{v_i\}_{i \in I}$. Betrachte $y \in \operatorname{Im} f$. Dann gibt es ein $x \in V$ mit $f(x) = y$. Man hat eine endliche Teilmenge $J \subset I$ mit

$$x = \sum_{j \in J} \alpha_j v_j.$$

Also gilt

$$y = f(x) = \sum_{j \in J} \alpha_j f(v_j).$$

Daher ist $f(E)$ Erzeugendensystem von $\operatorname{Im} f$.

(iv) Dies folgt unmittelbar aus dem obigen.

(vi) $\sum_{i=1}^n \alpha_i v_i = 0 \Rightarrow \sum_{i=1}^n \alpha_i f(v_i) = 0$.

(vii) Dies ist die logische Umkehrung von (vi).

(viii) Da V endlich erzeugt ist, gilt nach dem Äquivalenzsatz (4.8), daß $\dim V =: n < \infty$. Es sei $v_1, \dots, v_n$ eine Basis von V. Dann ist nach (v) die Menge $f(v_1), \dots, f(v_n)$ ein Erzeugendensystem von $f(V)$. Wiederum nach dem Äquivalenzsatz ist V endlich-dimensional und nach dem Schrankenlemma (4.4) gilt $\dim f(V) \leq n = \dim V$. $\square$

Beispiele.

(1) Der Homomorphismus

$$\begin{array}{rrcl} f: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x_1, x_2, x_3) & \longmapsto & (x_1 - x_2, x_1 - x_3, 0) \end{array}$$

hat folgenden Kern und folgendes Bild:

$$\begin{array}{rcl} \operatorname{Ker} f & = & \mathbb{R}(1,1,1) = \operatorname{Span}(e_1 + e_2 + e_3). \\ \operatorname{Im} f & = & \{(y_1, y_2, y_3);\ y_3 = 0\} = \operatorname{Span}(e_1, e_2). \end{array}$$

Es ist also

$$\dim \operatorname{Ker} f = 1, \quad \dim \operatorname{Im} f = 2.$$

(2) Wir betrachten den Raum der konvergenten Folgen

$$F_{\text{konv}} = \{(a_n)_{n\in\mathbb{N}};\ (a_n) \text{ ist eine konvergente Folge}\}.$$

Dies ist ein Unterraum des Raums aller Folgen. Die Abbildung

$$\begin{array}{rrcl} f: & F_{\text{konv}} & \longrightarrow & \mathbb{R} \\ & (a_n)_{n\in\mathbb{N}} & \longmapsto & \lim a_n \end{array}$$

ist ein Vektorraumhomomorphismus mit

$$\operatorname{Ker} f = \{(a_n)_{n\in\mathbb{N}};\ (a_n) \text{ ist Nullfolge}\}.$$

Dimensionssatz für Homomorphismen

Satz 4.18
Zwei endlich-dimensionale Vektorräume V, V' sind genau dann isomorph, wenn $\dim V = \dim V'$.

Beweis.
(i) $f : V \to V'$ sei ein Isomorphismus. Dann ist $f(V) = V'$ und nach Lemma (4.17) (viii) ist

$$\dim V' = \dim \operatorname{Im} f \leq \dim V.$$

Das analoge Argument gilt für $f^{-1} : V' \to V$. Dies ergibt

$$\dim V \leq \dim V',$$

also $\dim V = \dim V'$.

(ii) Sei $\dim V = \dim V' = n$. Ferner sei $v_1, \dots, v_n$ eine Basis von V und $v'_1, \dots, v'_n$ eine Basis von V'. Dann gibt es genau einen Homomorphismus $f : V \to V'$ mit

$$f(v_i) = v'_i \qquad (i = 1, \dots, n).$$

Nach Lemma (4.17) (iv) ist f surjektiv. f ist auch injektiv. Es sei nämlich

$$0 = f(x) = f(\sum_{i=1}^{n} \alpha_i v_i) = \sum_{i=1}^{n} \alpha_i f(v_i) = \sum_{i=1}^{n} \alpha_i v'_i.$$

Da die v'_i linear unabhängig sind, folgt $\alpha_1 = \cdots = \alpha_n = 0$, d. h. $x = 0$. Also ist $\operatorname{Ker} f = \{0\}$ und damit ist f injektiv nach Lemma (4.17) (iii). □

Bemerkung. Es gibt Vektorräume V, V' mit $\dim V = \dim V' = \infty$, aber $V \not\cong V'$. Beispielsweise gilt

$$\operatorname{Abb}[\mathbb{N}, \mathbb{Q}] \not\cong \operatorname{Abb}(\mathbb{N}, \mathbb{Q}).$$

Der Grund hierfür ist, daß $\operatorname{Abb}[\mathbb{N}, \mathbb{Q}]$ abzählbar ist, aber $\operatorname{Abb}(\mathbb{N}, \mathbb{Q})$ überabzählbar ist.

Korollar 4.19
Ist $0 < \dim V = n < \infty$, so ist $V \cong K^n$.

Satz 4.20
$f : V \to V'$ sei ein Homomorphismus. Dann gilt:

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim V.$$

Beweis. 1. Schritt: Wir können annehmen, daß $\dim \operatorname{Ker} f < \infty$, $\dim \operatorname{Im} f < \infty$. Ansonsten gilt

$$\begin{aligned} \dim \operatorname{Ker} f = \infty &\overset{\operatorname{Ker} f \subset V}{\Rightarrow} \dim V = \infty \\ \dim \operatorname{Im} f = \infty &\overset{(4.17)(\text{viii})}{\Rightarrow} \dim V = \infty. \end{aligned}$$

In diesen Fällen gilt die Aussage des Satzes offensichtlich.

2. Schritt: Wir betrachten $v_1, \dots, v_n, w_1, \dots, w_m \in V$ mit:

(1) $\quad v_1, \dots, v_n \in \operatorname{Ker} f \quad$ ist Basis von $\operatorname{Ker} f$
(2) $\quad f(w_1), \dots, f(w_m) \in \operatorname{Im} f \quad$ ist Basis von $\operatorname{Im} f$.

Dies schließt auch die Fälle $\operatorname{Ker} f = \{0\}$ oder $\operatorname{Im} f = \{0\}$ ein. In diesen Fällen ist $\{v_1, \dots, v_n\}$ oder $\{w_1 \dots, w_m\}$ die leere Menge.

Behauptung: $v_1, \dots, v_n, w_1, \dots, w_m$ ist eine Basis von V. Daraus folgt die vorherige Behauptung, denn dann gilt

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f = n + m = \dim V.$$

Lineare Unabhängigkeit: Es sei

$$\alpha_1 v_1 + \dots + \alpha_n v_n + \beta_1 w_1 + \dots + \beta_m w_m = 0.$$

Anwenden von f ergibt:

$$\beta_1 f(w_1) + \dots + \beta_m f(w_m) = 0 \overset{(2)}{\Rightarrow} \beta_1 = \dots = \beta_m = 0.$$

Also ist

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0 \overset{(1)}{\Rightarrow} \alpha_1 = \dots = \alpha_n = 0.$$

Erzeugendensystem: Es sei $x \in V$. Also ist $f(x) \in \operatorname{Im} f$, d. h.

$$\begin{aligned} f(x) &= \beta_1 f(w_1) + \dots + \beta_m f(w_m) \\ &\Rightarrow x - \beta_1 w_1 - \dots - \beta_m w_m \in \operatorname{Ker} f \\ &\Rightarrow x - \beta_1 w_1 - \dots - \beta_m w_m = \alpha_1 v_1 + \cdots \alpha_n v_n \\ &\Rightarrow x = \alpha_1 v_1 + \dots + \alpha_n v_1 + \beta_1 w_1 + \dots + \beta_m w_m. \end{aligned}$$

□

Korollar 4.21
Es sei $\dim V = \dim V' < \infty$. Dann sind für einen Homomorphismus $f : V \to V'$ äquivalent:

(i) f ist injektiv.

(ii) f ist surjektiv.

(iii) f ist bijektiv (d. h.ein Isomorphismus).

Beweis. (i)⇒(ii) Die Dimensionsformel ergibt

$$\dim \operatorname{Im} f = \dim V = \dim V'.$$

Daraus folgt $\operatorname{Im} f = V'$ nach Korollar (4.9).

(ii)⇒(iii) $\operatorname{Im} f = V'$. Also ist $\dim \operatorname{Im} f = \dim V' = \dim V$. Nach der Dimensionsformel folgt $\dim \operatorname{Ker} f = 0$, d. h.$\operatorname{Ker} f = \{0\}$, und damit ist f injektiv.

(iii)⇒(i) Trivial.

Bemerkung. Der Satz ist falsch in unendlicher Dimension:

(1) $\frac{d}{dx} : \mathbb{R}[x] \to \mathbb{R}[x], f \mapsto f'$ ist surjektiv, aber nicht injektiv.

(2) $\int : \mathbb{R}[x] \to \mathbb{R}[x], f \mapsto \int_0^x f(t)\,dt$ ist injektiv, aber nicht surjektiv.

Definition. Für einen Homomorphismus $f : V \to V'$ heißt

$$\operatorname{Rang} f := \dim \operatorname{Im} f$$

der *Rang* von f.

Wir halten nochmals fest:

$$\boxed{\dim \operatorname{Ker} f + \operatorname{Rang} f = \dim V}$$

Direkte Summen und Komplemente

V sei ein Vektorraum über K und $U_1, \dots, U_s$ seien Unterräume von V. Wir erweitern unsere früher gemachte Definition der Summe zweier Unterräume wie folgt:

Definition. Man nennt

$$U := U_1 + \cdots + U_s = \{u_1 + \cdots + u_s;\ u_i \in U_i;\ i = 1, \dots, s\}$$

die *Summe* von $U_1, \dots, U_s$.

Bemerkung. Es gilt

$$U_1 + \cdots + U_s = \operatorname{Span}(U_1 \cup \cdots \cup U_s).$$

Insbesondere ist $U_1 + \cdots + U_s$ ein Unterraum von V.

Schreibweise. $\sum_{i=1}^s U_i$.

Satz 4.22
$U_1, \dots, U_s$ seien Unterräume von V und $U = \sum_{i=1}^s U_i$. Dann sind äquivalent:

(i) $u_1 + \cdots + u_s = 0$ für $u_i \in U_i \Rightarrow u_1 = \cdots = u_s = 0$.

(ii) Jedes Element $u \in U$ besitzt eine eindeutige Darstellung $u = u_1 + \cdots + u_s$ mit $u_i \in U_i$.

(iii) Für $W_i := \sum_{j=1, j\neq i}^s U_i$ gilt $U_i \cap W_i = \{0\}$.

Definition. In diesem Fall nennt man U die *direkte Summe* von $U_1, \dots, U_s$.

Schreibweise. $U = U_1 \oplus \cdots \oplus U_s = \bigoplus_{i=1}^s U_i$.

Beweis. (i)$\Rightarrow$(ii) Es sei

$$\begin{aligned} u &= u_1 + \cdots + u_s \\ u &= u_1' + \cdots + u_s' \end{aligned} \qquad (u_i, u_i' \in U_i).$$

Dann folgt

$$0 = (u_1 - u_1') + \cdots + (u_s - u_s'). \qquad (u_i - u_i' \in U_i)$$

Aus (i) folgt dann

$$u_1 = u_1', \dots, u_s = u_s'.$$

(ii)$\Rightarrow$(iii) Es sei $W_i \cap U_i \neq \{0\}$. Also gibt es $0 \neq u_i \in U_i$ mit

$$u_i = u_1 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_s. \qquad (u_i \in U_i)$$

Dies ist ein Widerspruch zu (ii).

(iii)$\Rightarrow$(i) Es sei $u_1 + \cdots + u_s = 0 \quad (u_i \in U_i)$. Also folgt

$$u_i = -(u_1 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_s)$$

d. h. $u_i \in U_i \cap W_i$. Daher folgt $u_i = 0$ für $i = 1, \dots, s$. $\square$

Korollar 4.23
$\dim(U_1 \oplus \cdots \oplus U_s) = \dim U_1 + \cdots + \dim U_s$.

Beweis. Es genügt, dies für $\dim U_i < \infty$ zu zeigen, sonst stünde auf beiden Seiten ∞. Es genügt ferner die Aussage für $s = 2$ zu zeigen. Die Aussage für allgemeines s folgt dann durch Induktion, da

$$U_1 \oplus \cdots \oplus U_s = U_1 \oplus (U_2 \oplus \cdots \oplus U_s).$$

Sei also $u_1, \dots, u_n$ eine Basis von U_1, sowie $v_1, \dots, v_m$ eine Basis von U_2. Wir werden zeigen, daß $u_1, \dots, u_n, v_1, \dots, v_m$ eine Basis von $U_1 \oplus U_2$ ist. Hieraus folgt die Behauptung, da dann

$$\dim(U_1 \oplus U_2) = n + m = \dim U_1 + \dim U_2.$$

Offensichtlich erzeugen $u_1, \dots, u_n, v_1, \dots, v_m$ den Raum $U_1 \oplus U_2$. Diese Vektoren sind auch linear unabhängig. Es sei nämlich

$$\underbrace{(\alpha_1 u_1 + \cdots + \alpha_n u_n)}_{\in U_1} + \underbrace{(\beta_1 v_1 + \cdots + \beta_m v_m)}_{\in U_2} = 0.$$

Nach Satz (4.22) (i) folgt daraus

$$\begin{aligned} \alpha_1 u_1 + \cdots + \alpha_n u_n &= 0 \\ \beta_1 v_1 + \cdots + \beta_m v_m &= 0 \end{aligned}$$

Da $u_1, \dots, u_n$ und $v_1, \dots, v_m$ linear unabhängig sind, folgt $\alpha_1 = \cdots = \alpha_n = \beta_1 = \cdots = \beta_m = 0$. □

Beispiel. Es sei $V = \mathbb{R}^3$. Ist $U_1 = \text{Span}(e_1, e_3)$, $U_2 = \text{Span}(e_2)$, so ist $V = U_1 \oplus U_2$.

Ist dagegen $U_3 = \text{Span}(e_2, e_3)$, so ist zwar $V = U_1 + U_3$, die Summe ist aber nicht direkt, da $U_1 \cap U_3 = \text{Span}(e_3)$.

Definition. $U \subset V$ sei ein Unterraum. Ein Unterraum W heißt *Komplement* von U in V, falls

$$U \oplus W = V.$$

Bemerkung. Zu einem Unterraum U ist das Komplement W nicht eindeutig bestimmt.

Definition. Ist V endlich-dimensional, so heißt

$$\operatorname{co\,dim} U := \dim V - \dim U$$

die *Kodimension* von U in V.

Bemerkung. Ist W ein Komplement von U in V, so gilt $\operatorname{co\,dim} U = \dim W$.

Satz 4.24
Jeder Unterraum U besitzt ein Komplement in V.

Beweis. Es sei $\{u_i\}_{i \in I}$ eine Basis von U. Dann gibt es nach Satz (4.6), bzw. Satz (4.12) eine Indexmenge $J \supset I$, sowie eine Familie $\{u_i\}_{i \in J}$, die eine Basis von V ist. Es sei

$$W := \text{Span}\{u_j;\ j \in J \setminus I\}.$$

Dann ist offensichtlich $U + W = V$. Diese Summe ist sogar direkt, d. h.es gilt $V = U \oplus W$. Es sei nämlich $u \in U \cap W$. Dann ist

$$\begin{aligned} u &= \alpha_{i_1} u_{i_1} + \cdots + \alpha_{i_n} u_{i_n} \qquad \{i_1, \dots, i_n\} \subset I, \\ u &= \beta_{j_1} u_{j_1} + \cdots + \beta_{j_m} u_{j_m} \qquad \{j_1, \dots, j_m\} \subset J \setminus I \end{aligned}$$

also erhält man durch Subtraktion

$$\alpha_{i_1} u_{i_1} + \cdots + \alpha_{i_n} u_{i_n} - \beta_{j_1} u_{j_1} - \cdots - \beta_{j_m} u_{j_m} = 0.$$

Aus der linearen Unabhängigkeit folgt, $\alpha_{i_1} = \cdots = \alpha_{i_n} = \beta_{j_1} = \cdots = \beta_{j_m} = 0$, d. h.$u = 0$. Also ist die Summe direkt. □

Satz 4.25 (Dimensionsformel)
$U_1, U_2 \subset V$ seien Unterräume von V. Dann gilt

$$\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim U_1 + \dim U_2.$$

Beweis. Ist $\dim U_1 = \infty$ oder $\dim U_2 = \infty$, so ist auch $\dim(U_1 + U_2) = \infty$. Es sei daher $\dim U_i < \infty$ für $i = 1, 2$. Dann gilt auch $\dim(U_1 + U_2) < \infty$ (verwende den Äquivalenzsatz (4.8)) und $\dim(U_1 \cap U_2) < \infty$ (Korollar (4.9)). Es sei

$$\begin{aligned} U &:= U_1 + U_2 \\ W &:= U_1 \cap U_2. \end{aligned}$$

Nach Satz (4.24) gibt es Komplemente W_1, W_2 von W in U_1 und U_2; d. h.

$$(1) \qquad U_1 = W \oplus W_1$$

$$(2) \qquad U_2 = W \oplus W_2.$$

Behauptung. $U = W \oplus W_1 \oplus W_2$. Dann folgt nach Korollar (4.23):

$$\begin{aligned} \dim(U_1 + U_2) &= \dim(U_1 \cap U_2) + \dim U_1 - \dim(U_1 \cap U_2) + \dim U_2 - \dim(U_1 \cap U_2) \\ &= \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2). \end{aligned}$$

Beweis der Behauptung. $U = W + W_1 + W_2$ ist klar. Es bleibt zu zeigen, daß diese Summe direkt ist. Dazu sei

$$w + w_1 + w_2 = 0. \qquad (w \in W;\ w_i \in W_i;\ i = 1, 2)$$

Dann gilt

$$w_2 = -\underbrace{(w + w_1)}_{\in U_1} \in U_1 \cap W_2 \subset U_1 \cap U_2 = W.$$

Also gilt

$$w_2 \in W_2 \cap W \overset{(2)}{\Rightarrow} w_2 = 0.$$

Analog folgt $w_1 = 0$, und damit auch $w = 0$. Die Summe ist also direkt. □

Satz 4.26
V sei ein endlich-dimensionaler Vektorraum, und $f : V \to V$ sei ein Endomorphismus. Dann sind äquivalent:

(i) $V = \operatorname{Im} f \oplus \operatorname{Ker} f$.

(ii) $\operatorname{Im} f \cap \operatorname{Ker} f = \{0\}$.

Beweis. (i)$\Rightarrow$(ii) Dies ist die Definition von direkter Summe.

(ii)$\Rightarrow$(i) Aus der Dimensionsformel (4.25) und aus Satz (4.20) folgt

$$\dim(\operatorname{Im} f + \operatorname{Ker} f) = \dim \operatorname{Im} f + \dim \operatorname{Ker} f = \dim V.$$

Also ist $\operatorname{Im} f + \operatorname{Ker} f = V$ und nach Voraussetzung ist die Summe auch direkt. $\square$

§ 5 Matrizen

K sei ein fest gewählter Körper.

In dem linearen Gleichungssystem

$$\begin{array}{lllll} \alpha_{11}\xi_1 & + & \alpha_{12}\xi_2 & + & \cdots & +\alpha_{1n}\xi_n = 0 \\ \alpha_{21}\xi_1 & + & \cdots & & & +\alpha_{2n}\xi_n = 0 \\ \vdots & & & & & \\ \alpha_{m1}\xi_1 & + & \cdots & + & \cdots & +\alpha_{mn}\xi_n = 0 \end{array}$$

wollen wir die Koeffizienten zusammenfassen zu einer *Matrix*

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix}.$$

Die naive Definition einer $(m \times n)$-Matrix ist also ein rechteckiges Schema mit m Zeilen und n Spalten, dessen Einträge Elemente in K sind. Formal definieren wir:

Definition. Es seien $m, n \geq 1$ natürliche Zahlen. Eine $(m \times n)$-*Matrix* ist eine Abbildung

$$A : \{1, \dots, m\} \times \{1, \dots, n\} \longrightarrow K.$$

Schreibweise.
(i) $\alpha_{ij} := A(i, j) \in K$,

$$A = (\alpha_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix}.$$

(ii) Man sagt, A hat m *Zeilen* und n *Spalten*. Die α_{ij} heißen *Komponenten* (*Einträge*) der Matrix A. Die Indizes i bzw. j heißen der *Zeilenindex* bzw. der *Spaltenindex*.

Definition.
(i) Eine Matrix heißt *quadratisch*, falls $m = n$.

(ii) Ist A eine quadratische Matrix, so heißen die Einträge α_{ii} mit $1 \leq i \leq n$ die *Diagonalelemente*.

(iii) Eine quadratische Matrix A heißt eine *Diagonalmatrix*, falls $\alpha_{ij} = 0$ für $i \neq j$, d. h.falls

$$A = \begin{pmatrix} \alpha_{11} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \alpha_{nn} \end{pmatrix}.$$

Definition.

(i)

$$K^{(m\times n)} := \mathrm{Mat}(m,n;K) := \{A;\ A = (\alpha_{ij}) \text{ ist } (m\times n)\text{-Matrix über } K\}$$

heißt der *Raum der* $(m\times n)$ *Matrizen.*

(ii) $\mathrm{Mat}(n;K) := \mathrm{Mat}(n,n;K)$.

Bemerkung. $\mathrm{Mat}(m,n;K)$ ist in natürlicher Weise ein K-Vektorraum (vgl. Beispiel (2) auf Seite 47).

Addition:

$$\begin{array}{rcl} \mathrm{Mat}(m,n;K)\times\mathrm{Mat}(m,n;K) & \longrightarrow & \mathrm{Mat}(m,n;K) \\ (A,B) & \longmapsto & A+B \end{array}$$

ist wie folgt erklärt: Für $A = (\alpha_{ij}), B = (\beta_{ij})$ ist

$$(A+B) := (\alpha_{ij}+\beta_{ij}).$$

Neutrales Element:

$$0 = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

Inverses Element:

$$-A = \begin{pmatrix} -\alpha_{11} & \cdots & -\alpha_{1n} \\ \vdots & \ddots & \\ -\alpha_{m1} & \cdots & -\alpha_{mn} \end{pmatrix}.$$

Skalarmultiplikation: Die Abbildung

$$\begin{array}{rcl} K\times\mathrm{Mat}(m,n;K) & \longrightarrow & \mathrm{Mat}(m,n;K) \\ (\lambda,A) & \longmapsto & \lambda A \end{array}$$

ist definiert durch

$$(\lambda A) := (\lambda\alpha_{ij}).$$

Das Nachprüfen der Vektorraumaxiome ist trivial.

Bemerkung.

(i)

$$\begin{array}{rcl} K^{(1\times n)} = \mathrm{Mat}(1,n;K) & \cong & K^n \\ (\alpha_{11},\ldots,\alpha_{1n}) & \mapsto & (\alpha_{11},\ldots,\alpha_{1n}). \end{array}$$

(ii)

$$\begin{array}{rcl} K^{(m\times 1)} = \mathrm{Mat}(m,1;K) & \cong & {}^mK \\ \begin{pmatrix}\alpha_{11}\\ \vdots \\ \alpha_{m1}\end{pmatrix} & \mapsto & \begin{pmatrix}\alpha_{11}\\ \vdots \\ \alpha_{m1}\end{pmatrix}. \end{array}$$

Im folgenden seien m, n fest gewählt.

Definition. Das *Kroneckersche δ-Symbol* ist definiert durch

$$\delta_{ij} := \begin{cases} 1 & \text{falls} \quad i = j \\ 0 & \text{falls} \quad i \neq j. \end{cases}$$

Definition. Für $1 \leq k \leq m$; $1 \leq l \leq n$ definieren wir

$$E_{kl} := (e_{ij}^{kl})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} \in \mathrm{Mat}(m,n;K)$$

durch

$$(e_{ij}^{kl})_{ij} := \delta_{ik}\delta_{jl}.$$

Die Matrizen E_{kl} heißen *Elementarmatrizen.*

Bemerkung. E_{kl} ist die Matrix, die in der k-ten Zeile und der l-ten Spalte den Eintrag 1, und sonst überall den Eintrag 0 hat:

$$E_{kl} = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & 1 & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}. \begin{array}{l} \leftarrow k\text{-te Zeile} \\ \\ \\ \\ \end{array}$$

$$\uparrow l\text{-te Spalte.}$$

Satz 5.1
Die Matrizen E_{kl}; $1 \leq k \leq m, 1 \leq l \leq n$ bilden eine Basis von $\mathrm{Mat}(m,n;K)$. Insbesondere ist $\dim_K \mathrm{Mat}(m,n;K) = m \cdot n$.

Beweis. Es liegt ein *Erzeugendensystem* vor, denn für

$$A = (\alpha_{ij}) = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix}$$

hat man
$$A = \sum_{l=1}^{n} \sum_{k=1}^{m} \alpha_{kl} E_{kl}.$$

Die *lineare Unabhängigkeit* gilt, da aus
$$0 = \sum_{l=1}^{n} \sum_{k=1}^{m} \alpha_{kl} E_{kl} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix}$$
sofort $\alpha_{kl} = 0$ für alle k, l folgt. □

Definition. Die Matrix
$$E_n := (\delta_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$
heißt die *Einheitsmatrix* der Größe n.

Schreibweise. $E = E_n = E^{(n)}$.

Transposition einer Matrix

Sei $A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} \in \mathrm{Mat}(m, n; K)$.

Definition. Die zu A *transponierte Matrix* ist definiert durch
$${}^t A := ({}^t\alpha_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le m}} \in \mathrm{Mat}(n, m; K)$$
wobei
$${}^t\alpha_{ij} := \alpha_{ji}.$$

Beispiel.

(i)
$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix}; \qquad {}^t A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{pmatrix}.$$

(ii) Im quadratischen Fall hat man
$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix}; \qquad {}^t A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{n1} \\ \vdots & & \vdots \\ \alpha_{1n} & \cdots & \alpha_{nn} \end{pmatrix}$$

d. h.man erhält die transponierte Matrix durch Spiegeln an der Hauptdiagonalen. Transponieren einer Matrix bedeutet, daß man die Rolle der Zeilen und Spalten vertauscht.

Lemma 5.2

(i) Für $A, B \in \mathrm{Mat}(m, n; K); \alpha, \beta \in K$ gilt:

$$^t(\alpha A + \beta B) = \alpha^t A + \beta^t B.$$

(ii) $^t(^tA) = A$.

Beweis.

(i) $(\alpha A + \beta B) = (\alpha\alpha_{ij} + \beta\beta_{ij})$ Dann gilt:
$^t(\alpha A + \beta B) = (\alpha\alpha_{ji} + \beta\beta_{ji}) = \alpha(\alpha_{ji}) + \beta(\beta_{ji}) = \alpha^t A + \beta^t B$.

(ii) $A = (\alpha_{ij}) \Rightarrow {}^tA = (\alpha_{ji}) \Rightarrow {}^t({}^tA) = (\alpha_{ij})$. □

Korollar 5.3
Die Abbildung

$$\begin{array}{rcl} \mathrm{Mat}(m, n; K) & \longrightarrow & \mathrm{Mat}(n, m; K) \\ A & \longmapsto & {}^tA. \end{array}$$

ist ein Vektorraumisomorphismus.

Beweis. Die ist ein Homomorphismus wegen Lemma (5.2) (i). Wegen Lemma (5.2) (ii) ist

$$\begin{array}{rcl} \mathrm{Mat}(n, m; K) & \longrightarrow & \mathrm{Mat}(m, n; K) \\ B & \longmapsto & {}^tB. \end{array}$$

eine Umkehrabbildung. □

Definition.

(i) Eine Matrix $A \in \mathrm{Mat}(n; K)$ heißt *symmetrisch*, falls ${}^tA = A$. Man nennt

$$\mathrm{Sym}(n; K) := \{A \in \mathrm{Mat}(n; K);\ {}^tA = A\}$$

den *Raum der symmetrischen* $(n \times n)$ *Matrizen.*

(ii) A heißt *schiefsymmetrisch (alternierend)*, falls ${}^tA = -A$. Man nennt

$$\mathrm{Alt}(n, K) := \{A \in \mathrm{Mat}(n, K);\ {}^tA = -A\}$$

den *Raum der alternierenden* $(n \times n)$ *Matrizen* über K.

Bemerkung. $\mathrm{Sym}(n, K)$ und $\mathrm{Alt}(n, K)$ sind Unterräume von $\mathrm{Mat}(n; K)$.

Spalten- und Zeilenrang

Wir betrachten eine Matrix

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1j} & \cdots & \alpha_{1n} \\ \vdots & & \vdots & & \vdots \\ \alpha_{i1} & \cdots & \alpha_{ij} & \cdots & \alpha_{in} \\ \vdots & & \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mj} & \cdots & \alpha_{mn} \end{pmatrix} \in \mathrm{Mat}(m, n; K).$$

Definition.
(i) Die Vektoren

$$a_j := \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix} \in {}^mK \quad (j = 1, \dots, n)$$

heißen die *Spaltenvektoren* von A.
(ii) Die Vektoren

$$b_i := (\alpha_{i1}, \dots, \alpha_{in}) \in K^n \quad (i = 1, \dots, m)$$

heißen die *Zeilenvektoren* von A.

Schreibweise. $A = (a_1, \dots, a_n) = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$.

Bemerkung. ${}^tA = ({}^tb_1, \dots, {}^tb_m) = \begin{pmatrix} {}^ta_1 \\ \vdots \\ {}^ta_n \end{pmatrix}$.

Definition.
(i) Der *Spaltenrang* von A ist definiert durch

$$\text{Spaltenrang}\, A := \dim_K \mathrm{Span}(a_1, \dots, a_n).$$

(ii) Der *Zeilenrang* von A ist definiert durch

$$\text{Zeilenrang}\, A := \dim_K \mathrm{Span}(b_1, \dots, b_m).$$

Bemerkung.
(i) Der Spaltenrang ist also die maximale Anzahl von linear unabhängigen Vektoren in der Menge $\{a_1, \dots, a_n\}$ der Spaltenvektoren von A. Analog ist der Zeilenrang die maximale Anzahl von linear unabhängigen Vektoren in der Menge $\{b_1, \dots, b_m\}$ der Zeilen von A.

(ii) Nach Definition gilt

$$\begin{aligned} \text{Spaltenrang}\, A &= \text{Zeilenrang}\, {}^tA \\ \text{Zeilenrang}\, A &= \text{Spaltenrang}\, {}^tA. \end{aligned}$$

Beispiel.

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

Hier gilt Spaltenrang $A =$ Zeilenrang $A = 3$.

Elementare Umformungen

Es sei A eine Matrix gegeben durch ihre Spaltenvektoren, d. h.

$$A = (a_1, \dots, a_n) \in \mathrm{Mat}(m, n; K) \qquad (a_i \in {}^mK).$$

Wir betrachten folgende *elementare Spaltenumformungen:*

(ES1) Addition einer Spalte zu einer *anderen* Spalte:

$$(a_1, \dots, a_i, \dots, a_j, \dots, a_n) \curvearrowright (a_1, \dots, a_i + a_j, \dots, a_j, \dots, a_n).$$

(ES2) Multiplikation einer Spalte mit $\alpha \in K^*$:

$$(a_1, \dots, a_i, \dots, a_n) \curvearrowright (a_1, \dots, \alpha a_i, \dots, a_n).$$

(ES3) Addition einer Linearkombination von Spalten zu einer *weiteren* Spalte

$$(a_1, \dots, a_i, \dots, a_n) \curvearrowright (a_1, \dots, a_i + \sum_{j \neq i} \alpha_j a_j, \dots, a_n).$$

(ES4) Vertauschen zweier Spalten:

$$(a_1, \dots, a_i, \dots, a_j, \dots, a_n) \curvearrowright (a_1, \dots, a_j, \dots, a_i, \dots, a_n).$$

Lemma 5.4
Die Operationen (ES3) und (ES4) ergeben sich durch endliche Anwendung der Operationen (ES1) und (ES2).

