# 1 Vorbemerkungen

In den ersten beiden Abschnitten werden einige grundlegende Bezeichnungen und Begriffe eingeführt, auf die überall in der Mathematik zurückgegriffen wird. Außerdem formulieren wir logische Schlussweisen, die bereits aus der Schule bekannt sind, als mathematische Prinzipien.
Wenn Ihnen diese Dinge jetzt zu abstrakt erscheinen, können Sie auch darüber hinweg lesen und im Laufe der Stoffentwicklung immer mal auf diese Seiten zurück schauen.
Am Ende der Vorbemerkungen geben wir dann eine Einführung in Aufgabenstellungen der **linearen Algebra**, die sich wie ein roter Faden durch viele andere Gebiete der reinen und vor allem auch der angewandten Mathematik zieht.

## 1.1 Grundbegriffe aus der naiven Logik und Mengenlehre

### 1.1.1 Logische Zeichen und Schlussweisen

In der modernen mathematischen Sprache haben sich folgende **logische Zeichen** durchgesetzt, die viele Darstellungen verkürzen und abstrahieren:

1) <u>Logische Verbindungszeichen</u>

$$\left.\begin{array}{ll} \wedge & \text{und} \\ \vee & \text{oder} \\ \neg & \text{nicht} \\ \Rightarrow & \text{daraus folgt} \end{array}\right\} \quad \text{sind aussagenlogische Verknüpfungen}$$

$$\left.\begin{array}{ll} = & \text{gleich} \\ (\neq & \text{ungleich}) \end{array}\right\} \quad \text{gehören zur “Prädikatenlogik”}$$

2) <u>Logische Quantoren</u>

$\forall$ für alle (für jedes)
$\exists$ es existiert
$\exists!$ es existiert genau ein

3) <u>Variablenzeichen</u>

z.B. Buchstaben $A, B, C, \ldots,$
stehen in der naiven Logik für Aussagen, d. h. für Ihren Wahrheitsgehalt. Der konkrete Inhalt ist dabei unwichtig, es geht nur um “richtig” oder “falsch”.



4) <u>Klammern</u>

werden zur Kennzeichnung der Reihenfolge des logischen Schließens eingeführt.

Diese Zeichen benutzt man nun im Zusammenspiel, wie wir gleich an Beispielen sehen werden.

Zur Vereinfachung verwendet man noch folgende abgeleitete Symbole:

$A \Leftrightarrow B$ ($A$ gilt genau dann, wenn $B$ richtig ist)
steht für: $(A \Rightarrow B) \wedge (B \Rightarrow A)$
(aus $A$ folgt $B$ und aus $B$ folgt $A$)

$A \not\Rightarrow B$ (aus $A$ folgt nicht $B$)
steht für: $\neg(A \Rightarrow B)$.

Oft werden auch gewisse Zusatzzeichen benutzt:

a) Bei der Definition eines neuen Begriffes $A$ durch vorher eingeführte Sachverhalte $B$ schreibt man
$A :\Leftrightarrow B$ (z. B.: reelle Zahl heißt gerade $:\Leftrightarrow$ durch 2 teilbar).

b) $a := b$ beschreibt die Einführung eines neuen Symbols $a$ mit Hilfe alter Symbole $b$ (z. B.: $x^2 := x \cdot x$)

c) $\square$ kennzeichnet das Ende eines Beweises
$\bullet$ wird in unserer Vorlesung eine Aufgabe hervorheben.

**Grundlegende Beweisprinzipien (logische Schlussweisen)** in der Mathematik sind:

1) $A \vee \neg A$ das **Prinzip vom ausgeschlossenen Dritten**

2) $\neg(A \wedge \neg A)$ die **Widerspruchsfreiheit**
(wir schreiben auch $\lightning$ (Widerspruch) für $A \wedge \neg A$)

3) $\neg\neg A \Leftrightarrow A$ die **Negation der Negation**

4) $\neg(A \vee B) = \neg A \wedge \neg B$
$\neg(A \wedge B) = \neg A \vee \neg B$
die **de Morganschen Regeln**

5) $(A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow (A \Rightarrow C)$ die **Transitivität**

6) $(A \vee B) \wedge C \Leftrightarrow (A \wedge C) \vee (B \wedge C)$
$(A \wedge B) \vee C \Leftrightarrow (A \vee C) \wedge (B \vee C)$
die **Distributivgesetze** der Aussagenlogik.

**Bemerkung.** Die hier aufgeführten Regeln sind überbestimmt, z. B. sind unter Voraussetzung von 4) die Prinzipien 1) und 2) äquivalent.

Schließlich erinnern wir an das **Prinzip der vollständigen Induktion**:

Für jedes natürliche $n$ sei $A_n$ eine Aussage. $A_n$ ist für alle $n$ richtig, falls

a) $A_1$ richtig ist und

b) für jedes $n$ die Aussage $A_{n+1}$ aus $A_n$ folgt
(in Symbolen: $\forall n(A_n \Rightarrow A_{n+1})$ oder $\bigwedge\limits_n (A_n \Rightarrow A_{n+1})$.

Würden nämlich a) und b) gelten und $A_n$ nicht für alle $n$ richtig sein, so gäbe es ein **kleinstes** $n_0$, für das $A_{n_0}$ falsch ist. D. h. $A_{n_0-1}$ wäre richtig und wegen b) dann aber auch $A_{n_0}$, was offensichtlich einen Widerspruch ergibt. Das Prinzip der Widerspruchsfreiheit führt uns zu dem Schluss, dass unsere Annahme falsch war, also das Induktionsprinzip richtig ist.

### 1.1.2 Mengen und Abbildungen

Der Mengenbegriff wird heute überall in der Mathematik gebraucht. Nach Georg Cantor (1845–1918) ist ein **Menge** "eine Zusammenfassung bestimmter wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens – welche die **Elemente** der Menge genannt werden – zu einem Ganzen". Bei dieser "naiven" Begriffsbildung werden wir es belassen. (Es gibt jedoch auch eine axiomatische Mengenlehre, bei der der Mengenbegriff mit Hilfe von Axiomen eingeführt wird. Den axiomatischen Zugang zu einer mathematischen Theorie werden wir im Rahmen der linearen Algebra kennen lernen.)

**Beispiele für Mengen**

1) **Zahlbereiche** (ausführlich wird darauf in der Vorlesung Analysis 1 eingegangen):
$\mathbb{N}$ bezeichnet die Menge der natürlichen Zahlen,
$\mathbb{Z}$ die Menge der ganzen Zahlen,
$\mathbb{Q}$ die Menge der rationalen Zahlen,
$\mathbb{R}$ die Menge der reellen Zahlen,
$\mathbb{C}$ die Menge der komplexen Zahlen.

2) $\emptyset$ steht für die **leere Menge**.

3) $\{a, b, c\}$ ist die Menge der 3 Symbole $a$, $b$, $c$.

Für eine Menge aus **endlich** oder **abzählbar unendlich vielen Elementen** schreibt man dann z. B. $\{x_1, \ldots, x_n\}$ bzw. $\{x_1, x_2, \ldots\}$. Insbesondere bezeichnet $\{x\}$ die **einelementige Menge** bestehend aus $x$.
Die Zugehöringkeit des Elements $x$ zur Menge $M$ in Symbolen lautet: $x \in M$ ($x$ ist Element von $M$, $x$ liegt in $M$). $x \notin M$ heißt "$x$ liegt nicht in $M$".



**Teilmengen** $N$ einer Menge $M$ werden oft durch Eigenschaften ihrer Elemente beschrieben:
$$N := \{x \in M \,:\, x \text{ hat Eigenschaft } E\}\,,$$
z. B., $\mathbb{N} = \{z \in \mathbb{Z} : z > 0\}$, d. h. hier wird $\mathbb{N}$ als Menge der positiven ganzen Zahlen aufgefasst. Man verwendet aber auch die Schreibweise

$$\{x :\ x \text{ hat Eigenschaft } E\},$$

falls keine Obermenge gegeben ist. (Anstelle des Doppelpunktes wird oft auch ein senkrechter Strich benutzt.)

**Mengenoperationen**
lassen sich mit Hilfe logischer Zeichen und Schlussweisen einführen:

1) **Inklusion**[*] $A \subset B :\Leftrightarrow (x \in A \Rightarrow x \in B)$

2) **Durchschnitt** $A \cap B := \{x : x \in A \text{ und } x \in B\} = \{x : x \in A,\ x \in B\}$
$(x \in A \cap B :\Leftrightarrow (x \in A) \wedge (x \in B))$

3) **Vereinigung** $A \cup B := \{x : x \in A \text{ oder } x \in B\}$
$(x \in A \cup B :\Leftrightarrow (x \in A) \vee (x \in B))$

4) **Differenz** $A \setminus B := \{x : x \in A,\ x \notin B\}$
$(x \in A \setminus B :\Leftrightarrow (x \in A) \wedge \neg(x \in B))$

5) **Komplement** bzgl. einer einheitlichen Obermenge $X$
$A^c := \{x \in X : x \notin A\}$
$(x \in A^c :\Leftrightarrow (x \in X) \wedge \neg(x \in A))$

Für diese Mengenoperationen gelten die
**de Morganschen Regeln**

$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$

und die **Distributivgesetze**

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

Sie folgen unmittelbar aus den entsprechenden Regeln der naiven Logik. •

[*]Für $\subset$ schreibt man oft auch $\subseteq$, und $\subset$ steht dann für die echte Inklusion, wo noch $A \neq B$ gilt.

Allgemein definiert man für eine Familie von Mengen $A_\alpha$ mit dem Index $\alpha$ aus einer beliebigen Indexmenge $I$ den Durchschnitt bzw. die Vereinigung als

$$\bigcap_{\alpha\in I} A_\alpha := \{x : x \in A_\alpha,\ \forall\alpha\in I\} = \{x : \bigwedge_{\alpha\in I}(x\in A_\alpha)\}$$
$$\bigcup_{\alpha\in I} A_\alpha := \{x : x \in A_\alpha \text{ für ein } \alpha\in I\} = \{x : \bigvee_{\alpha\in I}(x\in A_\alpha)\}.$$

Die Distributivgesetze sind dann analog zu den obigen.
Eine weitere Mengenoperation ist die Paarbildung:

6) **Kartesisches Produkt (Kreuzprodukt)**
$A\times B := \{(a,b) : a\in A,\ b\in B\}$
sowie die $n$-fache Verallgemeinerung

$$\begin{aligned} A_1\times\ldots\times A_n &:= \{(a_1,\ldots,a_n) : a_1\in A_1,\ldots,a_n\in A_n\}\\ &= \{(a_1,\ldots,a_n) : \bigwedge_{i=1}^n (a_i\in A_i)\}.\end{aligned}$$

## Abbildungen

Der in der Schule verwendete Funktionsbegriff läßt sich auf beliebige Mengen $X$ und $Y$ verallgemeinern:

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

Man benutzt die Kurzschreibweisen $f : X\to Y$ ($X$ heißt **Definitionsbereich** und $Y$ **Wertevorrat** von $f$) und $x\to f(x)$ ($f(x)$ heißt **Bild** des Elementes $x$).

**Definition.** Für $A\subset X$ heißt $f(A) := \{f(x) : x\in A\}$ **Bild(menge)** von $A$ unter $f$ und Bild $f := f(X)$ **Bild** von $f$.

**Definition.** Für $B\subset Y$ heißt
$f^{-1}(B) := \{x\in X : f(x)\in B\}$ **(volles) Urbild** von $B$ unter $f$.
$f^{-1}(y) := f^{-1}(\{y\})$ ist dann das volle Urbild des Elementes $y\in Y$.

**Beispiel.** Wir setzen $X := X_1\times X_2,\ Y := X_1,\ f := \pi_1$ für die **Projektion** $\pi_1$ auf die erste Komponente, d. h. $\pi_1 : X_1\times X_2\to X_1,\ \pi(x_1,x_2) = x_1$. Dann gilt für $B\subset X_1$ und $y\in Y_1$:

$$\begin{aligned}\pi_1^{-1}(B) &= B\times X_2\\ \pi_1^{-1}(y) &= \{y\}\times X_2\,.\end{aligned}$$



## Beziehungen zwischen Mengenoperationen und Abbildungen

Seien $(A_\alpha)_{\alpha\in I}$ eine Familie von Teilmengen von $X,\ f : X\to Y$ eine Abbildung und $(B_\alpha)_{\alpha\in I}$ eine Familie von Teilmengen von $Y$. Dann gilt

1) $f(\bigcup_{\alpha\in I} A_\alpha) = \bigcup_{\alpha\in I} f(A_\alpha)$

2) $f(\bigcap_{\alpha\in I} A_\alpha) \subset \bigcap_{\alpha\in I} f(A_\alpha)$
(im Allgemeinen aber nicht die Gleichheit!)

3) $f^{-1}(\bigcup_{\alpha\in I} B_\alpha) = \bigcup_{\alpha\in I} f^{-1}(B_\alpha)$

4) $f^{-1}(\bigcap_{\alpha\in I} B_\alpha) = \bigcap_{\alpha\in I} f^{-1}(B_\alpha)$.

*Beweis von 1).* Aus der Definition der Mengenvereinigung folgt unmittelbar

$$\begin{aligned} f\Big(\bigcup_{\alpha\in I} A_\alpha\Big) &= \Big\{f(x) : x\in\bigcup_{\alpha\in I} A_\alpha\Big\} = \Big\{f(x) : \bigvee_{\alpha\in I}(x\in A_\alpha)\Big\}\\ &= \bigcup_{\alpha\in I}\{f(x) : x\in A_\alpha\} = \bigcup_{\alpha\in I} f(A_\alpha).\end{aligned}$$

□

**Aufgabe.** Beweisen Sie 2) – 4) und finden Sie ein Gegenbeispiel für die Gleichheit in 2). •

Es folgen drei wichtige Typen von Abbildungen:

**Definition.** $f : X\to Y$ heißt
**injektiv** (eineindeutig) $:\Leftrightarrow (f(x_1) = f(x_2) \Rightarrow x_1 = x_2)$,
**surjektiv** (auf) $:\Leftrightarrow$ Bild $f = Y$
(d. h. $\forall y\in Y\ \exists x\in X$ mit $f(x) = y$), bzw.
**bijektiv** (umkehrbar) $:\Leftrightarrow f$ ist injektiv und surjektiv.

Man spricht von **Injektionen, Surjektionen** bzw. **Bijektionen**.

**Bemerkung.** $f$ ist genau dann eine Bijektion, wenn für jedes $y$ aus $Y$ ein eindeutig bestimmtes Urbild $x$ existiert, d. h. $f^{-1}(\{y\}) = \{x\}$ in der Mengensprache. Man schreibt dann aber $x = f^{-1}(y)$ , und die Zuordnung $f^{-1} : Y\to X$ mit $f^{-1}(y) = x$ wird **inverse Abbildung (Umkehrabbildung)** genannt.

**Definition.** Für $f : X \to Y$ und $g : Y \to Z$ wird die Abbildung $g \circ f : X \to Z$ mit

$$g \circ f(x) := g(f(x))$$

**Komposition** der Abbildungen $f$ und $g$ genannt.

Die identische Abbildung (**Identität**) von $X$ werden wir mit $\mathrm{Id}_X$ bezeichnen, d. h.

$$\mathrm{Id}_X : X \to X\ , \quad \mathrm{Id}_X(x) := x\ .$$

Für eine Bijektion $f : X \to Y$ gilt dann

$$f^{-1} \circ f = \mathrm{Id}_X$$
$$f \circ f^{-1} = \mathrm{Id}_Y\ .$$

**Aufgabe.** Die Komposition ist assoziativ, d. h.

$$h \circ (g \circ f) = (h \circ g) \circ f\ ,$$

wobei die Klammern die Reihenfolge der Ausführung (von rechts nach links) kennzeichnen. •

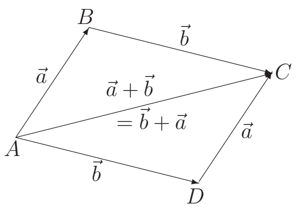
**Bemerkung.** Für den Spezialfall $X = Y = \mathbb{R}$ spricht man auch von (reellen) Funktionen. Im Rahmen der Analysis werden Sie viele Beispiele kennenlernen. Für die Lineare Algebra sind die linearen Funktionen und ihre natärliche Verallgemeinerung – die sogenannten linearen Abbildungen – von besonderer Bedeutung.

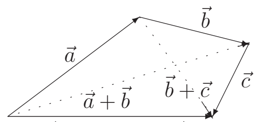
## 1.2 Einführung in die lineare Algebra

Zu Beginn wollen wir uns ein wichtiges Kapitel aus dem Schulstoff Mathematik ins Gedächtnis zurückrufen: die geometrische Vektorrechnung in der Ebene und im Raum. Vektoren werden dort als Pfeilklassen $\vec{a}$ eingeführt, deren Abtragung an einem Punkt $A$ jeweils einen eindeutig bestimmten Punkt $B$ ergibt. Der Pfeil $\overrightarrow{AB}$ heißt dann Repräsentant des Vektors $\vec{a}$. Mit Hilfe der Repräsentanten läßt sich über Parallelogrammbildung eine **Vektoraddition** einführen, die **kommutativ** und **assoziativ** ist:

$$\vec{a} + \vec{b} = \vec{b} + \vec{a}$$
$$(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c}).$$

Wir veranschaulichen uns dies an Spezialfällen:




$(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$



**Aufgabe.** Erläutern Sie diese Zeichnungen. •

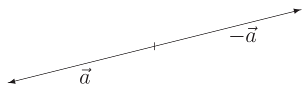
**Bemerkung.** Bei der eben gestellten Aufgabe sollte man sich noch einmal klarmachen, dass die Definition der Vektoraddition sowie die Rechenregeln unabhängig von der Wahl der Repräsentaten sind, d. h. es ist gleich, wo wir die Vektoren abtragen. Dies gilt auch für die nächsten Betrachtungen.