Beweis. (ES3): klar.

(ES4):

$$\begin{array}{rcl} A & = & (a_1, \dots, a_i, \dots, a_j, \dots, a_n) \\ & \overset{\text{(ES1)}}{\curvearrowright} & (a_1, \dots, a_i, \dots, a_j + a_i, \dots, a_n) \\ & \overset{\text{(ES2)}}{\curvearrowright} & (a_1, \dots, -a_i, \dots, a_j + a_i, \dots, a_n) \\ & \overset{\text{(ES1)}}{\curvearrowright} & (a_1, \dots, -a_i + (a_j + a_i), \dots, a_j + a_i, \dots, a_n) \\ & \overset{\text{(ES2)}}{\curvearrowright} & (a_1, \dots, -a_j, \dots, a_j + a_i, \dots, a_n) \\ & \overset{\text{(ES1)}}{\curvearrowright} & (a_1, \dots, -a_j, \dots, a_j + a_i - a_j, \dots, a_n) \\ & \overset{\text{(ES2)}}{\curvearrowright} & (a_1, \dots, a_j, \dots, a_i, \dots, a_n). \end{array}$$

□

Analog führt man elementare *Zeilenumformungen* (EZ1), ... , (EZ4) ein.

Satz 5.5
Es sei $A \in \mathrm{Mat}(m, n; K)$. Dann gilt

$$\boxed{\text{Zeilenrang } A = \text{Spaltenrang } A.}$$

Damit erhält man folgende wichtige

Definition. Der *Rang* einer Matrix A ist definiert durch

$$\text{Rang } A := \text{Zeilenrang } A \quad (\overset{(5.5)}{=} \text{Spaltenrang } A).$$

Bemerkung. $\text{Rang } A = \text{Rang } {}^tA$.

Dies folgt aus:

$$\begin{array}{rcl} \text{Rang } A & = & \text{Zeilenrang } A \overset{(5.5)}{=} \text{Spaltenrang } A \\ & = & \text{Zeilenrang } {}^tA = \text{Rang } {}^tA. \end{array}$$

Beweis von Satz (5.5). Es sei

$$\begin{array}{rcl} r & := & \text{Spaltenrang } A \\ s & := & \text{Zeilenrang } A. \end{array}$$

$$(1) \quad A = (a_1, \dots, a_n) = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Wir betrachten das Vertauschen von Spalten (ES4):

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1i} & \cdots & \alpha_{1j} & \cdots & \alpha_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mi} & \cdots & \alpha_{mj} & \cdots & \alpha_{mn} \end{pmatrix} \curvearrowright \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1j} & \cdots & \alpha_{1i} & \cdots & \alpha_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mj} & \cdots & \alpha_{mi} & \cdots & \alpha_{mn} \end{pmatrix}.$$

Dies ändert weder den Spaltenrang noch den Zeilenrang von A. Nach einer geeigneten Umordnung kann man dann annehmen, daß die ersten r Spalten $a_1, \dots, a_r$ linear unabhängig sind. Ebenso kann man durch Vertauschen der Zeilen erreichen, daß die ersten s Zeilen $b_1, \dots, b_s$ linear unabhängig sind.

(2) $s \leq r$

Wir beweisen dies durch Widerspruch und nehmen $s > r$ an. Dann betrachten wir folgendes Gleichungssystem

$$(1) \qquad \sum_{i=1}^{s} \alpha_{ik}\xi_i = 0 \qquad (k = 1, \dots, r).$$

Dies ist ein homogenes Gleichungssystem mit r Gleichungen und s Unbekannten. Nach dem Fundamentallemma (4.2) gibt es aufgrund der Annahme $s > r$ eine nicht-triviale Lösung $x_1, \dots, x_s \in K$.

Behauptung. $x_1 b_1 + \cdots + x_s b_s = 0$.

Dies ist ein Widerspruch zu der linearen Unabhängigkeit von $b_1, \dots, b_s$.

Beweis der Behauptung. $a_1, \dots, a_r$ erzeugen $\mathrm{Span}(a_1, \dots, a_n)$. D. h.es gibt $\rho_{jk} \in K$ mit

$$(2) \qquad a_j = \sum_{k=1}^{r} \rho_{jk} a_k. \qquad (j = 1, \dots, n)$$

Da

$$a_j = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix}$$

kann man dies auch schreiben als

$$(3) \qquad \alpha_{ij} = \sum_{k=1}^{r} \rho_{jk}\alpha_{ik}. \qquad (i = 1, \dots, m; j = 1, \dots, n).$$

Also gilt

$$\begin{aligned}\sum_{i=1}^{s} x_i\alpha_{ij} &\overset{(3)}{=} \sum_{i=1}^{s} x_i \sum_{k=1}^{r} \rho_{jk}\alpha_{ik}\\ &= \sum_{k=1}^{r} \rho_{jk} \underbrace{\sum_{i=1}^{s} x_i\alpha_{ik}}_{=0 \text{ nach } (1)} = 0.\end{aligned}$$

D. h.

$$\sum_{i=1}^{s} x_i\alpha_{ij} = 0. \qquad (j = 1, \dots, n) \tag{4}$$

Wegen

$$b_i = (\alpha_{i1}, \dots, \alpha_{in})$$

bedeutet dies

$$\sum_{i=1}^{s} x_i b_i = 0. \tag{5}$$

(3) $s \geq r$. Dies folgt nun formal aus:

$$\begin{aligned} r &= \text{Spaltenrang } A = \text{Zeilenrang } {}^tA \overset{\text{2. Schritt}}{\leq} \text{Spaltenrang } {}^tA\\ &= \text{Zeilenrang } A = s.\end{aligned}$$

□

Korollar 5.6
Für eine $(m \times n)$-Matrix A gilt

$$\text{Rang } A \leq \text{Min}(m, n).$$

Satz 5.7
Elementare Spaltenumformungen und Zeilenumformungen ändern den Rang einer Matrix nicht.

Beweis. Wegen Lemma (5.4) genügt es, dies für (ES1) und (ES2) zu zeigen.

(ES2): Für $\alpha \neq 0$ gilt offenbar $\text{Span}(a_1, \dots, a_i, \dots, a_n) = \text{Span}(a_1, \dots, \alpha a_i, \dots, a_n)$.

(ES1): Zu zeigen ist:

$$\text{Span}(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = \text{Span}(a_1, \dots, a_i, \dots, a_i + a_j, \dots, a_n).$$

Hierzu genügt es zu zeigen, daß

$$\mathrm{Span}(a,b) = \mathrm{Span}(a, a+b).$$

Dies gilt, da

$$\begin{aligned}
\mathrm{Span}(a,b) &= \{\alpha a + \beta b;\ \alpha,\beta \in K\} \\
&= \{\alpha a + \beta(a+b) - \beta a;\ \alpha,\beta \in K\} \\
&= \{(\alpha-\beta)a + \beta(a+b);\ \alpha,\beta \in K\} \\
&= \{\gamma a + \delta(a+b);\ \gamma,\delta \in K\} \\
&= \mathrm{Span}(a, a+b).
\end{aligned}$$

$\square$

Beispiel. $A = \begin{pmatrix} 1 & 2 & 4 \\ 4 & 5 & 6 \end{pmatrix} \in \mathrm{Mat}(2,3;\mathbb{Q})$. Dann gilt:

$$\begin{aligned}
A &= \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \overset{\text{EZ3}}{\curvearrowright} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \end{pmatrix} \\
&\overset{\text{EZ2}}{\curvearrowright} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix} \overset{\text{EZ3}}{\curvearrowright} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix} \\
&\overset{\text{ES1}}{\curvearrowright} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix} \overset{\text{ES3}}{\curvearrowright} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.
\end{aligned}$$

Insbesondere gilt

$$\mathrm{Rang}\, A = 2.$$

Kästchenschreibweise

Es sei $A \in \mathrm{Mat}(m,n;K)$. Ferner sei $1 \leq p < m$, $1 \leq q < n$. Man schreibt dann oft die Matrix

$$A = \begin{array}{c} \left(\begin{array}{ccc|ccc}
\alpha_{11} & \cdots & \alpha_{1q} & \alpha_{1,q+1} & \cdots & \alpha_{1n} \\
\vdots & & \vdots & \vdots & & \vdots \\
\alpha_{p1} & \cdots & \alpha_{pq} & \alpha_{p,q+1} & \cdots & \alpha_{pn} \\
\hline
\alpha_{p+1,1} & \cdots & \alpha_{p+1,q} & \alpha_{p+1,q+1} & \cdots & \alpha_{p+1,n} \\
\vdots & & \vdots & \vdots & & \vdots \\
\alpha_{m1} & \cdots & \alpha_{mq} & \alpha_{m,q+1} & \cdots & \alpha_{mn}
\end{array}\right) \\ \underbrace{\hspace{8em}}_{q}\ \underbrace{\hspace{8em}}_{n-q} \end{array}
\begin{array}{l} \left.\vphantom{\begin{array}{c}a\\b\\c\end{array}}\right\} p \\ \left.\vphantom{\begin{array}{c}a\\b\\c\end{array}}\right\} m-p \\ \vphantom{x} \end{array}$$

in der Form

$$A = \left(\begin{array}{c|c} A_1 & A_2 \\ \hline A_3 & A_4 \end{array}\right).$$

mit

$$\begin{array}{ll} A_1 \in \mathrm{Mat}(p,q;K); & A_2 \in \mathrm{Mat}(p,n-q;K) \\ A_3 \in \mathrm{Mat}(m-p,q;K); & A_4 \in \mathrm{Mat}(m-p,n-q;K) \end{array}$$

Spezialfall: p=q=1:

$$A = \left(\begin{array}{c|c} \alpha & b \\ \hline c & D \end{array}\right) \qquad \begin{array}{ll} \alpha \in K, & b \in K^{n-1} \\ c \in {}^{m-1}K, & D \in \mathrm{Mat}(m-1,n-1;K) \end{array}$$

Satz 5.8
Jede Matrix $A \in \mathrm{Mat}(m,n;K)$ vom Rang r kann durch elementare Zeilen- und Spaltenumformungen auf folgende Gestalt gebracht werden:

$$A = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Beweis. Wir beweisen dies per Induktion nach r.

$r = 0$: Dann ist $A = 0$ und es ist nichts zu zeigen.

$r - 1 \mapsto r$: Wir können annehmen, daß $A \neq 0$ ist. Nach endlich vielen Umformungen vom Typ (EZ4), (ES4) kann man annehmen, daß

$$A = \left(\begin{array}{c|c} \alpha & b \\ \hline c & A' \end{array}\right) \qquad \text{mit } \alpha \neq 0.$$

Mit (EZ2) erhält man

$$A \curvearrowright \left(\begin{array}{c|c} 1 & b \\ \hline c & A' \end{array}\right).$$

Mit (EZ3), (ES3) kann man diese Matrix überführen in

$$A \curvearrowright \left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & A'' \\ 0 & \end{array}\right).$$

Es ist

$$\mathrm{Rang}\, A = 1 + \mathrm{Rang}\, A''.$$

Nach Induktionsvoraussetzung erhält man schließlich:

$$\left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & A'' \\ 0 & \end{array}\right) \curvearrowright \left(\begin{array}{c|ccc|c} 1 & 0 & \cdots & & 0 \\ \hline 0 & & & & 0 \\ \vdots & & E_{r-1} & & \vdots \\ & & & & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 \end{array}\right) = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

□

Matrizenmultiplikation

Zur Motivation kommen wir nochmals auf homogene Gleichungssysteme zurück:

$$\begin{array}{ccccccc} \alpha_{11}\xi_1 & + & \cdots & + & \alpha_{1n}\xi_n & = & 0 \\ \vdots & & & & \vdots & & \\ \alpha_{i1}\xi_1 & + & \cdots & + & \alpha_{in}\xi_n & = & 0 \\ \vdots & & & & \vdots & & \\ \alpha_{m1}\xi_1 & + & \cdots & + & \alpha_{mn}\xi_n & = & 0. \end{array} \tag{1}$$

Die Koeffizientenmatrix lautet dann

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{i1} & \cdots & \alpha_{in} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_m \end{pmatrix},$$

wobei $a_1, \ldots, a_n$ die Zeilenvektoren von A sind. Führt man den Vektor

$$\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

ein, so ist die i-te Zeile des Gleichungssystems (1) gerade das „Skalarprodukt“

$$a_i\xi = (\alpha_{i1}, \ldots, \alpha_{in}) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \alpha_{i1}\xi_1 + \cdots + \alpha_{in}\xi_n.$$

Damit schreibt man (1) abkürzend in der Form

$$A\xi = 0.$$

Wir haben damit das „Produkt“ der $(m \times n)$-Matrix A mit dem Vektor ξ definiert. Dies wird verallgemeinert zur Matrizenmultiplikation. Wir haben bereits die folgende Notation eingeführt:

$$K^{(m\times n)} = \mathrm{Mat}(m, n; K).$$

Damit ist insbesondere

$$K^{(1\times n)} = K^n; \quad K^{(m\times 1)} = {}^mK.$$

Die *Matrizenmultiplikation* ist eine Abbildung

$$\begin{array}{rcl} K^{(m\times n)} \times K^{(n\times p)} & \longrightarrow & K^{(m\times p)} \\ (A, B) & \longmapsto & AB. \end{array}$$

Definition. Sei $A \in K^{(m\times n)}, B \in K^{(n\times p)}$. Dann ist das Produkt $C := AB \in K^{(m\times p)}$ erklärt durch

$$C = (\gamma_{ij}), \quad 1 \le i \le m, 1 \le j \le p$$

mit

$$\gamma_{ij} := \sum_{k=1}^{n} \alpha_{ik}\beta_{kj}.$$

Man erhält also des Eintrag γ_{ij} von C als das Skalarprodukt der i-ten Zeile von A mit der j-ten Spalte von B.

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \boxed{\alpha_{i1} \;\; \cdots \;\; \alpha_{in}} & & \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} \beta_{11} & \cdots & \boxed{\begin{matrix}\beta_{1j} \\ \vdots \\ \beta_{nj}\end{matrix}} & \cdots & \beta_{1p} \\ \vdots & & & & \vdots \\ \beta_{n1} & \cdots & & \cdots & \beta_{np} \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1j} & \cdots & \gamma_{1p} \\ \vdots & & \vdots & & \vdots \\ \gamma_{i1} & \cdots & \boxed{\gamma_{ij}} & \cdots & \gamma_{ip} \\ \vdots & & \vdots & & \vdots \\ \gamma_{m1} & \cdots & \gamma_{mj} & \cdots & \gamma_{mp} \end{pmatrix}$$

mit

$$\gamma_{ij} = (\alpha_{i1}, \ldots, \alpha_{in}) \begin{pmatrix} \beta_{1j} \\ \vdots \\ \beta_{nj} \end{pmatrix} = \sum_{k=1}^{n} \alpha_{ik}\beta_{kj}.$$

Beispiele.

(1)
$$\begin{pmatrix}\alpha_{11} & \alpha_{12}\\ \alpha_{21} & \alpha_{22}\end{pmatrix}\begin{pmatrix}\beta_{11} & \beta_{12}\\ \beta_{21} & \beta_{22}\end{pmatrix}=\begin{pmatrix}\alpha_{11}\beta_{11}+\alpha_{12}\beta_{21} & \alpha_{11}\beta_{12}+\alpha_{12}\beta_{22}\\ \alpha_{21}\beta_{11}+\alpha_{22}\beta_{21} & \alpha_{21}\beta_{12}+\alpha_{22}\beta_{22}\end{pmatrix}.$$

(2)
$$\begin{pmatrix}3 & 0 & 4 & 0\\ 2 & 0 & 1 & 0\end{pmatrix}\begin{pmatrix}0 & 2 & 1\\ 1 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & 0\end{pmatrix}=\begin{pmatrix}0 & 18 & 3\\ 0 & 7 & 2\end{pmatrix}.$$

(3) $E=(\delta_{ij}),\ A\in K^{(n\times n)}$:
$$\begin{pmatrix}1 & & \\ & \ddots & \\ & & 1\end{pmatrix}\begin{pmatrix}\alpha_{11} & \cdots & \alpha_{1n}\\ \vdots & & \vdots\\ \alpha_{n1} & \cdots & \alpha_{nn}\end{pmatrix}=\begin{pmatrix}\alpha_{11} & \cdots & \alpha_{1n}\\ \vdots & & \vdots\\ \alpha_{n1} & \cdots & \alpha_{nn}.\end{pmatrix},$$

d. h.
$$EA=A.$$

Lemma 5.9
Es gilt:

(i) $(\alpha A)B=\alpha(AB)=A(\alpha B)$,

(ii) $A(B+C)=AB+AC$,

(iii) $(A+B)C=AC+BC$,

(iv) $(AB)C=A(BC)$,

(v) ${}^t(AB)={}^tB\,{}^tA$.

Beweis.

(i) Klar.

(ii) Es sei $A=(\alpha_{ij}),\ B=(\beta_{jk}),\ C=(\gamma_{jk})$. Dann gilt:
$$\begin{aligned}(A(B+C))_{ik} &= \sum_j \alpha_{ij}(\beta_{jk}+\gamma_{jk})=\sum_j \alpha_{ij}\beta_{jk}+\sum_j \alpha_{ij}\gamma_{jk}\\ &= (AB)_{ik}+(AC)_{ik}.\end{aligned}$$

(iii) Zeigt man analog.

(iv) Es sei $A\in K^{(m\times n)},\ B\in K^{(n\times p)},\ C\in K^{(p\times r)}$ mit
$$\begin{aligned}A &= (\alpha_{ij}) \qquad 1\le i\le m; 1\le j\le n,\\ B &= (\beta_{jk}) \qquad 1\le j\le n; 1\le k\le p,\\ C &= (\gamma_{kl}) \qquad 1\le k\le p; 1\le l\le r.\end{aligned}$$

Wir betrachten das Produkt

$$C' := AB = (\gamma'_{ik}); \qquad 1 \le i \le m; 1 \le k \le p$$

mit

$$\gamma'_{ik} = \sum_{j=1}^{n} \alpha_{ij}\beta_{jk}.$$

Dann ist $(AB)C = C'C$ mit

$$(C'C)_{il} = \sum_{k=1}^{p} \gamma'_{ik}\gamma_{kl} = \sum_{k=1}^{p}(\sum_{j=1}^{n} \alpha_{ij}\beta_{jk})\gamma_{kl}.$$

Setzen wir

$$C'' := BC = (\gamma''_{jl}); \quad 1 \le j \le n; 1 \le l \le r$$

mit

$$\gamma''_{jl} = \sum_{k=1}^{p} \beta_{jk}\gamma_{kl},$$

so gilt für $A(BC) = AC''$:

$$(AC'')_{il} = \sum_{j=1}^{n} \alpha_{ij}\gamma''_{jl} = \sum_{j=1}^{n} \alpha_{ij}(\sum_{k=1}^{p} \beta_{jk}\gamma_{kl}).$$

Also folgt

$$(AB)C = C'C = AC'' = A(BC)$$

und dies ist gerade die Behauptung.

(v) Es gilt

$$({}^t(AB))_{ij} = (AB)_{ji} = \sum_{k=1}^{n} \alpha_{jk}\beta_{ki}.$$

Andererseits ist

$$({}^tB\,{}^tA)_{ij} = \sum_{k=1}^{n} ({}^tB)_{ik}({}^tA)_{kj} = \sum_{k=1}^{n} \beta_{ki}\alpha_{jk}.$$

Dies zeigt die Behauptung. □

Bemerkung. Im allgemeinen ist die Matrizenmultiplikation nicht kommutativ, wie folgendes Beispiel zeigt:

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

Spezielle Fälle der Matrizenmultiplikation

Im allgemeinen ist die Matrizenmultiplikation eine Abbildung

$$K^{(m\times n)} \times K^{(n\times p)} \longrightarrow K^{(m\times p)}.$$

1. Fall: $m = n = 1$. Dann ist $A = (\alpha)$

$$AB = (\alpha)(\beta_1, \ldots, \beta_p) = (\alpha\beta_1, \ldots, \alpha\beta_p).$$

2. Fall: $n = p = 1$. Dann ist $B = (\beta)$

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} \beta = \begin{pmatrix} \alpha_1\beta \\ \vdots \\ \alpha_m\beta \end{pmatrix} = \beta \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}.$$

3. Fall: $p = 1$, d. h. $B = b \in K^{(n\times 1)}$ ist ein Spaltenvektor. Dann gilt:

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n \alpha_{1k}\beta_k \\ \vdots \\ \sum_{k=1}^n \alpha_{mk}\beta_k \end{pmatrix} \in K^{(m\times 1)} = {}^mK.$$

Wir halten also fest:

$$\boxed{\text{Matrix} \cdot \text{Spaltenvektor} = \text{Spaltenvektor}}$$

4. Fall: $m = 1$, d. h. $A = a \in K^{(1\times n)} = K^n$ ist ein Zeilenvektor. Dann gilt:

$$(\alpha_1, \ldots, \alpha_n) \begin{pmatrix} \beta_{11} & \cdots & \beta_{1p} \\ \vdots & & \vdots \\ \beta_{n1} & \cdots & \beta_{np} \end{pmatrix} = (\sum_{i=1}^n \alpha_i\beta_{i1}, \ldots, \sum_{i=1}^n \alpha_i\beta_{ip}) \in K^{(1\times p)} = K^p.$$

Damit haben wir

$$\boxed{\text{Zeilenvektor} \cdot \text{Matrix} = \text{Zeilenvektor}}$$

5. Fall: $n = 1$, $A = a \in K^{(m\times 1)}$ ist ein Spaltenvektor und $B = b \in K^{(1\times p)}$ ist ein Zeilenvektor. Dann ist

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} (\beta_1, \ldots, \beta_p) = \begin{pmatrix} \alpha_1\beta_1 & \cdots & \alpha_1\beta_p \\ \vdots & & \vdots \\ \alpha_m\beta_1 & \cdots & \alpha_m\beta_p \end{pmatrix} \in K^{(m\times p)},$$

$$\boxed{\text{Spaltenvektor} \cdot \text{Zeilenvektor} = \text{Matrix}}$$

6. Fall: $m = p = 1$. In diesem Fall ist $a \in K^{(1\times n)}$ ein Zeilenvektor und $b \in K^{(n\times 1)}$ ein Spaltenvektor. Hier schließlich haben wir

$$(\alpha_1, \dots, \alpha_n) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \alpha_1\beta_1 + \dots + \alpha_n\beta_n = \sum_{i=1}^{n} \alpha_i\beta_i.$$

$$\boxed{\text{Zeilenvektor} \cdot \text{Spaltenvektor} = \text{Skalar}}$$

Als Spezialfall von Fall 3 halten wir hier noch fest:

$$Ae_j = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1j} & \cdots & \alpha_{1n} \\ \vdots & & \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mj} & \cdots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix}.$$

D. h.für $A = (a_1, \dots, a_n)$ gilt:

$$Ae_j = a_j.$$

Damit gilt für $x = \sum_{j=1}^{n} x_j e_j$:

$$Ax = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A(\sum_{j=1}^{n} x_j e_j) = \sum_{j=1}^{n} x_j a_j.$$

Schreibweise. Wir verzichten im folgenden endgültig auf die Unterscheidung von Zeilen- und Spaltenvektoren und identifizieren ${}^nK$ mit K^n.

Definition. Das *(Standard) Skalarprodukt* auf K^n ist definiert durch

$$\begin{array}{rcl} \langle\, ,\, \rangle : \; K^n \times K^n & \longrightarrow & K \\ \langle x, y \rangle & := & {}^t x y = (x_1, \dots, x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i. \end{array}$$

Eigenschaften: Das Skalarprodukt ist

(i) *bilinear*, d. h.

$$\begin{aligned} \langle \alpha x' + \beta x'', y \rangle &= \alpha\langle x', y\rangle + \beta\langle x'', y\rangle, \\ \langle x, \alpha y' + \beta y'' \rangle &= \alpha\langle x, y'\rangle + \beta\langle x, y''\rangle, \end{aligned}$$

(ii) *symmetrisch*, d. h.

$$\langle x, y \rangle = \langle y, x \rangle,$$

(iii) *nicht ausgeartet*, d. h.

$$\langle x, y \rangle = 0 \quad \text{für alle } y \in K^n \;\Rightarrow\; x = 0.$$

Dies folgt wegen $\langle x, e_k \rangle = x_k = 0$ für alle k, und damit $x = 0$.

Lineare Abbildungen zwischen Standardräumen

Es sei $A \in \mathrm{Mat}(m,n;K) = K^{(m\times n)}$.

Definition. Die Abbildung h_A ist definiert durch

$$\begin{aligned} h_A : K^n &\longrightarrow K^m \\ x &\longmapsto Ax. \end{aligned}$$

Bemerkung. h_A ist linear:

$$\begin{aligned} h_A(x+y) &= A(x+y) \overset{(5.9)(\mathrm{ii})}{=} Ax + Ay = h_A(x) + h_A(y), \\ h_A(\alpha x) &= A(\alpha x) \overset{(5.9)(\mathrm{i})}{=} \alpha(Ax) = \alpha h_A(x). \end{aligned}$$

Satz 5.10
Ist $f : K^n \to K^m$ linear, so gibt es genau eine Matrix A mit $f = h_A$.

Beweis. $\{e_j\}_{1\le j\le n}$, $\{e_i\}_{1\le i\le m}$, seien die Standardbasen von K^n und K^m. Wir betrachten

$$f(e_j) = \sum_{i=1}^m \alpha_{ij} e_i = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix} =: a_j \in K^m.$$

Setze

$$A := (a_1, \dots, a_n) \in K^{(m\times n)}.$$

Dann ist

$$h_A(e_j) = Ae_j = a_j = f(e_j).$$

D. h.es folgt, daß $h_A = f$ ist.

Eindeutigkeit: Sei $A' = (a'_1, \dots, a'_n)$ eine weitere Matrix mit $f = h_{A'}$. Dann gilt

$$a_j = f(e_j) = h_{A'}(e_j) = A'e_j = a'_j$$

und damit $A = A'$. □

Satz 5.11

$$h_{AB} = h_A \circ h_B.$$

Beweis.

$$\begin{aligned} h_{AB}(x) &= (AB)(x) = A(Bx) = h_A(h_B(x)) \\ &= (h_A \circ h_B)(x). \end{aligned}$$

□

Anwendung auf Linearformen

Definition. Eine *Linearform* auf K^n ist eine lineare Abbildung $\lambda : K^n \to K$.

Es sei $a \in K^n$. Dann ist

$$\begin{array}{rcl} \lambda : K^n & \longrightarrow & K \\ x & \longmapsto & {}^t\!ax = \langle a, x\rangle \end{array}$$

eine Linearform.

Korollar 5.12
Für jede Linearform λ auf K^n gibt es genau ein $a \in K^n$ mit

$$\lambda(x) = {}^t\!ax = \langle a, x\rangle \quad (\text{für alle } x \in K^n).$$

Beweis. Nach Satz (5.10) gibt es eine Matrix $A \in K^{(1\times n)}$ mit $\lambda = h_A$. Setze $a := {}^tA \in K^{(n\times 1)} = K^n$. Dann gilt

$$\lambda(x) = h_A(x) = Ax = {}^t\!ax.$$

Aus Satz (5.10) folgt auch die Eindeutigkeit. □

Dimensionsformel

Sei $A \in K^{(m\times n)}$. Dies definiert eine lineare Abbildung

$$\begin{array}{rcl} h_A : K^n & \longrightarrow & K^m \\ x & \longmapsto & Ax. \end{array}$$

Ist $A = (a_1, \dots , a_n)$, so gilt $Ax = \sum_{j=1}^n x_j a_j$. Damit folgt:

$$\boxed{\operatorname{Im} h_A = \{Ax;\ x \in K^n\} = \operatorname{Span}(a_1, \dots , a_n)}$$

Hieraus ergibt sich die Beziehung

$$\boxed{\operatorname{Rang} h_A = \operatorname{Rang} A}$$

Man setzt:

$$\begin{array}{rcl} \operatorname{Im} A & := & \{Ax;\ x \in K^n\} = \operatorname{Im} h_A \\ \operatorname{Ker} A & := & \{x;\ Ax = 0\} = \operatorname{Ker} h_A. \end{array}$$

Damit nimmt die Dimensionsformel die folgende Gestalt an:

$$\boxed{\dim \operatorname{Ker} A + \operatorname{Rang} A = n}$$

Schließlich ergibt sich noch für quadratische Matrizen $A \in K^{(n\times n)}$:

$$\boxed{\operatorname{Ker} A = \{0\} \Leftrightarrow \operatorname{Im} A = K^n \Leftrightarrow \operatorname{Rang} A = n}$$

Die allgemeine lineare Gruppe

Ist $E = E_n$ die Einheitsmatrix, so gilt für alle $A \in K^{(n\times n)}$:

$$AE = EA = A.$$

Definition. $U \in K^{(n\times n)}$ heißt *invertierbar*, falls es eine Matrix $V \in K^{(n\times n)}$ gibt mit

$$UV = VU = E.$$

Bemerkung. V ist eindeutig bestimmt. Man schreibt

$$V = U^{-1}.$$

Definition. $\mathrm{GL}(n, K) := \{U \in K^{(n\times n)};\ U \text{ ist invertierbar}\}$.

Lemma 5.13

(i) $U, V \in \mathrm{GL}(n, K) \Rightarrow UV \in \mathrm{GL}(n, K)$ mit $(UV)^{-1} = V^{-1}U^{-1}$.

(ii) $U \in \mathrm{GL}(n, K) \Rightarrow U^{-1} \in \mathrm{GL}(n, K)$ mit $(U^{-1})^{-1} = U$.

(iii) $U \in \mathrm{GL}(n, K), \alpha \in K^* \Rightarrow \alpha U \in \mathrm{GL}(n, K)$ mit $(\alpha U)^{-1} = \alpha^{-1}U^{-1}$.

(iv) $U \in \mathrm{GL}(n, K) \Rightarrow^t U \in \mathrm{GL}(n, K)$ mit $({}^tU)^{-1} =^t (U^{-1})$.

Beweis.

(i)

$$\begin{aligned} (UV)(V^{-1}U^{-1}) &= U(VV^{-1})U^{-1} = UU^{-1} = E, \\ (V^{-1}U^{-1})(UV) &= V^{-1}(U^{-1}U)V = V^{-1}V = E. \end{aligned}$$

(ii) Sofort aus der Definition.

(iii)

$$\begin{aligned} (\alpha U)(\alpha^{-1}U^{-1}) &= \alpha\alpha^{-1}UU^{-1} = E, \\ (\alpha^{-1}U^{-1})(\alpha U) &= \alpha^{-1}\alpha U^{-1}U = E. \end{aligned}$$

(iv)

$$\begin{aligned} UU^{-1} &= E \overset{(5.9)(\mathrm{v})}{\Rightarrow} {}^t(U^{-1})\,{}^tU = {}^tE = E, \\ U^{-1}U &= E \overset{(5.9)(\mathrm{v})}{\Rightarrow} {}^tU\,{}^t(U^{-1}) = {}^tE = E. \end{aligned}$$

Korollar 5.14

$\mathrm{GL}(n, K)$ ist zusammen mit der Matrizenmultiplikation eine Gruppe.

Bemerkung. Für $n \geq 2$ ist die Gruppe $GL(n,K)$ nicht abelsch.

Definition. $\mathrm{GL}(n,K)$ heißt die *allgemeine lineare Gruppe* (general linear group) in Dimension n.

Satz 5.15
Für $U \in \mathrm{Mat}(n \times n;K)$ sind äquivalent:

(i) $U \in \mathrm{GL}(n,K)$, d. h. U ist invertierbar.

(ii) Die Abbildung $h_U : K^n \to K^n, x \mapsto Ux$ ist bijektiv.

(iii) h_U ist injektiv.

(iv) h_U ist surjektiv.

(v) Die Spalten von U bilden eine Basis von K^n.

(vi) Die Zeilen von U bilden eine Basis von K^n.