Für die Vektoraddition gibt es einen eindeutig bestimmten **Nullvektor** (neutrales Element) $\vec{0}$ mit

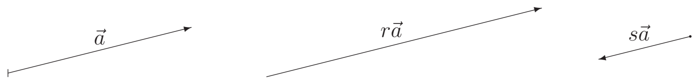
$$\vec{a} + \vec{0} = \vec{0} + \vec{a} = \vec{a}, \quad \forall \vec{a}.$$

Mit seiner Hilfe läßt sich dann zu jedem Vektor $\vec{a}$ ein eindeutig bestimmter **entgegengesetzter Vektor** (Gegenvektor) $-\vec{a}$ einführen:

$$\vec{a} + (-\vec{a}) = (-\vec{a}) + \vec{a} = \vec{0} \qquad \text{(s. Zeichung).}$$




Wir sehen, dass die geometrische Vektoraddition ähnliche Eigenschaften hat, wie die Addition von Zahlen innerhalb der einzelnen Zahlbereiche. Außerdem ist noch eine Operation erlaubt, die bei reellen und komplexen Zahlen ebenfalls verwendet wird: die **Multipikation** von Vektoren **mit reellen Zahlen**:




- In welchem Bereich liegen die reellen Zahlen $r$ und $s$ in der Zeichnung?

Hierbei gelten folgende Regeln:

| | |
|---|---|
| $1\vec{a} = \vec{a}$ | **Neutralität der Eins** |
| $(rs)\vec{a} = r(s\vec{a})$ | **Assoziativgesetz** |

Die Vektoraddition und die Multiplikation mit reellen Zahlen sind durch die **Distributivgesetze** miteinander verbunden.

$$(r + s)\vec{a} = r\vec{a} + s\vec{a}$$
$$r(\vec{a} + \vec{b}) = r\vec{a} + r\vec{b}.$$

Ein ganz direkter Zusammenhang zwischen Vektoren und reellen Zahlen ergibt sich durch die Koordinatendarstellung

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \qquad \text{in der Ebene}$$

$$\text{und} \qquad \vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \qquad \text{im Raum,} \quad a_i \in \mathbb{R}.$$

Damit läßt sich die geometrische Vektorrechung in eine analytische übersetzen, d. h. in das Rechnen mit Zahlen. Man begibt sich so in das Gebiet der **analytischen Geometrie**, wo die Vektorrechung erfolgreich Anwendung findet. Insbesondere gelingt als einfachster Spezialfall eine elegante Darstellung von **Geraden** und **Ebenen** im Raum mit Hilfe von Richtungsvektoren.
Die Geraden und Ebenen wiederum liefern eine geometrische Interpretation von **linearen Gleichungssystemen**, die in vielen mathematischen Gebieten von selbständigem Interesse sind. In engem Zusammenhang dazu stehen die **linearen Funktionen**, die in der Schule zumindest für eine Variable ausführlich behandelt werden.

Allen aufgezählten mathematischen Objekten liegt eine gemeinsame Idee zugrunde: das **Prinzip der Linearität**. Die allgemeine Mathematik dazu wurde zu Beginn des 20. Jahrhunderts entwickelt. Es zeigt sich, dass die linearen Abbildungen (Funktionen) bei allem eine tragende Rolle spielen. Da sich differenzierbare Funktionen lokal wie lineare verhalten, zieht sich das Prinzip der Linearität tief in die Differential– und Integralrechnung hinein. Es ist jedoch kein Allheilmittel, wie z. B. die Anforderungen der modernen Physik zeigen. Die **nichtlineare Analysis** und das allgemeinere **Prinzip der lokalen Skaleninvarianz** (fraktale Geometrie) gewinnen mehr und mehr an Bedeutung. Auf dem gegenwärtigen Entwicklungstand sthen lineare Modelle in der Praxis aber noch im Vordergrund.

Für viele Belange reichen zwei– und dreidimensionale reelle Koordinatenvektoren, die in der Schule eingeführt werden, nicht aus. Wir betrachen einige Beispiele:

1) Wie sie bereits wissen, ist die Vektorrechnung ein unentbehrliches mathematisches Hilfsmittel in der Physik.So werden beispielsweise die Bewegungsgesetze der klassischen Mechanik aus den sogenannten Lagrange–Gleichungen für die Hamilton–Funktion abgeleitet. In sie gehen sowohl die Ortsvektoren als auch die Impulsvektoren der physikalischen Körper ein. Man rechnet also gleichzeitig mit **6 Koordinaten**.

2) Bei der Beschreibung komplizierter technologischer Prozesse trifft man oft auf eine große Anzahl von Parametern. Hieraus ergibt sich eine praktische Notwendigkeit für das Rechnen mit **n Koordinaten**, wobei $n$ eine beliebige natürliche Zahl ist. Innermathematische Gründe dafür gab es schon wesentlich früher.



3) Nicht immer sind reelle Koordinatenvektoren zweckmäßig. In der Informatik, beispielsweise, wird die Signalübertragung durch **binäre Folgen** der Gestalt $(\varepsilon_1, \ldots, \varepsilon_n)$ mit $\varepsilon_i \in \{0, 1\}$ beschrieben. In der Koordinatenmenge $\{0, 1\}$ werden eine Addition und eine Multiplikation folgendermaßen eingeführt:

$$0 + 0 := 0\,, \quad 1 + 0 := 1\,, \quad 0 + 1 := 1\,, \quad 1 + 1 := 0;$$
$$0 \cdot 0 := 0\,, \quad 0 \cdot 1 := 0\,, \quad 1 \cdot 0 := 0\,, \quad 1 \cdot 1 := 1.$$

Anstelle reeller oder binärer Koordinaten verwendet man oft auch **komplexe Zahlen**, für die ebenfalls eine Addition und eine Multiplikation erklärt sind. Ein wesentliches Anwendungsgebiet hierfür ist wiederum die Physik, insbesondere die Elektrodynamik.

4) Zum Schluss sei noch einmal die Analysis erwähnt. Auf der Menge aller reellen Funktionen von einer reellen Variablen läßt sich eine ähnliche lineare Struktur einführen wir für geometrische Vektoren. Wir addieren zwei Funktionen $f_1$ und $f_2$ durch
$$(f_1 + f_2)(x) := f_1(x) + f_2(x)$$
Diese Addition ist kommutativ und assoziativ. Die Rolle des Nullvektors (neutrales Element) spielt hier die Funktion $f(x) := 0,\ \forall x$, und die zu einer Funktion $f$ entgegengesetzte Funktion $-f$ ist durch $(-f)(x) := -f(x),\ \forall x$, gegeben. Die Multiplikation mit reellen Zahlen $r$ erfolgt durch
$$(rf)(x) := rf(x),\ \forall x\,.$$
Offenbar ist dann $1f = f,\ (rs)f = r(sf),\ r, s \in \mathbb{R}$, und die Distributivgesetze übertragen sich ebenfalls von den reellen Zahlen auf die Funktionen. In diesem Sinne verhalten sich Funktionen wie geometrische Vektoren. Es wird Ihnen jedoch nicht gelingen erstere mit Hilfe von Koordinatenvektoren zu beschreiben.
- Was könnte dafür die Ursache sein ?

Unser Ziel ist nun die Erarbeitung eines allgemeinen mathematischen Zugangs zu Vektorräumen, der die erwähnten Beispiele und Problemstellungen umfasst. Wir führen dazu einen abstrakten Vektorbegriff zusammen mit einer Vektoraddition ein. Die Multiplikation mit reellen Zahlen wird zur Multiplikation mit Elementen eines algebraischen Körpers $K$ verallgemeinert. Für $K$ können wir uns z.B. $\mathbb{R}, \mathbb{Q}$ und $\mathbb{C}$, aber auch die binäre Menge $\{0, 1\}$ denken.

# 2 Lineare Räume

## 2.1 Gruppen und Körper

Um zu dem oben angedeuteten Vektorraumbegriff zu kommen, benötigen wir zwei algebraische Strukturen: die Gruppen und Körper. Hier werden sie nur die Rolle eines Hilfsmittels für die lineare Algebra spielen. Sie haben jedoch eine sehr große eigenständige Bedeutung für die gesamte höhere Algebra.

Ausgangspunkt für die Entwicklung der Gruppentheorie waren die Zahlbereiche mit ihren Additionseigenschaften. Anstelle der Addition setzt man allgemein eine abstrakte Verknüpfung, die im jeweiligen Beispiel konkretisiert wird.

**Definition.** Eine Menge $G$ heißt **Gruppe**, falls auf ihr eine binäre Operation $\circ$ (d.h. eine Verknüpfung von je zwei Elementen zu einem neuen) gegeben ist mit folgenden drei Eigenschaften:

($G$I) $(a \circ b) \circ c = a \circ (b \circ c) \quad , \quad a, b, c \in G$
(**Assoziativität**)

($G$II) $\exists e \in G$ mit $a \circ e = e \circ a = a \quad , \quad a \in G$
(**Existenz eines neutralen Elements**, auch **Einselement** genannt)

($G$III) $\forall a \in G \;\; \exists u \in G$ mit $a \circ u = u \circ a = e$
(**Existenz eines Umkehrelementes**)

Unmittelbar aus der Definition folgt:

**2.1.1 Satz.** *Einselement und Umkehrelement sind eindeutig bestimmt.*

*Beweis.* Sei $e'$ ein weiteres Einselement. Dann gilt

$$e' = e \circ e' = e \; ,$$

wobei ($G$II) zunächst linksseitig auf die Eins $e$ und dann rechtsseitig auf die Eins $e'$ angewendet wurde, d.h. $e' = e$.
Unter Benutzung von ($G$I) wird damit in ($G$III) für ein weiteres Umkehrelement $u'$ zu $a$

$$u' = u' \circ e = u' \circ (a \circ u) = (u' \circ a) \circ u = e \circ u = u \; ,$$

d.h. $u' = u$. □

**Bezeichnung.** Für das Umkehrelement zu $a$ schreiben wir wegen seiner Eindeutigkeit nunmehr im allgemeinen $a^{-1}$.

**Bemerkung.** Die Gruppenaxiome sind in der obigen Form eigentlich überbestimmt. Es genügt nämlich, die Existenz eines linken Einselementes und eines Linksinversen zu fordern, d.h. ($G$I)–($G$III) sind äquivalent zu:



($G$I) wie oben

($G$II$^l$) $\exists e \in G$ mit $e \circ a = a \quad , \quad a \in G$
(**Existenz eines linken Einselements**)

($G$III$^l$) $\forall a \in G \;\; \exists u \in G$ mit $u \circ a = e$
(**Existenz eines Linksinversen**)

*Beweis.* Zu einem linken Umkehrelement $u$ von $a$ existiert dann ebenfalls ein Linksinverses $u'$, d.h. $u' \circ u = e$. Damit erhalten wir aus ($G$I), ($G$II$^l$) und ($G$III$^l$)

$$\begin{aligned} a \circ u &= e \circ (a \circ u) = (u' \circ u) \circ (a \circ u) = u' \circ \big((u \circ a) \circ u\big) \\ &= u' \circ (e \circ u) = u' \circ u = e \end{aligned}$$

d.h. $a \circ u = e$ ($u$ ist auch Rechtsinverses zu $a$) und somit ($G$III).
($G$II) ergibt sich mit

$$a \circ e = a \circ (u \circ a) = (a \circ u) \circ a = e \circ a = a$$

d.h. $e$ ist auch rechtes Einselement. □

Die Rollen von "rechts" und "links" können wir dabei überall vertauschen. Ihre Unterscheidung ist aber notwendig, da wir von der Verknüpfung $\circ$ bisher nicht gefordert haben, dass sie kommutativ ist.

**2.1.2 Aufgabe.**

(1) $(a^{-1})^{-1} = a$

(2) $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$

(3) $\forall a, b \in G \quad \exists!\; x \in G$ mit $a \circ x = b$ ($\exists!\; y \in G$ mit $y \circ a = b$).
Geben Sie $x$ (bzw. $y$) explizit an! •

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

$$a \circ b = b \circ a \quad , \quad a, b \in G.$$

**Bemerkung.** Die Gruppenoperation $\circ$ läßt sich auf endlich viele Elemente ausdehnen: Wir definieren induktiv für $a_1, a_2, \ldots \in G$

$$a_1 \circ \ldots \circ a_n := (a_1 \circ \ldots \circ a_{n-1}) \circ a_n \, .$$

Wegen der Assoziativität kann man dann die Klammern beliebig setzen. Für kommutative Gruppen sind die $a_i$ auch beliebig vertauschbar.

**Beispiele für Gruppen.**

1. $G := \mathbb{R}$ mit der gewöhnlichen Addition + der reellen Zahlen als Verknüpfung. Die Zahl 0 spielt dabei die Rolle des neutralen Elementes $e$, und $-r$ ist das Umkehrelement zu $r \in \mathbb{R}$.
   Welche anderen Zahlbereiche bilden bzgl. der Addition eine Gruppe? •
   Für die endliche Addition schreibt man hier auch kurz
   $$a_1 + \ldots + a_n =: \sum_{i=1}^{n} a_n \,.$$
2. $G := \mathbb{R} \setminus \{0\}$ mit der gewöhnlichen Multiplikation als Verknüpfung. Hier fällt die reelle Zahl 1 mit dem Einselement $e$ zusammen, und $\frac{1}{r}$ ist das Umkehrelement zu $r \in \mathbb{R}$.
   Welche Zahlbereiche ohne die 0 besitzen die Gruppeneigenschaft bzgl. der Multiplikation? •
   Man setzt $a_1 \cdot \ldots \cdot a_n =: \prod_{i=1}^{n} a_i$.
3. $G :=$ Menge der geometrischen Vektoren in der Ebene mit der Vektoraddition als Verknüpfung. •
   Wieder schreibt man $\vec{v_1} + \ldots + \vec{v_n} =: \sum_{i=1}^{n} \vec{v_i}$.

Alle bisher genannten Gruppen sind abelsch.
Im Folgenden lernen wir ein Beispiel einer nicht kommutativen Gruppe ausführlich kennen, das eine wichtige Rolle bei der Behandlung von Determinanten spielen wird, nämlich die Permutationsgruppe der Ordnung $n \geq 3$. Zunächst sind wir jedoch etwas allgemeiner und betrachten die **Abbildungsgruppen**:

4. $G := S(X)$, **Menge der Bijektionen** eines Raumes $X$ mit der Abbildungskomposition $\circ$ als Verknüpfung. Die Rolle des neutralen Elements $e$ spielt hier die identische Abbildung $Id_X$, und das inverse Element zu $f \in G$ ist die Umkehrabbildung $f^{-1}$. $S(X)$ heißt auch **symmetrische Gruppe** von $X$.

**Aufgabe.** Wenn $X$ mehr als 2 Elemente enthält, so ist $S(X)$ nicht kommutativ. •

Wir untersuche nun den oben erwähnten Spezialfall.

5. Permutationsgruppen

**Definition.** Für $X := \{1, 2, \ldots, n\}$ heißt $S_n := S(X)$ **Permutationsgruppe** der Ordnung $n$.

**Aufgabe.** $S_n$ besteht aus $n!$ Elementen. •



Permutationen $\pi \in S_n$ beschreibt man auch durch ihre Wertetabellen
$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \pi(1) & \pi(2) & \ldots & \pi(n) \end{pmatrix}.$$
Die speziellen Permutationen
$$(i\,j) := \begin{pmatrix} 1 & \ldots & i & \ldots & j & \ldots & n \\ 1 & \ldots & j & \ldots & i & \ldots & n \end{pmatrix},$$
bei denen das $i$–te mit dem $j$–ten Element vertauscht wird, nennt man **Transpositionen** (oder **Vertauschungen**).
Offenbar läßt sich jede Permutaion durch Komposition von endlich vielen Transpositionen erzeugen.

**Definition.** Die **Charakteristik** einer Permutation $\pi$ ist die Anzahl $\chi(\pi)$ ihrer "Inversionen", d.h.
$$\chi(\pi) := \text{ Anzahl } \{(i,j) : 1 \leq i < j \leq n,\ \pi(i) > \pi(j)\},$$
und
$$sgn\,\pi := (-1)^{\chi(\pi)}$$
heißt **Vorzeichen** (oder **Signum**) von $\pi$.

**2.1.3 Satz.**

(i) *Jede Transposition $(i\,j)$ läßt sich als Komposition einer ungeraden Zahl von Transpositionen benachbarter Elemente darstellen.*

(ii) $sgn(i\,j) \circ \pi = -sgn\,\pi\,, \quad \pi \in S_n.$

(iii) $sgn(\sigma \circ \pi) = sgn\,\sigma \cdot sgn\,\pi\,, \quad \pi,\, \sigma \in S_n.$

*Beweis.* (i) Sei $r$ die Anzahl der Elemente zwischen $i$ und $j$ in der natürlichen Reihenfolge. Dann befördert man mit $r+1$ Nachbarschaftstranspositionen das Element $i$ auf den Platz von $j$ und dann mit $r$ solchen Vertauschungen $j$ auf den Ausgangsplatz von $i$. Folglich erhält man die Transposition $(i\ j)$ durch Komposition von $2r+1$ Vertauschungen benachbarter Elemente.

(ii) Für den Fall einer Nachbarschaftstransposition ändert sich die Zahl der Inversionen um 1, d.h.$\chi((i\ i+1) \circ \pi) = \chi(\pi) \pm 1$ und $sgn((i\ i+1) \circ \pi) = -sgn\,\pi$. Wegen (i) folgt die Behauptung durch Iteration für alle Transpositionen.

(iii) Als Spezialfall von (ii) für die Identität $\pi = Id_{S_n}$ ergibt sich $sgn(i\ j) = -1$. Da sich jede Permutation als Komposition von Transpositionen schreiben läßt, ist mit (ii) auch (iii) bewiesen.

□

**Bemerkung.** Die Gleichung (iii) widerspiegelt eine interessante algebraische Eigenschaft des Vorzeichens, aufgefaßt als Abbildung $sgn : S_n \to \{1, -1\}$. Die Bildmenge $\{1, -1\}$, versehen mit der gewöhnlichen Multiplikation von Zahlen, kann nämlich auch als Gruppe interpretiert werden. Dann besagt (iii), daß die Gruppenoperation $\circ$ von $S_n$ mit Hilfe des Signums in die Gruppenoperation $\cdot$ (mal) von $\{1, -1\}$ überführt wird. Eine solche Abbildung nennt man auch einen **Gruppenhomomorphismus**.

Nach diesem kleinen Ausflug kehren wir wieder zu den für die lineare Algebra wichtigen Grundstrukturen zurück. Wir betrachten im folgenden kommutative Gruppen, bei denen wir die Verknüpfung als Addition $+$ schreiben und für die noch eine weitere binäre Operation definiert ist, welche wir Multiplikation nennen.

**Definition.** Ein **Körper** $K$ ist eine Menge mit zwei binären Operationen $+$ (**Addition**) und $\cdot$ (**Multiplikation**), die folgenden Bedingungen genügen:

($K$I) $K$ ist bzgl. $+$ eine **kommutative Gruppe**. (Für das neutrale Element schreibt man 0 (Null), und $-\lambda$ bezeichnet das Umkehrelement zu $\lambda \in K$.)

($K$II) $\lambda \cdot (\mu + \nu) = \lambda \cdot \mu + \lambda \cdot \nu \quad , \quad \lambda, \mu, \nu \in K$
(**Distributivität**).

($K$III) Die Multiplikation ist **assoziativ** und **kommutativ**.

($K$IV) Es gibt ein **Einselement** 1 bzgl. der Multiplikation.

($K$V) Für jedes $\lambda \in K \setminus \{0\}$ gibt es ein **Umkehrelement** bzgl. der Multiplikation.

**Bemerkung.** Offenbar ist $K \setminus \{0\}$ bzgl. der Multiplikation eine kommutative Gruppe, und Einselement und Umkehrelement zu $\lambda$ sind eindeutig bestimmt. Für letzteres schreiben wir wieder $\lambda^{-1}$. Das Multiplikationszeichen $\cdot$ läßt man meist weg.

Wir schreiben

$$\sum_{i=1}^{k} \lambda_i := \lambda_1 + \ldots + \lambda_k$$

für die endliche Addition und

$$\prod_{i=1}^{k} \lambda_i := \lambda_1 \cdot \ldots \cdot \lambda_k$$

für die endliche Multiplikation sowie

$$\lambda^k := \prod_{i=1}^{k} \lambda \,.$$

Klassische Beispiele für Körper sind die reellen Zahlen $\mathbb{R}$ und die rationalen Zahlen $\mathbb{Q}$. Einen besonders einfachen Körper liefert die binäre Menge $\{0, 1\}$ mit der in der



Logik verwendeten Addition und Multiplikation (s. Einführung).

Überprüfen Sie dafür die Körperaxiome ($K$I)–($K$V). •

Ein weiteres wichtiges Beispiel erhalten wir durch die Erweiterung von $\mathbb{R}$ zu den **komplexen Zahlen** $\mathbb{C}$. Diese sind wie folgt definiert:

($\mathbb{C}$I) $\mathbb{C}$ ist ein Körper.

($\mathbb{C}$II) $\mathbb{C}$ enthält $\mathbb{R}$ als Teilkörper, d.h. die Einschränkung von Addition und Multiplikation von $\mathbb{C}$ auf die Teilmenge $\mathbb{R}$ fällt mit den entsprechenden Operationen auf den reellen Zahlen zusammen.

($\mathbb{C}$III) Es gibt eine komplexe Zahl $i$ mit $i^2 := i \cdot i = -1$.

($\mathbb{C}$IV) Jede komplexe Zahl läßt sich eindeutig in der Form

$$z = a + bi \text{ mit } a, b \in \mathbb{R}$$

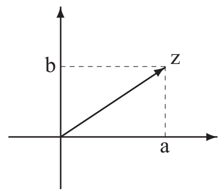
darstellen. ($a =: \operatorname{Re} z$ heißt **Realteil** und $b =: \operatorname{Im} z$ **Imaginärteil** von $z$.)

**Eigenschaften der komplexen Zahlen**

1. Die reellen Zahlen 0 und 1 sind auch Null– bzw. Einselement von $\mathbb{C}$.
2. Bei der Summierung zweier komplexer Zahlen addieren sich Realteile und Imaginärteile:

$$(a_1 + b_1 i) + (a_2 + b_2 i) = (a_1 + a_2) + (b_1 + b_2) i \,.$$

Die Zuordnumg $z \to (a, b)$ kann man deshalb als Koordinatendarstellung auffassen und $z$ als Vektor in der Ebene interpretieren:




Dabei bezeichnet man die horizontale Achse als **reelle Achse** und die vertikale als **imaginäre Achse**.

3. Für das Produkt $z = z_1 \cdot z_2$ von zwei komplexen Zahlen gilt

$$a + ib = (a_1 + ib_1)(a_2 + ib_2) = a_1 a_2 - b_1 b_2 + i(a_1 b_2 + a_2 b_1), \quad \text{d.h.}$$

$$a = a_1 a_2 - b_1 b_2 \text{ und } b = a_1 b_2 + a_2 b_1 .$$

4. Die zu $z = a + ib$ **konjugierte** komplexe Zahl ist definiert durch

$$\bar{z} := a - ib\,.$$

(Zeichnen sie $\bar{z}$ in der Koordinatendarstellung von $z$.) Dann gilt
$z\bar{z} = (a+ib)(a-ib) = a^2 + ib - ib - i^2b^2 = a^2 + b^2\,,$ d.h.

$$z\bar{z} = (\operatorname{Re} z)^2 + (\operatorname{Im} z)^2\,.$$

5. Die reelle Zahl

$$|z| := \sqrt{z\bar{z}}$$

heißt **Betrag** von $z$. Im Falle einer reellen Zahl $z$ stimmt sie mit dem alten Betrag überein. Ihre geometrische Interpretation in der Koordinatendarstellung ist die **Länge** des "Vektors" $z$. Der Betrag hat folgende Eigenschaften:

$|z| \geq 0\,,\ |z| = 0 \Leftrightarrow z = 0\,,\ |z_1 + z_2| \leq |z_1| + |z_2|$ sowie $|z_1z_2| = |z_1|\,|z_2|$. •

6. Die Inversenbildung bzgl. der Multiplikation kann mit den obigen Begriffen wie folgt beschrieben werden:

$$\frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{\bar{z}}{|z|^2} \quad , \quad z \neq 0.$$

(Wie bei den reellen Zahlen sind $\frac{1}{z}$ und $z^{-1}$ äquivalente Schreibweisen für das Umkehrelement von $z$.)



## 2.2 Der Begriff des Vektorraumes

Wir kommen jetzt zum wichtigsten Grundobjekt der linearen Algebra.

**Definition.** Ein **Vektorraum** (oder auch **linearer Raum**) $V$ über einem Körper $K$ ist eine Menge von Elementen (**Vektoren**), auf der eine **Addition** (+) von Vektoren und eine Multiplikation von Vektoren mit beliebigen Elementen von $K$ (**skalare Multiplikation**) erklärt sind mit folgenden Eigenschaften:

($V$I) $(V, +)$ ist eine kommutative Gruppe mit neutralem Element $\mathcal{O}$ (**Nullvektor**).

($V$II) Für die Multiplikation mit Elementen des Körpers gilt:

(1) $1v = v\,,\ v \in V$, wobei 1 das Einselement des Körpers ist.

(2) $(\lambda\mu)v = \lambda(\mu v)\,,\ \lambda, \mu \in K\,,\ v \in V\,,$
(**Assoziativgesetz**)

(3) $(\lambda + \mu)v = \lambda v + \mu v$
$\lambda(u + v) = \lambda u + \lambda v\,,\ \lambda, \mu \in K\,,\ u, v \in V$
(**Distributivgesetze**)

**Beispiele.**

1. Siehe Einführung.
2. Jeder Körper ist ein Vektorraum über sich selbst.
3. $V = \{\mathcal{O}\}$ trivialer Raum.
4. Die Menge $C(\mathbb{R})$ aller stetigen Abbildungen $f : \mathbb{R} \to \mathbb{R}$ mit der Addition

$$(f+g)(x) := f(x) + g(x)\,,\ f, g \in C(\mathbb{R})$$

und der skalaren Multiplikation

$$(\lambda f)(x) := \lambda f(x)\,,\ f \in C(\mathbb{R})\,,\ \lambda \in \mathbb{R}\,.$$

**Aufgabe.** Überprüfen Sie jeweils die Vektorraum–Axiome. (Insbesondere muß bei 4. gezeigt werden, daß die Funktionen $f + g$ und $\lambda f$ ebenfalls stetig sind.) •

### 2.2.1 Differenz von Vektoren

Mit Hilfe des Umkehrelements zu $v \in V$ bezüglich der Addition (wir schreiben hierfür wieder $-v$ und sprechen auch vom entgegengesetzten Vektor) lässt sich eine Vektorsubtraktion einführen:

$$u - v := u + (-v)\,, \quad u, v \in V\,.$$

(Analog sei

$$\lambda - \mu := \lambda + (-\mu)\,, \quad \lambda, \mu \in K\,.)$$

*Dann gelten die Distributivgesetze*

(1) $$\begin{aligned} \lambda(u-v) &= \lambda u - \lambda v\,, \quad u, v \in K\,,\, \lambda \in K \\ (\lambda-\mu)v &= \lambda v - \mu v\,, \quad v \in K\,,\, \lambda, \mu \in K. \end{aligned}$$

*Beweis.* Aus $(VII)$ (3) folgt $\lambda(u-v)+\lambda v = \lambda((u-v)+v) = \lambda(u+((-v)+v)) = \lambda u$. Die Addition von $-(\lambda v)$ auf beiden Seiten ergibt $\lambda(u-v) = \lambda u - \lambda v$.
Die zweite Gleichung zeigt man analog. □

(2) $$\forall(\lambda \in K \setminus \{0\}\,,\, u, v \in V)\ \exists!\, x \in V \text{ mit } \lambda x + u = v$$

Geben Sie den unbekannten Vektor $x$ explizit an. •

#### 2.2.2 Eigenschaften der skalaren Multiplikation

(1) $0v = \mathcal{O}\,,\ v \in V$
(*wobei* $0 \in K$, $\mathcal{O} \in V$ *die neutralen Elemente der jeweiligen Addition sind*).

*Beweis.* Wir setzen zunächst $u := 0v$ und erhalten wegen der Distributivität $u + u = 0v + 0v = 0v = u$, d.h. $u + u = u$. Die Subtraktion von $u$ auf beiden Seiten dieser Gleichung liefert $u = \mathcal{O}$. □

(2) $$\lambda\mathcal{O} = \mathcal{O}\,, \quad \lambda \in K.$$

Der Beweis ist ähnlich zum vorhergehenden. •

(3) $$v \in V \setminus \{\mathcal{O}\}\,, \quad \lambda, \mu \in K,\ \lambda \neq \mu \ \Rightarrow\ \lambda v \neq \mu v$$

*Beweis.* Wir nehmen an, daß $\lambda v = \mu v$, d.h. $\mathcal{O} = (\lambda - \mu)v$. Da $\lambda - \mu \neq 0$, existiert das Umkehrelement $(\lambda - \mu)^{-1}$. Wir multiplizieren beide Seiten der letzten Gleichung damit und erhalten $\mathcal{O} = 1v = v$, d.h. einen Widerspruch zur Annahme $v \neq \mathcal{O}$. □

Für den Fall, daß $K$ unendlich viele Elemente enthält und $V$ nicht nur aus dem Nullvektor besteht,können wir damit auf unendlich viele Elemente von $V$ schließen.

(4) $$(-\lambda)v = -\lambda v\,, \quad v \in V\,,\, \lambda \in K.$$ •



#### 2.2.3 Endliche Summen

Für $v_1, \ldots, v_n \in V$, $n \in \mathbb{N}$, schreiben wir wieder

$$\sum_{i=1}^{n} v_i := v_1 + \ldots + v_n\,.$$

Die Distributivität der Vektoroperationen überträgt sich unmittelbar auf diese endliche Addition (die wieder assoziativ und kommutativ ist).

### 2.3 Linearkombination und lineare Unabhängigkeit

Die unter 2.2.3 genannten Beziehungen lassen folgende Begriffsbildung zu:

**Definition.** Ein $v \in V$ ist **Linearkombination** von $a_1, \ldots, a_n \in V$, falls Elemente $\lambda_1, \ldots, \lambda_n$ des zugehörigen Körpers $K$ existieren derart, daß

$$v = \sum_{i=1}^{n} \lambda_i a_i\,.$$

Die Linearkombination $\sum\limits_{i=1}^{n} \lambda_i a_i$ heißt **nichttrivial**, falls eines der $\lambda_i$ verschieden von 0 ist.
Die Vektoren $a_1, \ldots, a_n \in V$ heißen **linear abhängig**, falls eine nichttriviale Linearkombination existiert mit $\sum\limits_{i=1}^{n} \lambda_i a_i = \mathcal{O}$. Ansonsten heißen sie **linear unabhängig**.

**Beispiel.** Angewendet auf die Menge der geometrischen Vektoren in der Ebene, bedeutet dies folgende Kette von äquivalenten Aussagen: $\vec{a}, \vec{b}$ sind *linear abhängig*

$$\begin{aligned} &\Leftrightarrow \exists r, s \in \mathbb{R} \text{ mit } r \neq 0 \quad (\text{oder } s \neq 0), \text{ so dass } r\vec{a} + s\vec{b} = \vec{0} \\ &\Leftrightarrow \vec{a} = (-r^{-1}s)\vec{b} \quad \text{für gewisse } r, s \in \mathbb{R} \text{ mit } r \neq 0 \end{aligned}$$

$\Leftrightarrow$ $\vec{a}$ und $\vec{b}$ sind *kollinear*, d.h., ein Vektor ist ein Vielfaches des anderen.

Welche geometrische Deutung gibt es für die lineare Abhängigkeit von drei geometrischen Vektoren im Raum? •

**Folgerungen.**

**2.3.1**
*Wenn ein endliches Vektorsystem ein linear abhängiges Teilsystem enthält, so ist es selbst linear abhängig.*
*Alle Teilsysteme von endlichen linear unabhängigen Vektorsystemen sind linear unabhängig.*

### 2.3.2

*$a_1, \ldots, a_n \in V$ sind genau dann linear abhängig, wenn eines der $a_j$ als Linearkombination aller übrigen darstellbar ist.*

*Beweis.* $a_1, \ldots, a_n$ sind linear abhängig

$\Leftrightarrow \exists \lambda_1, \ldots, \lambda_n \in K$ mit $\lambda_j \neq 0$ für ein $j$, so dass $\sum_{i=1}^{n} \lambda_i a_i = \mathcal{O}$

$\Leftrightarrow a_j = \lambda_j^{-1}(-\sum_{i \neq j} \lambda_i a_i) = \sum_{i \neq j} (-\lambda_j^{-1}\lambda_i) a_i = \sum_{i \neq j} \mu_i a_i$ für gewisse $\mu_i \in K$

(wir setzen $\mu_i := -\lambda_j^{-1}\lambda_i$ und die Summation läuft über alle $i$ verschieden von $j$)

$\Leftrightarrow a_j$ ist Linearkombination der restlichen $a_i$. □

### 2.3.3

*Wenn $a_1, \ldots, a_n$ linear unabhängig sind und $a_1, \ldots, a_n, a_{n+1}$ linear abhängig , so folgt die Darstellbarkeit $a_{n+1} = \sum_{i=1}^{n} \mu_i a_i$ für gewisse $\mu_i \in K$.*

*Beweis.* Sei $\sum_{i=1}^{n+1} \lambda_i a_i = \mathcal{O}$ für eine nichttriviale Linearkombination. Dann ist $\lambda_{n+1} \neq 0$, da ansonsten $a_1, \ldots, a_n$ linear abhängig wären. Demzufolge gilt:

$$a_{n+1} = \sum_{i=1}^{n} (-\lambda_{n+1}^{-1}\lambda_i) a_i.$$

Mit der Bezeichnung $\mu_i = -\lambda_{n+1}^{-1}\lambda_i$ ergibt sich die Behauptung. □

Wir kommen nun zu einer natürlichen Verallgemeinerung der Begriffsbildung:
Eine Zusammenfassung $S$ von Elementen von $V$, wo jeder Vektor mehrfach vorkommen kann, wollen wir **System** von Vektoren aus $V$ nennen. (Wir haben oben bereits von endlichen Vektorsystemen gesprochen.)

**Definition.** $S$ heißt **linear abhängiges System**, falls es ein endliches Teilsystem von linear abhängigen Vektoren enthält, und ansonsten **linear unabhängiges System**.

Demnach ist $S$ genau dann linear unabhängig, wenn jedes endliche Teilsystem linear unabhängig ist. (Insbesonders ist $S$ dann eine Teil**menge** von $V$, d.h., kein Vektor kommt mehrfach vor.)



## 2.4 Basis, Koordinaten, Austauschsatz und Dimension

Für geometrische Vektoren in der Ebene und im Raum wurden die Koordinatenvektoren mit Hilfe der Zerlegung nach einer “Basis” eingeführt. In unseren abstrakten Vektorräumen $V$ benutzen wir folgenden allgemeinen Begriff:

**Definition.** $B \subset V$ heißt **Basis** von $V$, falls $B$ linear unabhängig ist und jeder Vektor $v \in V$ als Linearkombination von Elementen von $B$ dargestellt werden kann.