(vii) $\mathrm{Rang}\, U = n$.

(viii) Es gibt $V \in \mathrm{Mat}(n,n;K)$ mit $UV = E$.

(ix) Es gibt $V \in \mathrm{Mat}(n,n;K)$ mit $VU = E$.

In den beiden letzten Fällen ist $V = U^{-1}$.

Beweis. (i)$\Rightarrow$(ii) Es gibt V mit $UV = VU = E$. Also ist $h_U \circ h_V = \mathrm{id}_E = h_V \circ h_U$ und damit ist h_U bijektiv.

(ii)$\Rightarrow$(i) Es sei
$$h_U^{-1} : K^n \longrightarrow K^n.$$
Diese Abbildung ist linear, also gibt es $V \in \mathrm{Mat}(n \times n;K)$ mit $h_U^{-1} = h_V$. Dann ist $h_{UV} = h_U h_V = \mathrm{id} = h_E$. Aus Satz (5.10) folgt hieraus $UV = E$. Analog zeigt man $VU = E$.

(ii)$\Leftrightarrow$(iii)$\Leftrightarrow$(iv) Folgt aus Satz (4.20).

(iv)$\Leftrightarrow$(v) Dies folgt, da für $U = (u_1, \ldots, u_n)$ gilt
$$\mathrm{Im}\, h_U = \mathrm{Span}(u_1, \ldots, u_n).$$

(v)$\Leftrightarrow$(vi)$\Leftrightarrow$(vii) Dies folgt sofort aus Satz (5.5).

(i)$\Rightarrow$(viii) Folgt aus der Definition einer invertierbaren Matrix.

(viii)$\Rightarrow$(iv)($\Rightarrow$(i)) Sei $a \in K^n, x := Va$. Dann ist $Ux = UVa = Ea = a$, also ist h_U surjektiv.

(i)$\Rightarrow$(ix) Nach Definition.

(ix)$\Rightarrow$(iii)($\Rightarrow$(i)) Aus $h_U(x) = Ux = 0$ folgt $VUx = 0$, also $x = Ex = 0$. Damit ist h_U injektiv. Die Äquivalenz von (i), (viii) und (ix) kann auch direkt aus Lemma (2.1) geschlossen werden. □

Satz 5.16
Es sei $A \in \mathrm{Mat}(m,n;K)$. Für $U \in \mathrm{GL}(m,K)$, $V \in \mathrm{GL}(n,K)$ gilt

$$\mathrm{Rang}(UAV) = \mathrm{Rang}\, A.$$

Lemma 5.17
Es sei $A \in \mathrm{Mat}(m,n;K), B \in \mathrm{Mat}(n,p;K)$. Dann gilt

$$\mathrm{Rang}(AB) \leq \mathrm{Min}(\mathrm{Rang}\, A, \mathrm{Rang}\, B).$$

Beweis. Wir betrachten die Abbildungen

$$K^p \xrightarrow{h_B} K^n \xrightarrow{h_A} K^m.$$

Es gilt $h_A \circ h_B = h_{AB}$ und damit folgt

$$\begin{aligned} \mathrm{Rang}\, AB &= \dim \mathrm{Im}\, h_{AB} = \dim(\mathrm{Im}(h_A | \mathrm{Im}\, h_B)). \\ &\leq \mathrm{Min}(\dim \mathrm{Im}\, h_B, \dim \mathrm{Im}\, h_A) = \mathrm{Min}(\mathrm{Rang}\, B, \mathrm{Rang}\, A). \end{aligned}$$

□

Beweis von Satz (5.16).
(i) Aus

$$\mathrm{Rang}(UA) \overset{(5.17)}{\leq} \mathrm{Rang}\, A = \mathrm{Rang}(U^{-1}(UA)) \overset{(5.17)}{\leq} \mathrm{Rang}(UA)$$

folgt

$$\mathrm{Rang}\, UA = \mathrm{Rang}\, A.$$

(ii) Ebenso zeigt man

$$\mathrm{Rang}\, AV = \mathrm{Rang}\, A.$$

(iii) Damit gilt dann (wobei (ii) auf UA angewendet wird)

$$\mathrm{Rang}(UAV) = \mathrm{Rang}((UA) \cdot V) \overset{\text{(ii)}}{=} \mathrm{Rang}\, UA \overset{\text{(i)}}{=} \mathrm{Rang}\, A.$$

□

Das Zentrum von $\mathbf{Mat(n;K)}$ und $\mathbf{GL(n,K)}$

Definition. Man nennt

$$Z(\mathrm{Mat}(n;K)) = \{A \in \mathrm{Mat}(n;K);\ AT = TA \text{ für alle } T \in \mathrm{Mat}(n;K)\}$$

das *Zentrum* von $M(n;K)$.

Satz 5.18

$$Z(\mathrm{Mat}(n;K)) = \{\lambda E;\ \lambda \in K\}.$$

Beweis. Die Matrizen

$$E_{kl} = \begin{pmatrix} 0 & \cdots & & 0 \\ \vdots & \ddots & 1 & \vdots \\ & & \ddots & \\ 0 & \cdots & & 0 \end{pmatrix} \in \mathrm{Mat}(n;K),$$

die in der k-ten Zeile und der l-ten Spalte eine 1 und sonst nur den Eintrag 0 haben, liefern eine Basis von $\mathrm{Mat}(n;K)$. Es sei

$$A = \sum_{k,l} \alpha_{kl} E_{kl}$$

im Zentrum. Dann gilt

$$E_{rs}A = \sum_{k,l} \alpha_{kl} E_{rs} E_{kl} = \sum_{k,l} \alpha_{kl} \delta_{sk} E_{rl} = \sum_{l} \alpha_{sl} E_{rl}.$$

Andererseits gilt

$$AE_{rs} = \sum_{k,l} \alpha_{kl} E_{kl} E_{rs} = \sum_{k,l} \alpha_{kl} \delta_{lr} E_{ks} = \sum_{k} \alpha_{kr} E_{ks}.$$

Aus

$$AE_{rs} = E_{rs}A$$

folgt nun

(1) $\alpha_{sl} = 0 \quad (l \neq s)$,

(2) $\alpha_{ss} = \alpha_{rr} \quad (l = s, k = r)$.

Insgesamt folgt daher

$$A = \alpha E.$$

□

Bemerkung. In obigem Beweis genügt es, den Fall $r \neq s$ zu betrachten.

Analog definiert man für die allgemeine lineare Gruppe:

Definition. Man nennt

$$\mathrm{Z}(\mathrm{GL}(n,K)) = \{A \in \mathrm{GL}(n,K);\ AT = TA \text{ für alle } T \in \mathrm{GL}(n,K)\}$$

das *Zentrum der Gruppe* $\mathrm{GL}(n,K)$.

Satz 5.19

$$\mathrm{Z}(\mathrm{GL}(n, K)) = \{\alpha E;\ \alpha \neq 0\}.$$

Beweis. Es sei $U \in \mathrm{Z}(\mathrm{GL}(n, K))$. Es genügt zu zeigen, daß $U \in \mathrm{Z}(\mathrm{Mat}(n; K))$ ist. Es genügt also zu zeigen (vgl. die Bemerkung nach Satz (5.18)), daß

$$UE_{kl} = E_{kl}U. \qquad (1 \leq k, l \leq n, k \neq l)$$

Nun ist

$$E_{kl} = (E + E_{kl}) - E.$$

Da $E + E_{kl}$ den Rang n hat (vgl. auch Lemma 5.20), ist $E + E_{kl} \in \mathrm{GL}(n, K)$. Also folgt

$$UE_{kl} = U(E + E_{kl}) - UE = (E + E_{kl})U - EU = E_{kl}U.$$

$\square$

Normalformensatz

Wir setzen

$$F_{kl} := E + E_{kl} = \begin{array}{cl} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & \cdots & 1 & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \leftarrow k \\ \\ \leftarrow l \\ \\ \\ \end{array} \\ \begin{array}{ccc} \uparrow k & \quad & \uparrow l \end{array} & \end{array} \qquad (k \neq l)$$

$$F_k(\alpha) := E + (\alpha - 1)E_{kk} = \begin{array}{cl} \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \leftarrow k \\ \\ \\ \end{array} \\ \uparrow k & \end{array} \quad (\alpha \in K^*).$$

Definition. Die Matrizen $F_{kl}, F_k(\alpha)$ heißen *Elementarmatrizen.*

Lemma 5.20
Die Elementarmatrizen sind invertierbar, d. h. $F_{kl}, F_k(\alpha) \in \mathrm{GL}(n, K) (\alpha \neq 0)$. Ferner sind die Inversen der Matrizen $F_{kl}, F_k(\alpha)$ wieder Produkte von Elementarmatrizen.

Beweis.

(i) Es ist

$$F_k(\alpha)\,F_k(\frac{1}{\alpha}) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha^{-1} & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} = E.$$

(ii) Es ist

$$F_{kl}(E - E_{kl}) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & 1 & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & -1 & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}.$$

Also ist

$$F_{kl}^{-1} = (E - E_{kl}).$$

Nun ist

$$F_l(-1)F_{kl}F_l(-1) =$$

$$= \begin{array}{c} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & -1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \\ \uparrow k \qquad \uparrow l \end{array} \begin{array}{c} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & 1 & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \\ \uparrow k \qquad \uparrow l \end{array} \begin{array}{c} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & -1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \\ \uparrow k \qquad \uparrow l \end{array}$$

$$= \begin{array}{c} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & -1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \\ \uparrow k \qquad \uparrow l \end{array} \begin{array}{c} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & -1 & & \\ & & & \ddots & \vdots & & \\ & & & & -1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \\ \uparrow k \qquad \uparrow l \end{array} = \begin{array}{c} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & -1 & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \\ \uparrow k \qquad \uparrow l \end{array}$$

$$= \quad E - E_{kl}.$$

□

Zusammenhang mit elementaren Umformungen

(1) $$F_k(\alpha)A = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{k1} & \cdots & \alpha_{kn} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix} \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha\alpha_{k1} & \cdots & \alpha\alpha_{kn} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix}$$

Multiplizieren von A mit $F_k(\alpha)$ von links bedeutet also Multiplizieren der k-ten Zeile von A mit α (EZ2).

Analog bedeutet Multiplizieren von A mit $F_k(\alpha)$ von rechts Multiplizieren der k-ten Spalte von A mit α (ES2).

(2) $$\begin{aligned} F_{kl}A &= \begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & \cdots & 1 & \\ & & & \ddots & \vdots & \\ & & & & 1 & \\ & 0 & & & & \ddots \\ & & & & & 1 \end{pmatrix} \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{k1} & \cdots & \alpha_{kn} \\ \vdots & & \vdots \\ \alpha_{l1} & \cdots & \alpha_{ln} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix} \\ &= \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{k1}+\alpha_{l1} & \cdots & \alpha_{kn}+\alpha_{ln} \\ \vdots & & \vdots \\ \alpha_{l1} & \cdots & \alpha_{ln} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix}. \end{aligned}$$

Multiplizieren von A mit F_{kl} von links bedeutet also Addition der l-ten Zeile von A zur k-ten Zeile von A (EZ1).

Analog bedeutet Multiplizieren von A mit F_{kl} von rechts Addition der k-ten Spalte von A zur l-ten Spalte von A (ES1).

Zusammenfassend kann man sagen, daß Multiplizieren mit Elementarmatrizen von links (rechts) gerade den elementaren Zeilenumformungen (Spaltenumformungen) entspricht.

Satz 5.21 (Normalenformensatz)
Zu jeder Matrix $0 \neq A \in K^{(m\times n)}$ gibt es Matrizen $U \in \mathrm{GL}(m,K), V \in \mathrm{GL}(n,K)$, die Produkte von Elementarmatrizen sind, so daß

$$UAV = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right) \qquad (r = \mathrm{Rang}\, A).$$

Beweis. Dies folgt nun aus der entsprechenden Aussage über elementare Umformungen (Satz (5.8)). $\square$

Satz 5.22
Sei $A \in K^{(n \times n)}$. Dann sind äquivalent

(i) $A \in \mathrm{GL}(n, K)$

(ii) A ist Produkt von Elementarmatrizen.

Beweis. (ii)$\Rightarrow$(i) Folgt, da Elementarmatrizen nach Lemma (5.20) invertierbar sind und $\mathrm{GL}(n, K)$ eine Gruppe ist.

(i)$\Rightarrow$(ii) Es gibt nach Satz (5.21) Matrizen U, V, die Produkte von Elementarmatrizen sind, so daß $UAV = E$ ist. Da U, V invertierbar sind, folgt $A = U^{-1}V^{-1}$. Die Aussage folgt dann aus dem zweiten Teil von Lemma (5.20). $\square$

Satz 5.23
Sei $A \in \mathrm{GL}(n, K)$. Dann kann A allein durch elementare Spaltenumformungen in E übergeführt werden.

Beweis. Sei $A \in \mathrm{GL}(n, K)$. Dann ist auch $A^{-1} \in \mathrm{GL}(n, K)$. Nach obigem Satz (5.22) gibt es Elementarmatrizen $F_1, \ldots, F_k$ mit

$$A^{-1} = F_1 \cdots F_k.$$

Hieraus folgt

$$AF_1 \cdots F_k = E.$$

Da Multiplikation mit Elementarmatrizen von rechts elementaren Spaltenumformungen entspricht, ergibt dies die Aussage. $\square$

Bemerkung. A sei invertierbar. Es sei $AF_1 \cdots F_k = E$. Dann ist

$$A^{-1} = E \cdot F_1 \cdots F_k.$$

Das heißt, führt man dieselben Spaltenumformungen, die A in E überführen, für E durch, so erhält man die inverse Matrix A^{-1}. Dies liefert ein Verfahren zur Berechnung der inversen Matrix.

Beispiel.

$$\begin{array}{c|c}
\begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\begin{pmatrix} 3 & 0 \\ 1 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \\
\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \\
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} \frac{1}{3} & 0 \\ -\frac{1}{6} & \frac{1}{2} \end{pmatrix}
\end{array}$$

Definition. Zwei Matrizen $A, B \in K^{(m\times n)}$ heißen *äquivalent*, falls es $U \in \mathrm{GL}(m,K), V \in \mathrm{GL}(n,K)$ gibt mit

$$A = UBV.$$

Schreibweise.

$$A \overset{\text{äquiv}}{\sim} B. \quad \text{Lies: } A \text{ ist äquivalent } B.$$

Bemerkung. $\overset{\text{äquiv}}{\sim}$ ist Äquivalenzrelation, d. h.

(i) $A \overset{\text{äquiv}}{\sim} A$.

(ii) $A \overset{\text{äquiv}}{\sim} B \Leftrightarrow B \overset{\text{äquiv}}{\sim} A$.

(iii) $A \overset{\text{äquiv}}{\sim} B, B \overset{\text{äquiv}}{\sim} C \Rightarrow A \overset{\text{äquiv}}{\sim} C$.

Beweis. (i) Mit $U = E_m \in \mathrm{GL}(m,K)$ und $V = E_n \in \mathrm{GL}(n,K)$ folgt $A = UAV$, also $A \overset{\text{äquiv}}{\sim} A$.

(ii) Ist $A = UBV$, so folgt $B = U^{-1}AV^{-1}$ und umgekehrt.

(iii) Aus $A = UBV$ und $B = U'CV'$ mit $U, U' \in \mathrm{GL}(m,K)$ und $V, V' \in \mathrm{GL}(n,K)$ folgt

$$A = UBV = U(U'CV')V = (UU')C(VV').$$

Da $UU' \in \mathrm{GL}(m,K)$ und $VV' \in \mathrm{GL}(n,K)$ ist, folgt $A \overset{\text{äquiv}}{\sim} C$. $\square$

Also kann man den Normalenformensatz (5.21) wie folgt formulieren:

$$\operatorname{Rang} A = r \Leftrightarrow A \overset{\text{äquiv}}{\sim} \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Folgerung 5.24

$$A \overset{\text{äquiv}}{\sim} B \Leftrightarrow \operatorname{Rang} A = \operatorname{Rang} B$$

Beweis. „$\Rightarrow$“ Folgt aus Satz (5.16).

„$\Leftarrow$“ Nach Voraussetzung folgt

$$A \overset{\text{äquiv}}{\sim} \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right), \qquad B \overset{\text{äquiv}}{\sim} \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Da $\sim$ eine Äquivalenzrelation ist, folgt $A \sim B$. $\square$

§ 6 Lineare Gleichungssysteme

Unter einem linearen Gleichungssystem verstehen wir ein System von Gleichungen

$$\begin{array}{ccccccc} \alpha_{11}\xi_1 & + & \cdots & + & \alpha_{1n}\xi_n & = & \beta_1 \\ \vdots & & & & & & \vdots \\ \alpha_{m1}\xi_1 & + & \cdots & + & \alpha_{mn}\xi_n & = & \beta_m \end{array}$$

mit Koeffizienten $\alpha_{ij}, \beta_i \in K$ und Unbekannten ξ_i. Dieses Gleichungssystem hat m Gleichungen mit n Unbekannten. Wir setzen

$$A = (\alpha_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in K^{(m \times n)}$$

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^n, \quad b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \in K^m.$$

Dann ist obiges Gleichungssystem gerade

$$(*) \qquad \boxed{Ax = b}$$

Schreibt man A als Matrix von Spaltenvektoren

$$A = (a_1, \ldots, a_n); \quad a_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} \in K^m$$

so lautet das Gleichungssystem auch

$$\xi_1 a_1 + \cdots + \xi_n a_n = b.$$

Definition. Die Matrix A heißt die *Koeffizientenmatrix* des Gleichungssystems, b nennt man auch die *rechte Seite*. Das Gleichungsystem heißt *homogen*, falls $b = 0$, ansonsten *inhomogen*.

Fragestellungen bei linearen Gleichungssystemen

(1) *Lösbarkeit:* D. h. unter welchen Bedingungen an A, b ist $(*)$ lösbar?

(2) *Universelle Lösbarkeit:* D. h. unter welchen Bedingungen an A ist $(*)$ für alle b lösbar?

(3) *Lösungsraum:* Beschreibung der Lösungsmenge.

(4) *Eindeutigkeit:* Unter welchen Bedingungen an A, b ist $(*)$ eindeutig lösbar?

(5) *Berechenbarkeit:* Man gebe einen Algorithmus zur Lösung von $(*)$ an.

Satz 6.1
Es sind äquivalent:

(i) $Ax = b$ ist lösbar.

(ii) $b \in \operatorname{Im} A$.

(iii) $\operatorname{Rang} A = \operatorname{Rang}(A, b)$.

Definition. (A, b) heißt die *erweiterte Koeffizientenmatrix*.

Beweis. (i)$\Leftrightarrow$(ii)

$$Ax = b \text{ lösbar } \Leftrightarrow \text{ Es gibt } x \in K^n \text{ mit } Ax = b \Leftrightarrow b \in \operatorname{Im} A.$$

(ii)$\Rightarrow$(iii)

$$\begin{aligned} b \in \operatorname{Im} A \quad &\Rightarrow \quad \text{Es gibt } x_1, \ldots, x_n \in K \text{ mit } x_1 a_1 + \cdots + x_n a_n = b \\ &\Rightarrow \quad b \in \operatorname{Span}(a_1, \ldots, a_n) \\ &\Rightarrow \quad \operatorname{Span}(a_1, \ldots, a_n, b) = \operatorname{Span}(a_1, \ldots, a_n) \\ &\Rightarrow \quad \operatorname{Rang}(A, b) = \operatorname{Rang} A. \end{aligned}$$

(iii)$\Rightarrow$(i)

$$\begin{aligned} \operatorname{Rang}(A, b) = \operatorname{Rang} A \quad &\overset{\text{Abh.lemma}}{\Rightarrow} \quad b \in \operatorname{Span}(a_1, \ldots, a_n) \\ &\Rightarrow \quad \text{Es gibt } x_1, \ldots, x_n \in K \text{ mit } b = x_1 a_1 + \cdots + x_n a_n \\ &\Rightarrow \quad (*) \text{ ist lösbar.} \end{aligned}$$

□

Satz 6.2
Es sind äquivalent:

(i) $(*)$ ist universell lösbar.

(ii) $\operatorname{Rang} A = m$.

Beweis. Es gilt für die Abbildung

$$A : K^n \longrightarrow K^m, \ x \longmapsto Ax$$

das folgende:

$$\begin{aligned} \operatorname{Rang} A = m \quad &\Leftrightarrow \quad A \text{ ist surjektiv} \\ &\Leftrightarrow \quad \text{Zu jedem } b \in K^m \text{ gibt es } x \text{ mit } Ax = b \\ &\Leftrightarrow \quad (*) \text{ ist universell lösbar.} \end{aligned}$$

□

Satz 6.3

(i) Die Lösungsmenge L_0 von $Ax = 0$ ist ein linearer Unterraum. Genauer gilt

$$L_0 = \operatorname{Ker} A, \quad \dim L_0 = n - \operatorname{Rang} A.$$

(ii) Die Lösungsmenge L_b von $Ax = b$ ist ein *affiner* Unterraum von K^n. D. h. ist u_0 eine spezielle Lösung von $Ax = b$, so gilt

$$L_b = u_0 + L_0.$$

Beweis. (i) Es gilt

$$Ax = 0 \Leftrightarrow x \in \operatorname{Ker} A.$$

Aus der Dimensionsformel folgt dann

$$\dim \operatorname{Ker} A + \operatorname{Rang} A = n.$$

(ii) $u_0 + L_0 \subset L_b$: Es sei $x \in L_0$. Dann gilt

$$A(u_0 + x) = Au_0 + \underbrace{Ax}_{=0} = Au_0 = b \Rightarrow u_0 + x \in L_b.$$

$L_b \subset u_0 + L_0$: Es sei $y \in L_b$. Wir setzen $x := (y - u_0)$. Dann ist $y = x + u_0$. Ferner gilt

$$Ax = A(y - u_0) = Ay - Au_0 = b - b = 0,$$

also $x \in L_0$. D. h. $y \in u_0 + L_0$. □

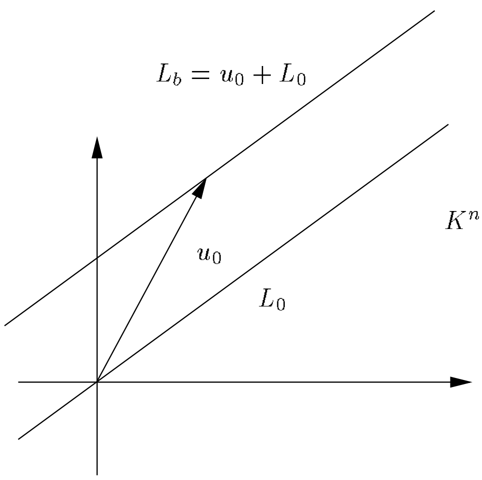


Abb. 32: Lösungsmenge L_b als affiner Unterrraum

Satz 6.4

$Ax = b$ sei lösbar. Dann sind äquivalent:

(i) $Ax = b$ ist eindeutig lösbar.

(ii) $\operatorname{Ker} A = \{0\}$.

(iii) $\operatorname{Rang} A = n$.

Beweis. Da $L_b = u_0 + L_0$ ist, gilt:

$$\begin{aligned} Ax = b \text{ ist eindeutig lösbar} &\Leftrightarrow Ax = 0 \text{ ist eindeutig lösbar} \\ &\Leftrightarrow \operatorname{Ker} A = \{0\} \\ &\Leftrightarrow \operatorname{Rang} A = n. \end{aligned}$$

Hierbei folgt die letzte Äquivalenz wieder aus der Dimensionsformel. □

Quadratische Gleichungssysteme $(m = n)$

Wir betrachten das Gleichungssystem

$$(*) \qquad Ax = b, \qquad A \in \operatorname{Mat}(n; K);\ b \in K^n.$$

Satz 6.5
Für quadratische Gleichungssysteme $(m = n)$ sind äquivalent:
(i) $Ax = b$ ist für jedes $b \in K^n$ lösbar.
(ii) $Ax = b$ ist für ein $b \in K^n$ eindeutig lösbar.
(iii) $Ax = 0$ besitzt nur die Lösung $x = 0$.
(iv) $Ax = b$ ist für jedes $b \in K^n$ eindeutig lösbar.
(v) $A \in \operatorname{GL}(n, K)$.

Beweis. Es gilt:

(i) $\Leftrightarrow Ax = b$ ist universell lösbar $\overset{(6.2)}{\Leftrightarrow} \operatorname{Rang} A = m = n \Leftrightarrow$ (v).

(iii) $\Leftrightarrow \operatorname{Ker} A = \{0\} \Leftrightarrow \operatorname{Rang} A = n \Leftrightarrow$ (v).

(ii) $\overset{(6.4)}{\Rightarrow} \operatorname{Rang} A = n \overset{(6.2)}{\Rightarrow} \left\{ \begin{matrix} Ax = b \text{ ist stets lösbar} \\ \operatorname{Ker} A = \{0\} \end{matrix} \right\} \overset{(6.4)}{\Rightarrow}$ (iv)

(iv)$\Rightarrow$(ii) ist trivial.

(ii) $\overset{(6.4)}{\Leftrightarrow} \operatorname{Ker} A = 0 \Leftrightarrow A \in \operatorname{GL}(n, K) \Leftrightarrow$ (v). □

Bemerkung. Es sei $A \in \operatorname{GL}(n, K)$. Dann erhält man die Lösung von

$$Ax = b$$

durch

$$x = A^{-1}b.$$

Der Gauß-Algorithmus

Wir beschreiben nun einen Algorithmus zur Lösung eines linearen Gleichungssystems

$$(*) \qquad Ax = b.$$

Es besteht aus drei Schritten:

I. Vorwärtselimination

II. Lösbarkeitsentscheidung (nur für $b \neq 0$)

III. Rückwärtssubstitution.

I. Vorwärtselimination

1. Eliminationsschritt Wir fragen zunächst, ob $\alpha_{11} \neq 0$ ist. Falls nicht, suchen wir in der ersten Spalte ein Element $\alpha_{k1} \neq 0$ und vertauschen die erste mit der k-ten Zeile. Falls alle $\alpha_{k1} = 0$ sind, fahren wir mit der nächsten Spalte fort, bis wir ein Element $\alpha_{ij} \neq 0$ finden. (Falls $A = 0$ ist, so ist die Aufgabe trivial, weil dann entweder jedes $x \in K^n$ ein Lösung ist, falls $b = 0$ ist, oder die Lösungsmenge leer ist, falls $b \neq 0$.) Wir sind dann in der Situation, daß wir annehmen können, daß die ersten $k-1$ Spalten von A gleich 0 sind, und daß $\alpha_{1k} \neq 0$ ist. Wir ziehen dann das $\frac{\alpha_{ik}}{\alpha_{1k}}$-fache der ersten Zeile von der i-ten Zeile ab. Das Ergebnis sieht für die erweiterte Koeffizientenmatrix dann wie folgt aus:

$$\left(\begin{array}{ccccccc|ccc|c}
0 & \cdots & 0 & \Box & * & \cdots & * & * & \cdots & * & * \\
\cline{8-11}
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & * & \cdots & * & * \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & * & \cdots & * & *
\end{array}\right).$$

Dabei bezeichnet $\Box$ eine von Null verschiedene Zahl und $*$ eine beliebige Zahl.

2. Eliminationsschritt Wir wenden dasselbe Verfahren nun auf die eingezeichnete Restmatrix an. Das Ergebnis ist eine Matrix in *Zeilenstufenform*:

$$\left(\begin{array}{cccccccccccccc|c}
0 & \cdots & 0 & \Box & * & \cdots & * & * & \cdots & & & & & * & \tilde{\beta}_1 \\
0 & \cdots & & & & & 0 & \Box & * & \cdots & * & \cdots & & * & \tilde{\beta}_2 \\
0 & \cdots & & & & & & & & 0 & \Box & * & \cdots & * & \tilde{\beta}_3 \\
\vdots & & & & & & & & & & \ddots & & \ddots & & \vdots \\
0 & \cdots & & & & & & & & 0 & \Box & * & \cdots & * & \tilde{\beta}_r \\
0 & \cdots & & & & & & & & & 0 & 0 & \cdots & 0 & \tilde{\beta}_{r+1} \\
\vdots & & & & & & & & & & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & & & & & & & & & 0 & 0 & \cdots & 0 & \tilde{\beta}_m
\end{array}\right).$$

II. Lösbarkeitsentscheidung

Ist einer der Einträge $\tilde{\beta}_{r+1}, \dots, \tilde{\beta}_m \neq 0$, so ist $\operatorname{Rang}(A, b) > \operatorname{Rang} A = r$, und das Gleichungssystem ist nicht lösbar.

III. Rückwärtssubstitution

Allgemeines Verfahren

(i) Die zu Spalten ohne $\square$-Stelle gehörenden Unbekannten sind die *freien Variablen*. Sie werden der Reihe nach gleich $\lambda_1, \dots, \lambda_{n-r}$ gesetzt.

(ii) Man löst dann das Gleichungssystem nach den zu den $\square$-Stellen gehörenden *abhängigen Variablen* aus und bestimmt diese nacheinander in Abhängigkeit von $\lambda_1, \dots, \lambda_{n-r}$.

Alternatives Verfahren

(i) Man ermittelt eine spezielle Lösung u_0 von $(*)$, indem man $\lambda_1 = \dots = \lambda_{n-r} = 0$ setzt.

(ii) Man ermittelt den Lösungsraum von $Ax = 0$: Setze $\tilde{\beta}_1 = \dots = \tilde{\beta}_m = 0$ und wähle für $j = 1, \dots, n-r$:

$$\lambda_i^{(j)} = \begin{cases} 0 & \text{für } i \neq j \\ 1 & \text{für } i = j. \end{cases}$$

Die Auflösung nach den abhängigen Variablen gibt dann linear unabhängige Lösungen $v_1, \dots, v_{n-r}$. Die allgemeine Lösung ergibt sich dann nach Satz (6.3) durch

$$x = u_0 + c_1 v_1 + \dots + c_{n-r} v_{n-r} \qquad (c_1, \dots, c_{n-r} \in K).$$

Beispiele.

(1) Wir betrachten das lineare Gleichungssystem ($K = \mathbb{R}$).

$$\begin{aligned} \xi_1 + \xi_2 + \xi_3 &= 1 \\ 2\xi_1 - \xi_2 + \xi_3 &= 0 \\ -4\xi_1 + 2\xi_2 - \xi_3 &= 0. \end{aligned}$$

Dann kann man das Gauß-Verfahren etwa in der folgenden Form aufschreiben:

		ξ_1	ξ_2	ξ_3		Regie
$\xi_1 + \xi_2 + \xi_3$	$= 1$	$\boxed{1}$	1	1	1	-2 4
$2\xi_1 - \xi_2 + \xi_3$	$= 0$	2	-1	1	0	1
$-4\xi_1 + 2\xi_2 - \xi_3$	$= 0$	-4	2	-1	0	1
$\xi_1 + \xi_2 + \xi_3$	$= 1$	1	1	1	1	
$-3\xi_2 - \xi_3$	$= -2$	0	$\boxed{-3}$	-1	-2	2
$6\xi_2 + 3\xi_3$	$= 4$	0	6	3	4	1
$\xi_1 + \xi_2 + \xi_3$	$= 1$	1	1	1	1	
$-3\xi_2 - \xi_3$	$= -2$	0	-3	-1	-2	
ξ_3	$= 0$	0	0	1	0	

Rückwärtssubstitution ergibt

$$x_3 = 0, \quad x_2 = \frac{2}{3}, \quad x_1 = \frac{1}{3}.$$

Das Gleichungssystem ist eindeutig gelöst durch

$$x = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ 0 \end{pmatrix}.$$

(2) Wir betrachten das lineare Gleichungssystem ($K = \mathbb{R}$):

$$\begin{aligned} \xi_1 + \xi_2 + \xi_3 + \xi_4 + \xi_5 &= 1 \\ \xi_3 - \xi_4 + \xi_5 &= 1 \\ \xi_4 - \xi_5 &= 2 \\ 0 &= 0. \end{aligned}$$

Die erweiterte Koeffizientenmatrix ist daher

$$(A, b) = \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Im Sinne des oben erläuterten Schemas hat diese Matrix die Gestalt

$$(A, b) = \left(\begin{array}{ccccc|c} \Box & * & * & * & * & * \\ 0 & 0 & \Box & * & * & * \\ 0 & 0 & 0 & \Box & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Wir haben also folgende freie Variable:

$$x_2 = \lambda_1, \quad x_5 = \lambda_2.$$

Rückwärtssubstitution ergibt dann:

$$\begin{aligned} x_4 &= 2 + x_5 = 2 + \lambda_2 \\ x_3 &= 1 + x_4 - x_5 = 3 \\ x_1 &= 1 - x_2 - x_3 - x_4 - x_5 = -4 - \lambda_1 - 2\lambda_2. \end{aligned}$$

Alternativ können wie zuerst eine *spezielle* Lösung bestimmen. Für $\lambda_1 = \lambda_2 = 0$ erhält man

$$u_0 = \begin{pmatrix} -4 \\ 0 \\ 3 \\ 2 \\ 0 \end{pmatrix}.$$

Lösungen des *homogenen Systems* erhält man wie folgt:

$\lambda_1 = 1,\ \lambda_2 = 0:$ Rückwärtssubstitution ergibt

$$v_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

$\lambda_1 = 0,\ \lambda_2 = 1:$ Rückwärtssubstitution ergibt

$$v_2 = \begin{pmatrix} -2 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Damit erhält man den Lösungsraum des homogenen Systems als

$$L_0 = \mathrm{Span}(v_1, v_2).$$

(Man beachte, daß $\dim L_0 = 2 = 5 - \mathrm{Rang}\, A$ ist.) Die Lösungsmenge des linearen Gleichungssystems $(*)$ ist damit

$$L = u_0 + L_0 = \{ \begin{pmatrix} -4 \\ 0 \\ 3 \\ 2 \\ 0 \end{pmatrix} + c_1 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -2 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} ;\ c_1, c_2 \in \mathbb{R}\}.$$

§ 7 Determinanten

Vorbereitungen

Für zwei Vektoren $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$; $x, y \in \mathbb{R}^2$ definieren wir die *Determinante* provisorisch als

$$\det(x, y) := \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} := \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} := x_1 y_2 - x_2 y_1.$$

Bemerkung. Es gilt

$$x, y \text{ linear abhängig } \Leftrightarrow \det(x, y) = 0.$$

Beweis. „$\Rightarrow$" Ist $x = 0$, so ist offensichtlich $\det(x, y) = 0$. Ist $x \neq 0$, gibt es ein $\lambda \in \mathbb{R}$ mit $y = \lambda x$. Dann ist

$$\det(x, y) = \det \begin{pmatrix} x_1 & \lambda x_1 \\ x_2 & \lambda x_2 \end{pmatrix} = \lambda x_1 x_2 - \lambda x_2 x_1 = 0.$$

„$\Leftarrow$" Es sei $x_1 y_2 - x_2 y_1 = 0$. Ist $x_1 = x_2 = 0$ sind x, y linear abhängig. Wir nehmen nun $x_1 \neq 0$ an (der Fall $x_2 \neq 0$ kann analog behandelt werden). Ist $y_1 = 0$, so folgt $y_2 = 0$, also $y = 0$. Ist $y_1 \neq 0$ und $y_2 = 0$, so folgt $x_2 = 0$. In diesem Fall sind x und y linear abhängig. Ist $y_1, y_2 \neq 0$, so erhalten wir $x_1/y_1 = x_2/y_2$. Also gilt für $\lambda := x_1/y_1 = x_2/y_2$, daß $x = \lambda y$ ist. □

Bemerkung. Das von x, y aufgespannte Parallelogramm

$$S = \{\lambda x + \mu y;\ 0 \leq \lambda, \mu \leq 1\}$$

hat die Fläche

$$|S| = |\det(x, y)|.$$

Beweis. Es sei $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. Der Vektor $y' = \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix}$ steht auf y senkrecht.

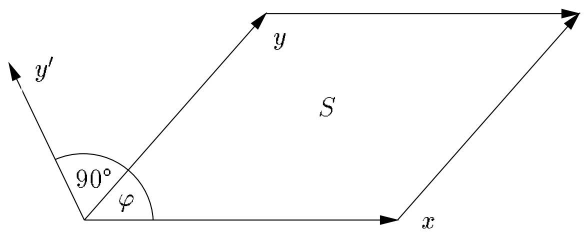


Abb. 33: Flächeninhalt eines Parallelogramms

Es gilt

$$\begin{aligned} |S| &= \|x\|\,\|y\|\,|\sin\varphi| = \|x\|\,\|y\|\,|\cos(\frac{\pi}{2}+\varphi)| \\ &= |\langle x, y'\rangle| = |x_1(-y_2)+x_2y_1| \\ &= |\det\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}|. \end{aligned}$$

Im $\mathbb{R}^3$ betrachten wir in diesem Zusammenhang das *Vektorprodukt* von zwei Vektoren.

Definition. Für zwei Vektoren $x, y \in \mathbb{R}^3$ ist das *Vektorprodukt* definiert durch

$$x \times y := \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

Beispiel. Es gilt

$$e_1 \times e_2 = \begin{pmatrix} 0-0 \\ 0-0 \\ 1-0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = e_3$$

Analog erhält man

$$e_2 \times e_3 = e_1, \quad e_1 \times e_3 = -e_2.$$

Man kann sich die Definition des Vektorprodukts formal auch wie folgt merken

$$x \times y = \begin{vmatrix} e_1 & x_1 & y_1 \\ e_2 & x_2 & y_2 \\ e_3 & x_3 & y_3 \end{vmatrix} := \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} e_1 - \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} e_2 + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} e_3.$$

Wir werden das später als das „Entwickeln einer (3×3)-Matrix nach der ersten Spalte“ erkennen.

Lemma 7.1
Für Vektoren $x, y, z \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$ gilt:

(i) $(x+y)\times z = x\times z + y\times z, \quad x\times(y+z) = x\times y + x\times z, \quad \lambda(x\times y) = (\lambda x)\times y = x\times(\lambda y).$

(ii) $\langle x \times y, x\rangle = \langle x \times y, y\rangle = 0.$

(iii) $\|x \times y\|^2 = \|x\|^2 \, \|y\|^2 - \langle x, y\rangle^2.$

(iv) $x \times y = -y \times x, \quad x \times x = 0.$

(v) $x \times y = 0 \Leftrightarrow x, y$ sind linear abhängig.

Beweis. (i), (ii) und (iv) rechnet man sofort nach.

(iii) Es gilt

$$\begin{aligned}
\|x \times y\|^2 &= (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - x_2y_1)^2 \\
&= (x_1y_2)^2 + (x_1y_3)^2 + (x_2y_1)^2 + (x_2y_3)^2 + (x_3y_1)^2 + (x_3y_2)^2 \\
&\quad -2(x_1x_2y_1y_2 + x_1x_3y_1y_3 + x_2x_3y_2y_3) \\
&= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2 \\
&= \|x\|^2\|y\|^2 - \langle x, y\rangle^2.
\end{aligned}$$

(iv) Die Aussage ist klar für $y = 0$. Es sei also $y \neq 0$. Nach (iii) ist $x \times y = 0$ äquivalent zu

$$\|x\|^2\|y\|^2 = \langle x, y\rangle^2.$$

Nach der Cauchy-Schwarzschen Ungleichung (Satz(1.8)) ist dies wiederum äquivalent dazu, daß x und y linear abhängig sind. □

Korollar 7.2
Es gilt

$$\|x \times y\| = \|x\| \, \|y\| \sin \sphericalangle(x, y).$$

Beweis. Es sei $\varphi := \sphericalangle(x, y)$. Dann gilt nach der Definition des Winkels

$$\langle x, y\rangle = \|x\| \, \|y\| \cos \varphi.$$

Aus Lemma (7.1) (ii) folgt daher

$$\begin{aligned}
\|x \times y\|^2 &= \|x\|^2\|y\|^2 - \langle x, y\rangle^2 \\
&= \|x\|^2\|y\|^2(1 - \cos^2 \varphi) \\
&= \|x\|^2\|y\|^2 \sin^2 \varphi.
\end{aligned}$$

Da $0 \leq \varphi \leq \pi$ ist, gilt $\sin \varphi \geq 0$, also erhält man die Behauptung durch Wurzelziehen. □

Folgerung 7.3
Der Vektor $x \times y$ steht senkrecht auf x und y, und die Länge von $x \times y$ ist gleich dem Flächeninhalt des von x und y aufgespannten Parallelogramms.

Beweis. Dies folgt sofort aus Lemma (7.1) (ii) und Korollar (7.2). □

Die in Folgerung (7.3) genannten Eigenschaften bestimmen den Vektor $x \times y$ bis auf das Vorzeichen. Letzteres kann man durch die „Dreifingerregel“ der rechten Hand bestimmen: Zeigen Daumen und Zeigefinger der rechten Hand in Richtung von x und y, so zeigt der Mittelfinger in Richtung von $x \times y$.

Schließlich führen wir noch das *Spatprodukt* ein.

Definition. Für Vektoren $x, y, z \in \mathbb{R}^3$ ist das *Spatprodukt* die reelle Zahl

$$[x, y, z] := \langle x \times y, z \rangle.$$

Lemma 7.4
Es sind äquivalent:
(i) $[x, y, z] = 0$
(ii) x, y, z sind linear abhängig.

Beweis. Wir hatten bereits gesehen, daß $x \times y = 0$ genau dann gilt, wenn x und y linear abhängig sind. Die Aussage folgt dann aus der

Behauptung. Es sei $x \times y \neq 0$. Dann gilt

$$[x, y, z] = 0 \;\Leftrightarrow\; z \in \mathbb{R}x + \mathbb{R}y.$$

Beweis der Behauptung. Ist $x \times y \neq 0$, so ist

$$E := \{z \in \mathbb{R}^3;\ \langle x \times y, z \rangle = 0\}$$

nach Satz (1.12) eine Ebene durch den Nullpunkt, die wegen Lemma (7.1) (ii) die Vektoren x und y enthält. Da x und y linear unabhängig sind, gilt

$$E = \mathbb{R}x + \mathbb{R}y.$$

Also ist $z \in E$ genau dann, wenn

$$[x, y, z] = \langle x \times y, z \rangle = 0.$$

□

Wir definieren wieder provisorisch die *Determinante* durch

$$\det(x, y, z) := [x, y, z] = \langle x \times y, z \rangle.$$

Bemerkung. Direktes Nachrechnen zeigt sofort, daß

$$\det(x, y, z) = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} := \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} z_1 - \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} z_2 + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} z_3.$$

Schließlich betrachten wir für drei Vektoren x, y, z den durch diese Vektoren aufgespannten *Spat*

$$S = \{\lambda x + \mu y + \nu z;\ 0 \leq \lambda, \mu, \nu \leq 1\}.$$

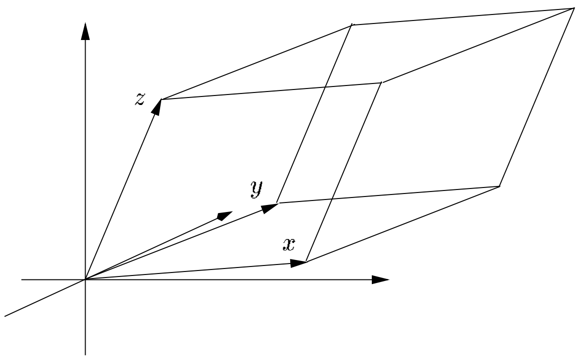


Abb. 34: Von drei Vektoren aufgespannter Spat

Lemma 7.5
Es gilt für das Volumen des Spates:

$$|S| = |[x, y, z]| = |\det(x, y, z)|.$$

Beweis. Es gilt

$$\begin{aligned} |[x, y, z]| &= |\langle x \times y, z\rangle| = \|x \times y\|\,\|z\|\,|\cos \sphericalangle(x \times y, z)| \\ &= |S|. \end{aligned}$$

□

Wir haben gesehen, daß Determinanten bei zwei Typen von Problemen natürlich auftreten, und zwar einmal bei der Frage, ob Vektoren im $\mathbb{R}^n$ linear unabhängig sind, und zum anderen bei der Volumenberechnung. Der erste Aspekt wird uns in der linearen Algebra interessieren, der zweite Aspekt spielt in der Integrationstheorie eine Rolle. Bevor wir zur allgemeinen Theorie übergehen, betrachten wir noch zwei Anwendungen für Geraden und Ebenen.

Ebenengleichungen im $\mathbb{R}^3$

Es seien w_1, w_2 linear unabhängig und

$$E = v + \mathbb{R}w_1 + \mathbb{R}w_2$$

eine Ebene im $\mathbb{R}^3$. Wegen Folgerung (7.3) steht der Vektor $s = w_1 \times w_2$ auf E senkrecht. Ist $s = (a_1, a_2, a_3)$, so gibt es nach Satz (1.6) ein $b \in \mathbb{R}$, so daß E durch die Gleichung

$$a_1 x_1 + a_2 x_2 + a_3 x_3 = b$$

beschrieben wird. Die Zahl b kann man durch Einsetzen eines Punktes auf E bestimmen.

Beispiel. Gesucht ist die Gleichung der Ebene E durch die Punkte

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

Als Richtungsvektoren können wir

$$w_1 = v_2 - v_1 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \quad w_2 = v_3 - v_1 = \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}$$

verwenden. Dann ist

$$s = w_1 \times w_2 = \begin{pmatrix} 5 \\ -1 \\ 2 \end{pmatrix}.$$

Damit ist

$$E: \quad 5x_1 - x_2 + 2x_3 = b.$$

Die Zahl b erhalten wir durch Einsetzen, etwa von v_1, d. h.

$$b = 3.$$

Also ist

$$E: \quad 5x_1 - x_2 + 2x_3 = 3.$$

Abstand zweier windschiefer Geraden

Gegeben seien zwei windschiefe Geraden

$$A_1 = v_1 + \mathbb{R}w_1, \quad A_2 = v_2 + \mathbb{R}w_2.$$

Der kürzeste Abstand zwischen zwei Punkten auf A_1 und A_2 ist (wegen des Satzes von Pythagoras) der senkrechte Abstand.

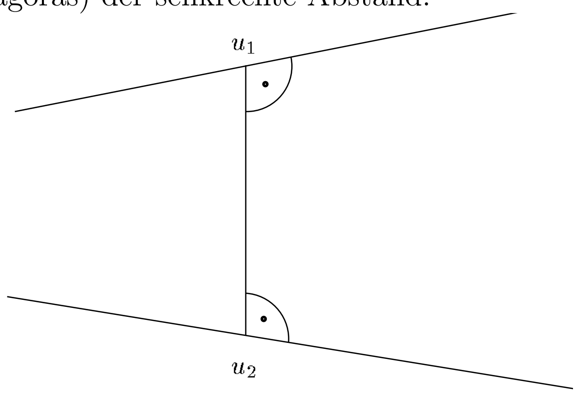


Abb. 35: Abstand zweier windschiefer Geraden

Die Aufgabe besteht also darin, Skalare $\lambda_1, \lambda_2 \in \mathbb{R}$ so zu finden, daß für

$$u_1 = v_1 + \lambda_1 w_1, \quad u_2 = v_2 + \lambda_2 w_2$$

der Vektor

$$d := u_2 - u_1 = v_2 - v_1 + \lambda_2 w_2 - \lambda_1 w_1$$

auf A_1 und A_2 senkrecht steht. Da A_1 und A_2 windschief sind, ist $w_1 \times w_2 \neq 0$ und $w_1, w_2, w_1 \times w_2$ bilden eine Basis von $\mathbb{R}^3$. Also gibt es, da $w_1 \times w_2$ senkrecht auf w_1 und w_2 steht, und d dieselbe Eigenschaft hat, einen Skalar α mit $d = \alpha(w_1 \times w_2)$. Für dieses α gilt also

$$\alpha(w_1 \times w_2) = (v_2 - v_1) + \lambda_2 w_2 - \lambda_1 w_1 = d. \tag{1}$$

Dies ist ein lineares Gleichungssystem für λ_1, λ_2 und α und da w_1, w_2 und $w_1 \times w_2$ eine Basis von $\mathbb{R}^3$ ist, besitzt dieses Gleichungssystem eine eindeutige Lösung. Wir betrachten

$$x := v_2 - v_1.$$

Da A_1 und A_2 windschiefe Geraden sind, sind w_1, w_2 und x linear unabhängig, bilden also eine Basis des $\mathbb{R}^3$. Also gibt es eindeutige bestimmte Skalare ϱ, μ_1, μ_2 mit

$$w_1 \times w_2 = \varrho x + \mu_1 w_1 + \mu_2 w_2. \tag{2}$$

Um (2) zu berechnen, wenden wir die Linearformen

$$\langle w_1 \times w_2, - \rangle, \langle w_1 \times x, - \rangle, \langle w_2 \times x, - \rangle \in (\mathbb{R}^3)^*$$

auf (2) an. (Es sei daran erinnert, daß jeder Vektor $a \neq 0$ nach Korollar (5.12) eine Linearform

$$\begin{array}{rccc} \langle a, - \rangle : & \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \langle a, x \rangle \end{array}$$

definiert.) Dies ergibt wegen Lemma (7.1) (ii):

$$\begin{array}{rcl} \langle w_1 \times w_2, w_1 \times w_2 \rangle & = & \langle w_1 \times w_2, \varrho x \rangle \\ \langle w_1 \times x, w_1 \times w_2 \rangle & = & \langle w_1 \times x, \mu_2 w_2 \rangle \\ \langle w_2 \times x, w_1 \times w_2 \rangle & = & \langle w_2 \times x, \mu_1 w_1 \rangle. \end{array}$$

Hieraus folgt dann sofort, daß

$$\begin{array}{rcl} \varrho & = & \dfrac{\|w_1 \times w_2\|^2}{\langle w_1 \times w_2, x \rangle} \\ \mu_2 & = & \dfrac{\langle w_1 \times x, w_1 \times w_2 \rangle}{\langle w_1 \times x, w_2 \rangle} = -\dfrac{\langle w_1 \times x, w_1 \times w_2 \rangle}{\langle w_1 \times w_2, x \rangle} \\ \mu_1 & = & \dfrac{\langle w_2 \times x, w_1 \times w_2 \rangle}{\langle w_2 \times x, w_1 \rangle} = \dfrac{\langle w_2 \times x, w_1 \times w_2 \rangle}{\langle w_1 \times w_2, x \rangle}. \end{array}$$

Da w_1, w_2 und $x = v_2 - v_1$ gegeben sind, sind ϱ, μ_1, μ_2 hieraus eindeutig berechenbar. Wir behaupten, daß dies die eindeutig bestimmte Lösung von (2) ist. Nach unserer Rechnung ist (ϱ, w_1, w_2) die einzig mögliche Lösung von (2). Andererseits ist nach Satz (6.5) das lineare Gleichungssystem eindeutig lösbar. Also muß (ϱ, w_1, w_2) die eindeutig bestimmte Lösung sein. Damit erhält man den Abstand $\|d\|$ und die Skalare, die zu den Fußpunkten u_1, u_2 gehören, wie folgt. Durch Vergleich von (1) und (2) ergibt sich, daß $\alpha = \frac{1}{\varrho}$, $\lambda_1 = -\frac{\mu_1}{\varrho}$, $\lambda_2 = \frac{\mu_2}{\varrho}$ ist. Also erhält man:

$$\begin{aligned}
\|d\| &= \frac{1}{|\varrho|}\|w_1 \times w_2\| = \frac{|\langle w_1 \times w_2, v_2 - v_1\rangle|}{\|w_1 \times w_2\|} \\
\lambda_1 &= -\frac{\mu_1}{\varrho} = \frac{\langle w_1 \times w_2, (v_2 - v_1) \times w_2\rangle}{\|w_1 \times w_2\|^2} \\
\lambda_2 &= \frac{\mu_2}{\varrho} = \frac{\langle w_1 \times w_2, (v_2 - v_1) \times w_1\rangle}{\|w_1 \times w_2\|^2}.
\end{aligned}$$

In diese Formeln kann man bei gegebenen Geraden A_1 und A_2 direkt einsetzen.

Wir wenden uns nun der allgemeinen Theorie der Determinanten zu.

Die symmetrische Gruppe

Es sei $n \in \mathbb{N}$ und M_n die Menge

$$M_n := \{1, 2, \ldots, n\}.$$

Definition. Die Gruppe

$$S_n := \mathrm{Bij}(M_n, M_n) = \{\sigma;\ \sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\} \text{ ist bijektiv}\}$$

heißt die *symmetrische Gruppe* (*Permutationsgruppe*) von n Elementen.

Lemma 7.6
$\#S_n = n!$.

Beweis. Wir beweisen dies durch Induktion nach n.

$n = 1$: Klar.

$n \mapsto n + 1$: Für $i = 1, \ldots, n + 1$ betrachten wir die Menge

$$S^i_{n+1} := \{\sigma \in S_{n+1};\ \sigma(1) = i\}.$$

Dann ist S_{n+1} die disjunkte Vereinigung

$$S_{n+1} = S^1_{n+1} \cup \cdots \cup S^{n+1}_{n+1}.$$

Die Menge S^i_{n+1} ist bijektiv zu der Menge der bijektiven Abbildungen

$$\{\sigma';\ \sigma' : M_{n+1}\backslash\{1\} \longrightarrow M_{n+1}\backslash\{i\} \text{ bijektiv}\} \xleftrightarrow{1:1} S_n$$

und damit selbst bijektiv zur symmetrischen Gruppe S_n.

Also gilt

$$\#S_{n+1} = \#S^1_{n+1} + \cdots + \#S^{n+1}_{n+1} \overset{\text{IV}}{=} (n+1)n! = (n+1)!.$$

□

Schreibweise. $\sigma \in S_n$; $i \mapsto \sigma(i)$ schreibt man als

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

Beispiel.

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Dann ist

$$\sigma_1 \circ \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \sigma_2 \circ \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \neq \sigma_1 \circ \sigma_2.$$

Bemerkung. Wir hatten bereits gesehen, daß die Gruppe S_n für $n \geq 3$ nicht abelsch ist.

Definition. Ein Element $\tau \in S_n$ heißt eine *Transposition*, falls es zwei Elemente vertauscht und alle anderen festläßt, d. h. falls es $k \neq l$ gibt mit

(1) $\tau(k) = l, \quad \tau(l) = k$

(2) $\tau(i) = i$ für $i \neq k, l$.

Bemerkung. $\tau^2 = \text{id}$, $\tau = \tau^{-1}$.

Lemma 7.7
Jede Permutation $\sigma \in S_n$ läßt sich als endliche Hintereinanderschaltung von Transpositionen schreiben, d. h. die Transpositionen erzeugen die Gruppe S_n.

Beweis. Ist $\sigma = \text{id}$, so gilt $\sigma = \tau^2$ für jede Permutation τ. Sei nun $\sigma \neq \text{id}$. Wir setzen

$$i_1 := \min\{i;\ \sigma(i) \neq i\}.$$

Dann gilt $\sigma(i_1) > i_1$. Es sei τ_1 diejenige Transposition, die i_1 und $\sigma(i_1)$ vertauscht.

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

Dann gilt

$$\sigma_1(i) = i \quad \text{für } i \leq i_1.$$

Sei
$$i_2 := \min\{i;\ \sigma_1(i) \neq i\}.$$
Es ist
$$i_2 > i_1.$$
Dann wiederhole man obiges Verfahren. Dies muß nach endlich vielen Schritten abbrechen, d. h. es gibt Permutationen $\tau_1, \ldots, \tau_k$ mit
$$\tau_k \circ \cdots \circ \tau_1 \circ \sigma = \mathrm{id},$$
d. h.
$$\sigma = (\tau_k \circ \cdots \circ \tau_1)^{-1} = \tau_1^{-1} \circ \cdots \circ \tau_k^{-1} = \tau_1 \circ \cdots \circ \tau_k.$$
□

Definition.
$$\tau_0 := \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ 2 & 1 & 3 & \cdots & n \end{pmatrix}.$$

Lemma 7.8
Sei τ eine Transposition. Dann gibt es eine Permutation $\sigma \in S_n$ mit $\tau = \sigma \circ \tau_0 \circ \sigma^{-1}$.

Beweis. Die Transposition τ vertausche k und l. Wähle σ mit $\sigma(1) = k$, $\sigma(2) = l$. Ist $i \neq k, l$, so ist $\sigma^{-1}(i) \neq 1, 2$. Also gilt:
$$\begin{aligned} i \neq k, l: &\quad (\sigma \circ \tau_0 \circ \sigma^{-1})(i) = \sigma \circ \tau_0(\sigma^{-1}(i)) = \sigma(\sigma^{-1}(i)) = i. \\ i = k: &\quad (\sigma \circ \tau_0 \circ \sigma^{-1})(k) = \sigma \circ \tau_0(1) = \sigma(2) = l. \\ i = l: &\quad (\sigma \circ \tau_0 \circ \sigma^{-1})(l) = \sigma \circ \tau_0(2) = \sigma(1) = k. \end{aligned}$$
Insgesamt folgt daher, daß
$$\sigma \circ \tau_0 \circ \sigma^{-1} = \tau.$$
□

Definition. Es sei $\sigma \in S_n$. Ein *Fehlstand* von σ ist ein Paar (i, j) mit $i < j$, $\sigma(i) > \sigma(j)$.

Beispiel.
$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$
Fehlstände:
$$\begin{aligned} (i, j) &= (1, 2), \quad (\sigma(i), \sigma(j)) = (2, 1) \\ (i, j) &= (3, 4), \quad (\sigma(i), \sigma(j)) = (4, 3). \end{aligned}$$

Definition. Das *Signum* einer Permutation $\sigma \in S_n$ ist definiert durch
$$\operatorname{sign} \sigma := \begin{cases} +1, & \text{falls \# Fehlbestände gerade} \\ -1, & \text{falls \# Fehlbestände ungerade.} \end{cases}$$

Beispiele.
(i)

$$\operatorname{sign}\begin{pmatrix}1 & 2 & 3 & 4\\ 2 & 1 & 4 & 3\end{pmatrix} = 1,$$

(ii)

$$\operatorname{sign}\tau_0 = \operatorname{sign}\begin{pmatrix}1 & 2 & 3 & \cdots & n\\ 2 & 1 & 3 & \cdots & n\end{pmatrix} = -1.$$

Lemma 7.9

$$\operatorname{sign}\sigma = \prod_{i<j}\frac{\sigma(j)-\sigma(i)}{j-i}.$$

Beweis. Ist m die Anzahl der Fehlstände, so gilt

$$\begin{aligned}\prod_{i<j}(\sigma(j)-\sigma(i)) &= (\prod_{\substack{i<j\\ \sigma(i)<\sigma(j)}}(\sigma(j)-\sigma(i)))(\prod_{\substack{i<j\\ \sigma(i)>\sigma(j)}}|\sigma(j)-\sigma(i)|)(-1)^m\\ &= (-1)^m\prod_{i<j}|\sigma(j)-\sigma(i)| = (-1)^m\prod_{i<j}(j-i).\end{aligned}$$

Hierbei folgt das letzte Gleichheitszeichen aus der Bijektivität von σ. Insgesamt folgt:

$$\prod_{i<j}\frac{\sigma(j)-\sigma(i)}{j-i} = (-1)^m = \left\{\begin{array}{rl} 1 & \text{falls } m \text{ gerade}\\ -1 & \text{falls } m \text{ ungerade}\end{array}\right\} = \operatorname{sign}\sigma.$$

□

Lemma 7.10
Für $\sigma, \tau \in S_n$ gilt: $\operatorname{sign}(\tau\circ\sigma) = \operatorname{sign}\tau\cdot\operatorname{sign}\sigma$.

Bemerkung. Obiges Lemma besagt gerade, daß die Abbildung

$$\operatorname{sign} : S_n \longrightarrow \{\pm 1\}$$

ein Gruppenhomomorphismus ist.

Korollar 7.11
$\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}\sigma$.

Beweis. $\operatorname{sign}(\sigma^{-1})\cdot\operatorname{sign}\sigma = \operatorname{sign}(\sigma^{-1}\sigma) = \operatorname{sign}(\mathrm{id}) = 1$. Also gilt

$$\operatorname{sign}(\sigma^{-1}) = \frac{1}{\operatorname{sign}\sigma} = \frac{1}{(\pm 1)} = \pm 1 = \operatorname{sign}\sigma.$$

□

Beweis von (7.10). Nach Lemma (7.9) gilt:

$$\begin{aligned}
\operatorname{sign}(\tau\circ\sigma) &= \prod_{i<j}\frac{\tau\circ\sigma(j)-\tau\circ\sigma(i)}{j-i}\\
&= \prod_{i<j}\frac{\tau\circ\sigma(j)-\tau\circ\sigma(i)}{\sigma(j)-\sigma(i)}\prod_{i<j}\frac{\sigma(j)-\sigma(i)}{j-i}\\
&= \prod_{i<j}\frac{\tau\circ\sigma(j)-\tau\circ\sigma(i)}{\sigma(j)-\sigma(i)}\cdot\operatorname{sign}\sigma.
\end{aligned}$$

Das Lemma folgt dann aus der

Behauptung.

$$\prod_{i<j}\frac{\tau\circ\sigma(j)-\tau\circ\sigma(i)}{\sigma(j)-\sigma(i)}=\operatorname{sign}\tau.$$

Die gilt, da

$$\begin{aligned}
\prod_{i<j}\frac{\tau\circ\sigma(j)-\tau\circ\sigma(i)}{\sigma(j)-\sigma(i)} &= \prod_{\substack{i<j\\ \sigma(i)<\sigma(j)}}\frac{\tau\circ\sigma(j)-\tau\circ\sigma(i)}{\sigma(j)-\sigma(i)}\prod_{\substack{i<j\\ \sigma(i)>\sigma(j)}}\frac{\tau\circ\sigma(j)-\tau\circ\sigma(i)}{\sigma(j)-\sigma(i)}\\
&= \prod_{\substack{i<j\\ \sigma(i)<\sigma(j)}}\frac{\tau\circ\sigma(j)-\tau\circ\sigma(i)}{\sigma(j)-\sigma(i)}\prod_{\substack{j<i\\ \sigma(j)>\sigma(i)}}\frac{\tau\circ\sigma(j)-\tau\circ\sigma(i)}{\sigma(j)-\sigma(i)}\\
&= \prod_{\substack{\sigma(i)<\sigma(j)\\ i<j}}\frac{\tau\circ\sigma(j)-\tau\circ\sigma(i)}{\sigma(j)-\sigma(i)}\prod_{\substack{\sigma(i)<\sigma(j)\\ i>j}}\frac{\tau\circ\sigma(j)-\tau\circ\sigma(i)}{\sigma(j)-\sigma(i)}\\
&= \prod_{\sigma(i)<\sigma(j)}\frac{\tau\circ\sigma(j)-\tau\circ\sigma(i)}{\sigma(j)-\sigma(i)}\\
&= \prod_{i<j}\frac{\tau(j)-\tau(i)}{j-i}=\operatorname{sign}\tau.
\end{aligned}$$

Hierbei folgt das zweite Gleichheitszeichen durch Umbenennen im zweiten Faktor und das letzte Gleichheitszeichen wieder aus der Bijektivität von σ. □

Korollar 7.12

(i) Für jede Permutation τ gilt $\operatorname{sign}\tau=-1$.

(ii) Ist $\sigma=\tau_1\circ\cdots\circ\tau_k$ wobei die τ_j Permutationen sind, so gilt $\operatorname{sign}\sigma=(-1)^k$.