Unmittelbar aus der Definition ergibt sich folgende **Charakterisierung** von Basen:

### 2.4.1

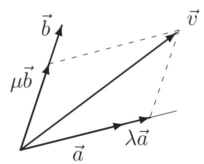
*Eine Teilmenge $B \subset V$ ist genau dann Basis, wenn sie linear unabhängig und für jedes $v \in V$ das System $\{B, v\}$ linear abhängig ist.*

*Beweis.* Aus der Basiseigenschaft erhalten wir für $v \in V$ die Darstellung

$$v = \sum_{i=1}^{n} \lambda_i b_i$$

für ein $n \in \mathbb{N}$ und gewisse $b_i \in B$, $\lambda_i \in K$. Deshalb bilden $\{b_1, \ldots, b_n, v\}$ ein linear abhängiges System. Daß die im Kriterium formulierte Bedingung auch hinreichend ist, schließt man aus 2.3.3. □

**Beispiel.** Wir erinnern uns, daß zwei geometrische Vektoren $\vec{a}, \vec{b}$ in der Ebene genau dann linear unabhängig sind, wenn sie nicht kollinear sind. Jeder andere Vektor $\vec{v}$ läßt sich nach ihnen zerlegen:




$$\vec{v} = \lambda\vec{a} + \mu\vec{b}$$

Demnach bildet jedes Paar nichtkollinearer Vektoren eine Basis, und jede Basis für die geometrischen Vektoren der Ebene hat diese Gestalt.

**Aufgabe.** Formulieren Sie das entsprechende Resultat für den Raum. •

**2.4.2 Satz.** *In jedem Vektorraum existiert eine Basis.*

Diese Aussage ist von grundlegender Bedeutung für alles weitere. Ihr Beweis stützt sich auf das sogenannte Zornsche Lemma, welches auf die axiomatische Begründung der Mathematik zurückgeht (Auswahlaxiom). Da uns hier die entsprechenden Hilfsmittel nicht zur Verfügung stehen, wollen wir auf die Literatur verweisen.

**2.4.3. Eindeutigkeit der Basisdarstellung von Vektoren**

*Die Darstellung eines Vektors als Linearkombination von Basisvektoren ist eindeutig.*

*Beweis.* Sei $\sum\limits_{i=1}^{k} \lambda_i b_i = v = \sum\limits_{i=1}^{k+l} \mu_i b_i$ für gewisse $b_i \in B,\ \lambda_i,\ \mu_i \in K$. Dann gilt

$$\sum_{i=1}^{k} (\mu_i - \lambda_i) b_i + \sum_{i=k+1}^{k+l} \mu_i b_i = \mathcal{O}.$$

Da die $b_1, \ldots, b_{k+l}$ linear unabhängig sind, erhalten wir $\mu_i = \lambda_i$ für $i = 1, \ldots, k$ und $\mu_i = 0$ für $i > k$. D.h., die Koeffizienten in den beiden Basisdarstellungen stimmen überein. □

**Definition.** Die Koeffizienten $\lambda_i, \ldots, \lambda_n$ in der Darstellung $v = \sum\limits_{i=1}^{n} \lambda_i b_i$ heißen **Koordinaten** des Vektors $v$ bzgl. der Basis $B$.

Um die Struktur linearer Räume besser verstehen zu können, werden wir nun ein wichtiges Hilfsmittel herleiten.

**2.4.4. Austauschsatz** (Steinitz)

*Sei $B$ eine Basis. Dann gibt es für jedes System $\{a_1, \ldots, a_k\}$ linear unabhängiger Vektoren Basiselemente $b_1, \ldots, b_k$ derart, daß $(B \setminus \{b_1, \ldots, b_k\}) \cup \{a_1, \ldots, a_k\}$ wieder eine Basis bildet.*

*Beweis.* Wir benutzen die vollständige Induktion nach $k$. Für $k = 0$ ist die Aussage trivialerweise richtig. (Es wird nichts ausgetauscht.) Für $k = l$ sei nun $(B \setminus \{b_1, \ldots, b_l\}) \cup \{a_1, \ldots, a_l\} =: B'$ eine Basis.
Wir suchen einen Vektor $b_{l+1} \in B$, der sich durch $a_{l+1}$ ersetzen läßt. Zunächst gilt die Darstellung

(1) $$a_{l+1} = \sum_{i=1}^{l} \lambda_i a_i + \sum_{i=l+1}^{l+m} \mu_i b_i$$

für gewisse $b_i \in B \setminus \{b_1, \ldots, b_l\}$.Eines der $\mu_i$ muß verschieden von 0 sein, da sonst die $a_1, \ldots, a_{l+1}$ linear abhängig wären. Durch Umnumerierung können wir erreichen, daß dies $\mu_{l+1}$ ist. Wir zeigen, daß dann $b_{l+1}$ den gewünschten Bedingungen genügt. Zunächst stellen wir (1) nach $b_{l+1}$ um:

(2) $$b_{l+1} = \sum_{i=1}^{l} (-\mu_{l+1}^{-1} \lambda_i) a_i + \mu_{l+1}^{-1} a_{l+1} + \sum_{i=l+2}^{l+m} (-\mu_{l+1}^{-1} \mu_i) b_i.$$

Sei $B'' = (B \setminus \{b_1, \ldots, b_{l+1}\}) \cup \{a_1, \ldots, a_{l+1}\}$.



$B''$ ist *linear unabhängig*: Falls nämlich

(3) $$\sum_{i=1}^{l+1} \nu_i a_i + \sum_{i=l+2}^{l+p} \nu_i b_i = \mathcal{O}$$

für gewisse $b_i \in B \setminus \{b_1, \ldots, b_{l+1}\}$, so erhalten wir mit (1)

$$\sum_{i=1}^{l} \nu_i a_i + \sum_{i=1}^{l} \nu_{l+1} \lambda_i a_i + \sum_{i=l+1}^{l+m} \nu_{l+1} \mu_i b_i + \sum_{i=l+2}^{l+p} \nu_i b_i = \mathcal{O}$$

und nach Zusammenfassen der Koeffizienten bei gleichen Vektoren

$$\sum_{i=1}^{l} (\nu_i + \nu_{l+1} \lambda_i) a_i + \nu_{l+1} \mu_{l+1} b_{l+1} + \sum_{i=l+2}^{l+\max(m,p)} \omega_i b_i = \mathcal{O}$$

für gewisse $\omega_i \in K$. Da $B'$ Basis ist, folgt $\nu_i + \nu_{l+1} \lambda_i = 0\,,\ i = 1, \ldots, l$, und $\nu_{l+1} \mu_{l+1} = 0$. Nach Konstruktion ist $\mu_{l+1} \neq 0$, d.h. $\nu_{l+1} = 0$ und deshalb auch $\nu_i = 0\,,\ i = 1, \ldots, l$. Demnach verschwindet in (3) die erste Summe. Die lineare Unabhängigkeit der $b_i$ ergibt nun $\nu_i = 0$ für $i = l+2, \ldots, l+p$, d.h. alle Koeffizienten in (3) sind $0$.

Für jedes $v \in V$ können wir eine Darstellung nach der Basis $B'$ angeben:

$$v = \sum_{i=1}^{l} \alpha_i a_i + \sum_{i=l+1}^{l+q} \alpha_i b_i\,.$$

Die Substitution von $b_{l+1}$ durch den Ausdruck in (2) führt zu

$$\begin{aligned} v = {} & \sum_{i=1}^{l} \alpha_i a_i + \sum_{i=1}^{l} \alpha_{l+1} (-\mu_{l+1}^{-1} \lambda_i) a_i + \alpha_l \mu_{l+1}^{-1} a_{l+1} \\ & + \sum_{i=l+2}^{l+m} \alpha_{l+1} (-\mu_{l+1}^{-1} \mu_i) b_i + \sum_{i=l+2}^{l+q} \alpha_i b_i\,. \end{aligned}$$

Nach Zusammenfassung der Koeffizienten bei gleichen Vektoren liefert dies eine *Darstellung* von $v$ nach $B''$.
Wir haben somit die Basiseigenschaften für $B''$ gezeigt. □

**2.4.5. Folgerung.**

(i) Jedes System $\{a_1, \ldots, a_k\}$ von linear unabhängigen Vektoren in $V$ läßt sich zu einer Basis ergänzen (**Basisergänzungssatz**).

(ii) Sind $\{a_1, \ldots, a_n\}$ und $B$ Basen von $V$, so ist die Anzahl der Elemente von $B$ gleich $n$.

*Beweis.* (i) **Aufgabe.** •

(ii) Nach dem Austauschsatz existieren Elemente $b_1, \ldots, b_n$ von $B$, so daß $\{a_1, \ldots, a_n\} \cup (B \setminus \{b_1, \ldots, b_n\})$ ebenfalls eine Basis ist . Aus der Charakterisierung 2.4.1 einer Basis schließen wir auf $B \setminus \{b_1 \ldots, b_n\} = \emptyset$, da das obenstehende System sonst linear abhängig wäre.

□

Diese Eigenschaft rechtfertigt den folgenden *Dimensionsbegriff*:

**Definition.** Ein Vektorraum $V$ heißt **n–dimensional** $(\dim V = n)$, falls in ihm eine Basis aus $n$ Elementen existiert.

Aufgrund von 2.4.5 ist die Definition korrekt, jede Basis in $V$ besteht dann aus $n$ Elementen.

**Beispiele**

1. Der Raum der geometrischen Vektoren in der Ebene (im Raum) hat die Dimension 2 (bzw. 3).

2. Sei $K^n := \{(\lambda_1, \ldots, \lambda_n) : \lambda_i \in K\} = K \times \ldots \times K$
   (oder $K^n := \left\{ \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} : \lambda_i \in K \right\}$) der Vektorraum der **Zeilen-** (bzw. **Spalten**)-**vektoren** über dem Körper $K$. Die Vektoraddition ist dabei komponentenweise durch
   $$(\lambda_1, \ldots, \lambda_n) + (\mu_1, \ldots, \mu_n) := (\lambda_1 + \mu_1, \ldots, \lambda_n + \mu_n)$$
   gegeben und die skalare Multiplikation durch
   $$\mu(\lambda_1, \ldots, \lambda_n) := (\mu\lambda_1, \ldots, \mu\lambda_n).$$
   Das System
   $$e_1 := (1, 0, \ldots, 0),\ e_2 := (0, 1, 0, \ldots, 0), \ldots, e_n := (0, \ldots, 0, 1)$$
   bildet die sogenannte **kanonische** (oder **Standard**) **Basis** in $K^n$. (Analoges gilt für die Spaltenvektoren.) Folglich ist $\dim K^n = n$.

Wir wollen nun zeigen, dass $\boldsymbol{K^n}$ **als Prototyp eines $\boldsymbol{n}$–dimensionalen Vektorraumes über $\boldsymbol{K}$** aufgefasst werden kann. Dazu benötigen wir den Begriff der **Isomorphie** von Vektorenräumen.

**Definition.** Zwei Vektorräume $V$ und $W$ über $K$ heißen **isomorph**, falls es eine Bijektion $\mathbf{F} : V \to W$ gibt, die linear ist, d.h.

$$\begin{aligned} \mathbf{F}(u + v) &= \mathbf{F}(u) + \mathbf{F}(v) &&, u, v \in V \\ \mathbf{F}(\lambda v) &= \lambda \mathbf{F}(v) &&, v \in V\ ,\ \lambda \in K. \end{aligned}$$



Nach dieser Definition wird also die lineare Struktur von $V$ auf die in $W$ übertragen. Die Umkehrung gilt ebenfalls, denn die inverse Abbildung $\mathbf{F}^{-1} : W \to V$ ist auch linear: Für $w_1, w_2 \in W$ gilt nämlich

$$\mathbf{F}(\mathbf{F}^{-1}(w_1) + \mathbf{F}^{-1}(w_2)) = \mathbf{F}(\mathbf{F}^{-1}(w_1)) + \mathbf{F}(\mathbf{F}^{-1}(w_2)) = w_1 + w_2\ ,$$

so daß
$$\mathbf{F}^{-1}(w_1) + \mathbf{F}^{-1}(w_2) = \mathbf{F}^{-1}(w_1 + w_2).$$

Analog zeigt man $\lambda \mathbf{F}^{-1}(w) = \mathbf{F}^{-1}(\lambda w)\ ,\ \lambda \in K\ ,\ w \in W$. Demnach überträgt sich auch die lineare Struktur von $W$ auf die von $V$.
Wir kommen nun zu der erwähnten Bedeutung von $K^n$.

**2.4.6 Satz.** *Jeder $n$–dimensionale Vektorraum $V_n$ über $K$ ist isomorph zu $K^n$.*

*Beweis.* Für eine beliebige Basis $b_1, \ldots, b_n$ in $V_n$ definieren wir $\mathbf{F} : V_n \to K^n$ durch

$$\mathbf{F}(v) = (\lambda_1, \ldots, \lambda_n)\ ,\ \text{falls } v = \sum_{i=1}^{n} \lambda_i b_i\ ,\ \text{ d.h.}$$

$V$ wird auf seinen Koordinatenvektor abgebildet. Die Injektivität von $\mathbf{F}$ folgt sofort aus der Definition. Außerdem liegt für jedes $(\lambda_1, \ldots, \lambda_n) \in K^n$ der Vektor $\sum\limits_{i=1}^{n} \lambda_i b_i$ in $V_n$, so daß $\mathbf{F}$ auch surjektiv ist. Wir überprüfen nun die Linearität: Sei $u = \sum\limits_{i=1}^{n} \lambda_i b_i\ ,\ v = \sum\limits_{i=1}^{n} \mu_i b_i$. Dann gilt für $\nu \in K$

$$\nu u + v = \sum_{i=1}^{n} (\nu\lambda_i + \mu_i) b_i \ \text{ und folglich}$$
$$\begin{aligned} \mathbf{F}(\nu u + v) &= (\nu\lambda_1 + \mu_1, \ldots, \nu\lambda_n + \mu_n) \\ &= \nu(\lambda_1, \ldots, \lambda_n) + (\mu_1, \ldots, \mu_n) = \nu\mathbf{F}(u) + \mathbf{F}(v). \end{aligned}$$

Für $\nu = 1$ erhält man die Additivität von $\mathbf{F}$ und für $v = \mathcal{O}$ die Homogenität bzgl. der skalaren Multiplikation. □

**Bemerkung.** $K^n$ heißt deshalb auch **Koordinatenraum** zu $V_n$, und $V_n$ wird oft mit $K^n$ identifiziert.

**2.4.7. Folgerung.** *Zwei endlichdimensionale Vektorräume über $K$ sind genau dann isomorph, wenn sie die gleiche Dimension besitzen.* •

## 2.5 Lineare Unterräume, Durchschnitt, lineare Hülle und Rang eines Vektorsystems

**Definition.** Eine nichtleere Teilmenge $V'$ eines Vektorraums $V$ heißt (**linearer**) **Unterraum**, falls

$(U1)$ $u', v' \in V' \Rightarrow u' + v' \in V'$

$(U2)$ $v' \in V$ , $\lambda \in K \Rightarrow \lambda v' \in V'$

**2.5.1 Satz.** *Jeder lineare Unterraum $V'$ von $V$ ist bzgl. der gegebenen Operationen wieder ein Vektorraum.*

*Beweis.* Wir überprüfen das Axiom ($V'$I): Die Assoziativität imf Kommutativität der Addition sind in $V$ gegeben, also auch in $V'$. Da insbesondere $0v' \in V'$, so enthält $V'$ das neutrale Element $\mathcal{O}$. Außerdem ist $-v' = (-1)v' \in V$, so daß mit jedem Vektor $v'$ auch sein entgegengesetzter Vektor $-v'$ in $V'$ liegt, d.h. $V'$ ist bzgl. $+$ eine kommutative Gruppe. Alle Bedingungen von ($V'$II) kann man als Spezialfälle von ($V$II) auffassen. □

**Beispiele**

1. $\{\mathcal{O}\}$ ist trivialer Unterraum jedes Vektorraums $V$.
2. $V$ sei der Raum der geometrischen Vektoren in der Ebene (im Raum) und $V'$ die Menge der Vektoren, die kollinear (komplanar) zu einer Geraden (Ebene) sind.
3. $V = K^n$ , $V' = \{(\lambda_1, \ldots, \lambda_n) \in K^n : \lambda_1 = \ldots = \lambda_m = 0\}$, $m < n$.

**Definition.** Wenn $\mathcal{U}$ ein beliebiges System von Unterräumen von $V$ ist, so wird $\bigcap\limits_{U\in\mathcal{U}} U$ der **Durchschnittsraum** dieser Unterräume genannt, falls diese Menge nicht leer ist.

**2.5.2 Satz.** *Der Durchschnitt von Unterräumen ist wieder ein Unterraum.*

*Beweis.* Wenn $u, v \in \bigcap\limits_{U\in\mathcal{U}} U$, so folgt $u, v \in U$ für jeden Unterraum $U$ aus dem System $\mathcal{U}$. Dann ist aber wegen der Unterraumstruktur von $U$ auch $u + v \in U$. Da dies für alle $U$ gilt, können wir auf $u + v \in \bigcap\limits_{U\in\mathcal{U}} U$ schließen. Analog zeigt man, dass mit $u$ auch $\lambda u$ im Durchschnitt liegt. □

Eine wichtige Operation zur Vektorraumerzeugung ist folgende.