Beweis. (i) Nach Lemma (7.8) gibt es eine Permutation $\sigma\in S_n$ mit

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

Also gilt

$$\begin{aligned}\operatorname{sign}\tau &= \operatorname{sign}\sigma\cdot\operatorname{sign}\tau_0\cdot\operatorname{sign}(\sigma^{-1})\\ &\overset{(7.11)}{=} (\operatorname{sign}\sigma)^2\cdot\operatorname{sign}\tau_0=\operatorname{sign}\tau_0=-1.\end{aligned}$$

(ii) Wegen Lemma (7.10) gilt

$$\operatorname{sign}\sigma=\operatorname{sign}\tau_1\cdots\operatorname{sign}\tau_k=(-1)^k.$$

□

Definition. Die alternierende Gruppe ist definiert durch

$$A_n:=\{\sigma\in S_n;\ \operatorname{sign}\sigma=1\}.$$

Lemma 7.13
Die alternierende Gruppe A_n ist eine Untergruppe von S_n.

Beweis. Es ist zu zeigen: Sind $\sigma,\tau\in A_n$, so ist auch $\sigma\circ\tau^{-1}\in A_n$. Dies gilt, da

$$\operatorname{sign}(\sigma\circ\tau^{-1})=\operatorname{sign}\sigma\operatorname{sign}(\tau^{-1})=\operatorname{sign}\sigma\operatorname{sign}\tau=1\cdot1=1.$$

(Alternativ kann man auch argumentieren, daß A_n der Kern des Homomorphismus sign : $S_n\to\{\pm1\}$ ist.) □

Lemma 7.14
Für die Anzahl der Elemente der alternierenden Gruppe gilt $\#A_n=\frac{1}{2}n!$ für $n\geq 2$.

Beweis. Wir setzen

$$A_n\tau_0:=\{\sigma\circ\tau_0;\ \sigma\in A_n\}.$$

Dann ist

$$A_n\longrightarrow A_n\tau_0,\ \sigma\longmapsto\sigma\circ\tau_0$$

eine Bijektion. Die Behauptung folgt dann aus:
(i) $S_n=A_n\cup A_n\tau_0$,

(ii) $A_n\cap A_n\tau_0=\emptyset$.

Zu (i): Sei $\sigma\in S_n$. Ist $\operatorname{sign}\sigma=1$, so ist $\sigma\in A_n$. Sei $\operatorname{sign}\sigma=-1$. Dann ist $\operatorname{sign}(\sigma\circ\tau_0)=1$. Es ist $\sigma=(\sigma\circ\tau_0)\circ\tau_0\in A_n\tau_0$.

Zu (ii): Ist $\sigma\in A_n\tau_0$, so ist $\operatorname{sign}\sigma=-1$. Also folgt (ii). □

Determinanten

Es sei

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathrm{Mat}(n; K) = K^{(n \times n)}$$

eine quadratische Matrix, gegeben durch die Zeilenvektoren a_i.

Definition. Eine *Determinante* (oder *Determinantenfunktion*) ist eine Abbildung

$$\det : \mathrm{Mat}(n; K) = K^{(n \times n)} \longrightarrow K$$

mit folgenden Eigenschaften:
(D1) det ist linear in jeder Zeile, d. h.

$$\det \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i' + \mu a_i'' \\ \vdots \\ a_n \end{pmatrix} = \lambda \det \begin{pmatrix} a_1 \\ \vdots \\ a_i' \\ \vdots \\ a_n \end{pmatrix} + \mu \det \begin{pmatrix} a_1 \\ \vdots \\ a_i'' \\ \vdots \\ a_n \end{pmatrix}.$$

(D2) det ist *alternierend*, d. h. gibt es $i \neq j$ mit $a_i = a_j$, so ist $\det A = 0$.

(D3) det ist *normiert*, d. h. $\det E = 1$.

Satz 7.15
Es gibt genau eine Determinante.

Bevor wir den Beweis erbringen, betrachten wir das folgende

Beispiel. $K = \mathbb{R}, n = 2$.

$$\begin{array}{rccl} \det : & \mathrm{Mat}(2; \mathbb{R}) & \longrightarrow & \mathbb{R} \\ & M = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} & \longmapsto & a_1 b_2 - a_2 b_1 =: \det M. \end{array}$$

Wir zeigen, daß die Eigenschaften (D1) – (D3) erfüllt sind:
(D1):

$$\det \begin{pmatrix} \lambda a \\ b \end{pmatrix} = \det \begin{pmatrix} \lambda a_1 & \lambda a_2 \\ b_1 & b_2 \end{pmatrix} = \lambda (a_1 b_2 - a_2 b_1) = \lambda \det \begin{pmatrix} a \\ b \end{pmatrix}.$$

$$\begin{aligned} \det \begin{pmatrix} a + a' \\ b \end{pmatrix} &= \det \begin{pmatrix} a_1 + a_1' & a_2 + a_2' \\ b_1 & b_2 \end{pmatrix} = (a_1 + a_1') b_2 - (a_2 + a_2') b_1 \\ &= (a_1 b_2 - a_2 b_1) + (a_1' b_2 - a_2' b_1) = \det \begin{pmatrix} a \\ b \end{pmatrix} + \det \begin{pmatrix} a' \\ b \end{pmatrix}. \end{aligned}$$

Analog zeigt man die Linearität in der zweiten Zeile.

(D2):

$$\det\begin{pmatrix} a \\ a \end{pmatrix} = \det\begin{pmatrix} a_1 & a_2 \\ a_1 & a_2 \end{pmatrix} = a_1a_2 - a_2a_1 = 0.$$

(D3):

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

Satz 7.16
Es sei $\det : \mathrm{Mat}(n; K) \to K$ eine Determinante. Dann gilt
(D4) $\det(\lambda A) = \lambda^n \det A$.

(D5) Ist $a_i = 0$ für ein i, so ist $\det A = 0$.

(D6)

$$\det\begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} = -\det\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix}.$$

(D7)

$$\det\begin{pmatrix} \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = \det\begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix}.$$

(D8) Ist σ eine Permutation, so gilt:

$$\det\begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} = \operatorname{sign}\sigma.$$

(D9)

$$\det\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \lambda_1 \cdots \lambda_n.$$

(D10) $\det A = 0 \Leftrightarrow a_1, \ldots, a_n$ sind linear abhängig.

(D11) $\det A \neq 0 \Leftrightarrow A \in \mathrm{GL}(n, K)$.

(D12) $\det(A \cdot B) = \det A \cdot \det B$.

Bemerkung. Die Formeln (D6) und (D7) beschreiben das Verhalten der Determinante bei elementaren Zeilenumformungen vom Typ (EZ4) und (EZ3).

Beweis. (D4): Mehrfaches Anwenden von (D1).

(D5): Aus (D1) mit $\lambda = \mu = 0$.

(D6): Es gilt:

$$0 \overset{\text{(D2)}}{=} \det \begin{pmatrix} \vdots \\ a_i + a_j \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} \overset{\text{(D1)}}{=} \underbrace{\det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \end{pmatrix}}_{=0 \text{ (D2)}} + \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} + \underbrace{\det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix}}_{=0 \text{ (D2)}}.$$

Also folgt:

$$\det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = -\det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix}.$$

(D7):

$$\det \begin{pmatrix} \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} \overset{\text{(D1)}}{=} \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \lambda \underbrace{\det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix}}_{=0 \text{ (D2)}}.$$

(D8): Ist τ eine Transposition, so gilt nach (D6)

$$\det \begin{pmatrix} e_{\tau(1)} \\ \vdots \\ e_{\tau(n)} \end{pmatrix} = -\det \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \overset{\text{(D3)}}{=} -1 = \operatorname{sign} \tau.$$

Sei nun $\sigma \in S_n$ beliebig. Nach Lemma (7.7) gibt es eine Darstellung

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

mit Transpositionen $\tau_1, \ldots, \tau_k$. Dann folgt mit (D6)

$$\det \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} = (-1)^k \det \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \overset{\text{(D3)}}{=} (-1)^k = \operatorname{sign} \sigma.$$

(D9): Es sei

$$A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

1. Fall: Es gibt ein i mit $\lambda_i = 0$. Wähle dieses maximal, d. h.

$$A = \begin{pmatrix} \lambda_1 & * & \cdots & & & \cdots & * \\ & \ddots & & & & & \\ & & \lambda_{i-1} & * & \cdots & \cdots & * \\ & & & 0 & * & \cdots & * \\ & & & & \lambda_{i+1} & \cdots & \\ & & & & & \ddots & \\ & & & & & & \lambda_n \end{pmatrix}, \quad \lambda_{i+1}, \dots, \lambda_n \neq 0.$$

Durch Zeilenumformungen vom Typ (EZ3) kann man dann erreichen:

$$A \curvearrowright B = \begin{pmatrix} \lambda_1 & * & \cdots & & & \cdots & * \\ & \ddots & & & & & \\ & & \lambda_{i-1} & * & \cdots & \cdots & * \\ 0 & \cdots & \cdots & 0 & \cdots & \cdots & 0 \\ & & & & \lambda_{i+1} & \cdots & \\ & & & & & \ddots & \\ & & & & & & \lambda_n \end{pmatrix}.$$

Die Eigenschaft (D7) besagt gerade, daß elementare Zeilenoperationen vom Typ (EZ3) die Determinante nicht ändern. Also folgt:

$$\det A = \det B \overset{\text{(D5)}}{=} 0 = \lambda_1 \cdots \lambda_i \cdots \lambda_n.$$

2. Fall: Alle $\lambda_i \neq 0$. Dann gilt wegen (D1)

$$\det A = (\lambda_1 \cdots \lambda_n) \cdot \underbrace{\det \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix}}_{=:B}.$$

Man erhält:

$$B \overset{\text{(EZ3)}}{\curvearrowright} E = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

Also

$$\det A = \lambda_1 \cdots \lambda_n \cdot \det B \overset{\text{(D7)}}{=} \lambda_1 \cdots \lambda_n \underbrace{\det E}_{=1 \text{ (D3)}}.$$

(D10): Durch Umformungen vom Typ (EZ3) und (EZ4) kann man erreichen (vgl. den Gauß-Algorithmus), daß

$$A \curvearrowright \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} =: B.$$

Wegen (D6) und (D7) ändert dies die Determinante höchstens um das Vorzeichen, also

$$\det A = \pm \det B \overset{(D9)}{=} \pm\lambda_1 \cdots \lambda_n.$$

Behauptung. Rang A = Rang $B < n \Leftrightarrow$ ein $\lambda_i = 0$.

„$\Rightarrow$": Es seien alle $\lambda_i \neq 0$. Dann erhalten wir jeweils durch Zeilenumformungen

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \curvearrowright \begin{pmatrix} 1 & & * \\ & \ddots & \\ & & 1 \end{pmatrix} \curvearrowright \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

Also ist Rang A = Rang $B = n$.

„$\Leftarrow$": Es sei $\lambda_i = 0$. Dann gilt für die Zeilenvektoren von B:

$$\begin{aligned} b_i &= (\underbrace{0\ldots 0}_{i}, b_i') \\ \vdots & \qquad \vdots \\ b_n &= (\underbrace{0\ldots 0}_{i}, b_n') \end{aligned}$$

mit $b_i', \ldots, b_n' \in K^{n-i}$. Nach dem Fundamentallemma sind $b_i', \ldots, b_n'$ linear abhängig, also auch $b_i, \ldots, b_n$. Damit gilt, daß Rang $B < n$.

(D11):

$$\begin{aligned} A \in \mathrm{GL}(n,K) &\overset{(5.15)}{\Leftrightarrow} \text{ Rang } A = n \overset{(5.15)}{\Leftrightarrow} (a_1, \ldots, a_n) \text{ sind linear unabhängig} \\ &\overset{(D10)}{\Leftrightarrow} \det A \neq 0. \end{aligned}$$

(D12): *1. Fall:* Rang $A < n$. Dann ist $\det A = 0$ nach (D11). Nach Lemma (5.17) ist Rang$(AB) \leq$ Rang $A < n$. Also ist nach (D11) $\det(AB) = 0$, und es gilt

$$0 = \det(AB) = \det A \cdot \det B.$$

2. Fall: Rang $A = n$. Dann ist $A \in \mathrm{GL}(n,K)$. Nach Satz (5.22) gibt es Elementarmatrizen $C_1, \ldots, C_s$ mit

$$A = C_1 \cdots C_s.$$

Also folgt

$$\det(AB) = \det(C_1 \cdots C_s B).$$

Es genügt daher zu zeigen:

Behauptung. Ist B eine beliebige Matrix, und ist C_i eine Elementarmatrix, so gilt $\det(C_i B) = \det C_i \det B$.

1. Fall: $C_i = F_k(\alpha) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$. Dann ist $C_iB = \begin{pmatrix} b_1 \\ \vdots \\ \alpha b_k \\ \vdots \\ b_n \end{pmatrix}$.

Damit folgt:

$$\det(C_iB) \overset{\text{(D1)}}{=} \alpha \cdot \det B \overset{\text{(D9)}}{=} \det(C_i) \cdot \det B.$$

2. Fall:

$$C_i = F_{kl} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & \cdots & 1 & \\ & & & \ddots & \vdots & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

Aus (D9) folgt

$$\det C_i = \det F_{kl} = 1.$$

Andererseits ist

$$C_iB = \begin{pmatrix} b_1 \\ \vdots \\ b_k + b_l \\ \vdots \\ b_l \\ \vdots \\ b_n \end{pmatrix}.$$

Also erhalten wir

$$\det(C_iB) \overset{\text{(D7)}}{=} \det B = 1 \cdot \det B = \det C_i \cdot \det B.$$

Satz 7.15'
Es gibt genau eine Determinante

$$\det : \ \mathrm{Mat}(n; K) \longrightarrow K$$

und zwar gilt für $A = (\alpha_{ij})$:

$$(*) \qquad \det A = \sum_{\sigma \in S_n} \operatorname{sign} \sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)}.$$

Beweis. *Eindeutigkeit:* Falls es eine Determinantenfunktion gibt, muß $(*)$ gelten. Dazu sei

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \text{ mit } a_i = \alpha_{i1}e_1 + \cdots + \alpha_{in}e_n.$$

Also gilt

$$\begin{aligned}
\det A &\overset{(D1)}{=} \sum_{i_1=1}^{n} \alpha_{1i_1} \det \begin{pmatrix} e_{i_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \\
&\overset{(D1)}{=} \sum_{i_1=1}^{n} \alpha_{1i_1} \sum_{i_2=1}^{n} \alpha_{2i_2} \det \begin{pmatrix} e_{i_1} \\ e_{i_2} \\ a_3 \\ \vdots \\ a_n \end{pmatrix} \\
&= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} \alpha_{1i_1}\alpha_{2i_2} \cdots \alpha_{ni_n} \det \begin{pmatrix} e_{i_1} \\ \vdots \\ e_{i_n} \end{pmatrix}.
\end{aligned}$$

Nun ist

$$\det \begin{pmatrix} e_{i_1} \\ \vdots \\ e_{i_n} \end{pmatrix} = \begin{cases} 0, & \text{falls } i_k = i_l \text{ für ein } k \neq l \text{ (D2)} \\ \operatorname{sign}\sigma, & \begin{array}{l}\text{falls es eine Permutation } \sigma \in S_n \text{ gibt} \\ \text{mit } i_1 = \sigma(1), \dots, i_n = \sigma(n) \text{ (D8).}\end{array} \end{cases}$$

Also folgt

$$\det A = \sum_{\sigma \in S_n} \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)} \operatorname{sign}\sigma.$$

Existenz: Wir müssen zeigen, daß $(*)$ die Eigenschaften (D1) – (D3) erfüllt.

(D1): Es gilt

$$\begin{aligned}
\det\begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i' + \mu a_i'' \\ \vdots \\ a_n \end{pmatrix} &\overset{(*)}{=} \sum_{\sigma\in S_n} \operatorname{sign}\sigma\,\alpha_{1\sigma(1)}\cdots(\lambda\alpha'_{i\sigma(i)}+\mu\alpha''_{i\sigma(i)})\cdots\alpha_{n\sigma(n)} \\
&= \lambda\sum_{\sigma\in S_n}\operatorname{sign}\sigma\,\alpha_{1\sigma(1)}\cdots\alpha'_{i\sigma(i)}\cdots\alpha_{n\sigma(n)} + \mu\sum_{\sigma\in S_n}\alpha_{1\sigma(1)}\cdots\alpha''_{i\sigma(i)}\cdots\alpha_{n\sigma(n)} \\
&\overset{(*)}{=} \lambda\det\begin{pmatrix} a_1 \\ \vdots \\ a_i' \\ \vdots \\ a_n \end{pmatrix} + \mu\det\begin{pmatrix} a_1 \\ \vdots \\ a_i'' \\ \vdots \\ a_n \end{pmatrix}.
\end{aligned}$$

(D2): Es sei $k \neq l$ mit $a_k = a_l$. Ferner sei $\tau \in S_n$ diejenige Transposition, die k und l vertauscht. Dann gilt

$$S_n = A_n \cup A_n\tau.$$

Nun ist

$$\begin{aligned}
\alpha_{1\sigma\circ\tau(1)}\cdots\alpha_{k\sigma\circ\tau(k)}\cdots\alpha_{l\sigma\circ\tau(l)}\cdots\alpha_{n\sigma\circ\tau(n)} &= \\
\alpha_{1\sigma(1)}\cdots\alpha_{k\sigma(l)}\cdots\alpha_{l\sigma(k)}\cdots\alpha_{n\sigma(n)} &\overset{a_l=a_k}{=} \\
\alpha_{1\sigma(1)}\cdots\alpha_{l\sigma(l)}\cdots\alpha_{,\sigma(k)}\cdots\alpha_{n\sigma(n)} &= \alpha_{1\sigma(1)}\cdots\alpha_{n\sigma(n)}.
\end{aligned}$$

Damit folgt, da $\operatorname{sign}(\sigma\circ\tau) = -\operatorname{sign}\sigma$ gilt, daß

$$\det A = 0.$$

(D3): Wir betrachten die Einheitsmatrix $E = (\delta_{ij})$. Dann gilt:

$$\det E \overset{(*)}{=} \sum_{\sigma\in S_n}\operatorname{sign}\sigma\,\delta_{1\sigma(1)}\cdots\delta_{n\sigma(n)} = \operatorname{sign}(\mathrm{id})\delta_{11}\cdots\delta_{nn} = 1.$$

□

Berechnung von Determinanten

$n = 1$: $A = (\alpha)$, $\det A = \alpha$.

$n = 2$: $A = \begin{pmatrix}\alpha_{11} & \alpha_{12}\\ \alpha_{21} & \alpha_{22}\end{pmatrix}$, $\det A = \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}$.

$n = 3$: $\quad A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}$. Es ist $\# S_3 = 3! = 6$. Aus $(*)$ ergibt sich

$$\begin{aligned} \det A &= \alpha_{11}\alpha_{22}\alpha_{33} - \alpha_{11}\alpha_{23}\alpha_{32} - \alpha_{12}\alpha_{21}\alpha_{33} \\ &\quad + \alpha_{12}\alpha_{23}\alpha_{31} + \alpha_{13}\alpha_{21}\alpha_{32} - \alpha_{13}\alpha_{22}\alpha_{31}. \end{aligned}$$

Die Determinante von (3×3)-Matrizen kann man auch mit folgender Regel berechnen:

Regel von Sarrus

$$\left(\begin{array}{ccc|cc} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{31} & \alpha_{32} \end{array} \right) = \begin{array}{l} \alpha_{11}\alpha_{22}\alpha_{33} + \alpha_{12}\alpha_{23}\alpha_{31} + \alpha_{13}\alpha_{21}\alpha_{32} \\ -\alpha_{31}\alpha_{22}\alpha_{13} - \alpha_{32}\alpha_{23}\alpha_{11} - \alpha_{33}\alpha_{21}\alpha_{12}. \end{array}$$

Achtung: Diese Regel ist *nicht* auf $n \geq 4$ verallgemeinerbar.

Satz 7.17
Es gilt

$$\det A = \det {}^t\!A$$

Beweis. ${}^t\!A = ({}^t\alpha_{ij})$ mit ${}^t\alpha_{ij} = \alpha_{ji}$. Also gilt

$$\begin{aligned} \det {}^t\!A &= \sum_{\sigma \in S_n} \operatorname{sign} \sigma \, {}^t\alpha_{1\sigma(1)} \cdots {}^t\alpha_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sign} \sigma \alpha_{\sigma(1)1} \cdots \alpha_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} \operatorname{sign} \sigma \, \alpha_{1\sigma^{-1}(1)} \cdots \alpha_{n\sigma^{-1}(n)} \\ &\overset{(7.11)}{=} \sum_{\sigma \in S_n} \operatorname{sign} \sigma^{-1} \, \alpha_{1\sigma^{-1}(1)} \cdots \alpha_{n\sigma^{-1}(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sign} \sigma \, \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)} \\ &= \det A. \end{aligned}$$

□

Bemerkung. Dies bedeutet, daß die Aussagen, die wir über die Zeilenumformungen vom Typ (EZ3) und (EZ4) gemacht haben, analog auch für die Spaltenumformungen vom Typ (ES3) und (ES4) gelten.

Satz 7.18
Es sei

$$A = \left(\begin{array}{c|c} A_1 & C \\ \hline 0 & A_2 \end{array} \right), \quad (A_1, A_2 \text{ quadratisch.})$$

Dann gilt

$$\det A = \det A_1 \cdot \det A_2.$$

Beweis. Durch elementare Zeilenumformungen vom Typ (EZ3) und (EZ4) erreicht man

(1) $A_1 \curvearrowright B_1 = \begin{pmatrix} \beta_1 & & * \\ & \ddots & \\ 0 & & \beta_{n_1} \end{pmatrix}$

Hierbei sei k_1 die Anzahl der Zeilenvertauschungen.

(2) $A_2 \curvearrowright B_2 = \begin{pmatrix} \beta'_1 & & * \\ & \ddots & \\ 0 & & \beta'_{n_2} \end{pmatrix}$ und es sei k_2 die Anzahl der hier durchgeführten Zeilenvertauschungen. Dann gilt nach (D6) und (D7):

$$\det A_i = (-1)^{k_i} \det B_i, \quad i = 1, 2.$$

Dieselben Operationen bewirken

$$A \curvearrowright \left(\begin{array}{c|c} B_1 & D \\ \hline 0 & B_2 \end{array}\right) = \left(\begin{array}{cccc|cccc} \beta_1 & * & \cdots & * & & & & \\ & \ddots & \ddots & \vdots & & * & & \\ & & \ddots & * & & & & \\ & & & \beta_{n_1} & & & & \\ \hline & & & & \beta'_1 & * & \cdots & * \\ & & & & & \ddots & \ddots & \vdots \\ & & & & & & \ddots & * \\ & & & & & & & \beta'_{n_2} \end{array}\right) =: B.$$

Hierbei wurden $k_1 + k_2$ Zeilenvertauschungen vorgenommen. Damit folgt

$$\begin{aligned} \det A &= (-1)^{k_1+k_2} \det B \overset{\text{(D9)}}{=} (-1)^{k_1+k_2} (\beta_1 \cdots \beta_{n_1})(\beta'_1 \cdots \beta'_{n_2}) \\ &\overset{\text{(D9)}}{=} (-1)^{k_1} \det B_1 (-1)^{k_2} \det B_2 \\ &= \det A_1 \cdot \det A_2. \end{aligned}$$

□

Entwicklung von Determinanten

Es sei $A = (\alpha_{ij})$ eine quadratische $(n \times n)$-Matrix.

Definition.

(i)

$$A_{ij} := \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1,j-1} & 0 & \alpha_{1,j+1} & \cdots & \alpha_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{i-1,1} & \cdots & \alpha_{i-1,j-1} & 0 & \alpha_{i-1,j+1} & \cdots & \alpha_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \alpha_{i+1,1} & \cdots & \alpha_{i+1,j-1} & 0 & \alpha_{i+1,j+1} & \cdots & \alpha_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{n,j-1} & 0 & \alpha_{n,j+1} & \cdots & \alpha_{nn} \end{pmatrix} \leftarrow i\text{-te Zeile.}$$

$$\uparrow j\text{-te Spalte.}$$

(ii) Die Matrix

$$\tilde{A} := (\tilde{\alpha}_{ij})$$

mit

$$\tilde{\alpha}_{ij} := \det A_{ji}$$

heißt die zu A *komplementäre Matrix.*

(iii) Die Matrix

$$A'_{ij} := \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1j} & \cdots & \alpha_{1n} \\ \vdots & & \vdots & & \vdots \\ \alpha_{i1} & \cdots & \alpha_{ij} & \cdots & \alpha_{in} \\ \vdots & & \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nj} & \cdots & \alpha_{nn} \end{pmatrix} \leftarrow i \quad \in \mathrm{Mat}(n-1; K)$$

$$\uparrow j\text{-te Spalte.}$$

heißt *Streichungsmatrix* (zur Stelle (i, j)).

(iv) Ist $A = (a_1, \ldots, a_n)$ setzt man schließlich

$$A^{ij} := (a_1, \ldots, a_{j-1}, e_i, a_{j+1}, \ldots, a_n).$$

Lemma 7.19

(i) $\det A_{ij} = (-1)^{i+j} \det A'_{ij}$.

(ii) $\det A^{ij} = \det A_{ij}$.

Beweis. (i) Durch $(i-1)$ Zeilenvertauschungen und $(j-1)$ Spaltenvertauschungen erreicht man

$$A_{ij} \curvearrowright \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A'_{ij} & \\ 0 & & & \end{pmatrix} =: A''_{ij}.$$

Dann ist unter Verwendung von (D6) und Satz (7.17)

$$\begin{aligned} \det A_{ij} &= (-1)^{i-1}(-1)^{j-1} \det A''_{ij} \overset{(7.18)}{=} (-1)^{i+j-2} \det A'_{ij} \\ &= (-1)^{i+j} \det A'_{ij}. \end{aligned}$$

(ii) Mit Spaltenumformungen vom Typ (ES3) erreicht man

$$A^{ij} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1,j-1} & 0 & \alpha_{1,j+1} & \cdots & \alpha_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{i1} & \cdots & \alpha_{i,j-1} & 1 & \alpha_{i,j+1} & \cdots & \alpha_{in} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{n,j-1} & 0 & \alpha_{n,j+1} & \cdots & \alpha_{nn} \end{pmatrix} \quad \curvearrowright A_{ij}.$$

Also folgt mit (D7) und der Bemerkung nach Satz (7.17):

$$\det A^{ij} = \det A_{ij}.$$

Satz 7.20

$$A\tilde{A} = \tilde{A}A = (\det A)E.$$

Beweis.

$$\begin{aligned} (\tilde{A}A)_{ik} &= \sum_{j=1}^{n} \tilde{\alpha}_{ij}\alpha_{jk} = \sum_{j=1}^{n} \det A_{ji}\alpha_{jk} \overset{(7.19)(\mathrm{ii})}{=} \sum_{j=1}^{n} \det A^{ji}\alpha_{jk} \\ &= \sum_{j=1}^{n} \alpha_{jk} \det(a_1, \dots, a_{i-1}, e_j, a_{i+1}, \dots, a_n) \\ &\overset{(\mathrm{D1})}{=} \det(a_1, \dots, a_{i-1}, \sum_{j=1}^{n} \alpha_{jk}e_j, a_{i+1}, \dots, a_n) \\ &= \det(a_1, \dots, a_{i-1}, a_k, a_{i+1}, \dots, a_n) \\ &= \delta_{ik} \det A. \end{aligned}$$

Also gilt

$$\tilde{A}A = (\det A)E.$$

Analog zeigt man

$$A\tilde{A} = (\det A)E.$$

□

Korollar 7.21
Ist A invertierbar ($\det A \neq 0$), so gilt

$$A^{-1} = \frac{1}{\det A}\tilde{A}.$$

Beispiel. $(n = 2)$:

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$$

$$\begin{aligned} A_{11} &= \begin{pmatrix} 1 & 0 \\ 0 & \alpha_{22} \end{pmatrix}, \quad A_{12} = \begin{pmatrix} 0 & 1 \\ \alpha_{21} & 0 \end{pmatrix} \\ A_{21} &= \begin{pmatrix} 0 & \alpha_{12} \\ 1 & 0 \end{pmatrix}, \quad A_{22} = \begin{pmatrix} \alpha_{11} & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Damit ist

$$\tilde{A} = \begin{pmatrix} \det A_{11} & \det A_{21} \\ \det A_{12} & \det A_{22} \end{pmatrix} = \begin{pmatrix} \alpha_{22} & -\alpha_{12} \\ -\alpha_{21} & \alpha_{11} \end{pmatrix}.$$

Also folgt, falls $\det A \neq 0$ ist:

$$A^{-1} = \frac{1}{\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}} \begin{pmatrix} \alpha_{22} & -\alpha_{12} \\ -\alpha_{21} & \alpha_{11} \end{pmatrix}.$$

Satz 7.22 (Entwicklungssatz von Laplace)
Es sei $A \in \mathrm{Mat}(n; K)$. Dann gilt für festes i, bzw. j:
(i) $\det A = \sum_{j=1}^{n} (-1)^{i+j} \alpha_{ij} \det A'_{ij}$.
(ii) $\det A = \sum_{i=1}^{n} (-1)^{i+j} \alpha_{ij} \det A'_{ij}$.
(Man sagt, daß man die Determinante nach der i-ten Zeile, bzw. der j-ten Spalte entwickelt.)

Beispiel.

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

Entwickeln nach der zweiten Zeile:

$$\begin{aligned} \det A &= (-1)^{2+1}\, 1 \det \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} + (-1)^{2+2}\, 0 \det \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} + (-1)^{2+3}\, 2 \det \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \\ &= -(4-3) + 0 - 2 \cdot (1) = -3. \end{aligned}$$

Merkregel: Das Vorzeichen kann man sich nach dem „Schachbrettmuster“ merken:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \end{pmatrix}$$

Beweis von (7.22)
(ii): Es gilt

$$\begin{aligned}\det A &\overset{(7.20)}{=} (\tilde{A}A)_{jj} = \sum_{i=1}^{n} \tilde{\alpha}_{ji}\alpha_{ij} = \sum_{i=1}^{n} \alpha_{ij} \det A_{ij} \\ &\overset{(7.19)(\mathrm{i})}{=} \sum_{i=1}^{n} (-1)^{i+j} \alpha_{ij} \det A'_{ij}.\end{aligned}$$

(i): Zeigt man analog. □

Cramersche Regel

Wir betrachten ein *quadratisches* Gleichungssystem

$$(*) \qquad Ax = b \qquad (A = (a_1, \dots, a_n)).$$

Ist $A \in \mathrm{GL}(n, K)$, so hat $(*)$ genau eine Lösung, nämlich

$$x = A^{-1}b.$$

Satz 7.23 (Cramersche Regel)
Ist A invertierbar, so gilt für die eindeutig bestimmte Lösung $x = (x_1, \dots, x_n)$ von $(*)$:

$$x_i = \frac{\det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n)}{\det A}.$$

Beweis. Es gilt

$$x = A^{-1}b.$$

D. h.

$$(1) \qquad x_i = \sum_{j=1}^{n} (A^{-1})_{ij} b_j.$$

Nun ist nach (7.21)

$$(A^{-1})_{ij} = \frac{\det A_{ji}}{\det A} \overset{(7.19)(\mathrm{ii})}{=} \frac{\det A^{ji}}{\det A} = \frac{\det(a_1, \dots, a_{i-1}, e_j, a_{i+1}, \dots, a_n)}{\det A}.$$

Wegen (1) folgt also:

$$\begin{aligned}x_i &= \frac{1}{\det A}\Big(\sum_{j=1}^{n} b_j \det(a_1, \dots, a_{i-1}, e_j, a_{i+1}, \dots, a_n)\Big) \\ &\overset{(\mathrm{D1})}{=} \frac{1}{\det A} \det(a_1, \dots, a_{i-1}, \sum_{j=1}^{n} b_j e_j, a_{i+1}, \dots, a_n) \\ &= \frac{1}{\det A} \det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n).\end{aligned}$$

$\square$

Beispiel.

$$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - 2x_2 &= -1. \end{aligned}$$

$$\begin{aligned} A &= \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}, \qquad \det A = -3. \\ x_1 &= -\frac{1}{3} \det \begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix} = -\frac{1}{3}(-4+1) = 1, \\ x_2 &= -\frac{1}{3} \det \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = -\frac{1}{3}(-1-2) = 1. \end{aligned}$$

Bemerkung. Die Lösung hängt algebraisch (rational) von A, b ab.