**Definition.** Sei $S$ eine beliebige Teilmenge des Vektorraums $V$. Dann heißt die daraus erzeugte Teilmenge von $V$

$$L(S) := \left\{ \sum_{i=1}^{k} \lambda_i v_i : v_i \in S \ , \ \lambda_i \in K \ , \ k \in \mathbb{N} \right\}$$

**lineare Hülle** von $S$.



Unmittelbar aus der Definition schlußfolgert man, daß $L(S)$ ein linearer Unterraum von $V$ ist. Mehr noch, man kann $L(S)$ als den kleinsten Unterraum betrachten, der $S$ enthält:

**2.5.3 Satz.**
$$L(S) = \bigcap \{U : U \textit{ ist Unterraum} \ , S \subset U\}$$

*Beweis.* Aufgrund der Linearität enthält jeder Unterraum mit $S$ auch $L(S)$. Da $L(S)$ selbst Unterraum ist, folgt die Behauptung. □

**2.5.4. Folgerung.** *$L(S) = S$ gilt genau dann, wenn $S$ selbst ein Unterraum von $V$ ist.*

**Beispiel.** Für die geometrischen Vektoren im Raum besteht die lineare Hülle $L(\{\vec{u}, \vec{v}\})$ zweier nicht kollinearer Vektoren aus dem Unterraum der Vektoren, die komplanar zu einer von $\vec{u}, \vec{v}$ aufgespannten Ebene sind.

Wir betrachten nun ein endliches Vektorsystem.

**Definition.** Die maximale Anzahl linear unabhängiger Vektoren von $\{a_1, \ldots, a_k\}$ heißt **Rang** des Vektorsystems.

Dafür werden wir die Schreibweise Rang$\{a_1, \ldots, a_k\}$ benutzen. Der Rang besitzt folgende algebraische Interpretation.

**2.5.5 Satz.**
$$\text{Rang}\{a_1, \ldots, a_k\} = \dim L(\{a_1, \ldots, a_k\}).$$

*Beweis.* Sei Rang$\{a_1, \ldots, a_k\} = r$. Dann können wir – bis auf Permutation der Indizes – schließen, daß $\{a_1, \ldots, a_r\}$ linear unabhängig und für jedes $j > r$ das System $\{a_1, \ldots, a_r, a_j\}$ linear abhängig ist. Nach 2.3.3 gilt dann $a_j \in L(\{a_1, \ldots, a_r\})$ auch für $j > r$. Folglich sind alle Linearkombinationen der $a_1, \ldots, a_k$ bereits in $L(\{a_1, \ldots, a_r\})$ enthalten, so daß

$$L(\{a_1, \ldots, a_k\}) = L(\{a_1, \ldots, a_r\}),$$

d.h. die $a_1, \ldots, a_r$ bilden eine Basis für $L(\{a_1, \ldots, a_k\})$. □

## 2.6 Direkte Summe von Unterräumen und Dimensionssatz

Neben dem Durchschnitt und der linearen Hülle kann man Unterräume auch durch Summenbildung erzeugen. Seien $V_1, \ldots, V_k$ Unterräume von $V$.

**Definition.** Die Menge

$$\sum_{i=1}^{k} V_i = V_1 + \ldots + V_k := \{v_1 + \ldots + v_k : v_i \in V_i \ , \ i = 1, \ldots, k\}$$

heißt **Summe der Vektorräume**.

Man überlegt sich leicht, daß diese Summe wieder ein Unterraum ist. •

**Definition.** Die Summe $V_1 + \ldots + V_k$ heißt **direkte**, falls für jedes $v \in V_1 + \ldots + V_k$ die Darstellung $v = v_1 + \ldots + v_k$ mit $v_i \in V_i$, $i = 1, \ldots, k$, eindeutig ist.

Für die direkte Summe der Vektorräume schreibt man dann $V_1 \oplus \ldots \oplus V_k$.

**Beispiel.** Sei $\dim V = n$ und $b_1, \ldots, b_n$ eine beliebige Basis in $V$. Dann läßt sich $V$ darstellen als

$$V = L(\{b_1, \ldots, b_m\}) \oplus L(\{b_{m+1}, \ldots, b_n\}) \,,\; m < n \,,$$

oder auch als

$$V = L(\{b_1\}) \oplus \ldots \oplus L(\{b_n\}).$$

Allgemein kann man direkte Summen folgendermaßen charakterisieren.

**2.6.1 Satz.** *$V_1 + \ldots + V_k$ ist genau dann direkte Summe, wenn für jedes $i$ der Durchschnitt von $V_i$ und $V_1 + \ldots + V_{i-1} + V_{i+1} + \ldots + V_k$ trivial ist.*

*Beweis.* Sei $V_1 + \ldots + V_k$ direkte Summe und $V_i \cap \sum_{j\neq i} V_j \neq \{\mathcal{O}\}$. Dann gibt es ein $v_i \in V \setminus \{\mathcal{O}\}$, so daß $v_i = \sum_{j\neq i} v_j$ für gewisse $v_j \in V_j$ und somit

$$v_i + \sum_{j\neq i}(-v_j) = \mathcal{O} = \mathcal{O} + \ldots + \mathcal{O}.$$

Dies steht aber im Widerspruch zur Eindeutigkeit der Darstellung des Nullvektors bzgl. $V_1 \oplus \ldots \oplus V_k$. Deshalb folgt $V_i \cap \sum_{j\neq i} V_j = \{\mathcal{O}\}$.

Wenn wir nun die letzte Beziehung voraussetzen und $v_1 + \ldots + v_k = v_1' + \ldots + v_k'$ mit $v_j, v_j' \in V_j$, $j = 1, \ldots, k$, und $v_i \neq v_i'$ für ein $i$ annehmen, so erhalten wir in $\mathcal{O} \neq v_i - v_i' = \sum_{j\neq i}(v_j' - v_j)$ einen Vektor aus $V_i \cap \sum_{j\neq i} V_j$, d.h. einen Widerspruch. Also muß $v_i = v_i'$ für alle $i$ gelten. □

Wenn $V$ und $W$ Unterräume endlicher Dimension sind, kann man folgenden **Dimensionssatz** zeigen.

**2.6.2 Satz.**

$$\dim(V + W) = \dim V + \dim W - \dim(V \cap W)$$

*Beweis.* Sei $\dim V = m$, $\dim W = n$ und $\dim(V \cap W) = p$. Aus dem Basisergänzungssatz folgt, daß man jede Basis $\{u_1, \ldots, u_p\}$ in $V \cap W$ (die leere Menge, falls $V \cap W = \{\mathcal{O}\}$) zu einer Basis $\{u_1, \ldots, u_p, v_{p+1}, \ldots, v_m\}$ in $V$ bzw. $\{u_1, \ldots, u_p, w_{p+1}, \ldots, w_n\}$ in $W$ ergänzen kann. Es genügt nun zu zeigen, daß

$$\{u_1, \ldots, u_p, v_{p+1}, \ldots, v_m, w_{p+1}, \ldots, w_n\}$$

eine Basis in $V + W$ ist, denn dann gilt $\dim(V + W) = m + n - p$.



1. *Lineare Unabhängigkeit:* Es sei

$$\sum_{i=1}^{p} \lambda_i\, u_i + \sum_{i=p+1}^{m} \mu_i v_i + \sum_{i=p+1}^{n} \nu_i\, w_i = \mathcal{O}.$$

Dann ist

$$x := \sum_{i=p+1}^{m} \mu_i v_i = \sum_{i=1}^{p}(-\lambda_i)u_i + \sum_{i=p+1}^{n}(-\nu_i)w_i$$

ein Vektor aus $V \cap W$ und läßt sich nach der Basis $\{u_1, \ldots, u_p\}$ zerlegen: $x = \sum_{i=1}^{p} \lambda_i' u_i$. Es folgt

$$\sum_{i=1}^{p}(\lambda_i + \lambda_i')u_i + \sum_{i=p+1}^{n} \nu_i w_i = \mathcal{O}.$$

Die lineare Unabhängigkeit von $\{u_1, \ldots, u_p, w_{p+1}, \ldots, w_n\}$ ergibt insbesondere $\nu_i = 0$, $i = p+1, \ldots, n$. Wir setzen dies in die erste Gleichung ein und erhalten $\sum_{i=1}^{p} \lambda_i u_i + \sum_{i=p+1}^{m} \mu_i v_i = \mathcal{O}$. Da $\{u_1, \ldots, u_p, v_{p+1}, \ldots, v_m\}$ ebenfalls linear unabhängig ist, verschwinden alle Koeffizienten.

2. *Basisdarstellung:* Sei $v = \sum_{i=1}^{p} \lambda_i u_i + \sum_{i=p+1}^{m} \mu_i v_i \in V$ und $w = \sum_{i=1}^{p} \lambda_i' u_i + \sum_{i=p+1}^{n} \nu_i w_i \in W$. Dann ergibt sich die Basiszerlegung von $v + w \in V + W$ durch

$$v + w = \sum_{i=1}^{p}(\lambda_i + \lambda_i')u_i + \sum_{i=p+1}^{m} \mu_i v_i + \sum_{i=p+1}^{n} \nu_i w_i.$$ □

# 3 Lineare Abbildungen

## 3.1 Koordinatentransformation und quadratische Matrizen

In Abschnitt 2.4 haben wir einen Isomorphismus $\mathbb{F}$ zwischen $V_n$ und $K^n$ konstruiert, der jedem Vektor aus $V_n$ seinen Koordinatenvektor bezüglich einer festen Basis zuordnet. Wir wollen nun den Übergang von einer Basis zu einer anderen als Koordinatentransformation darstellen. Seien $\{a_1, \ldots, a_n\}$ und $\{b_1, \ldots, b_n\}$ Basen in $V_n$ und $(\alpha_1, \ldots, \alpha_n)$ bzw. $(\beta_1, \ldots, \beta_n)$ die zugehörigen Koordinaten eines Vektors $v$, d.h.

(1) $$v = \sum_{i=1}^{n} \alpha_i \, a_i$$

(2) $$v = \sum_{j=1}^{n} \beta_j \, b_j \, .$$

Jeder Vektor $b_j$ der "neuen" Basis besitzt bzgl. der "alten" Basis Koordinaten, sagen wir $(\gamma_{1j}, \ldots, \gamma_{nj})$, d.h.

(3) $$b_j = \sum_{i=1}^{n} \gamma_{ij} \, a_i \, .$$

Wir setzen (3) in (2) ein und erhalten

$$v = \sum_{j=1}^{n} \beta_j \sum_{i=1}^{n} \gamma_{ij} \, a_i = \sum_{i=1}^{n} \left( \sum_{j=1}^{n} \gamma_{ij} \, \beta_j \right) a_i \, .$$

Da die Koordinaten von $v$ bzgl. der alten Basis eindeutig bestimmt sind, folgt durch Koordinatenvergleich

(4) $$\boxed{\alpha_i = \sum_{j=1}^{n} \gamma_{ij} \, \beta_j} \quad , \qquad i = 1, \ldots, n \, .$$

Durch diese **Koordinatentransformation** werden die alten Koordinaten mit Hilfe der neuen ausgedrückt.

Für das Gleichungssystem (4) führt man nun eine kompakte und nützliche Schreibweise ein. Zunächst ordnet man die Koordinaten $\gamma_{ij}$ in einer sogenannten **quadratischen $n \times n$–Matrix** an:

(5) $$C := \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \vdots & & & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} \end{pmatrix} .$$



Dabei gibt der Index $i$ von $\gamma_{ij}$ die Nummer der Zeile und $j$ die Nummer der Spalte von $C$ an, in der $\gamma_{ij}$ steht.

Für die Koordinatenvektoren als Spalten schreiben wir allgemein

$$\bar{\lambda} := \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in K^n \, .$$

Die Gleichungen (4) sind dann äquivalent zu

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1n} \\ \vdots & & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} ,$$

in Kurzschreibweise,

(6) $$\boxed{\bar{\alpha} = C\bar{\beta}} \, ,$$

wobei auf der rechten Seite das **Produkt einer $n \times n$–Matrix mit einem Spaltenvektor** der Länge $n$ steht. Dieses ist per **Definition** gerade der Spaltenvektor $\bar{\alpha}$, dessen Komponenten $\alpha_i$ durch die Vorschrift (4) gegeben sind. (Die Definition dieses Produkts gilt für jede Wahl der $n \times n$–Matrix $C$ und des Spaltenvektors $\bar{\beta}$.)

Die Transformationsmatrix $C$ induziert eine Abbildung $\mathbf{C} : K^n \to K^n$ durch $\mathbf{C}(\bar{\lambda}) = C\bar{\lambda}$. Wir zeigen, daß $\mathbf{C}$ linear ist. Es gilt

$$\mathbf{C}(\bar{x} + \bar{\mu}) = C(\bar{\lambda} + \bar{\mu}) = \begin{pmatrix} \vdots \\ \sum\limits_{j=1}^{n} \gamma_{ij} \, (\lambda_j + \mu_j) \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ \sum\limits_{j=1}^{n} \gamma_{ij}\lambda_j + \sum\limits_{j=1}^{n} \gamma_{ij}\mu_j \\ \vdots \end{pmatrix}$$

$$= \begin{pmatrix} \vdots \\ \sum\limits_{j=1}^{n} \gamma_{ij}\lambda_j \\ \vdots \end{pmatrix} + \begin{pmatrix} \vdots \\ \sum\limits_{j=1}^{n} \gamma_{ij}\mu_j \\ \vdots \end{pmatrix} = C\bar{\lambda} + C\bar{\mu} = \mathbf{C}(\bar{\lambda}) + \mathbf{C}(\bar{\mu})$$

und analogerweise $\mathbf{C}(\mu\bar{\lambda}) = \mu\mathbf{C}(\bar{\lambda})$ , $\mu \in K$.

Aus allgemeineren Prinzipien können wir bald ableiten, daß $\mathbf{C}$ auch bijektiv ist. Wir haben es also wieder mit einem Vektorraumisomorphismus zu tun, der hier **Automorphismus** genannt wird, da Ausgangsraum und Bildraum übereinstimmen.

## 3.2 Allgemeine lineare Abbildungen

Wir betrachten nun beliebige **lineare Abbildungen** $\mathbf{A}$ zwischen zwei Vektorräumen $V$ und $W$ über $K$, d.h. $\mathbf{A}$ besitzt die Eigenschaften

$$\begin{aligned} \mathbf{A}(u+v) &= \mathbf{A}(u) + \mathbf{A}(v) \quad , \quad u, v \in V \\ \mathbf{A}(\lambda v) &= \lambda\, \mathbf{A}(v) \quad , \quad v \in V\ ,\ \lambda \in K. \end{aligned}$$

Diese werden auch **lineare Operatoren** oder **Vektorraumhomomorphismen** genannt. In Zukunft schreiben wir kurz $\mathbf{A}v := \mathbf{A}(v)$. Ehe wir die linearen Abbildungen in ihrer Wechselwirkung mit den zugehörigen Vektorräumen genauer untersuchen, wollen wir noch einige Beispiele aufzählen.

1. **Nullabbildung:** $\mathbf{O} : V \to W$ mit $\mathbf{O}v = \mathcal{O}$
2. **Identische Abbildung:** $\mathbf{I} : V \to V$ mit $\mathbf{I}v := v$
3. **Lineare Funktionale** oder **Linearformen:** Jede lineare Abbildung von $V$ nach $K$.
4. **Koordinatenabbildung:** $\mathbf{F} : V_n \to K^n$ aus Abschnitt 2.4
5. **Koordinatentransformation:** $\mathbf{C} : K^n \to K^n$ aus Abschnitt 3.1
6. **Projektionen:** Sei $V = V_1 \oplus \ldots \oplus V_k$. Dann heißt $\mathbf{\Pi}_i : V \to V_i$ mit $\mathbf{\Pi}_i v = v_i$, falls $v = v_1 + \ldots + v_k$, Projektion von $V$ auf den Unterraum $V_i$ entlang dem Unterraum $V_1 \oplus \ldots \oplus V_{i-1} \oplus V_{i+1} \oplus \ldots \oplus V_k$.

**Aufgabe.** Zeigen Sie, daß die Projektionen $\mathbf{\Pi}_i$ sämtlich linear sind. Geben Sie eine geometrische Veranschaulichung im Raum der geometrischen Vektoren an. •

Der Zusammenhang zwischen der linearen Struktur der Räume und der linearen Abbildungen kommt in den folgenden Sachverhalten noch deutlicher zum Ausdruck.

**3.2.1 Satz.**

(1) $$\mathbf{A}(L(S)) = L(\mathbf{A}(S)) \quad , \quad S \subset V$$

(2) $$\mathbf{A}^{-1}(L(S')) \supset L(\mathbf{A}^{-1}(S')) \quad , \quad S' \subset W$$

*Beweis.*

(1) Jedes $w \in \mathbf{A}(L(S))$ ist darstellbar als

$$w = \mathbf{A}\left(\sum_{i=1}^{k} \lambda_i v_i\right)$$

für gewisse $v_i \in S$ , $\lambda_i \in K$. Die Linearität von $\mathbf{A}$ impliziert

$$w = \sum_{i=1}^{k} \lambda_i\, \mathbf{A} v_i\ ,$$

was offenbar ein Element von $L(\mathbf{A}(S))$ ist. Die Umkehrung dieser logischen Kette ist ebenfalls richtig.



(2) $v \in L(\mathbf{A}^{-1}(S))$ heißt, daß $v = \sum_{i=1}^{l} \mu_i v_i$ für gewisse $v_i \in \mathbf{A}^{-1}(S'), \mu_i \in K$ . Es gilt also $\mathbf{A}\, v_i \in S'$ , $i = 1, \ldots, l$. Wegen der Linearität von $\mathbf{A}$ ist

$$\mathbf{A}\, v = \sum_{i=1}^{l} \mu_i \mathbf{A} v_i \in L(S') \ \text{, d.h. } v \in \mathbf{A}^{-1}(L(S'))\ .$$

□

**Aufgabe.** Weisen Sie anhand eines Gegenbeispiels nach, daß in (2) die Mengengleichheit nicht gilt. (* Für welche $S'$ läßt sich diese erreichen?)

**3.2.2 Folgerung.** $V'$ ist Unterraum von $V \Rightarrow \mathbf{A}(V')$ ist Unterraum von $W$. •

**3.2.3 Aufgabe.** $W'$ ist Unterrraum von $W \Rightarrow \mathbf{A}^{-1}(W')$ ist leer oder Unterraum von $V$. •

## 3.3 Bild, Kern, Rang und Defekt linearer Abbildungen

Wie oben sei $\mathbf{A} : V \to W$ linear.

**Definition.**
Bild $\mathbf{A} = \mathbf{A}(V)$ heißt **Bildraum** von $\mathbf{A}$.
Kern $\mathbf{A} := \mathbf{A}^{-1}(\{\mathcal{O}\})$ heißt **Kern** von $\mathbf{A}$.

Aufgrund der Aussagen 3.2.2 und 3.2.3 sind Bild $\mathbf{A}$ und Kern $\mathbf{A}$ Unterräume von $W$ bzw. $V$. Für ihre Dimensionen gibt es spezielle Begriffe:

**Definition.**
Rang $\mathbf{A} := \dim$(Bild $\mathbf{A}$) heißt **Rang** von $\mathbf{A}$.
Defekt $\mathbf{A} := \dim$(Kern $\mathbf{A}$) heißt **Defekt** von $\mathbf{A}$.

**Beispiele.**

1. $\mathbf{O} : V \to W$ , Bild $\mathbf{O} = \{\mathcal{O}\}$ , Rang $\mathbf{O} = 0$
   Kern $\mathbf{O} = V$ , Defekt $\mathbf{O} = \dim V$
2. $\mathbf{I} : V \to V$ , Bild $\mathbf{I} = V$ , Rang $\mathbf{I} = \dim V$
   Kern $\mathbf{I} = \{\mathcal{O}\}$ , Defekt $\mathbf{I} = 0$
3. $V = V_1 \oplus V_2$ , $\mathbf{\Pi}_i : V \to V_i$ Projektion auf $V_i$ entlang $V_j$ , $i, j = 1, 2$ , $i \neq j$
   Bild $\mathbf{\Pi}_i = V_i$ , Rang $\mathbf{\Pi}_i = \dim V_i$
   Kern $\mathbf{\Pi}_i = V_j$ , Defekt $\mathbf{\Pi}_i = \dim V_j$

**3.3.1 Satz.** *Für jede lineare Abbildung* $\mathbf{A} : V \to W$ *sind die folgenden drei Aussagen äquivalent:*

(1) $\mathbf{A}$ *ist injektiv*

(2) Defekt $\mathbf{A} = 0$

(3) $S \subset V$ *ist linear unabhängig* $\Rightarrow$ $\mathbf{A}(S)$ *ist linear unabhängig*

*Gilt außerdem* $\dim V < \infty$*, so ist*

(4) Rang $\mathbf{A} = \dim V$

*äquivalent zu* (1), (2), (3).

*Beweis.*

Aus (1) folgt (2): Wenn $\mathbf{A}$ injektiv ist, so gilt insbesondere Kern $\mathbf{A} = \{\mathcal{O}\}$, d.h. Defekt $\mathbf{A} = 0$.

Aus (2) folgt (1): Defekt $\mathbf{A} = 0$ bedeutet Kern $\mathbf{A} = \{\mathcal{O}\}$. Falls $\mathbf{A}v = \mathbf{A}v'$, so erhält man $\mathbf{A}(v - v') = \mathcal{O}$, d.h. $v - v' \in$ Kern $\mathbf{A}$ und demnach $v - v' = \mathcal{O}$.

Aus (2) folgt (3): Sei $\sum\limits_{i=1}^{k} \lambda_i \mathbf{A} v_i = \mathcal{O}$ für gewisse $v_i \in S$.

Dann ist $\mathbf{A}\left(\sum\limits_{i=1}^{k} \lambda_i v_i\right) = \mathcal{O}$ und demnach $\sum\limits_{i=1}^{k} \lambda_i v_i \in$ Kern $\mathbf{A} = \{\mathcal{O}\}$, d.h. $\sum\limits_{i=1}^{k} \lambda_i v_i = \mathcal{O}$. Die lineare Unabhängigkeit von $S$ ergibt $\lambda_1 = \ldots = \lambda_k = 0$.

Aus (3) folgt (2): Falls $\mathbf{A}v = \mathcal{O}$ für ein $v \neq \mathcal{O}$, so wird das linear unabhängige System $S = \{v\}$ in das linear abhängige System $\{\mathcal{O}\}$ überführt, was ein Widerspruch zu (3) ist.

Sei nun $\dim(V) = n$.

Aus (3) folgt dann (4): Für jede Basis $b_1, \ldots, b_n$ in $V$ sind die $\mathbf{A}b_1, \ldots, \mathbf{A}b_n$ linear unabhängig. Nach Satz 3.2.1 (1) gilt $\mathbf{A}(V) = L(\{\mathbf{A}b_1, \ldots, \mathbf{A}b_n\})$, so dass $\mathbf{A}b_1, \ldots, \mathbf{A}b_n$ Basis in $\mathbf{A}(V)$ ist, d.h. $\dim \mathbf{A}(V) = n$.



Aus (4) folgt (3): Wir nehmen an, daß es ein linear unabhängiges System $S = \{a_1, \ldots, a_k\} \subset V$ gibt, für das $\{\mathbf{A}a_1, \ldots, \mathbf{A}a_k\}$ linear abhängig wird. In diesem Fall sei o. E. $\mathbf{A}a_1 \in L(\{\mathbf{A}a_2, \ldots, \mathbf{A}a_k\})$. Wir ergänzen $a_1, \ldots, a_k$ zu einer Basis $a_1, \ldots, a_n$ in $V$ und erhalten

$$\begin{aligned}\mathbf{A}(V) &= \mathbf{A}(L(\{a_1, \ldots, a_n\})) = L(\{\mathbf{A}a_1, \ldots, \mathbf{A}a_n\}) \\ &= L(\{\mathbf{A}a_2, \ldots, \mathbf{A}a_n\}).\end{aligned}$$

Dies ergibt $\dim \mathbf{A}(V) \leq n - 1$, d.h. einen Widerspruch zu (4). Also ist (3) richtig.

□

**3.3.2 Folgerung.**

(i) Ein Vektorraumisomorphismus $\mathbf{A} : V_n \to W_n$ überträgt jede Basis in $V_n$ in eine Basis von $W_n$.

(ii) Wenn eine lineare Abbildung $\mathbf{A} : V_n \to W_n$ irgendeine Basis von $V_n$ in eine Basis von $W_n$ überträgt, so ist $\mathbf{A}$ ein Vektorraumisomorphismus. •

**3.3.3 Folgerung.** Die Koordinatentransformation $\mathbf{C} : K^n \to K^n$ ist ein Vektorraumisomorphismus.

*Beweis.* In 3.3.1 haben wir die Linearität der Abbildung $\mathbf{C}\bar{\lambda} = C \cdot \bar{\lambda}, \bar{\lambda} \in K^n$, bewiesen. Wegen 3.3.2 genügt es nun zu zeigen, dass $\mathbf{C}$ die Standardbasis $\bar{e_1}, \ldots, \bar{e_n}$ von $K^n$ in eine neue Basis überträgt. Es gilt aber

$$C \cdot \bar{e_k} = \bar{\gamma_k}$$

wobei $\bar{\gamma_k}$ der $k$-te Spaltenvektor der Koordinatentransformationsmatrix $C$ ist. Nach Konstruktion von $C$ sind die $\bar{\gamma_1}, \ldots, \bar{\gamma_n}$ als Koordinatenvektoren von Basisvektoren im gegebenen Vektorraum $V_n$ immer Basisvektoren in $K^n$. □

Wir können nun den **Dimensionssatz für lineare Abbildungen** zeigen.

**3.3.4 Satz.** *Sei* $\dim V < \infty$*. Dann gilt*

$$\text{Rang } \mathbf{A} + \text{Defekt } \mathbf{A} = \dim V\,.$$

*Beweis.*

1. Fall: Defekt $\mathbf{A} = 0$.
Nach dem obigen Satz ist dann Rang $\mathbf{A} = \dim V$.

2. Fall: Rang $\mathbf{A} = 0$.
Dann ist $\mathbf{A}(V) = \{\mathcal{O}\}$, d.h. Kern $\mathbf{A} = V$ und deshalb Defekt $\mathbf{A} = \dim V$.

3. Fall: Defekt $\mathbf{A} \neq 0$, Rang $\mathbf{A} \neq 0$.
Wir wählen eine Basis $b_1, \ldots, b_k$ in Kern $\mathbf{A}$, ergänzen sie zu einer Basis $b_1, \ldots, b_n$ in $V$ und zeigen, daß Rang $\mathbf{A} = n - k$, indem wir nachweisen, dass $\mathbf{A}b_{k+1}, \ldots, \mathbf{A}b_n$ eine Basis in $\mathbf{A}(V)$ ist. Sei $\sum\limits_{i=k+1}^{n} \lambda_i \mathbf{A} b_i = \mathcal{O}$, d.h. $\mathbf{A}\big(\sum\limits_{i=k+1}^{n} \lambda_i\, b_i\big) = \mathcal{O}$, also $\sum\limits_{i=k+1}^{n} \lambda_i\, b_i \in$ Kern $\mathbf{A}$. Dann läßt sich dieser Vektor nach der Basis $b_1, \ldots, b_k$ zerlegen: $\sum\limits_{i=k+1}^{n} \lambda\, b_i = \sum\limits_{i=1}^{k} \lambda\, b_i$. Da dann $\sum\limits_{i=1}^{k} \lambda_i\, b_i + \sum\limits_{i=k+1}^{n} (-\lambda_i) b_i = \mathcal{O}$ gilt, sind alle $\lambda_i$ gleich Null. Insbesondere ist die obenstehende Linearkombination der $\mathbf{A}b_{k+1}, \ldots, \mathbf{A}b_n$ trivial, d.h. diese Vektoren sind linear unabhängig. Außerdem erzeugen sie $\mathbf{A}(V)$, denn jedes $v \in V$ läßt sich darstellen als $v = \sum\limits_{i=1}^{n} \mu_i\, b_i$, so dass

$$\mathbf{A}v = \sum_{i=1}^{n} \mu_i\, \mathbf{A}b_i = \sum_{i=k+1}^{n} \mu_i\, \mathbf{A}b_i, \quad \text{da } \mathbf{A}b_i = \mathcal{O} \text{ für } i \leq k.$$

□

## 3.4 Abbildungsoperationen und Matrizendarstellung

Wir wollen nun der Menge aller linearen Abbildungen von $V$ nach $W$ selbst wieder die Struktur eines Vektorraumes geben. Dazu führen wir eine Addition und eine skalare Multiplikation ein. Je zwei solchen Abbildungen $\mathbf{A}$ und $\mathbf{B}$ ordnen wir ihre **Summe** $\mathbf{A} + \mathbf{B}$ als neue Abbildung von $V$ nach $W$ zu:

$$(\mathbf{A} + \mathbf{B})v := \mathbf{A}v + \mathbf{B}v\,.$$

Analog entsteht das **skalare Vielfache** $\lambda\mathbf{A}$ von $\mathbf{A}$:

$$(\lambda\mathbf{A})v := \lambda\mathbf{A}v$$

**Aufgabe.** $\mathbf{A} + \mathbf{B}$ und $\lambda\mathbf{A}$ sind wieder linear, und es gilt die folgende Aussage. •

**3.4.1 Satz.** *Die Menge $L(V, W)$ aller linearen Abbildungen von $V$ nach $W$ bildet bezüglich der Abbildungsaddition und der skalaren Multiplikation einen Vektorraum über $K$.*

Die Komposition von linearen Abbildungen spielt ebenfalls eine wichtige Rolle.

**3.4.2 Satz.** *Es sei $\mathbf{A} \in L(U, V)$ , $\mathbf{B} \in L(V, W)$. Dann ist $\mathbf{B} \circ \mathbf{A} \in L(U, W)$, und es gelten das Assoziativgesetz*

$$\mathbf{A} \circ (\mathbf{B} \circ \mathbf{C}) = (\mathbf{A} \circ \mathbf{B}) \circ \mathbf{C}$$



*und die Distributivgesetze*

$$\begin{aligned}(\mathbf{A} + \mathbf{B}) \circ \mathbf{C} &= \mathbf{A} \circ \mathbf{B} + \mathbf{B} \circ \mathbf{C}\\ \mathbf{A} \circ (\mathbf{B} + \mathbf{C}) &= \mathbf{A} \circ \mathbf{B} + \mathbf{A} \circ \mathbf{C}\\ \lambda(\mathbf{A} \circ \mathbf{B}) &= (\lambda\mathbf{A}) \circ \mathbf{B} = \mathbf{A} \circ (\lambda\mathbf{B})\,.\end{aligned}$$ •

**Matrizendarstellung linearer Abbildungen**

Wir wissen bereits, daß sich Vektoren mit Hilfe ihrer Koordinaten bezüglich einer Basis beschreiben lassen. Ähnlich kann man bei linearen Abbildungen vorgehen, die ja selbst als Vektoren von $L(V, W)$ betrachtet werden dürfen. Im Fall von endlich–dimensionalen $V$ und $W$ ordnet man die Koordinaten aus Zweckmäßigkeitsgründen hier jedoch matrizenförmig an. Wir wollen diesen Zusammenhang jetzt herleiten. Zunächst machen wir uns den folgenden Sachverhalt klar.

**3.4.3 Lemma.** *Jede lineare Abbildung ist durch ihre Werte für die Elemente einer Basis eindeutig bestimmt.*

*Beweis.* Sei $\mathbf{A} \in L(V, W)$ und $B$ Basis in $V$. Dann hat jedes $v \in V$ die eindeutige Darstellung $v = \sum\limits_{j=1}^{k} \lambda_j b_j$ für gewisse $b_j \in B$, so daß $\mathbf{A}v = \sum\limits_{j=1}^{k} \lambda_j\, \mathbf{A}b_j$. Es genügt also, die Werte $\mathbf{A}b$ , $b \in B$, zu kennen. □

**Bemerkung.** Im Sinne einer suggestiven Schreibweise bezeichnen wir die Koordinaten eines Vektors $v$ bzgl. einer gegebenen Basis jetzt auch mit $v_i$ anstelle von $\lambda_i$ usw. $\bar{v}$ ist dann der Koordinatenvektor zu $v$ .

Im weiteren setzen wir voraus $\dim V = m$ und $\dim W = n$. Jede Basis in $V$ hat dann die Gestalt $B = \{b_1, \ldots, b_m\}$. Die Bildvektoren $\mathbf{A}b_1, \ldots, \mathbf{A}b_m$ lassen sich nach einer zweiten Basis $\{e_1, \ldots, e_n\}$ in $W$ zerlegen. Ihre zugehörigen Koordinatenvektoren seien $\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1m} \\ \vdots \\ a_{nm} \end{pmatrix}$. Wir schreiben dafür auch kurz $\bar{a}_1, \ldots, \bar{a}_m$, d.h. $\bar{a}_j = \overline{\mathbf{A}b_j}$. Wir betrachten nun einen beliebigen Vektor $v \in V$ mit dem Koordinatenvektor $\bar{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$ bzgl. $\{b_1, \ldots, b_m\}$ und sein Bild $\mathbf{A}v =: w$. Der Koordinatenvektor von $w$ bzgl. $\{e_1, \ldots, e_n\}$ sei $\bar{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$. Wegen der Linearität von $\mathbf{A}$ gilt wieder $w = \sum\limits_{j=1}^{m} v_j \mathbf{A} b_j$, und die Linearität der Koordinatendarstellung $\mathbf{F}$ aus 2.4 ergibt

$$\bar{w} = \sum_{j=1}^{m} v_j\, \bar{a}_j\,.$$

Koordinatenweise aufgeschrieben, erhält man die Gleichungen

(1) $$\boxed{w_i = \sum_{j=1}^{m} a_{ij}\, v_j} \;,\quad i = 1, \ldots, n\,.$$

Ähnlich wie bei der Koordinatentransformation in 3.1 wollen wir dafür wieder eine kompakte Schreibweise einführen. Wir bilden die sogenannte $\boldsymbol{n \times m}$**–Matrix**

$$A := \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \ldots & & & \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix}$$

und schreiben (1) in der Form

$$\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nm} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} ,$$

kurz

(2) $$\boxed{\bar{w} = A\bar{v}}\,.$$

Dabei sei die Multiplikation von $A$ mit $\bar{v}$ durch die Vorschrift (1) erklärt. Gemäß (2) gibt also die Matrix $A$ die **zur linearen Abbildung A gehörende Koordinatenabbildung** an. (Bei festen Basen wollen wir diese Zuordnung immer durch die Wahl gleicher Buchstaben unterschiedlicher Stärke zum Ausdruck bringen.)
Parallel zu den Abbildungsoperationen gibt es nun einen passenden Matrizenkalkül. Wir wollen letzteren zunächst formal einführen und ihn dann anhand der Operationen für lineare Abbildungen deuten.

**Matrizenkalkül**
Eine $n \times m$–Matrix über dem Körper $K$ sei ein beliebiges Rechteckschema

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nm} \end{pmatrix} \qquad \text{mit } a_{ij} \in K\,.$$

Die horizontalen Elementefolgen heißen **Zeilen** und die vertikalen **Spalten** der Matrix $A$. Man schreibt auch kurz $A = (a_{ij})$, wenn $n$ und $m$ fest sind.