§ 8 Eigenwerttheorie I

Wir hatten bereits früher den Polynomring in einer Variablen über einem Körper K betrachtet:

$$K[x] = \mathrm{Abb}[\mathbb{N}, K] = \{P;\ P = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0;\ a_\nu \in K\}.$$

Definition. Eine Funktion $f : K \to K$ heißt eine *Polynomfunktion*, falls es ein Polynom $P \in K[x]$ gibt mit $f(a) = P(a)$ für alle $a \in K$.

Eigenwerte und Eigenvektoren

Es sei $A \in \mathrm{Mat}(n; K)$ eine quadratische Matrix.

Definition. Ein Skalar $\lambda \in K$ heißt *Eigenwert* von A, falls es einen Vektor $0 \neq v \in K^n$ gibt mit

$$Av = \lambda v.$$

v heißt dann auch *Eigenvektor* von A zum Eigenwert λ.

Die geometrische Deutung ist, daß für $\lambda \neq 0$ die Gerade Kv unter der Abbildung $h_A : K^n \to K^n$, $x \mapsto Ax$ fest bleibt.

Bemerkung. Es gilt

$$\begin{aligned} 0 \text{ ist Eigenwert von } A \quad &\Leftrightarrow \quad \text{Es gibt } 0 \neq v \in K^n \text{ mit } Av = 0 \cdot v = 0 \\ &\Leftrightarrow \quad \mathrm{Ker}\, A \neq \{0\} \Leftrightarrow A \notin \mathrm{GL}(n, K). \end{aligned}$$

Definition. Zwei Matrizen $A, B \in \mathrm{Mat}(n; K)$ heißen *ähnlich*, falls es ein Element $W \in \mathrm{GL}(n, K)$ gibt mit

$$B = W^{-1}AW.$$

Schreibweise. $A \sim B$.

Bemerkung.

(i) Ähnliche Matrizen sind insbesondere äquivalent.

(ii) $\sim$ ist Äquivalenzrelation, d. h.

$A \sim A$

$A \sim B \Leftrightarrow B \sim A$

$A \sim B, B \sim C \Rightarrow A \sim C.$

Die Begründung erfolgt genau wie bei äquivalenten Matrizen.

Wir werden später sehen, daß ähnliche Matrizen denselben Endomorphismus bezüglich verschiedener Basen beschreiben.

Definition. $A \in \mathrm{Mat}(n;K)$ heißt *diagonalisierbar*, falls A zu einer Diagonalmatrix ähnlich ist.

Satz 8.1
A ist genau dann diagonalisierbar, wenn es eine Basis $v_1, \dots, v_n$ von K^n gibt, die aus Eigenvektoren von A besteht.

Beweis. „$\Leftarrow$" Da $v_1, \dots, v_n$ eine Basis ist, ist

$$W := (v_1, \dots, v_n) \in \mathrm{GL}(n, K).$$

Dann gilt mit $Av_i = \lambda_i v_i$:

$$\begin{aligned} AW &= A(v_1, \dots, v_n) = (Av_1, \dots, Av_n) = (\lambda_1 v_1, \dots, \lambda_n v_n) \\ &= (v_1, \dots, v_n) \underbrace{\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}}_{=:D} = WD \end{aligned}$$

und damit

$$W^{-1}AW = D.$$

„$\Rightarrow$" Es sei $D := \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$ und $W \in \mathrm{GL}(n, K)$ mit

$$W^{-1}AW = D.$$

Dann ist

$$AW = WD.$$

Mit

$$W = (v_1, \dots, v_n)$$

gilt

$$\begin{aligned} A(v_1, \dots, v_n) &= (v_1, \dots, v_n) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \\ \Rightarrow\ & (Av_1, \dots, Av_n) = (\lambda_1 v_1, \dots, \lambda_n v_n) \\ \Rightarrow\ & Av_i = \lambda_i v_i\ ; \quad i = 1, \dots, n. \end{aligned}$$

Da W invertierbar ist, sind $v_1, \ldots, v_n$ eine Basis von K^n. $\square$

Das charakteristische Polynom

Es sei $A \in \mathrm{Mat}(n; K)$.

Definition. Die Abbildung

$$\chi_A : K \longrightarrow K; \quad \chi_A(x) := \det(xE - A)$$

heißt *charakteristische Polynomfunktion* von A.

Bemerkung. Es ist

$$\chi_A(x) = \det \begin{pmatrix} x - \alpha_{11} & -\alpha_{12} & \cdots & -\alpha_{1n} \\ -\alpha_{21} & x - \alpha_{22} & \cdots & -\alpha_{2n} \\ \vdots & & & \vdots \\ -\alpha_{n1} & \cdots & \cdots & x - \alpha_{nn} \end{pmatrix}.$$

Entwickelt man diese Determinante entsprechend der Formel

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign} \sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)},$$

so erhält man ein Polynom von Grad n. Man nennt dies das *charakteristische Polynom* von A.

Definition. Für $A = (\alpha_{ij}) \in \mathrm{Mat}(n; K)$ heißt

$$\mathrm{Spur}\, A := \sum_{i=1}^{n} \alpha_{ii}$$

die *Spur* von A.

Lemma 8.2
Es gilt

$$\chi_A(x) = x^n - \mathrm{Spur}\, A x^{n-1} + \cdots + (-1)^n \det A.$$

Beweis. Wir schreiben

$$\chi_A(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

Nun ist

$$a_0 = \chi_A(0) = \det(-A) = (-1)^n \det A.$$

Um die anderen Koeffizienten zu bestimmen, geht man so vor:

$$xE - A = \begin{pmatrix} x-\alpha_{11} & -\alpha_{12} & \cdots & -\alpha_{1n} \\ -\alpha_{21} & x-\alpha_{22} & \cdots & -\alpha_{2n} \\ \vdots & & & \vdots \\ -\alpha_{n1} & \cdots & \cdots & x-\alpha_{nn} \end{pmatrix} = \gamma_{ij}(x)$$

wobei

$$\begin{array}{rcl} \gamma_{ii}(x) & = & x - \alpha_{ii}, \\ \gamma_{ij}(x) & = & -\alpha_{ij} \quad (i \neq j). \end{array}$$

Nun ist

$$\chi_A(x) = \det(\gamma_{ij}(x)) = \sum_{\sigma \in S_n} \operatorname{sign} \sigma \gamma_{1\sigma(1)}(x) \cdots \gamma_{n\sigma(n)}(x).$$

Ist $\sigma \neq \mathrm{id}$, so ist $\sigma(i) \neq i$ für mindestens zwei i. D. h. der zu σ gehörige Summand enthält x höchstens zur Potenz $n-2$. Also erhält man a_n, a_{n-1} aus

$$(x-\alpha_{11})(x-\alpha_{22})\cdots(x-\alpha_{nn}) = x^n - x^{n-1}(\alpha_{11} + \cdots + \alpha_{nn}) + \text{Terme niedrigerer Ordnung.}$$

Damit folgt

$$a_n = 1, \quad a_{n-1} = -\operatorname{Spur} A.$$

□

Lemma 8.3
Sind A, B ähnliche Matrizen ($A \sim B$), so gilt $\chi_A = \chi_B$.

Beweis. Da $A \sim B$ gibt es ein $W \in \mathrm{GL}(n, K)$ mit

$$B = W^{-1}AW.$$

Also gilt

$$\begin{array}{rcl} xE - B & = & xE - W^{-1}AW = xW^{-1}EW - W^{-1}AW \\ & = & W^{-1}(xE - A)W. \end{array}$$

Nach dem Determinantenmultiplikationssatz folgt

$$\det(xE - B) = \det W^{-1} \det(xE - A) \det W = \det(xE - A)$$

und damit

$$\chi_B(x) = \chi_A(x).$$

□

Korollar 8.4
Sind A und B ähnlich, $(A \sim B)$, so gilt $\operatorname{Spur} A = \operatorname{Spur} B$.

Bemerkung. Die Umkehrung von Lemma 8.3 gilt im allgemeinen nicht, wie das folgende Beispiel zeigt:

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \not\sim B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

es gilt aber

$$\chi_A(x) = \det \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} = x^2, \qquad \chi_B(x) = \det \begin{pmatrix} x & -1 \\ 0 & x \end{pmatrix} = x^2.$$

Satz 8.5
Für $A \in \operatorname{Mat}(n; K)$ sind äquivalent:

(i) λ ist Eigenwert von A.

(ii) $\chi_A(\lambda) = 0$.

(iii) $A \sim \begin{pmatrix} \lambda & * \cdots * \\ 0 & \\ \vdots & B \\ 0 & \end{pmatrix}$ mit $B \in \operatorname{Mat}(n-1; K)$. Ferner gilt

$$\chi_A(x) = (x - \lambda)\chi_B(x).$$

Beweis. (i)$\Leftrightarrow$(ii): (i) ist äquivalent dazu, daß es ein $0 \neq v \in K^n$ gibt mit $Av = \lambda v$.

Nun ist

$$Av = \lambda v \Leftrightarrow Av = \lambda E v \Leftrightarrow (\lambda E - A)v = 0.$$

Also gilt

$$\begin{aligned} \text{(i)} \quad &\Leftrightarrow \quad \operatorname{Ker}(\lambda E - A) \neq \{0\} \Leftrightarrow \det(\lambda E - A) = 0 \\ &\Leftrightarrow \quad \chi_A(\lambda) = 0 \Leftrightarrow \text{(ii)}. \end{aligned}$$

(iii)$\Rightarrow$(ii): Wegen $A \sim \begin{pmatrix} \lambda & * \\ 0 & B \end{pmatrix} =: A'$ folgt aus Lemma (8.3)

$$\begin{aligned} \chi_A(x) &= \chi_{A'}(x) = \det(xE - A') = \\ &= \det \begin{pmatrix} x - \lambda & * \\ 0 & xE - B \end{pmatrix} = (x - \lambda) \det(xE - B) = (x - \lambda)\chi_B(x), \end{aligned}$$

und damit $\chi_A(\lambda) = 0$.

(i)$\Rightarrow$(iii): Es sei $0 \neq v_1 \in K^n$ mit

$$Av_1 = \lambda v_1.$$

Man kann v_1 zu einer Basis $v_1, \dots, v_n$ von K^n ergänzen. Dann ist

$$W := (v_1, \dots, v_n) \in \mathrm{GL}(n, K).$$

Damit gilt

$$AW = A(v_1, \dots, v_n) = (Av_1, \dots, Av_n) = (\lambda v_1, Av_2, \dots, Av_n).$$

Andererseits ist

$$(e_1, \dots, e_n) = E = W^{-1}W = W^{-1}(v_1, \dots, v_n) = (W^{-1}v_1, \dots, W^{-1}v_n).$$

Das heißt

$$W^{-1}(v_1) = e_1.$$

Damit ergibt sich

$$\begin{aligned} W^{-1}AW &= W^{-1}(\lambda v_1, Av_2, \dots, Av_n) \\ &= (W^{-1}\lambda v_1, W^{-1}Av_2, \dots, W^{-1}Av_n) \\ &= (\lambda e_1, W^{-1}Av_2, \dots, W^{-1}Av_n) = \begin{pmatrix} \lambda & * \cdots * \\ 0 & \\ \vdots & B \\ 0 & \end{pmatrix}. \end{aligned}$$

□

Eigenräume

Bemerkung. Die Eigenwerte von A sind also genau die Nullstellen des charakteristischen Polynoms von A.

Beispiel. Es sei $K = \mathbb{R}$. Die Matrix

$$A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \in \mathrm{Mat}\ (2, \mathbb{R})$$

beschreibt eine Drehung um den Winkel φ. Das charakteristische Polynom ist

$$\begin{aligned} \chi_A(x) &= \det \begin{pmatrix} x - \cos\varphi & \sin\varphi \\ -\sin\varphi & x - \cos\varphi \end{pmatrix} \\ &= x^2 - 2\cos\varphi x + 1. \end{aligned}$$

Dieses Polynom hat genau dann reelle Nullstellen, falls

$$4\cos^2\varphi - 4 \geq 0,$$

d.h. falls $\varphi = 0$ oder π ist, für $\varphi \in [0, 2\pi[$. Im allgemeinen besitzt eine Drehung also keine Eigenwerte und Eigenräume.

Über dem Grundkörper $K = \mathbb{C}$ besitzt dieses (wie jedes andere Polynom auch) aber stets Nullstellen.

Es sei $A \in \mathrm{Mat}(n; K)$ und λ sei Eigenwert von A.

Definition. Der Raum

$$\mathrm{Eig}(A, \lambda) = \{v \in K^n;\ Av = \lambda v\} = \mathrm{Ker}(A - \lambda E)$$

heißt der *Eigenraum* von A zum Eigenwert λ.

Lemma 8.6
$\lambda_1, \ldots, \lambda_r$ seien verschiedene Eigenwerte von A. Dann ist die Summe

$$\mathrm{Eig}(A, \lambda_1) + \ldots + \mathrm{Eig}(A, \lambda_r)$$

eine direkte Summe.

Beweis. Es ist zu zeigen:

$$v_1 + \ldots + v_r = 0 \text{ mit } v_i \in \mathrm{Eig}(A, \lambda_i) \ \Rightarrow\ v_1 = \ldots = v_r = 0.$$

Anwenden von A^m gibt:

$$0 = A^m v_1 + \ldots + A^m v_r = \lambda_1^m v_1 + \ldots + \lambda_r^m v_r.$$

Für ein Polynom

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

und eine Matrix $A \in \mathrm{Mat}\ (n; K)$ setzen wir

$$P(A) = a_m A^m + a_{m-1} A^{m-1} + \ldots + a_0 E \in \mathrm{Mat}\ (n, K).$$

Für jedes Polynom $P \in K[x]$ gilt dann:

$$0 = P(A)(v_1 + \ldots + v_r) = P(\lambda_1) v_1 + \ldots + P(\lambda_r) v_r.$$

Man wähle:

$$P_i(x) := \prod_{\substack{j=1 \\ j \neq i}}^{r} (x - \lambda_j)\ \in K[x].$$

Also ist $P_i(\lambda_j) = 0$ für $i \neq j$ und $P_i(\lambda_i) \neq 0$. Damit gilt

$$0 = P_i(A)(v_1 + \ldots + v_r) = \underbrace{P_i(\lambda_i)}_{\neq 0} v_i.$$

Daraus folgt

$$v_i = 0 \text{ für } i = 1, \ldots, r.$$

□

Korollar 8.7
Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

Homomorphismen und Matrizen

Es sei V ein K-Vektorraum der Dimension n.

Konvention. Im folgenden verstehen wir unter einer *Basis* von V eine *geordnete Basis*, d. h. ein n-Tupel

$$B = (b_1, \dots , b_n).$$

Dann besitzt jeder Vektor $x \in V$ eine eindeutige Darstellung

$$x = x_1 b_1 + \dots + x_n b_n.$$

Man definiert

$$q_B : \ V \longrightarrow K^n; \quad x \longmapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Es gilt $q_B(b_i) = e_i$ und die Abbildung q_B ist ein *Isomorphismus* von Vektorräumen.

Sei nun

$$C = (c_1, \dots , c_n)$$

eine weitere Basis. Da sich jeder Vektor b_i eindeutig durch die Vektoren c_j darstellen läßt, gibt es eine Matrix $A = (\alpha_{ij}) \in \mathrm{Mat}(n; K)$ mit

$$(*) \qquad b_j = \sum_{i=1}^{n} \alpha_{ij} c_i \qquad (j = 1, \dots , n).$$

Lemma 8.8
Das Diagramm

$$\begin{tikzcd} K^n \arrow[rr, "h_A"] & & K^n \\ & V \arrow[ul, "q_B"] \arrow[ur, "q_C"'] & \end{tikzcd}$$

kommutiert, d. h.

$$q_C = h_A \circ q_B.$$

Beweis. Es ist für $x \in V$:

$$x = \sum_{j=1}^{n} x_j b_j \overset{(*)}{=} \sum_{j=1}^{n} x_j \sum_{i=1}^{n} \alpha_{ij} c_i = \sum_{i=1}^{n} \sum_{j=1}^{n} (x_j \alpha_{ij}) c_i.$$

Es sei nun

$$x'_i := \sum_{j=1}^{n} \alpha_{ij} x_j.$$

Dann ist

$$x = \sum_{i=1}^{n} x'_i c_i.$$

Damit erhält man

$$(h_A \circ q_B)(x) = h_A(q_B(x)) = A q_B(x) = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} \alpha_{1j} x_j \\ \vdots \\ \sum_{j=1}^{n} \alpha_{nj} x_j \end{pmatrix} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = q_C(x).$$

□

Bemerkung. Da q_B und q_C Isomorphismen sind, gilt dies auch für h_A, d.h. $A \in \mathrm{GL}(n, K)$.

Definition. Die Matrix $A = (\alpha_{ij})$ heißt die *Übergangsmatrix* des Basiswechsels.

Darstellung von Homomorphismen

Gegeben sei ein n-dimensionaler Vektorraum V mit einer Basis $B = (b_1, \ldots, b_n)$ und ein m-dimensionaler Vektorraum V' mit einer Basis $B' = (b'_1, \ldots, b'_m)$. Ferner sei $f : V \to V'$ ein Homomorphismus. Dann betrachten wir das Diagramm

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "q_B"'] & V' \arrow[d, "q_{B'}"] \\ K^n & K^m. \end{tikzcd}$$

Da q_B ein Isomorphismus ist, kann man definieren

$$\boxed{\hat{f} := q_{B'} \circ f \circ q_B^{-1}}.$$

Dann ist $\hat{f}$ die lineare Abbildung, die folgendes Diagramm kommutativ macht:

$$\begin{tikzcd} V \arrow[r, "f"] & V' \arrow[d, "q_{B'}"] \\ K^n \arrow[u, "q_B^{-1}"] \arrow[r, "\hat{f}"] & K^m. \end{tikzcd}$$

Nach Satz (5.10) gibt es genau eine Matrix

$$M(f) := M_{B,B'}(f) \in \mathrm{Mat}(m,n;K)$$

mit

$$\boxed{\hat{f} = h_{M(f)}}.$$

Definition. Man sagt, f wird bezüglich der Basen B, B' durch die Matrix $M(f) = M_{B,B'}(f)$ dargestellt.

Satz 8.9
Es seien V, V' Vektorräume mit Basen B, B'. Dann gilt:
(i) Es gibt zu jedem Homomorphismus $f \in \mathrm{Hom}(V, V')$ genau eine Matrix $M(f) = M_{B,B'}(f) \in \mathrm{Mat}(m,n;K)$ mit

$$f = q_{B'}^{-1} \circ h_{M(f)} \circ q_B.$$

(ii) Es gilt $M(f) = (\mu_{ij})$ mit

$$f(b_j) = \sum_{i=1}^{m} \mu_{ij} b_i' \qquad (j = 1, \ldots, n).$$

(iii) Die Abbildung

$$\begin{array}{rrcl} M = M_{B,B'}: & \mathrm{Hom}_K(V,V') & \longrightarrow & \mathrm{Mat}(m,n;K) \\ & f & \longmapsto & M_{B,B'}(f) = M(f) \end{array}$$

ist ein Isomorphismus von Vektorräumen.

Beweis. (i) Es gilt

$$\hat{f} = q_{B'} \circ f \circ q_B^{-1}, \quad \hat{f} = h_{M(f)}.$$

Damit ist $\hat{f}$, und somit auch $M(f)$, durch f eindeutig festgelegt.

(ii) Es gilt

$$\begin{array}{rl} q_B: & V \longrightarrow K^n,\ q_B(b_j) = e_j\,; \quad j = 1, \ldots, n \\ q_{B'}: & V' \longrightarrow K^m,\ q_{B'}(b_i') = e_i\,; \quad i = 1, \ldots, m. \end{array}$$

Nun gilt

$$\begin{array}{ll} (q_{B'} \circ f \circ q_B^{-1})(e_j) & = \hat{f}(e_j) = M(f)e_j = \sum_{i=1}^m \mu_{ij} e_i \\ \quad\quad || & \\ q_{B'} \circ f(b_j). & \end{array}$$

Anwendung von $q_{B'}^{-1}$ von links ergibt:

$$f(b_j) = \sum_{i=1}^{m} \mu_{ij} q_{B'}^{-1}(e_i) = \sum_{i=1}^{m} \mu_{ij} b'_i.$$

(iii) Die Abbildung M ist nach dem bisher Gesagten bijektiv. Dies ist Homomorphismus: Da die Zuordnung $\hat{f} \mapsto M(f)$ linear ist, bleibt noch zu zeigen, daß die Zuordnung

$$f \longmapsto \hat{f} = q_{B'} \circ f \circ q_B^{-1}$$

linear ist. Dies folgt aus

$$(\widehat{\alpha f + \beta g}) = q_{B'} \circ (\alpha f + \beta g) \circ q_B^{-1} = \alpha(q_{B'} \circ f \circ q_B^{-1}) + \beta(q_{B'} \circ g \circ q_B^{-1}) = \alpha \hat{f} + \beta \hat{g}.$$

□

Basiswechsel

V sei ein n-dimensionaler K-Vektorraum mit Basis $B = (b_1, \ldots, b_n)$. Es sei $C = (c_1, \ldots, c_n)$ eine weitere Basis mit Übergangsmatrix $A \in \mathrm{GL}(n; K)$, d. h.

$$b_j = \sum_{i=1}^{n} \alpha_{ij} c_i. \quad (j = 1, \ldots, n).$$

Ebenso sei V' ein m-dimensionaler K-Vektorraum mit Basis $B' = (b'_1, \ldots, b'_m)$ und $C' = (c'_1, \ldots, c'_m)$ eine weitere Basis mit Übergangsmatrix $A' \in \mathrm{GL}(m; K)$, d. h.

$$b'_j = \sum_{i=1}^{m} \alpha'_{ij} c'_i. \quad (j = 1, \ldots, m).$$

Sei $f \in \mathrm{Hom}_K(V, V')$.

Satz 8.10

$$M_{C,C'}(f) = A' M_{B,B'}(f) A^{-1}.$$

Beweis. Man hat ein Diagramm:

$$\begin{tikzcd}
 & V \arrow[r, "f"] \arrow[d, "q_B"] \arrow[dl, phantom, "\circlearrowleft"] \arrow[dr, phantom, "\circlearrowleft"] \arrow[ddl, "q_C"'] & V' \arrow[d, "q_{B'}"'] \arrow[ddr, "q_{C'}"] \arrow[dr, phantom, "\circlearrowleft"] & \\
 & K^n \arrow[r, "h_{M_{B,B'}(f)}"] \arrow[dl, "h_A"] \arrow[r, phantom, shift right=3, "\circlearrowleft"] & K^m \arrow[dr, "h_{A'}"] & \\
K^n \arrow[rrr, "h_{M_{C,C'}(f)}"] & & & K^m.
\end{tikzcd}$$

Zunächst gilt nämlich nach Definition von $M_{B,B'}(f)$ und $M_{C,C'}(f)$: Hierbei bedeutet das Symbol $\circlearrowleft$, daß das entsprechende Diagramm kommutiert. Dann gilt:

$$h_{M_{B,B'}(f)} = q_{B'} \circ f \circ q_B^{-1},$$
$$h_{M_{C,C'}(f)} = q_{C'} \circ f \circ q_C^{-1}.$$

Ferner hatten wir schon in Lemma (8.8) gesehen, daß

$$q_C = h_A \circ q_B \ ; \quad q_{C'} = h_{A'} \circ q_{B'}.$$

Also gilt auch

$$h_{M_{C,C'}(f)} = h_{A'} \circ \underbrace{q_{B'} \circ f \circ q_B^{-1}}_{h_{M_{B,B'}(f)}} \circ h_A^{-1} = h_{A'} \circ h_{M_{B,B'}(f)} \circ h_{A^{-1}} = h_{A'M_{B,B'}(f)A^{-1}}.$$

Damit folgt

$$M_{C,C'}(f) = A'M_{B,B'}(f)A^{-1}.$$

$\square$

Folgerung. Zwei Matrizen sind genau dann *äquivalent*, wenn sie denselben Homomorphismus bezüglich zweier verschiedener Basen beschreiben.

Spezialfall. $V = V'$. Es sei $B = (b_1, \dots, b_n)$ eine Basis von V und $C = (c_1, \dots, c_n)$ eine weitere Basis von V mit

$$b_j = \sum_{i=1}^{n} \alpha_{ij} c_i \quad (j = 1, \dots, n),$$

d. h. $A = (\alpha_{ij})$ ist die Übergangsmatrix.

Wir setzen für $f \in \mathrm{End}(V)$:

$$M_B(f) := M_{B,B}(f).$$

Korollar 8.11

$$M_C(f) = AM_B(f)A^{-1}.$$

Folgerung. Zwei quadratische Matrizen sind genau dann *ähnlich*, wenn sie denselben Endomorphismus bezüglich verschiedener Basen darstellen.

Beispiel. Es sei $V = V' = \mathbb{R}^2$. Wir betrachten die Basen $B = (e_1, e_2) = (b_1, b_2)$ und $C = (e_1 + e_2, e_1 - e_2) = (c_1, c_2)$. Ferner sei $f : V \to V$ gegeben durch

$$f(e_1) = e_2, \quad f(e_2) = e_1.$$

Dann ist

$$M_B(f) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Wegen

$$f(e_1 + e_2) = e_1 + e_2, \quad f(e_1 - e_2) = -(e_1 - e_2)$$

d.h.

$$f(c_1) = c_1, \quad f(c_2) = -c_2,$$

gilt ferner

$$M_C(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Da

$$\begin{aligned} b_1 = e_1 &= \frac{1}{2}(e_1 + e_2) + \frac{1}{2}(e_1 - e_2) = \frac{1}{2}c_1 + \frac{1}{2}c_2 \\ b_2 = e_2 &= \frac{1}{2}(e_1 + e_2) - \frac{1}{2}(e_1 - e_2) = \frac{1}{2}c_1 - \frac{1}{2}c_2, \end{aligned}$$

folgt für die Matrix des Basiswechsels

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Es gilt

$$\begin{aligned} AM_B(f)A^{-1} &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = M_C(f). \end{aligned}$$

Wir können nun viele Begriffe von Matrizen auf Abbildungen übertragen. Dazu sei $f : V \to V$ ein Endomorphismus. Es sei B eine Basis von V und $M_B(f)$ die darstellende Matrix bezüglich der Basis B.

Definition.

(i) $\det f := \det M_B(f)$,

(ii) $\operatorname{Spur} f := \operatorname{Spur} M_B(f)$,

(iii) $\chi_f := \chi_{M_B(f)} = \det(xE - M_B(f))$.

Bemerkung. Diese Definitionen sind unabhängig von der Wahl von B, da die Größen auf der rechten Seite für alle Elemente in einer festen Ähnlichkeitsklasse gleich sind.

Definition. $0 \neq v \in V$ heißt *Eigenvektor* der Abbildung f zum *Eigenwert* $\lambda \in K$, falls

$$f(v) = \lambda v.$$

Bemerkung. Der Vektor v ist genau dann ein Eigenvektor von f, wenn $q_B(v)$ ein Eigenvektor von $M_B(f)$ ist. Dies folgt aus dem kommutativen Diagramm

$$\begin{tikzcd} K^n \arrow[r, "M_B(f)"] & K^n \\ V \arrow[u, "q_B"] \arrow[r, "f"] & V. \arrow[u, "q_B"'] \end{tikzcd}$$

Es gilt nämlich

$$f(v) = \lambda v \;\Leftrightarrow\; q_B^{-1}(M_B(f)(q_B(v))) = \lambda v \;\Leftrightarrow\; M_B(f)(q_B(v)) = \lambda q_B(v).$$

Definition. f heißt *diagonalisierbar*, falls f eine Basis von Eigenvektoren besitzt.

Bemerkung. f hat diese Eigenschaften genau dann, wenn $M_B(f)$ für ein (und damit für alle) Basen B diese Eigenschaften besitzt.

§ 9 Euklidische und unitäre Vektorräume

Im folgenden gelte für den Grundkörper stets $K \in \{\mathbb{R}, \mathbb{C}\}$. Obwohl manche der in diesem Abschnitt betrachteten Begriffe auch für unendlich dimensionale Vektorräume erklärt sind, beschränken wir uns in diesem Kapitel auf die Theorie der endlich-dimensionalen Vektorräume.

Definition. V sei ein $\mathbb{R}$-Vektorraum. Eine Abbildung

$$\begin{array}{rrcl} \langle\ ,\ \rangle: & V \times V & \longrightarrow & \mathbb{R} \\ & (v,w) & \longmapsto & \langle v,w\rangle \end{array}$$

heißt

(i) eine *Bilinearform*, falls stets gilt

(BF1) $$\langle \lambda v + \mu v', w\rangle = \lambda\langle v,w\rangle + \mu\langle v',w\rangle$$

(BF2) $$\langle v, \lambda w + \mu w'\rangle = \lambda\langle v,w\rangle + \mu\langle v,w'\rangle,$$

(ii) *symmetrisch*, falls

(S) $$\langle v,w\rangle = \langle w,v\rangle,$$

(iii) *positiv definit*, falls

$$\langle v,v\rangle > 0 \quad \text{für } v \neq 0.$$

Definition.

(i) Ist V ein $\mathbb{R}$-Vektorraum, so nennt man eine positiv definite symmetrische Bilinearform ein *Skalarprodukt*.

(ii) Ein *euklidischer Raum* ist ein Paar $(V, \langle\ ,\ \rangle)$, wobei V ein $\mathbb{R}$-Vektorraum und $\langle\ ,\ \rangle$ ein Skalarprodukt ist.

Die Motivation, ein Skalarprodukt einzuführen, besteht darin Winkel und Längen zu messen.

Beispiele.

(1) $V = \mathbb{R}^n$

$$\langle x,y\rangle := {}^t\!xy = (x_1,\dots,x_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n\end{pmatrix} = \sum_{i=1}^n x_i y_i.$$

Bilinearität:

$$\begin{aligned} \langle \lambda x + \mu x', y\rangle &= {}^t(\lambda x + \mu x')y = \lambda({}^t\!xy) + \mu({}^t\!x'y) \\ &= \lambda\langle x,y\rangle + \mu\langle x',y\rangle. \end{aligned}$$

Analog im zweiten Argument.

(2) Symmetrie:

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} y_i x_i = \langle y, x\rangle.$$

(3) Positiv definit:

$$\langle x, x\rangle = \sum_{i=1}^{n} x_i^2 > 0 \quad \text{für } x \neq 0.$$

(4) $V = C[0,1] = \{f;\ f : [0,1] \to \mathbb{R} \text{ stetig}\}$.

$$\langle f, g\rangle := \int_0^1 f(t)g(t)\,dt.$$

Diese Bilinearform ist positiv definit, da

$$\langle f, f\rangle = \int_0^1 f^2(t)\,dt > 0 \quad \text{für } f \neq 0, f \in C[0,1].$$

Sei nun $K = \mathbb{C}$, V sei ein $\mathbb{C}$-Vektorraum.

Definition. Eine Abbildung

$$\langle\ ,\ \rangle :\ V \times V \longrightarrow \mathbb{C}$$

heißt

(i) eine *Sesquilinearform*, falls

(SF1) $$\langle \lambda v + \mu v', w\rangle = \overline{\lambda}\langle v, w\rangle + \overline{\mu}\langle v', w\rangle$$

(SF2) $$\langle v, \lambda w + \mu w'\rangle = \lambda\langle v, w\rangle + \mu\langle v, w'\rangle,$$

(ii) *hermitesch*, falls

(H) $$\langle v, w\rangle = \overline{\langle w, v\rangle},$$

(iii) *positiv definit*, falls

$$\langle v, v\rangle > 0 \quad \text{für } v \neq 0.$$

Bemerkung. Aus (H) folgt stets: $\langle v, v\rangle \in \mathbb{R}$.