Die Menge $M(n \times m)$ aller solcher Matrizen läßt sich mit einer Vektorraumstruktur versehen. Dabei ist die **Summe** zweier Matrizen $A = (a_{ij})$ und $B = (b_{ij})$aus $M(n \times m)$ definiert als

$$A + B := (a_{ij} + b_{ij})$$

und die **skalare Multiplikation** von $A = (a_{ij})$ mit $\lambda \in K$ als

$$\lambda A := (\lambda a_{ij})\,.$$

**Aufgabe.** Überprüfen Sie die Axiome des Vektorraums und folgende Dimensionsformel:

**3.4.4.** $$\dim M(n \times m) = n \cdot m$$ •

Schließlich wollen wir noch die **Matrizenmultiplikation** einführen. Hierbei muß man die zugehörigen Ausmaße beachten. Seien $A = (a_{ij})$ eine $p \times q$–Matrix und $B = (b_{ij})$ eine $q \times r$–Matrix über $K$. Dann ist ihr **Produkt** eine $p \times r$–Matrix, die wie folgt definiert ist:

$$\boxed{AB := \left( \sum_{k=1}^{q} a_{ik}\, b_{kj} \right)}\,.$$

Den Spezialfall $r = 1$ (für $p = n,\ q = m$) haben wir bereits oben betrachtet, als wir eine $n \times m$–Matrix mit einem Spaltenvektor der Länge $m$ – also mit einer $m \times 1$–Matrix – multipliziert haben.

**Aufgabe.** Zeigen Sie an einem Beispiel, daß die Matrizenmultiplikation im Raum der $n \times n$–Matrizen nicht kommutativ ist. •

**Definition.** Die **transponierte Matrix** zu einer $m \times n$–Matrix $A = (a_{ij})$ ist die $n \times m$–Matrix

$$A^T = (a_{ij}^T) \;\text{ mit } a_{ij}^T := a_{ji}\,.$$

(Für quadratische Matrizen $A$ entsteht $A^T$ durch Spiegelung an der Diagonalen.) Produkt und Transposition sind wir folgt verbunden:

**3.4.5.** $$(AB)^T = B^T A^T.$$ •

Die Bedeutung des Matrizenkalküls wird entscheidend durch den folgenden Zusammenhang erhellt.

**3.4.6 Satz.** $\mathbf{A}, \mathbf{B}, \mathbf{C}$ *seien lineare Abbildungen und* $A, B, C$ *die zugehörigen Matrizen bzgl. fester Basen. Dann gelten die Aussagen*

(1) $$\mathbf{A} + \mathbf{B} = \mathbf{C} \Rightarrow A + B = C$$
(2) $$\lambda\mathbf{A} = \mathbf{C} \Rightarrow \lambda A = C$$
(3) $$\mathbf{B} \circ \mathbf{A} = \mathbf{C} \Rightarrow BA = C\,.$$

*Beweis.* (1) und (2) folgen unmittelbar aus den Definitionen und der Linearität der Koordinatendarstellung, denn die Spalten der Matrix $A$ von $\mathbf{A}$ sind gerade die Koordinatenvektoren $\overline{\mathbf{A}b_1}, \ldots, \overline{\mathbf{A}b_m}$ bzgl der Basis $e_1, \ldots, e_n$. Sei nun $\mathbf{A} \in L(U_m, V_n)$ , $\mathbf{B} \in L(V_n, W_p)$. Dann sind $\bar{v} = A\bar{u}$ und $\bar{w} = B\bar{v}$ die angehörigen Koordinatenabbildungen. Die Koordinatenabbildungen zur Komposition $\mathbf{B} \circ \mathbf{A}$ ist $\bar{w} = C\bar{u}$.
Andererseits gilt $\bar{w} = B(A\bar{u})$, d.h.

$$w_i = \sum_{k=1}^{n} b_{ik} \left( \sum_{j=1}^{m} a_{kj}\, u_j \right) = \sum_{j=1}^{m} \left( \sum_{k=1}^{n} b_{ik}\, a_{kj} \right) u_j \ , \ i = 1, \ldots, p,$$

also $\bar{w} = (BA)\bar{u}$. Wir erhalten $(BA)\bar{u} = C\,\bar{u}$ für alle $\bar{u} \in K^n$. Setzt man für $\bar{u}$ sukzessive die kanonischen Basisvektoren $\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ in $K^n$ ein, so ergibt sich die Gleichheit der entsprechenden Spaltenvektoren von $BA$ und $C$, so daß diese beiden Matrizen übereinstimmen. $\square$

Die Beziehung der linearen Abbildungen zu ihren Matrizen läß sich noch etwas schärfer als in 3.4.6 (1), (2) formulieren:

**3.4.7 Satz.** *Die Abbildung $f : L(V_m, W_n) \to M(n \times m)$ mit $f(\mathbf{A}) = A$ ist ein Vektorraumisomorphismus.*

**Bemerkung.** In diesem Sinne kann man eine lineare Abbildung $\mathbf{A}$ mit ihrer Matrix $A$ bzgl. fester Basen identifizieren.

*Beweis des Satzes.*
**1)** Die **Linearität** von $f$ ist in (1) und (2) von Satz 3.4.6 bewiesen.
**2)** $f$ ist **injektiv**: Falls $\mathbf{A} \neq \mathbf{B}$, so gibt es wegen 3.4.3 einen Basisvektor $b_j$ in $V_m$, für den $\mathbf{A}b_j \neq \mathbf{B}b_j$. Dann sind aber auch die zugehörigen Koordinatenvektoren verschieden, und diese stehen in der $j$–ten Spalte von $A$ bzw. $B$. Es gilt also $A \neq B$.
**3)** $f$ ist **surjektiv**: Für eine beliebige $n \times m$–Matrix $A$ über $K$ definieren wir die Abbildung $\mathbf{A} : V_m \to W_n$ durch $\mathbf{A}v := \sum_{i=1}^{n} w_i\, e_i$, falls $v = \sum_{j=1}^{m} v_j\, b_j$ und $A\bar{v} = \bar{w}$.
Es bleibt die Linearität von $\mathbf{A}$ zu zeigen, denn $A$ ist dann nach Konstruktion die zugehörige Matrix. Wenn $\bar{v}, \bar{v}'$ die Koordinatenvektoren von $v$ bzw. $v'$ sind, so ist $\bar{v} + \bar{v}'$ der Koordinatenvektor von $v + v'$, und $A(\bar{v} + \bar{v}') = A\bar{v} + A\bar{v}' = \bar{w} + \bar{w}'$, wobei $\bar{w}$ und $\bar{w}'$ die Koordinatenvektoren von $\mathbf{A}v$ bzw. $\mathbf{A}v'$ sind. $\bar{w} + \bar{w}'$ ist deshalb nach Konstruktion der Koordinatenvektor von $\mathbf{A}(v + v')$. Es muß also gelten $\mathbf{A}(v + v') = \mathbf{A}v + \mathbf{A}v'$. Analog zeigt man $\mathbf{A}(\lambda v) = \lambda \mathbf{A}v$. $\square$

Da $\dim M(n \times m) = nm$ (s. 3.4.4), kann man aufgrund von 3.4.7 und 2.4.7 auf die Dimension von $L(V_m, W_n)$ schließen.



**3.4.8. Folgerung.**

$$\dim L(V_m, W_n) = mn\,.$$

**Weitere Beziehungen**
**3.4.9.** Zur Nullabbildung $\mathbf{O} : V_m \to V_n$ gehört die $n \times m$–**Nullmatrix**

$$O := \begin{pmatrix} 0 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & 0 \end{pmatrix}.$$

**3.4.10.** Zur Identität $\mathbf{I} : V_n \to V_n$ gehört die $n \times n$–**Einheitsmatrix**

$$I := \begin{pmatrix} 1 & 0 & \ldots & & 0 \\ 0 & 1 & & & \vdots \\ \vdots & & \ddots & & \\ & & & 1 & 0 \\ 0 & \ldots & & 0 & 1 \end{pmatrix}.$$

**3.4.11.** Zur inversen Abbildung gehört die **inverse Matrix** :

Sei $\mathbf{A} \in L(V_n, W_n)$ ein Vektorraumisomorphismus. Wie wir bereits für Satz 2.4.6 gezeigt haben, ist die inverse Abbildung $\mathbf{A}^{-1}$ dann ebenfalls linear. Die zugehörige Matrix wollen wir mit $A^{-1}$ bezeichnen. Der Zusammenhang zwischen $A^{-1}$ und $A$ läßt sich jedoch auch im Matrizenkalkül beschreiben: Aus den Gleichungen $\mathbf{A}^{-1} \circ \mathbf{A} = \mathbf{I}_{V_n}$ und $\mathbf{A} \circ \mathbf{A}^{-1} = \mathbf{I}_{W_n}$ schließen wir mit Hilfe von Satz 3.4.6 (3) auf

$$A^{-1}A = I \quad , \quad AA^{-1} = I\,.$$

$A^{-1}$ spielt also die Rolle des inversen Elements zu $A$ bzgl. der Matrizenmultiplikation, denn $I$ kann dabei als neutrales Element interpretiert werden. Man nennt $A^{-1}$ deshalb **inverse Matrix** zu $A$.

**Bemerkung.** Die Menge $M(n \times n)^*$ aller invertierbaren Matrizen $A \in M(n \times n)$, d.h. solcher, für die es ein $A^{-1} \in M(n \times n)$ gibt mit $A^{-1}A = AA^{-1} = I$, ist in der Gruppentheorie von Bedeutung. Sie bildet nämlich bzgl. der Matrizenmultiplikation eine Gruppe. Für die Menge aller Vektorraumisomorphismen von $V_n$ auf sich selbst schreibt man auch $GL(V_n)$. Diese wiederum bildet bzgl. der Komposition eine Gruppe, und die Abbildung $f : GL(V_n) \to M(n \times n)^*$ mit $f(\mathbf{A}) = A$ ist ein sogenannter Gruppenisomorphismus. Dies bedeutet, daß $f$ bijektiv ist und die Gruppenstruktur überträgt:

$$\begin{aligned} f(\mathbf{A} \circ \mathbf{B}) &= AB \\ f(\mathbf{I}) &= I \\ f(\mathbf{A}^{-1}) &= A^{-1}. \end{aligned}$$

Ausführtlich werden wir uns mit solchen Dingen in der höheren Algebra beschäftigen. Ein Existenzkriterium und ein Verfahren zur Bestimmung der inversen Matrix werden wir später kennenlernen.

**Änderung der Matrix einer linearen Abbildung beim Übergang zu neuen Basen**

Für ein $\mathbf{A} \in L(V_m, W_n)$ betrachten wir jetzt die zugehörigen Matrizen $A$ und $A'$ bzgl. verschiedener Basen $\{b_1, \ldots, b_m\}, \{e_1, \ldots, e_n\}$ bzw. $\{b'_1, \ldots, b'_m\}, \{e'_1, \ldots, e'_m\}$. $C_V$ und $C_W$ seien die Koordinatentransformationsmatrizen beim Übergang von $\{b'_1, \ldots, b'_m\}$ nach $\{b_1, \ldots, b_m\}$ bzw. von $\{e'_1, \ldots, e'_n\}$ nach $\{e_1, \ldots, e_n\}$, d.h.

$$(1) \quad \bar{v} = C_V \bar{v}' \qquad\qquad (2) \quad \bar{w} = C_W \bar{w}' \,.$$

Die zu $\mathbf{A}$ gehörenden Koordinatenabbildungen lauten

$$(3) \quad \bar{w} = A\bar{v} \qquad\qquad (4) \quad \bar{w}' = A'\bar{v}' \,.$$

Wir setzen (1) und (2) in (3) ein und erhalten

$$C_W \bar{w}' = A\, C_V \bar{v}' \,.$$

Da Koordinatentransformationen bijektiv sind (vgl. 3.3.3), existieren die Inversen $C_V^{-1}$ und $C_W^{-1}$. Durch Multiplikation der letzten Gleichung mit $C_W^{-1}$ kommen wir zu

$$(5) \quad \bar{w}' = C_W^{-1}\, A\, V_V\, \bar{v}' \,.$$

Ein Vergleich von (4) und (5) ergibt die folgende **Transformationsregel** für den Übergang von $A$ zu $A'$.

**3.4.12 Lemma.**

$$A' = C_W^{-1}\, A\, C_V \,,$$

*oder umgekehrt*

$$A = C_W\, A'\, C_V^{-1} \,.$$

## 3.5 Rangbestimmung mittels Matrizen – Gaußscher Algorithmus

Der Rang einer linearen Abbildung kann mit Hilfe ihrer Matrizen bestimmt werden. Es genügt dabei, dies für eine feste Matrix zu tun: Für $\mathbf{A} \in L(V_m, W_n)$ gilt

$$\text{Rang}\ \mathbf{A} = \dim \mathbf{A}(V_m) = \text{Rang}\{\mathbf{A}b_1, \ldots, \mathbf{A}b_m\} = \text{Rang}\{\overline{\mathbf{A}b_1}, \ldots, \overline{\mathbf{A}b_m}\}$$

wobei $b_1, \ldots, b_m$ eine beliebige Basis in $V_m$ ist und $\overline{\mathbf{A}b_i}$ den Spaltenkoordinatenvektor von $\mathbf{A}b_i$ bzgl. eine beliebigen Basis $\{e_1, \ldots, e_n\}$ in $W_n$ bezeichnet. Die $\overline{\mathbf{A}b_1}, \ldots, \overline{\mathbf{A}b_m}$ bilden gleichzeitig die Spalten der zugehörigen Matrix $A$ von $\mathbf{A}$. Ihren Rang nennt man deshalb auch **Spaltenrang** von $A$. Analog heißt der Rang des Systems der Zeilenvektoren einer Matrix auch **Zeilenrang**. Weiter unten werden wir den nachfolgenden Satz beweisen.



**3.5.1 Satz.** *Der Spaltenrang einer Matrix $A$ ist gleich ihrem Zeilenrang (und wird deshalb mit **Rang** der Matrix, kurz* Rang *$A$ bezeichnet).*

**3.5.2 Folgerung.** Für jede Matrixdarstellung $A$ von $\mathbf{A} \in L(V_m, V_n)$ gilt Rang $\mathbf{A} =$ Rang $A$.

Zum Beweis von Satz 3.5.1 benutzen wir zwei Hilfsschritte:

**3.5.3 Lemma.** *Der Spalten- und der Zeilenrang einer Matrix sind invariant gegenüber*
(1) *Vertauschen zweier beliebigen Spalten oder Zeilen*
(2) *Addition eines Vielfachen einer Spalte oder Zeile zu einer anderen.*

*Beweis.* Zunächst sei bemerkt, daß es genügt, alle Aussagen für den Spaltenrang zu beweisen, denn beim Übergang zur transponierten Matrix werden die Zeilen zu Spalten und umgekehrt.
Wir zeigen zuerst die Invarianz des Spaltenranges gegenüber den genannten Spaltenoperationen. (1) ist in diesem Fall offensichtlich. Sei nun $A' = (\bar{a}'_1 \ldots \bar{a}'_n)$ die Matrix, die aus $A = (\bar{a}_1 \ldots \bar{a}_n)$ entsteht, wenn man zur $j$–ten Spalte $\bar{a}_j$ das $\mu$–fache der Spalte $\bar{a}_k$ addiert. Offenbar ist dann $L(\{\bar{a}'_1, \ldots, \bar{a}'_n\}) \subset L(\{\bar{a}_1, \ldots, \bar{a}_n\})$, d.h. der Spaltenrang von $A'$ ist nicht größer als der Spaltenrang von $A$. Andererseits gewinnt man $A$ aus $A'$ durch Subtraktion des $\mu$–fachen von $\bar{a}'_k$ von der Spalte $\bar{a}'_j$. Aus den gleichen Gründen wie oben ist nun der Spaltenrang von $A$ nicht größer als der Spaltenrang von $A'$. Deshalb fallen die beiden Ränge zusammen, d.h. (2) ist für den Spaltenfall bewiesen.
Als nächstes betrachten wir die Zeilenoperationen. Diesmal entsteht bei (1) die Matrix $A' = (\bar{a}'_1 \ldots \bar{a}'_n)$ aus $A = (\bar{a}_1 \ldots \bar{a}_n)$ durch Vertauschen zweier Zeilen. Ein Teilsystem $\{\bar{a}_{j_1}, \ldots, \bar{a}_{j_k}\}$ ist genau dann linear abhängig, wenn

$$(3) \quad \lambda_1\, \bar{a}_{j_1} + \ldots + \lambda_k\, \bar{a}_{j_k} = \bar{0} \quad \text{mit } \lambda_l \neq 0 \quad \text{für ein } l \,.$$

Wenn wir diese Vektorgleichung komponentenweise aufschreiben, so ändert sich beim Übergang von $A$ zu $A'$ nur die Reihenfolge der Gleichungen des Systems, so daß folgendes dazu äquivalent ist:

$$(4) \quad \lambda_1\, \bar{a}'_{j_1} + \ldots + \lambda_k\, \bar{a}'_{j_k} = \bar{0} \quad \text{mit } \lambda_l \text{ wie oben.}$$

Dies ist aber gleichbedeutend mit der linearen Abhängigkeit von $\{\bar{a}'_{j_1}, \ldots, \bar{a}'_{j_k}\}$. Also stimmen die Spaltenränge von $A$ und $A'$ überein. Bei (2) gehen wir ähnlich vor. Diesmal entsteht das Gleichungssystem (4) aus (3) durch Addition des $\mu$–Fachen der $k$–ten Zeile zur $i$–ten Zeile. Umgekehrt gewinnt man (3) aus (4) hier durch Subtraktion des $\mu$–fachen der $k$–ten Zeile von der $i$–ten. □

Die in Lemma 3.5.3 verwendeten Umformmengen nennt man auch **elementare Zeilen- bzw. Spaltenoperationen**.