Definition.
(i) Eine positiv definite, hermitesche Sesquilinearform auf einem $\mathbb{C}$-Vektorraum V heißt auch *Skalarprodukt* auf V.

(ii) Ein *unitärer Raum* ist ein Paar $(V, \langle\ \rangle)$ bestehend aus einem $\mathbb{C}$-Vektorraum und einem Skalarprodukt.

Beispiele.
(1) $V = \mathbb{C}^n$.

$$\langle x, y\rangle := {}^t\overline{x}y = (\overline{x}_1, \ldots, \overline{x}_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^n \overline{x}_i y_i.$$

Es gilt

$$\begin{aligned} \langle \lambda x + \mu x', y\rangle &= {}^t\overline{(\lambda x + \mu x')}y = \overline{\lambda}\,{}^t\overline{x}y + \overline{\mu}\,{}^t\overline{x}'y \\ &= \overline{\lambda}\langle x, y\rangle + \overline{\mu}\langle x', y\rangle. \\ \langle x, y\rangle &= \sum_{i=1}^n \overline{x}_i y_i = \overline{\sum_{i=1}^n x_i \overline{y}_i} = \overline{\sum_{i=1}^n \overline{y}_i x_i} = \overline{\langle y, x\rangle}. \\ \langle x, x\rangle &= \sum_{i=1}^n \overline{x}_i x_i = \sum_{i=1}^n |x_i|^2 > 0 \text{ für } x \neq 0. \end{aligned}$$

(2) $V = C_{\mathbb{C}}[0,1] = \{f : [0,1] \to \mathbb{C} \text{ stetig}\}$. ($f$ stetig heißt hier, daß $f = f_1 + if_2$ mit $f_1, f_2 : [0,1] \to \mathbb{R}$ stetig.)
Wir setzen

$$\langle f, g\rangle := \int_0^1 \overline{f(t)}g(t)\,dt.$$

Letzteres ist so zu verstehen:

$$F(t) := \overline{f(t)}g(t) = F_1(t) + iF_2(t)$$

mit

$$F_1, F_2 :\ [0,1] \longrightarrow \mathbb{R}.$$

Dann ist

$$\int_0^1 F(t)\,dt := \int_0^1 F_1(t)\,dt + i\int_0^1 F_2(t)\,dt \in \mathbb{C}.$$

Dann definiert $\langle\ ,\ \rangle$ ein Skalarprodukt auf $C_{\mathbb{C}}[0,1]$.

Definition.

(i) Eine Matrix $A \in \mathrm{Mat}(n;\mathbb{R})$ heißt *symmetrisch*, falls

$$A = {}^t A.$$

(ii) Eine Matrix $A \in \mathrm{Mat}(n;\mathbb{C})$ heißt *hermitesch*, falls

$$A = {}^t\overline{A}.$$

Sei $V = K^n$, $K \in \{\mathbb{R},\mathbb{C}\}$. Wir setzen für A symmetrisch (hermitesch):

$$\begin{array}{rccl} \langle\,,\,\rangle_A: & V \times V & \longrightarrow & K \\ & \langle v,w\rangle_A & := & {}^t\overline{v}Aw. \end{array}$$

Dies macht im reellen ebenso wie im komplexen Fall Sinn.

Lemma 9.1
Ist A symmetrisch (hermitesch), so ist $\langle\,,\,\rangle_A$ eine symmetrische Bilinearform (hermitesche Sesquilinearform).

Beweis. (i):

$$\begin{aligned} \langle \lambda v + \mu v', w\rangle_A &= {}^t\overline{(\lambda v + \mu v')}Aw = \\ (\overline{\lambda}\,{}^t\overline{v}A + \overline{\mu}\,{}^t\overline{v}'A)w &= \overline{\lambda}\,{}^t\overline{v}Aw + \overline{\mu}\,{}^t\overline{v}'Aw \\ &= \overline{\lambda}\langle v,w\rangle_A + \overline{\mu}\langle v',w\rangle_A. \end{aligned}$$

Analog im zweiten Argument.

(ii):

$$\overline{\langle w,v\rangle_A} = \overline{{}^t\overline{w}Av} = {}^t w\overline{A}\overline{v} = {}^t\overline{v}\,{}^t\overline{A}w = {}^t\overline{v}Aw = \langle v,w\rangle_A.$$

□

Diesen Prozeß kann man verallgemeinern. Sei dazu V ein K-Vektorraum mit einer fest gewählten Basis $B = (b_1, \ldots, b_n)$. Diese Basis liefert einen Isomorphismus

$$\begin{array}{rccl} q_B: & V & \longrightarrow & K^n \\ & b_i & \longmapsto & e_i. \end{array}$$

Für A wie oben setzt man

$$\begin{array}{rccl} \langle\,,\,\rangle_A: & V \times V & \longrightarrow & K \\ & \langle v,w\rangle_A & := & {}^t\overline{q_B(v)}Aq_B(w). \end{array}$$

Lemma 9.2
Ist A symmetrisch (hermitesch), so ist $\langle\,,\,\rangle_A$ eine symmetrische Bilinearform (hermitesche Sesquilinearform).

Beweis. (i):

$$\begin{aligned}\langle \lambda v + \mu v', w\rangle &= {}^t\overline{q_B(\lambda v + \mu v')}Aq_B(w)\\ &= {}^t\overline{(\lambda q_B(v) + \mu q_B(v'))}Aq_B(w)\\ &= \overline{\lambda}\,{}^t\overline{q_B(v)}Aq_B(w) + \overline{\mu}\,{}^t\overline{q_B(v')}Aq_B(w)\\ &= \overline{\lambda}\langle v, w\rangle_A + \overline{\mu}\langle v', w\rangle_A.\end{aligned}$$

(ii):

$$\begin{aligned}\overline{\langle w, v\rangle}_A &= ({}^t\overline{\overline{q_B(w)}Aq_B(v)})\\ &= {}^tq_B(w)\overline{Aq_B(v)}\\ &= {}^t\overline{q_B(v)}\,{}^t\overline{A}q_B(w) = \langle v, w\rangle_A.\end{aligned}$$

□

Bemerkung. Für $A = (a_{ij})$ gilt: $\langle b_i, b_j\rangle_A = a_{ij}$.

Denn:

$$\langle b_i, b_j\rangle_A = {}^t\overline{q_B(b_i)}Aq_B(b_j) = {}^t\overline{e_i}Ae_j = {}^te_iAe_j = a_{ij}.$$

Es sei nun umgekehrt

$$\langle\ ,\ \rangle :\ V \times V \longrightarrow K$$

eine symmetrische Bilinearform (hermitesche Sesquilinearform). Dann ordnen wir dieser eine Matrix A zu durch

$$A = (a_{ij});\ a_{ij} := \langle b_i, b_j\rangle.$$

Definition. A heißt die $\langle\ ,\ \rangle$ *darstellende Matrix* bezüglich der Basis B.

Lemma 9.3
A ist symmetrisch (hermitesch).

Beweis. $\overline{a_{ji}} = \overline{\langle b_j, b_i\rangle} = \overline{\overline{\langle b_i, b_j\rangle}} = \langle b_i, b_j\rangle = a_{ij}$, also $A = {}^t\overline{A}$. □

Zusammenfassend hat man

Satz 9.4
V sei ein K-Vektorraum, $K \in \{\mathbb{R}, \mathbb{C}\}$ mit fest gewählter Basis B. Dann liefert die Zuordnung $A \mapsto \langle\ ,\ \rangle_A$ eine Bijektion.

$$\left\{\begin{array}{c}\text{symmetrische (hermitesche)}\\ \text{Matrizen}\end{array}\right\} \overset{1:1}{\longleftrightarrow} \left\{\begin{array}{c}\text{symmetrische Bilinearformen auf } V\\ \text{(hermitesche Sesquilinearformen auf } V\text{)}\end{array}\right\}.$$

Bemerkung. Sei $V = \mathbb{R}^n$ mit der Standardbasis $\{e_1, \dots, e_n\}$. Dann entspricht das „Standardskalarprodukt“

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

gerade der Einheitsmatrix E, da

$$\langle x, y \rangle = {}^t x y = {}^t x E y.$$

Transformationsformel

Es sei

$$\langle\,,\,\rangle : \quad V \times V \longrightarrow K$$

eine symmetrische Bilinearform (hermitesche Sesquilinearform). Gegeben seien zwei Basen von V:

$$B = (b_1, \dots, b_n), \quad B' = (b'_1, \dots, b'_n).$$

Bezüglich dieser Basen werde die Form $\langle\,,\,\rangle$ durch Matrizen A, bzw. A' dargestellt.

Satz 9.5
Ist $S \in \mathrm{GL}(n; K)$ die Transformationsmatrix von B nach B', d. h. kommutiert das Diagramm

$$\begin{tikzcd} K^n \arrow[rr, "h_S"] & & K^n \\ & V \arrow[ul, "q_B"] \arrow[ur, "q_{B'}"'] & \end{tikzcd}$$

dann gilt

$$A = {}^t\overline{S} A' S.$$

Beweis. Es seien $v, w \in V$.

$$x := q_B(v), \quad y := q_B(w).$$

Dann gilt:

(1) $$\langle v, w \rangle = {}^t\overline{q_B(v)} A q_B(w) = {}^t\overline{x} A y.$$

Ferner ist

$$q_{B'}(v) = Sx, \quad q_{B'}(w) = Sy.$$

Also gilt

(2) $$\langle v, w\rangle = {}^t\overline{q_{B'}(v)}A'q_{B'}(w) = {}^t\overline{Sx}A'Sy = {}^t\overline{x}({}^t\overline{S}A'S)y.$$

Speziell für $v = b_i$, $w = b_j$ erhält man

(1): $$\langle b_i, b_j\rangle = {}^t\overline{e_i}Ae_j = (A)_{ij}$$

(2): $$\langle b_i, b_j\rangle = {}^t\overline{e_i}({}^t\overline{S}A'S)e_j = ({}^t\overline{S}A'S)_{ij}$$

Es folgt:

$$A = {}^t\overline{S}A'S.$$

□

Definition. Die Abbildung

$$\begin{array}{rccl} q: & V & \longrightarrow & K \\ & v & \longmapsto & \langle v, v\rangle =: q(v) \end{array}$$

heißt die zu $\langle\ ,\ \rangle$ gehörige *quadratische Form.*

Bemerkung. Man kann $\langle\ ,\ \rangle$ aus q zurückgewinnen („Polarisierung"):

$K = \mathbb{R}$: $\quad \langle v, w\rangle = \frac{1}{4}(q(v + w) - q(v - w))$.

$K = \mathbb{C}$: $\quad \langle v, w\rangle = \frac{1}{4}(q(v + w) - q(v - w) + iq(v - iw) - iq(v + iw))$.

Diese Aussagen zeigt man durch direktes Nachrechnen.

Normen und Metriken

V sei ein Vektorraum über $K \in \{\mathbb{R}, \mathbb{C}\}$.

Definition. Eine *Norm* auf V ist eine Abbildung

$$\begin{array}{rccl} \|\ \|: & V & \longrightarrow & \mathbb{R} \\ & v & \longmapsto & \|v\| \end{array}$$

mit folgenden Eigenschaften:

(N1) $\quad \|\lambda v\| = |\lambda|\|v\|$

(N2) $\quad \|v + w\| \leq \|v\| + \|w\|$ (Dreiecksungleichung)

(N3) $\quad \|v\| = 0 \Leftrightarrow v = 0.$

Bemerkung. Es gilt stets $\|v\| \geq 0$, da

$$0 = \|0 \cdot v\| = \|v - v\| \leq \|v\| + \|-v\| = \|v\| + \|v\| = 2\|v\|.$$

Definition. X sei eine Menge. Eine *Metrik* auf X ist eine Abbildung

$$d: \quad X \times X \longrightarrow \mathbb{R}$$

mit

(M1) $\quad d(x,y) = d(y,x) \quad$ (Symmetrie)

(M2) $\quad d(x,z) \leq d(x,y) + d(y,z) \quad$ (Dreiecksungleichung)

(M3) $\quad d(x,y) = 0 \Leftrightarrow x = y.$

Bemerkung.

(i) Es gilt stets $d(x,y) \geq 0$, da:

$$0 = d(x,x) \leq d(x,y) + d(y,x) = 2d(x,y).$$

(ii) $d(x,y)$ kann man als „*Abstand*“ von x und y interpretieren. Die Dreiecksungleichung hat dann die Interpretation, daß die Summe der Längen zweier Seiten in einem Dreieck stets größer gleich der Länge der dritten Seite ist.

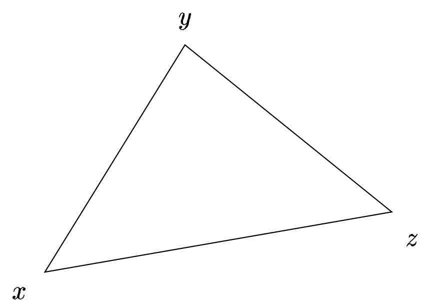


Abb. 36: Dreiecksungleichung

Lemma 9.6

Ist $\| \ \|$ eine Norm auf V, so wird durch

$$d(v,w) := \|v - w\|$$

eine Metrik auf V definiert.

Beweis. (i):

$$d(v,w) = \|v-w\| = \|(-1)(w-v)\| \overset{\text{(N1)}}{=} |-1|\|w-v\| = d(w,v).$$

(ii):

$$d(u,w) = \|u-w\| \overset{\text{(N2)}}{\leq} \|u-v\| + \|v-w\| = d(u,v) + d(v,w).$$

(iii):

$$d(u,v) = \|u-v\| = 0 \overset{\text{(N3)}}{\Leftrightarrow} u-v=0 \Leftrightarrow u=v.$$

Bemerkung. Nicht jede Metrik kommt von einer Norm. Betrachte etwa ($V = \mathbb{R}^n$):

$$d(v,w) = \begin{cases} 1 & \text{falls } v \neq w \\ 0 & \text{falls } v = w. \end{cases}$$

Die Metrik ist beschränkt, kommt also nicht von einer Norm.

Satz 9.7
Ist $(V, \langle\ ,\ \rangle)$ ein euklidischer (unitärer) Vektorraum, so wird durch

$$\|v\| := \sqrt{\langle v,v\rangle}$$

auf V eine Norm definiert.

Bemerkung. Es gilt stets $\langle v,v\rangle \geq 0$, da $\langle v,v\rangle > 0$ für $v \neq 0$, und $\langle 0,0\rangle = \langle 0, 0+0\rangle = \langle 0,0\rangle + \langle 0,0\rangle$, also $\langle 0,0\rangle = 0$. Damit ist die (nicht-negative) Wurzel stets definiert.

Beispiel. $V = \mathbb{R}^n$; $\langle x,y\rangle = \sum_{i=1}^n x_i y_i$. Dann ist

$$\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$$

und die dazugehörige Metrik ist die *euklidische Metrik*

$$d(x,y) = \sqrt{(x_1-y_1)^2 + \ldots + (x_n-y_n)^2}.$$

Bevor wir den letzten Satz beweisen können, benötigen wir noch

Satz 9.8 (Cauchy-Schwarz-Ungleichung)

$$|\langle v,w\rangle| \leq \|v\|\,\|w\|.$$

Beweis. Für $w = 0$ ist

$$0 = \langle v, 0 \rangle = \|v\| \, \|0\|.$$

Sei nun $w \neq 0$, also $\langle w, w \rangle > 0$. Für $\lambda \in K$ gilt:

$$0 \leq \langle v - \lambda w, v - \lambda w \rangle = \langle v, v \rangle - \overline{\lambda} \langle w, v \rangle - \lambda \langle v, w \rangle + \lambda \overline{\lambda} \langle w, w \rangle.$$

Setze nun

$$\lambda := \frac{\overline{\langle v, w \rangle}}{\langle w, w \rangle}.$$

Also gilt

$$0 \leq \langle v, v \rangle - \frac{\langle v, w \rangle \langle w, v \rangle}{\langle w, w \rangle} - \frac{\overline{\langle v, w \rangle} \langle v, w \rangle}{\langle w, w \rangle} + \frac{\langle v, w \rangle \overline{\langle v, w \rangle}}{\langle w, w \rangle}.$$

Es folgt

$$0 \leq \langle v, v \rangle \langle w, w \rangle - \langle v, w \rangle \underbrace{\langle w, v \rangle}_{= \overline{\langle v, w \rangle}}.$$

Also

$$0 \leq \|v\|^2 \|w\|^2 - |\langle v, w \rangle|^2$$

und damit gilt

$$|\langle v, w \rangle| \leq \|v\| \, \|w\|.$$

□

Bemerkung. Dies ist genau derselbe Beweis wie bei Satz (1.8), nur daß wir für die Skalare auch komplexe Zahlen zugelassen haben.

Beweis von Satz 9.7 (N1):

$$\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{\overline{\lambda} \lambda \langle v, v \rangle} = |\lambda| \sqrt{\langle v, v \rangle} = |\lambda| \|v\|.$$

(N2):

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\
&= \|v\|^2 + \|w\|^2 + \langle v, w \rangle + \overline{\langle v, w \rangle} \\
&= \|v\|^2 + \|w\|^2 + 2 \operatorname{Re} \langle v, w \rangle \\
&\leq \|v\|^2 + \|w\|^2 + 2 |\langle v, w \rangle| \\
&\overset{(9.8)}{\leq} \|v\|^2 + \|w\|^2 + 2 \|v\| \, \|w\| \\
&= (\|v\| + \|w\|)^2.
\end{aligned}$$

Also

$$\|v + w\| \le \|v\| + \|w\|.$$

(Vergleiche den Beweis von Korollar (1.9)).

(N3): $\|v\| = \sqrt{\langle v, v\rangle} = 0 \Leftrightarrow v = 0$. Dies folgt, da $\langle\, ,\, \rangle$ positiv definit ist. □

Bemerkung. Für Normen, die von einem Skalarprodukt kommen, gilt:

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + \langle v, w\rangle + \langle w, v\rangle \quad \text{(Pythagoras)}$$
$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2) \quad \text{(Parallelogrammgleichung).}$$

Bemerkung. Nicht jede Norm kommt von einem Skalarprodukt. Etwa:

$$V = \mathbb{R}^n, \quad \|x\|_\infty := \max\{|x_i|;\ i = 1, \dots, n\}.$$

Dies definiert eine Norm, welche die Parallelogrammgleichung nicht erfüllt:

$$x = (1, 0, \dots, 0);\ y = (0, 1, 0, \dots, 0).$$
$$\|x + y\|_\infty = 1;\ \|x - y\|_\infty = 1,$$
$$\|x\|_\infty = \|y\|_\infty = 1.$$

Definition. V sei ein euklidischer (unitärer) Vektorraum. Zwei Vektoren $v, w \in V$ heißen zueinander *orthogonal*, falls $\langle v, w\rangle = 0$.

Schreibweise. $v \perp w$.

Definition. Sei $(V, \langle\, ,\, \rangle)$ ein euklidischer (unitärer) Vektorraum.

(i) Zwei Unterräume $U, W \subset V$ heißen *orthogonal* ($U \perp W$), falls $u \perp w$ für alle $u \in U, w \in W$.

(ii) Ist W ein Unterraum, so heißt

$$W^\perp = \{v \in V;\ \langle v, w\rangle = 0 \text{ für alle } w \in W\}$$

das *orthogonale Komplement* von W.

(iii) Eine Familie $(v_i)_{i\in I}$ heißt *orthogonal*, falls $\langle v_i, v_j\rangle = 0$ für $i \neq j$.

(iv) Eine Familie $(v_i)_{i\in I}$ heißt *orthonormal*, falls die Familie orthogonal ist und $\|v_i\| = 1$ für alle $i \in I$.

(v) Eine Familie $(v_i)_{i\in I}$ heißt *Orthonormalbasis* (ONB, ONBasis), falls die Familie zugleich Basis und orthonormal ist.

Beispiel. $V = \mathbb{R}^n$ oder $\mathbb{C}^n$. $\langle\, ,\, \rangle$ sei das Standardskalarprodukt. Dann ist $e_1, \dots, e_n$ eine ONBasis, da

$$\langle e_i, e_j\rangle = \delta_{ij}.$$

Lemma 9.9
V sei ein euklidischer (unitärer) Vektorraum. Es sei $(v_i)_{i\in I}$ eine orthogonale Familie mit $v_i \neq 0$ für alle i. Sei $\tilde{v}_i := \frac{v_i}{\|v_i\|}$. Dann gilt

(i) $(\tilde{v}_i)_{i\in I}$ ist orthonormal.

(ii) $(\tilde{v}_i)_{i\in I}$ ist linear unabhängig.

Beweis. (i): Sei $c_i := \|v_i\|$. Dann gilt

$$\begin{aligned} \langle \tilde{v}_i, \tilde{v}_j \rangle &= \langle \frac{1}{c_i} v_i, \frac{1}{c_j} v_j \rangle = \frac{1}{\overline{c_i} c_j} \langle v_i, v_j \rangle = 0 \quad (i \neq j). \\ \|\tilde{v}_i\| &= \sqrt{\langle \tilde{v}_i, \tilde{v}_i \rangle} = \sqrt{\langle \frac{1}{c_i} v_i, \frac{1}{c_i} v_i \rangle} = \frac{1}{|c_i|} \sqrt{\langle v_i, v_i \rangle} \\ &= \frac{\|v_i\|}{c_i} = 1. \end{aligned}$$

(ii): Es sei für paarweise verschiedene Vektoren

$$\lambda_1 \tilde{v}_{i_1} + \ldots + \lambda_n \tilde{v}_{i_n} = 0.$$

Dann folgt

$$0 = \langle 0, \tilde{v}_{i_k} \rangle = \langle \lambda_1 \tilde{v}_{i_1} + \ldots + \lambda_n \tilde{v}_{i_n}, \tilde{v}_{i_k} \rangle = \overline{\lambda}_1 \langle \tilde{v}_{i_1}, \tilde{v}_{i_k} \rangle + \ldots + \overline{\lambda}_n \langle \tilde{v}_{i_1}, \tilde{v}_{i_k} \rangle = \overline{\lambda}_k.$$

□

Satz 9.10
Es sei V ein endlich-dimensionaler euklidischer (unitärer) Vektorraum. Es sei $W \subset V$ ein Unterraum und $w_1, \ldots, w_m$ eine ONBasis von W. Dann kann man diese zu einer ONBasis $w_1, \ldots, w_m, w_{m+1}, \ldots, w_n$ von V ergänzen.

Beweis. (Schmidtsches Orthonormalisierungsverfahren)
Ist $W = V$, so ist man fertig. Ansonsten wähle $v \in V \setminus W$. Sei

$$\tilde{v} := \overline{\langle v, w_1 \rangle} w_1 + \ldots + \overline{\langle v, w_m \rangle} w_m.$$

(Dies kann man als die senkrechte Projektion von v auf W interpretieren.) Wir setzen

$$w := v - \tilde{v}.$$

Da $\tilde{v} \in W, v \in V \setminus W$ ist $w \neq 0$. Es ist $w \in W^\perp$, da unter Verwendung von $\langle w_k, w_l \rangle = \delta_{kl}$ gilt:

$$\langle w, w_k \rangle = \langle v, w_k \rangle - \langle \tilde{v}, w_k \rangle = \langle v, w_k \rangle - \langle v, w_k \rangle = 0.$$

Wir setzen nun

$$w_{m+1} := \frac{w}{\|w\|}.$$

Dann ist $w_1, \dots, w_{m+1}$ eine orthogonale Familie und damit linear unabhängig. Diese Vektoren sind eine ONBasis von

$$W' := \operatorname{Span}(w_1, \dots, w_{m+1}) \supsetneqq W.$$

Ist $W' = V$, so ist man fertig. Ansonsten kommt man nach endlich vielen Schritten zum Ziel. $\square$

Beispiel. Wir betrachten $V := \mathbb{R}^3$ mit dem Standardskalarprodukt. Der Vektor

$$w_1 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

hat die Länge 1. Wir wählen $v := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \notin \mathbb{R}w_1$. Dann erhalten wir

$$\tilde{v} = 0, \quad w = v - \tilde{v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Da w die Länge 1 hat, ist w bereits normiert und wir erhalten

$$w_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Schließlich ist

$$v' := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \notin \operatorname{Span}(w_1, w_2) =: W.$$

Die orthogonale Projektion auf W ist

$$\tilde{v}' = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

und damit

$$w' = v' - \tilde{v}' = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{pmatrix}.$$

Wegen

$$\|w'\| = \frac{1}{\sqrt{2}}$$

setzen wir

$$w_3 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}.$$

Dann ist w_1, w_2, w_3 eine ONBasis von $\mathbb{R}^3$.

Korollar 9.11
Jeder endlich-dimensionale euklidische (unitäre) Vektorraum besitzt eine Orthonormalbasis.

Beweis. Obiges Verfahren mit $W = \{0\}$. $\square$

Definition. V sei ein euklidischer (unitärer) Vektorraum. Man sagt, V ist die *orthogonale Summe* der Unterräume $V_1, \ldots, V_k$, falls
(i) $V = V_1 + \ldots + V_k$.
(ii) $V_i \perp V_j$ für $i \neq j$.

Schreibweise. $V = V_1 \perp \ldots \perp V_k$.

Bemerkung. Eine orthogonale Summe ist stets direkt.

Denn: Wir haben zu zeigen, daß

$$V_i \cap (\underbrace{V_1 + \ldots + V_{i-1} + V_{i+1} + \ldots + V_k}_{=:V'}) = \{0\}.$$

Sei

$$0 \neq v \in V_i \cap V'.$$

Also

$$v = v_1 + \ldots + v_{i-1} + v_{i+1} + \ldots + v_k \quad (v_j \in V_j).$$

Damit gilt

$$\begin{aligned} 0 \neq \langle v, v \rangle &= \langle v, v_1 \rangle + \ldots + \langle v, v_{i-1} \rangle + \langle v, v_{i+1} \rangle + \ldots + \langle v, v_k \rangle \\ &= 0 \quad (\text{da } V_i \perp V_j \text{ für } i \neq j). \end{aligned}$$

Dies ist ein Widerspruch. $\square$

Satz 9.12
Es sei V ein endlich-dimensionaler euklidischer (unitärer) Vektorraum und $W \subset V$ ein Unterraum. Dann gilt

$$V = W \perp W^\perp.$$

Insbesondere ist

$$\dim V = \dim W + \dim W^\perp.$$

Beweis. Es sei $B = (b_1, \ldots, b_m)$ eine ONBasis von W. Diese ergänze man nach Satz (9.10) zu einer ONBasis $b_1, \ldots, b_m, b_{m+1}, \ldots, b_n$ von V. Es ist $b_{m+1}, \ldots, b_n \in W^\perp$. Sei nun $x \in V$. Also hat man eine Darstellung

$$x = (\underbrace{x_1 b_1 + \ldots + x_m b_m}_{\in W}) + (\underbrace{x_{m+1} b_{m+1} + \ldots + x_n b_n}_{\in W^\perp})$$

d. h.

$$V = W + W^\perp.$$

Nach Konstruktion ist $W \perp W^\perp$. □

Orthogonale und unitäre Endomorphismen

$(V, \langle\ ,\ \rangle)$ sei ein euklidischer oder unitärer Vektorraum.

Definition. Ein Endomorphimus $F : V \to V$ heißt *orthogonal (unitär)* wenn gilt:

$$\langle v, w\rangle = \langle F(v), F(w)\rangle \quad \text{(für alle } v, w \in V).$$

Lemma 9.13
F sei orthogonal (unitär). Dann gilt:
(i) $\|F(v)\| = \|v\|$ (für alle $v \in V$).

(ii) Ist λ ein Eigenwert von F, so ist $|\lambda| = 1$.

(iii) $v \perp w \Leftrightarrow F(v) \perp F(w)$.

(iv) F ist injektiv.

(v) Ist $\dim V < \infty$, so ist F ein Automorphismus und F^{-1} ist ebenfalls orthogonal (unitär).

Beweis. (i): $\|F(v)\| = \sqrt{\langle F(v), F(v)\rangle} = \sqrt{\langle v, v\rangle} = \|v\|$.

(ii): $F(v) = \lambda v \Rightarrow \|F(v)\| = |\lambda|\,\|v\| \underset{(v\neq 0)}{\overset{\text{(i)}}{\Rightarrow}} |\lambda| = 1$.

(iii): $v \perp w \Leftrightarrow \langle v, w\rangle = 0 \Leftrightarrow \langle F(v), F(w)\rangle = 0 \Leftrightarrow F(v) \perp F(w)$.

(iv): $F(v) = 0 \overset{\text{(i)}}{\Rightarrow} \|v\| = \|F(v)\| = 0 \Rightarrow v = 0$.

(v): F ist ein Automorphismus wegen (iv), falls $\dim V < \infty$. Die Abbildung F^{-1} ist ebenfalls orthogonal (unitär), da

$$\langle F^{-1}(v), F^{-1}(w)\rangle = \langle F(F^{-1}(v)), F(F^{-1}(w))\rangle = \langle v, w\rangle.$$

□

Satz 9.14
F ist genau dann orthogonal (unitär) wenn $\|F(v)\| = \|v\|$ für alle $v \in V$ gilt.

Beweis. „$\Rightarrow$" Siehe Lemma (9.13) (i).

„$\Leftarrow$" Es sei $K = \mathbb{R}$ (der Fall $K = \mathbb{C}$ kann analog bewiesen werden).

$$\begin{aligned}\langle v, w\rangle &= \frac{1}{4}(\langle v+w, v+w\rangle - \langle v-w, v-w\rangle) \quad \text{(Polarisierung)}\\ &= \frac{1}{4}(\langle F(v+w), F(v+w)\rangle - \langle F(v-w), F(v-w)\rangle) = \langle F(v), F(w)\rangle.\end{aligned}$$

□

Definition.
(i) $A \in \mathrm{GL}(n;\mathbb{R})$ heißt *orthogonal*, falls

$$A^{-1} = {}^tA \quad (\Leftrightarrow A\,{}^tA = E \Leftrightarrow {}^tAA = E),$$

(ii) $A \in \mathrm{GL}(n;\mathbb{C})$ heißt *unitär*, falls

$$A^{-1} = {}^t\overline{A} \quad (\Leftrightarrow A{}^t\overline{A} = E \Leftrightarrow {}^t\overline{A}A = E).$$

Bemerkung. A unitär (orthogonal) $\Rightarrow |\det A| = 1$.

Denn: $A{}^t\overline{A} = E \Rightarrow 1 = \det A \det {}^t\overline{A} = \det A \det \overline{A} = (\det A)(\overline{\det A}) = |\det(A)|^2$.

Definition.
(i) $O(n) := \{A \in \mathrm{GL}(n;\mathbb{R});\ A^{-1} = {}^tA\}$ (*orthogonale Gruppe*)

(ii) $SO(n) := \{A \in O(n);\ \det A = 1\}$ (*spezielle orthogonale Gruppe*)

(iii) $U(n) := \{A \in \mathrm{GL}(n;\mathbb{C});\ A^{-1} = {}^t\overline{A}\}$ (*unitäre Gruppe*)

(iv) $SU(n) := \{A \in U(n);\ \det A = 1\}$ (*spezielle unitäre Gruppe*).

Daß die angegebenen Mengen Untergruppen der $\mathrm{GL}(n;K)$ sind, rechnet man leicht nach. Wir betrachten als ein Beispiel die unitäre Gruppe:

Seien $A, B \in U(n)$. Dann gilt:

$$\begin{aligned}(AB)^{-1} &= B^{-1}A^{-1} = {}^t\overline{B}{}^t\overline{A} = {}^t(\overline{AB}) \Rightarrow AB \in U(n),\\ (A^{-1})^{-1} &= A = ({}^t\overline{A})^{-1} = {}^t(\overline{A}^{-1}) = {}^t(\overline{A^{-1}}) \Rightarrow A^{-1} \in U(n).\end{aligned}$$

Lemma 9.15
Sei $K \in \{\mathbb{R}, \mathbb{C}\}$. Für $A \in \mathrm{Mat}(n;K)$ sind äquivalent:
(i) $A \in O(n)$ (bzw. $U(n)$).

(ii) Die Spaltenvektoren von A bilden eine ONBasis von K^n (bezüglich des Standardskalarprodukts).

(iii) Die Zeilenvektoren von A bilden eine ONBasis von K^n (bezüglich des Standardskalarprodukts).

Beweis. Für $A = (a_1, \dots, a_n)$ gilt ${}^t\overline{A} = \begin{pmatrix} {}^t\overline{a}_1 \\ \vdots \\ {}^t\overline{a}_n \end{pmatrix}$.

(i)$\Leftrightarrow$ (ii):

$$\begin{aligned} A \in U(n) \Leftrightarrow {}^t\overline{A}A = E &\Leftrightarrow {}^t\overline{a}_i a_j = \delta_{ij} \Leftrightarrow \langle a_i, a_j \rangle = \delta_{ij} \\ &\Leftrightarrow a_1, \dots, a_n \text{ ist eine ONBasis.} \end{aligned}$$

(i)$\Leftrightarrow$(iii): Analog. □

Satz 9.16
Es sei $B = (v_1, \dots, v_n)$ eine ONBasis von X. Dann sind äquivalent:

(i) F ist ein orthogonaler (unitärer) Endomorphismus

(ii) Die darstellende Matrix $M_B(F)$ ist orthogonal (unitär).

Beweis. Es sei

$$v = x_1 v_1 + \dots + x_n v_n; \quad w = y_1 v_1 + \dots + y_n v_n$$

und

$$x = q_B(v) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = q_B(w) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Dann gilt

$$\begin{aligned} \langle v, w \rangle &= \langle \sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j \rangle = \sum_{i,j=1}^n \overline{x}_i y_j \underbrace{\langle v_i, v_j \rangle}_{\delta_{ij}} = \\ &= \sum_{i=1}^n \overline{x}_i y_i = \langle x, y \rangle = {}^t\overline{x} y. \end{aligned}$$

Wir setzen $A := M_B(F)$. Dann gilt:

$$\begin{aligned} F \text{ unitär (orthogonal)} &\Leftrightarrow \langle v, w \rangle = \langle F(v), F(w) \rangle \text{ für alle } v, w \in V \\ &\Leftrightarrow {}^t\overline{x} y = {}^t(\overline{Ax})(Ay) \text{ für alle } x, y \in K^n \\ &\Leftrightarrow {}^t\overline{x} E y = {}^t\overline{x}\, {}^t\overline{A} A y \text{ für alle } x, y \in K^n \\ &\Leftrightarrow {}^t\overline{A} A = E \\ &\Leftrightarrow A \in U(n) (\text{bzw. } 0(n)). \end{aligned}$$

□

Theorem 9.17
V sei ein unitärer Vektorraum endlicher Dimension. Ferner sei $F : V \to V$ ein unitärer Endomorphismus. Dann besitzt V eine ONBasis bestehend aus Eigenvektoren von F. Insbesondere ist F diagonalisierbar.

Beweis. V ist ein $\mathbb{C}$-Vektorraum. Nach dem Fundamentalsatz der Algebra (für eine eingehendere Diskussion siehe Lineare Algebra II) zerfällt jedes Polynom über $\mathbb{C}$ in Linearformen. Insbesondere gilt für das charakteristische Polynom des Endomorphismus F, daß

$$\chi_F(x) = (x - \lambda_1) \cdots (x - \lambda_n) \quad (\lambda_i \in \mathbb{C}).$$

Wir machen nun Induktion nach $n = \dim V$.

$n = 1$: Klar.

$n - 1 \mapsto n$: Es sei v_1 ein Eigenvektor. (Ein solcher existiert stets, da χ_F zerfällt.) Wir können annehmen, daß $\|v_1\| = 1$. Sei

$$W := \langle v_1 \rangle^\perp = \{w \in V;\ \langle w, v_1 \rangle = 0\}.$$

Dann ist nach Satz (9.12):

$$V = W \perp \operatorname{Span}\{v_1\}.$$

Behauptung. $F(W) \subset W$.

Denn: Sei $F(v_1) = \lambda_1 v_1$. Dann ist nach Lemma (9.13) (ii) $|\lambda_1| = 1$. Damit gilt für $w \in W$:

$$\lambda_1 \langle F(w), v_1 \rangle = \langle F(w), \lambda_1 v_1 \rangle = \langle F(w), F(v_1) \rangle = \langle w, v_1 \rangle = 0.$$

Da $\lambda_1 \neq 0$ folgt

$$F(w) \in W.$$

Nun ist $F|_W : W \to W$ unitär. Nach der Induktionsvoraussetzung besitzt W eine ONBasis $v_2, \ldots, v_n$ von Eigenvektoren. Setze

$$B = \{v_1, \ldots, v_n\}.$$

□

Korollar 9.18
Sei $A \in U(n)$. Dann gibt es eine unitäre Matrix $S \in U(n)$ mit

$${}^t\overline{S} A S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Hierbei gilt, daß $|\lambda_i| = 1$ für $i = 1, \ldots, n$.

Beweis. Es sei $B = (v_1, \dots, v_n)$ eine ONBasis von Eigenvektoren von A bezüglich des Standardskalarprodukts. Damit ist nach Lemma (9.15)

$$S := (v_1, \dots, v_n) \in U(n).$$

Also

$${}^t\overline{S}AS = S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Die Aussage $|\lambda_i| = 1$ folgt aus Lemma (9.13) (ii). □

Korollar 9.19
V sei unitär, $\dim V < \infty$. Es sei $F : V \to V$ ein unitärer Endomorphismus mit paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$. Dann ist

$$V = \mathrm{Eig}(F, \lambda_1) \perp \dots \perp \mathrm{Eig}(F, \lambda_k).$$

Beweis. Da F diagonalisierbar ist, ist V die Summe der Eigenräume. Es bleibt zu zeigen, daß

$$\mathrm{Eig}(F, \lambda_i) \perp \mathrm{Eig}(F, \lambda_j) \quad \text{für } i \neq j.$$

Sei dazu

$$v \in \mathrm{Eig}(F, \lambda_i); \quad w \in \mathrm{Eig}(F, \lambda_j).$$

Dann gilt

$$\langle v, w \rangle = \langle F(v), F(w) \rangle = \langle \lambda_i v, \lambda_j w \rangle = \overline{\lambda}_i \lambda_j \langle v, w \rangle.$$

Falls $\langle v, w \rangle \neq 0$ ist, gilt:

$$1 = \overline{\lambda}_i \lambda_j \Rightarrow \ \lambda_i = \underbrace{\lambda_i \overline{\lambda}_i}_{=1} \lambda_j \Rightarrow \ \lambda_i = \lambda_j.$$

Dies ist ein Widerspruch. □

Orthogonale Abbildungen

Orthogonale Abbildungen lassen sich nicht immer diagonalisieren. Im zwei-dimensionalen sind die orthogonalen Abbildungen gerade die Drehungen und Spiegelungen.

Satz 9.20
Es sei $A \in O(2)$. Dann gibt es $\alpha \in [0, 2\pi[$ mit

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \text{ oder } A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

Beweis. Wir schreiben

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in O(2).$$

Da A reell ist, gilt

$$a_{11}^2 + a_{21}^2 = 1, \text{ d. h. } a_{11} + ia_{21} \in S^1 = \{z \in \mathbb{C};\ |z| = 1\}.$$

Dann gibt es $\alpha \in [0, 2\pi[$ mit

$$a_{11} + ia_{21} = e^{i\alpha} = \cos\alpha + i\sin\alpha,$$

also

$$a_{11} = \cos\alpha;\ a_{21} = \sin\alpha.$$

Wegen

$$a_{11}a_{12} + a_{21}a_{22} = 0$$

gilt

$$\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = c \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}.$$

Da $A \in O(2)$ ist $a_{12}^2 + a_{22}^2 = 1$, also ist $c = \pm 1$. Für $c = 1$ erhalten wir

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Für $c = -1$ ergibt sich

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}.$$

□

Bemerkung. Die beiden Fälle unterscheiden sich dadurch, daß im ersten Fall $\det A = 1$, d.h. $A \in SO(2)$ gilt, während im zweiten Fall $\det A = -1$ ist.

Wir diskutieren nun die geometrische Bedeutung dieser Fälle.

Fall 1: $\quad A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$

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

$\alpha = \pi$: $\quad A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$

Sei nun $\alpha \neq k\pi$; $k \in \mathbb{Z}$:

$$\chi_A(x) = (x - \cos\alpha)^2 + \sin^2\alpha = x^2 - 2\cos\alpha x + 1.$$

Die Lösungen dieser quadratischen Gleichung sind nicht reell:

$$x_{1/2} = \frac{2\cos\alpha \pm \sqrt{4\cos^2\alpha - 4}}{2} \notin \mathbb{R}.$$

Es gibt also keine Eigenwerte über $\mathbb{R}$. A entspricht einer *Drehung* um den Winkel α.

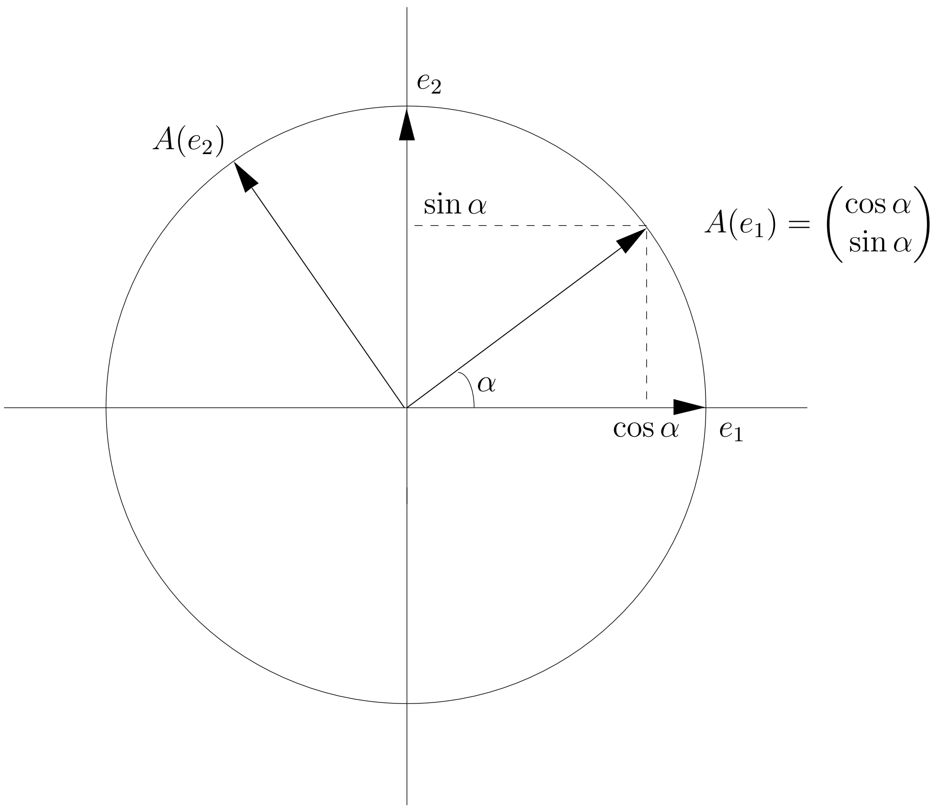


Abb. 37: Drehung um den Winkel α

Fall 2: $A = \begin{pmatrix}\cos\alpha & \sin\alpha\\ \sin\alpha & -\cos\alpha\end{pmatrix}$.

$$\chi_A(x) = (x - \cos\alpha)(x + \cos\alpha) - \sin^2\alpha = x^2 - 1 = (x-1)(x+1).$$

Es gibt also zwei Eigenvektoren v_1, v_2. Diese stehen aufeinander senkrecht und es gilt $Av_1 = v_1, Av_2 = -v_2$. A ist eine *Spiegelung* an $\mathbb{R}v_1$, $\det A = -1$, d. h. $A \notin SO(2)$.

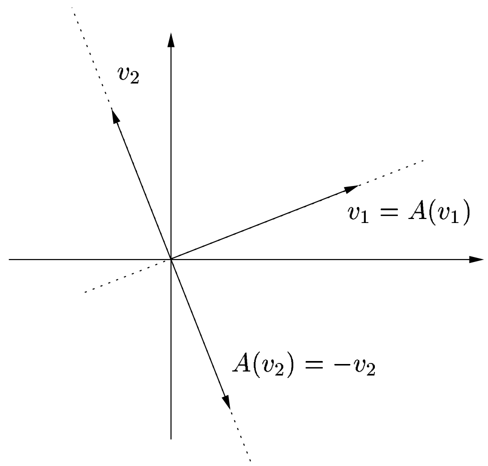


Abb. 38: Spiegelung

Mit

$$W = (v_1, v_2)$$

folgt

$$W^{-1}AW = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Theorem 9.21
V sei ein endlich-dimensionaler euklidischer Vektorraum und F ein orthogonaler Endomorphismus. Dann gibt es eine ONBasis von V bezüglich der F dargestellt wird durch eine Matrix der Form

$$A = \begin{pmatrix} 1 & & & & & & & 0 \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ & & & & & & A_1 & \\ & 0 & & & & & & \ddots \\ & & & & & & & & A_k \end{pmatrix}$$

mit $A_i = \begin{pmatrix} \cos\alpha_i & -\sin\alpha_i \\ \sin\alpha_i & \cos\alpha_i \end{pmatrix}$; $\alpha_i \neq k\pi$, $k \in \mathbb{Z}$.

Lemma 9.22
Jedes Polynom $P \in \mathbb{R}[x]$ besitzt eine Zerlegung

$$P = (x - \lambda_1) \cdots (x - \lambda_r)\, Q_1 \cdots Q_k$$

mit $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$; $Q_i \in \mathbb{R}[x]$ quadratisch ohne reelle Nullstelle; $i = 1, \ldots, k$.

Beweis. Wir betrachten $P \in \mathbb{R}[x] \subset \mathbb{C}[x]$. Ist $\lambda \in \mathbb{C}$ eine Nullstelle von P, so auch $\overline{\lambda}$:

$$\begin{aligned} P(\overline{\lambda}) &= a_0 + a_1\overline{\lambda} + \ldots + a_n\overline{\lambda}^n \overset{(a_i \in \mathbb{R})}{=} \overline{a}_0 + \overline{a}_1\overline{\lambda} + \ldots + \overline{a}_n\overline{\lambda}^n \\ &= \overline{P(\lambda)} = \overline{0} = 0. \end{aligned}$$

Also hat man eine Zerlegung

$$P(x) = (x - \lambda_1) \cdots (x - \lambda_r)(x - \lambda_{r+1})(x - \overline{\lambda}_{r+1}) \cdots (x - \lambda_s)(x - \overline{\lambda}_s).$$

wobei $\lambda_1, \ldots, \lambda_r$ reell sind. Setze

$$Q_i(x) := (x - \lambda_{r+i})(x - \overline{\lambda}_{r+i}) = x^2 - x(\underbrace{\lambda_{r+i} + \overline{\lambda}_{r+i}}_{=2\operatorname{Re}\lambda_{r+i}}) + |\lambda_{r+i}|^2.$$

Dann ist

$$Q_i \in \mathbb{R}[x]$$

und wir haben eine Zerlegung der gesuchten Form angegeben. □

Lemma 9.23
Es gibt einen Unterraum $W \subset V$, $\dim W = 1$ oder 2 mit $F(W) = W$.

Beweis. Die Behauptung ist klar, falls F einen Eigenwert und damit einen Eigenvektor besitzt. Wir können also annehmen, daß

$$\chi_F = Q_1 \cdots Q_k$$

wobei $Q_i(x) \in \mathbb{R}[x]$ quadratische Polynome ohne reelle Nullstelle sind. Es sei $0 \neq w \in V$. Nach dem Satz von Cayley-Hamilton, den wir in der Linearen Algebra II beweisen werden, gilt

$$\chi_F(F) = 0.$$

Es sei nun $i \in \{1, \ldots, k\}$ minimal mit:

$$Q_i(F)Q_{i-1}(F) \cdots Q_1(F)(w) = 0.$$

Dann ist

$$v := Q_{i-1}(F) \cdots Q_1(F)(w) \neq 0.$$

(Für $i = 1$ ist $v = w$.) Also gilt

$$v \neq 0, \quad Q_i(F)(v) = 0.$$

Sei

$$W := \mathrm{Span}(v, F(v)).$$

Es genügt zu zeigen, daß $F(W) \subset W$, da F nach Lemma (9.13) (iv) injektiv ist. Die Behauptung $F(W) \subset W$ folgt, falls $F^2(v) \in W$. Es sei nun

$$Q_i(x) = x^2 + \alpha x + \beta.$$

Es ist

$$0 = Q_i(F)(v) = F^2(v) + \alpha F(v) + \beta v \Rightarrow F^2(v) \in W.$$

□

Beweis des Theorems. Es sei W wie oben. Es genügt zu zeigen, daß

$$F(W^\perp) \subset W^\perp,$$

da wir dann Induktion nach $\dim V$ machen können. Hierzu sei $v \in W^\perp$, $w \in W$. Dann gilt, da aus $W = F(W)$ auch $F^{-1}(W) = W$ folgt:

$$\langle F(v), w\rangle \overset{F^{-1} \text{ orthogonal}}{=} \langle v, \underbrace{F^{-1}(w)}_{\in W}\rangle = 0.$$

□

Selbstadjungierte Endomorphismen

Es sei V ein unitärer (euklidischer) Vektorraum und $F : V \to V$ ein Endomorphismus.

Definition. Ein Endomorphismus F heißt *selbstadjungiert*, falls

$$\langle F(v), w\rangle = \langle v, F(w)\rangle \quad (\text{für alle } v, w \in V).$$

Bemerkung. Die Bezeichnung selbstadjungiert kommt von folgendem Sachverhalt: Ist $F : V \to V$ ein Endomorphismus, so kann man zeigen, daß es dann genau eine lineare Abbildung $F^{\mathrm{adj}} : V \to V$ gibt mit

$$\langle F(v), w\rangle = \langle v, F^{\mathrm{adj}}(w)\rangle \quad \text{für alle } v, w \in V.$$

Die Abbildung F^{adj} heißt die zu F *adjungierte Abbildung*. Wir betrachten hier den Fall, daß $F = F^{\mathrm{adj}}$ ist.

Satz 9.24
Sei $B = (v_1, \ldots, v_n)$ eine ONBasis von V und $A = M_B(F)$ sei die darstellende Matrix des Endomorphismus F bezüglich der Basis B. Dann sind äquivalent:

(i) F ist selbstadjungiert.

(ii) A ist hermitesch (symmetrisch).

Beweis. Für

$$v = x_1 v_1 + \ldots + x_n v_n, \quad w = y_1 v_1 + \ldots + y_n v_n$$

gilt, da B eine ONBasis ist, daß

$$\langle v, w \rangle = \overline{x}_1 y_1 + \ldots + \overline{x}_n y_n = {}^t\overline{x} y$$

mit

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = q_B(v); \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = q_B(w).$$

Also gilt:

$$\begin{aligned} F \text{ ist selbstadjungiert} \quad &\Leftrightarrow \quad \langle F(v), w \rangle = \langle v, F(w) \rangle \quad \text{für alle } v, w \in V \\ &\Leftrightarrow \quad \langle Ax, y \rangle = \langle x, Ay \rangle \quad \text{für alle } x, y \in K^n \\ &\Leftrightarrow \quad {}^t\overline{Ax} y = {}^t\overline{x} A y \quad \text{für alle } x, y \in K^n \\ &\Leftrightarrow \quad {}^t\overline{x}\, {}^t\overline{A} y = {}^t\overline{x} A y \quad \text{für alle } x, y \in K^n \\ &\Leftrightarrow \quad {}^t\overline{A} = A. \end{aligned}$$

□

Satz 9.25
Der Endomorphismus $F : V \to V$ sei selbstadjungiert. Dann zerfällt χ_F und alle Eigenwerte sind reell, d. h.

$$\chi_F(x) = (x - \lambda_1) \cdots (x - \lambda_n) \quad (\lambda_i \in \mathbb{R}).$$

Beweis. (i): $K = \mathbb{C}$. Dann zerfällt χ_F automatisch. Es sei λ ein Eigenwert, v ein zugehöriger Eigenvektor.

$$\lambda \langle v, v \rangle = \langle v, \lambda v \rangle = \langle v, F(v) \rangle = \langle F(v), v \rangle = \langle \lambda v, v \rangle = \overline{\lambda} \langle v, v \rangle.$$

Da $\langle v, v \rangle > 0$ ist, folgt hieraus $\lambda = \overline{\lambda}$.

(ii): $K = \mathbb{R}$. Es sei B eine ONBasis und A die darstellende Matrix. Man hat ein kommutatives Diagramm

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{h_A} & \mathbb{R}^n \\ \downarrow & & \downarrow \\ \mathbb{C}^n & \xrightarrow{h_A} & \mathbb{C}^n \end{array}$$

wobei wir lediglich die reelle Matrix A als komplexe Matrix auffassen. Als solche ist sie hermitesch und die Aussage folgt aus (i). □

Theorem 9.26
V sei ein unitärer (euklidischer) Raum. Die Abbildung $F : V \to V$ sei selbstadjungiert. Dann besitzt V eine ONBasis bestehend aus Eigenvektoren von F.

Beweis. Induktion nach $n = \dim V$.

$n = 1$: Klar.

$n-1 \mapsto n$: Nach Satz (9.25) gibt es einen Eigenwert λ_1 sowie einen dazugehörigen Eigenvektor v_1. Wir können $\|v_1\| = 1$ annehmen. Es sei

$$W := (Kv_1)^{\perp}.$$

Wir behaupten zunächst, daß W invariant unter F ist, d. h.

$$F(W) \subset W.$$

Sei hierzu $w \in W$, d. h. $\langle w, v_1\rangle = 0$. Dann ist

$$\langle F(w), v_1\rangle = \langle w, F(v_1)\rangle = \langle w, \lambda_1 v_1\rangle = \lambda_1\langle w, v_1\rangle = 0,$$

also

$$F(w) \in W.$$

Man wähle nun eine ONBasis von Eigenvektoren für $F|_W : W \to W$ etwa $B' = (v_2, \dots, v_n)$ und setze

$$B = (v_1, \dots, v_n).$$

□

Korollar 9.27 (Hauptachsentransformation)
Sei $A \in \mathrm{Mat}(n; K)$ eine hermitesche (bzw. symmetrische) Matrix. Dann gibt es eine Transformationsmatrix $S \in U(n)$ (bzw. $O(n)$) mit

$${}^t\overline{S}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Hierbei sind $\lambda_1, \dots, \lambda_n$ reell.

Beweis. Es gibt eine ONBasis $B = (v_1, \dots, v_n)$ von Eigenvektoren. Setze

$$S := (v_1, \dots, v_n) \in U(n)\,(\text{bzw. } O(n)).$$

Dann ist

$${}^t\overline{S}AS = S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

□

Korollar 9.28
V sei unitär (euklidisch) und $F : V \to V$ sei selbstadjungiert mit paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k$. Dann ist

$$V = \mathrm{Eig}(F, \lambda_1) \perp \ldots \perp \mathrm{Eig}(F, \lambda_k).$$

Beweis. Wir haben schon gesehen, daß V die Summe der Eigenräume ist. Es bleibt die Orthogonalität zu zeigen. Sei hierzu

$$v \in \mathrm{Eig}(F, \lambda_i), \quad w \in \mathrm{Eig}(F, \lambda_j) \quad (\lambda_i \neq \lambda_j).$$

Dann gilt

$$\begin{aligned} \lambda_j \langle v, w \rangle &= \langle v, \lambda_j w \rangle = \langle v, F(w) \rangle \\ &= \langle F(v), w \rangle = \langle \lambda_i v, w \rangle = \overline{\lambda}_i \langle v, w \rangle \\ &= \lambda_i \langle v, w \rangle. \end{aligned}$$

Also gilt

$$\underbrace{(\lambda_j - \lambda_i)}_{\neq 0} \langle v, w \rangle = 0 \;\Rightarrow\; \langle v, w \rangle = 0.$$

□

Geometrische Deutung der Hauptachsentransformation

Wir gehen hier kurz auf die geometrische Deutung der Hauptachsentransformation ein. Wir werden im nächsten Semester ausführlicher darauf zurückkommen.

Eine *quadratische Hyperfläche* im $\mathbb{R}^n$ ist die Menge aller Punkte $P = (x_1, \ldots, x_n) \in \mathbb{R}^n$, die eine Gleichung der Gestalt

$$\sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + c = 0 \qquad (a_{ij}, b_i, c \in \mathbb{R})$$

erfüllen.

Wir betrachten zunächst einmal den „homogenen“ Fall

$$\sum_{i,j=1}^{n} a_{ij} x_i x_j = 0.$$

Wegen

$$a_{ij} x_i x_j + a_{ji} x_j x_i = \frac{1}{2}(a_{ij} + a_{ji}) x_i x_j + \frac{1}{2}(a_{ij} + a_{ji}) x_j x_i$$

können wir annehmen, daß $a_{ij} = a_{ji}$ ist. D. h.

$$\sum_{i,j=1}^{n} a_{ij}x_i x_j = {}^t x A x \text{ mit } A = {}^t A = (a_{ij}).$$

Nach Korollar (9.27) gibt es einen Koordinatenwechsel

$$x = Sx', \text{ bzw. } x' = S^{-1}x$$

mit

$${}^t x A x = {}^t x' \, {}^t S A S x' = \lambda_1 (x'_1)^2 + \ldots + \lambda_n (x'_n)^2.$$

Beispiel.

(E) $$7x_1^2 + 6\sqrt{3}x_1x_2 + 13x_2^2 = 1.$$

Dann ist

$$A = \begin{pmatrix} 7 & 3\sqrt{3} \\ 3\sqrt{3} & 13 \end{pmatrix}.$$

Die Eigenwerte von A sind die Lösungen der Gleichung

$$\det(\lambda E - A) = \det \begin{pmatrix} \lambda - 7 & 3\sqrt{3} \\ 3\sqrt{3} & \lambda - 13 \end{pmatrix} = \lambda^2 - 20\lambda + 64 = 0,$$

d. h.

$$\lambda_1 = 4, \ \lambda_2 = 16.$$

Zugehörige normierte Eigenvektoren sind

$$v_1 = \frac{1}{2}\begin{pmatrix} \sqrt{3} \\ -1 \end{pmatrix}, \ v_2 = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}.$$

Es sei

$$S = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}.$$

Die Matrix S beschreibt eine Drehung um $-\pi/6$. Damit gilt

$${}^t S A S = \begin{pmatrix} 4 & 0 \\ 0 & 16 \end{pmatrix},$$

und mit

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = S \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

wird die Gleichung (E) zu

$$4y_1^2 + 16y_2^2 = 1.$$

Hier erkennen wir die Gleichung einer Ellipse.

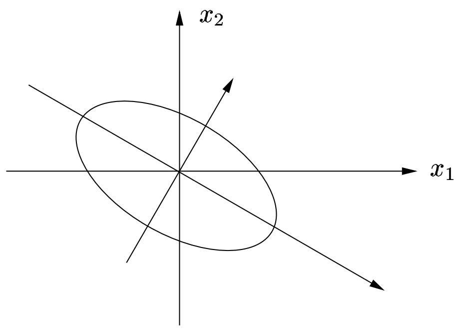


Abb. 39: Hauptachsentransformation einer Ellipse

Man kann mit diesem Verfahren jede quadratische Hyperfläche auf eine Gestalt der Form

$$\lambda_1 x_1^2 + \ldots + \lambda_n x_n^2 + \mu_1 x_1 + \ldots + \mu_n x_n + c = 0$$

bringen. Ist $\lambda_i \neq 0$, so kann man durch quadratische Ergänzung noch $\mu_i = 0$ erreichen. Wir werden hierauf zurückkommen.

Satz 9.29
V sei ein unitärer (euklidischer) Vektorraum. $F_1, \ldots, F_m : V \to V$ seien unitär oder selbstadjungiert. Es sind äquivalent:
(i) Die F_j sind simultan diagonalisierbar. D. h. es gibt eine ONBasis $B = (v_1, \ldots, v_n)$ so daß die v_i Eigenvektoren für alle F_j sind.
(ii) $F_i \circ F_j = F_j \circ F_i$ für alle i, j.

Bemerkung. Diese Aussage gilt nicht für orthogonale Abbildungen, da diese im allgemeinen nicht diagonalisierbar sind.

Beweis. (i)$\Rightarrow$(ii): Es sei $\lambda_j^{(i)}$ der Eigenwert von v_j bezüglich F_i. Dann gilt

$$\begin{aligned} F_i(F_j(v_k)) &= F_i(\lambda_k^{(j)} v_k) = \lambda_k^{(j)} F_i(v_k) = \lambda_k^{(j)} \lambda_k^{(i)} v_k \\ F_j(F_i(v_k)) &= F_j(\lambda_k^{(i)} v_k) = \lambda_k^{(i)} F_j(v_k) = \lambda_k^{(i)} \lambda_k^{(j)} v_k. \end{aligned}$$

Da $B = (v_1, \ldots, v_n)$ eine Basis ist, folgt $F_j \circ F_i = F_i \circ F_j$.

(ii)$\Rightarrow$(i): Induktion nach m.

$m = 1$: Dies ist der Diagonalisierungssatz.

$m - 1 \mapsto m$: Man betrachte die Zerlegung von V für F_1:

$$V = \underbrace{\mathrm{Eig}(F_1, \lambda_1^{(1)})}_{=:W_1} \perp \ldots \perp \underbrace{\mathrm{Eig}(F_1, \lambda_{k_1}^{(1)})}_{=:W_{k_1}}.$$

Es ist $F_1|_{W_i} = \lambda_i^{(1)}\mathrm{id}$, d. h. F_1 ist auf W_i in jeder Basis diagonalisiert. Um die Induktionsvoraussetzung anwenden zu können, genügt zu zeigen, daß

$$F_j(W_i) \subset W_i \qquad (j = 2, \ldots, m;\ i = 1, \ldots, k_1).$$

Sei hierzu $v \in W_i$:

$$F_1(F_j(v)) = F_j(F_1(v)) = F_j(\lambda_i^{(1)} v) = \lambda_i^{(1)} F_j(v).$$

Also gilt

$$F_j(v) \in \mathrm{Eig}(F_1, \lambda_i^{(1)}) = W_i.$$

$\square$

Korollar 9.30
$A_1, \ldots, A_m \in \mathrm{Mat}(n; K)$ seien unitäre oder hermitesche (bzw. symmetrische) Matrizen. Es sei

$$[A_i, A_j] := A_i A_j - A_j A_i = 0 \qquad (i, j = 1, \ldots, m).$$

Dann gibt es $S \in U(n)$ (bzw. $O(n)$), so daß

$${}^t\overline{S} A_i S = \text{Diagonalmatrix}$$

für $i = 1, \ldots, m$.

Bemerkung. Man nennt

$$[A, B] := AB - BA$$

den *Kommutator* von A und B.

Literaturverzeichnis

[Ar] M. Artin, *Algebra*, Birkhäuser Verlag, Basel, 1993.

[Fi] G. Fischer, *Lineare Algebra*, Vieweg, Braunschweig, 1975.

[Gr] W. H. Greub, *Lineare Algebra*, Springer Verlag, Berlin, 1967.

[Koe] M. Koecher, *Lineare Algebra und Analytische Geometrie*, Springer Verlag, Berlin, 1992.

[KM] H.–J. Kowalsky, G. Michler *Lineare Algebra*, W. de Gruyter Verlag, Berlin, 1997.

[Lo] F. Lorenz, *Lineare Algebra I und II*, BI-Verlag, Mannheim, 1982

Englischsprachige Literatur

[Ha] P. Halmos *Finite dimensional vector spaces*, Springer Verlag, 1974.

[La] P. Lax *Linear algebra*, A. Wiley, 1997.