**3.5.4 Lemma.** *Jede Matrix $A$ läßt sich durch elementare Zeilenoperationen und Vertauschen von Spalten auf folgende Dreiecksform bringen:*

$$A' := \left(\begin{array}{ccccc|ccc} b_{11} & \dots & & & b_{1r} & & \dots & \\ 0 & b_{22} & \dots & & b_{2r} & & & \\ 0 & 0 & \ddots & & & & & \\ \dots & & & \ddots & & & & \\ 0 & 0 & & 0 & b_{rr} & & \dots & \\ \hline 0 & 0 & \dots & & 0 & 0 & \dots & 0 \\ \dots & & & & & \dots & & \\ 0 & 0 & & & 0 & 0 & \dots & 0 \end{array}\right) \qquad \textit{mit } b_{ii} \neq 0\,,\ \forall i\,.$$

*Beweis.* Bei der Ausgangsmatrix $A = (a_{ij})$ vertauscht man nötigenfalls Zeilen bzw. Spalten, um an der Stelle $i = j = 1$ ein Element $b_{11} \neq 0$ zu erhalten. Dann addiert man geeignete Vielfache der ersten Zeile zu den anderen, um die Elemente, die unterhalb von $b_{11}$ liegen, zu Null zu machen. Anschließend vertauscht man, wenn nötig, nachfolgende Zeilen bzw. Spalten mit der zweiten, so daß an die Stelle $i = j = 2$ ein Element $b_{22} \neq 0$ kommt. Danach läßt man durch Addition von Vielfachen der zweiten Zeile zu den nachfolgenden die Elemente unter $b_{22}$ verschwinden, usw. □

Das im letzten Beweis benutzte Verfahren nennt man den **Gaußschen Algorithmus**. Er wird auch eine wichtige Rolle bei der Lösung von linearen Gleichungssystemen spielen.

*Beweis von Satz 3.5.1.* Nach 3.5.3 und 3.5.4 genügt es, Matrizen der Form $A'$ aus Satz 3.5.4 zu betrachten. Wir bezeichnen die Matrix oberhalb der waagerechten Linie von $A'$ mit $B$. Aufgrund der speziellen Struktur der Matrix $A'$ fallen ihr Zeilen- und ihr Spaltenrang mit den entsprechenden Rängen von $B$ zusammen. (Erläutern Sie dies genauer!) Es reicht deshalb zu zeigen, daß der Zeilenrang der Matrix

$$B = \left(\begin{array}{cccc|c} b_{11} & b_{12} & \dots & b_{1r} & \dots \\ 0 & b_{22} & & & \\ \dots & & & & \\ 0 & \dots & 0 & b_{rr} & \dots \end{array}\right)$$

gleich ihrem Spaltenrang ist. Da die $b_{ii}$ von Null verschieden sind, können wir auf die lineare Unabhängigkeit der $r$ Zeilenvektoren und die der ersten $r$ Spaltenvektoren von $B$ schließen. Es bleibt zu beweisen, daß der Spaltenrang von $B$ gleich $r$ ist. Dazu zeigen wir die Darstellbarkeit jedes Spaltenvektors $\bar{\mu}$ rechts von der senkrechten Linie von $B$ als Linearkombination der ersten $r$ Spaltenvektoren. Gesucht sind also Koeffizienten $\lambda_1, \dots, \lambda_r$ der Gleichung $\bar{\mu} = \sum\limits_{j=1}^{r} \lambda_j \bar{b}_j$ , welche in Koordinatenschreibweise lautet:

$$\begin{array}{rcl} \mu_1 & = & \lambda_1 b_{11} + \lambda_2 b_{12} + \dots + \lambda_r b_{1r} \\ \mu_1 & = & \phantom{\lambda_1 b_{11} +{}} \lambda_2 b_{22} + \dots + \lambda_r b_{2r} \\ \vdots & & \\ \mu_1 & = & \phantom{\lambda_1 b_{11} + \lambda_2 b_{12} + \dots +{}} \lambda_r b_{rr}\,. \end{array}$$



Dieses Gleichungssystem läßt sich rekursiv lösen: $\lambda_r = b_{rr}^{-1}\,\mu_r$,
$\lambda_k = b_{kk}^{-1}\left(\mu_k - \sum\limits_{j=k+1}^{r} \lambda_j\, b_{kj}\right)$. □

## 3.6 Ein Invertierungsverfahren für quadratische Matrizen

Ähnlich wie beim Gaußschen Algorithmus kann man die elementaren Zeilenoperationen zum Invertieren einer Matrix $A$ benutzen. Die theoretische Grundlage für das Verfahren wird durch folgendes gegeben.

**3.6.1 Satz.** *Es sei $AB = C$ ein Matrizenprodukt. Wenn die Matrizen $A$ und $C$ durch die gleichen elementaren Zeilenoperationen in $A'$ bzw $C'$ überführt werden, so gilt $A'B = C'$.*

*Beweis.* Die Elemente $a_{ij}, b_{jk}$ und $c_{ik}$ der Matrizen $A, B$ bzw. $C$ sind durch die Beziehung

$$\sum_{j=1}^{n} a_{ij} b_{jk} = c_{ik}$$

verbunden. Wenn wir in $A$ und $C$ die gleichen Zeilen vertauschen, ergeben sich die Gleichungen

$$\sum_{j=1}^{n} a'_{ij} b_{jk} = c'_{ik}$$

durch Änderung der Reihenfolge. Beim Addieren des Vielfachen der $l$–ten Zeile zur $i$–Zeile von $A$ und von $C$ müssen wir nur zu den Gleichungen des ersten Systems mit vorderem Index $i$ dieselben Vielfachen der Gleichungen mit vorderem Index $l$ dazuzählen, um auf das zweite Gleichungssystem zu kommen. □

Sei nun $A \in M(n \times n)$ eine invertierbare Matrix. Der Rang von $A$ muß dann gleich $n$ sein. (Warum?) Wir setzen in Satz 3.6.1 $B := A^{-1}$ und deshalb $C = I$. Ähnlich wie beim Gaußschen Alogrithmus kann man die Matrix $A$ durch elementare Zeilenoperationen zur Einheitsmatrix $I =: A'$ umformen. Die Matrix $C$ soll bei diesen Operationen in $C'$ übergehen. Nach Satz 3.6.1 gilt aber $C' = A'B = IA^{-1} = A^{-1}$. Praktisch geht man deshalb nach folgendem Schema vor: Man schreibt neben die Matrix $A$ die $n \times n$–Einheitsmatrix $I$ und formt die neue $n \times 2n$–Matrix solange durch elementare Zeilenoperationen um, bis aus $A$ die Einheitsmatrix entsteht. Gleichzeitig erhält man daneben die Matrix $A^{-1}$.

Die letzten beiden Abschnitte finden in der Theorie linearer Gleichungssysteme eine wichtige Anwendung.

# 4 Lineare Gleichungssysteme

Diese tauchten verschiedentlich schon im letzten Kapitel auf. Ein allgemeines **lineares Gleichungssystem** über dem Körper $K$ hat die Gestalt

$$(1)\qquad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\ldots \\ a_{m1}x_1 + a_{k2}x_2 + \ldots + a_{mn}x_n &= b_m. \end{aligned}$$

Dabei sind sämtliche Glieder Körperelemente. Die $a_{ij}$ heißen **Koeffizienten**, die $b_i$ **konstante Glieder**, und die $x_j$ sind die **Unbekannten**. In Matrizenschreibweise lautet das Gleichungssystem

$$(2)\qquad A\overline{x} = \overline{b}\,.$$

Falls $\overline{a}_j$ die Spaltenvektoren von $A = (a_{ij})$ sind, so können wir auch schreiben

$$(3)\qquad x_1\overline{a}_1 + \ldots + x_n\overline{a}_n = \overline{b}\,.$$

**Lösbarkeitskriterium**

Fügen wir zur Matrix $A$ noch die Spalte $\overline{b}$ hinzu, so entsteht die sogenannte **erweiterte Koeffizientenmatrix** $(A, \overline{b})$.

**4.1 Satz** (Lösbarkeitskriterium von Kronecker–Capelli)**.**
*Das System* (1) *besitzt genau dann eine Lösung, wenn* Rang $A = \mathrm{Rang}\,(A, \overline{b})$.

*Beweis.* Wenn (1) eine Lösung $(x_1, \ldots, x_n)$ besitzt, so ist wegen (3) $\overline{b} \in L(\{\overline{a}_1, \ldots, \overline{a}_n\})$ und demnach Rang $(A, \overline{b})$ = Rang $A$. Setzen wir andererseits die letzten Gleichungen voraus und $r :=$ Rang $A$, so gibt es linear unabhängige Spaltenvektoren $\overline{a}_{j_1}, \ldots, \overline{a}_{j_r}$, und das System $\{\overline{a}_{j_1}, \ldots, \overline{a}_{j_r}, \overline{b}\}$ ist linear abhängig. Deshalb muß $\overline{b}$ in der linearen Hülle der $\overline{a}_{j_1}, \ldots, \overline{a}_{j_r}$ liegen, d.h. $\sum\limits_{l=1}^{r} x_{j_l}\overline{a}_{j_l} = \overline{b}$ für gewisse $x_{j_l} \in K$. Wir setzen zusätzlich $x_j = 0$, falls $j \notin \{j_1, \ldots, j_r\}$ und bekommen $\sum\limits_{j=1}^{n} x_j\overline{a}_j = \overline{b}$, also (3). Folglich genügen die $x_1, \ldots, x_n$ dem System (1). □

**Lösungsstruktur**

Um die gesamte Lösungsmenge von (1) zu bestimmen, wollen wir die Vektorraum– und Abbildungssprache benutzen. Wie früher deuten wir die Matrix $A$ des Systems in (2) als lineare Abbildung $\mathbf{A} : K^n \to K^m$ mit $\mathbf{A}\overline{v} = A\overline{v}$. Dann ist

$$\mathbf{A}^{-1}(\{\overline{b}\}) = \{\overline{x} \in K^n : A\overline{x} = \overline{b}\}$$



die Lösungsmenge von (2), kurz $LM$, und

$$\mathbf{A}^{-1}(\{\overline{0}\}) = \{\overline{x} \in K^n : A\overline{x} = \overline{0}\}$$

die Lösungsmenge des sogenannten **homogenen Gleichungssystems**, kurz $LM(\overline{0})$.

**4.2 Satz.**
(i) $LM(\overline{0})$ *ist ein linearer Unterraum von* $K^n$ *der Dimension* $n$ – Rang $A$ .

(ii) *Jeder lineare Unterraum* $U$ *von* $K^n$ *ist Lösungsmenge eines homogenen linearen Gleichungssystems.*

*Beweis.* **(i)** $LM(\overline{0})$ ist als Kern der Abbildung $\mathbf{A}$ ein Unterraum, und der Dimensionssatz 3.3.4 führt zu

$$\dim LM(\overline{0}) = n - \text{Rang }\mathbf{A} = n - \text{Rang }A\,.$$

**(ii)** Wir ergänzen eine beliebige Basis $b_1, \ldots, b_k$ in $U$ zu einer Basis $b_1, \ldots, b_n$ in $K^n$ und benutzen die Darstellung

$$K^n = U \oplus V$$

mit $V = L(\{b_{k+1}, \ldots, b_n\})$. Seien $\mathbf{A} := \mathbf{\Pi}_V$ die Projektion auf $V$ entlang $U$ und $A$ die zugehörige Matrix bzgl. der Standardbasen in $K^n$ und $K^m$. Damit ist offenbar $U = \text{Kern }\mathbf{A} = \{\overline{x} \in K^n : A\overline{x} = \overline{0}\}$. □

Für die Bestimmung der Lösungsmenge des Ausgangsgleichungssystems benötigen wir folgenden Begriff.

**Definition.** Die **Verschiebung** eines Unterraums $U$ des Vektorraums $V$ um einen Vektor $v \in V$ ist gegeben durch

$$U + v := \{u + v : u \in U\}.$$

**4.3 Satz.**
(i) $LM = LM(\overline{0}) + \overline{v}$ *für jedes* $\overline{v} \in LM$.

(ii) *Jede Verschiebung eines Unterraums von* $K^n$ *ist Lösungsmenge eines linearen Gleichungssystems.*

*Beweis.* **(i)** Die Beziehungen $\overline{x} \in LM$ und $\overline{v} \in LM$, d.h. $A\overline{x} = \overline{b}$ und $A\overline{v} = \overline{b}$, sind äquivalent zu $A(\overline{x} - \overline{v}) = \overline{0}$ und $A\overline{v} = \overline{b}$, d.h. zu $\overline{u} := \overline{x} - \overline{v} \in LM(\overline{0})$ und $\overline{v} \in LM$. Umgeschrieben lautet dies $\overline{x} = \overline{u} + \overline{v}$ mit $\overline{u} \in LM(\overline{0})$ und $\overline{v} \in LM$, was gleichbedeutend mit $LM = LM(\overline{0}) + \overline{v}$ ist.
**(ii)** Sei nun $U + \overline{v}$ Verschiebung eines Unterraums. Wegen Satz 4.2 (ii) gibt es eine Matrix $A$ mit $U = \{\overline{x} \in K^n : A\overline{x} = \overline{0}\}$.Wir setzen nun $\overline{b} := A\overline{v}$ und betrachten die Gleichung $A\overline{x} = \overline{b}$. Jedes $\overline{x} \in U+\overline{v}$ gehört wegen der Darstellung $\overline{x} = \overline{u}+\overline{v}$ mit $\overline{u} \in U$ und $A\overline{x} = A\overline{u} + A\overline{v} = \overline{0} + \overline{b} = \overline{b}$ zur Lösungsmenge. Ist andererseits $\overline{x}$ eine Lösung des Gleichungssystems, so erhalten wir $A(\overline{x} - \overline{v}) = \overline{b} - \overline{b} = \overline{0}$, d.h. $\overline{u} := \overline{x} - \overline{v} \in U$ und damit $\overline{x} \in U + \overline{v}$. □

**Eindeutige Lösbarkeit**

**4.4 Satz.** *Das Gleichungssystem* (1) *besitzt genau dann eine eindeutig bestimmte Lösung* $\overline{x}$*, wenn gilt*

$$\text{Rang}\,(A,\overline{b}) = \text{Rang}\,A = n\,.$$

*Beweis.* Aufgrund der Darstellung $LM = LM(\overline{0}) + \overline{v}$ aus dem Satz 4.3 ist die Lösungsmenge $LM$ genau dann einelementig, wenn $LM(\overline{0}) = \{\overline{0}\}$ gilt und überhaupt eine Lösung $\overline{v}$ existiert. Das Lösbarkeitskriterium und der schon erwähnte Dimensionssatz ergeben die äquivalente Bedingung Rang $(A,\overline{b})$ = Rang $A = n$. □

**Lösungsalgorithmus (nach Gauß)**
Um die Lösungsmenge rechnerisch bestimmen zu können, betrachten wir zunächst den **Spezialfall**

$$(2') \qquad A'\overline{x} = \overline{b'}\,,$$

wo die Matrix $A'$ eine obere Dreiecksgestalt besitzt, d.h.

$$(A',\overline{b'}) = \left(\begin{array}{ccc|c|c} a'_{11} & & & & b'_1 \\ & \ddots & & & \vdots \\ \mathbf{0} & & a'_{rr} & & b'_r \\ \hline & & & & b'_{r+1} \\ & \mathbf{0} & & \mathbf{0} & \vdots \\ & & & & b'_m \end{array}\right) \qquad \text{mit } a'_{ii} \neq 0\,,\ i = 1,\ldots,r\,.$$

Eine Lösung $\overline{x}$ existiert hier höchstens dann, wenn $b'_{r+1} = \ldots = b'_m = 0$. In diesem Fall können wir die Unbekannten $x_{r+1},\ldots,x_n \in K$ beliebig vorgeben und in die $r$–te Gleichung des Systems einsetzen. $x_r$ läßt sich dort eliminieren:

$$x_r = (a'_{rr})^{-1}\,(b'_r - \sum_{j=r+1}^{n} a'_{rj}x_j)\,.$$

Anschließend werden $x_r,\ldots,x_n$ in die $(r-1)$–te Gleichung eingesetzt, und man stellt nach $x_{r-1}$ um, usw.

**Übungsaufgabe**
Demzufolge ist die Lösungsmenge gleich der Menge der $(x_1,\ldots,x_n)$, für die die $i$–te Koordinate, $i \leq r$ als durch das Verfahren bestimmte Linearkombination der Variablen $x_{r+1},\ldots,x_n$ und der Konstanten Glieder $b'_{i+1},\ldots,b'_r$ dargestellt werden kann:

$$x_i = L_i(x_{r+1},\ldots,x_n,b'_i,\ldots,b'_r)\,.$$



Der Zusammenhang zu Satz 4.3 entseht folgendermaßen. Eine spezielle Lösungs liefert nach Konstruktion z.B. $\overline{v}$ mit $v_{r+1} = \ldots = v_n = 0$ und $v_i = L_i(0,\ldots,0,b'_{i+1},\ldots,b'_r)$, $i \leq r$. Als nächstes wählen wir die Basis $\overline{e}_{r+1},\ldots,\overline{e}_n$ in $LM(\overline{0})$ mit $e_{ij} = \delta_{ij} = \left\{\begin{smallmatrix} 1 & i=j \\ 0 & i\neq j\end{smallmatrix}\right.$ für $i > r$ und $e_{ij} = L_i(0,\ldots,0,1,0\ldots0)$ mit 1 an $(j-r)$–ter Stelle für $i \leq r$.

Dann gilt
$$\overline{x} = \sum_{j=r+1}^{n} x_j\overline{e}_j + \overline{v}\,.$$
•

Wir kommen nun zum Lösungsverfahren für (1) im **Allgemeinfall**, wobei Rang $(A\overline{b})$ = Rang $A$. Dieses läßt sich auf den obigen Spezialfall zurückführen. Die Lösungsmenge von (1) ist nämlich invariant gegenüber

– Vertauschen von Zeilen von $(A\overline{b})$,

– Vertauschen von Spalten von $A$, falls die zugehörigen $x_j$ mit vertauscht werden,

– Addition des Vielfachen einer Zeile von $(A\overline{b})$ zu einer anderen.

Andererseits kann $(A\overline{b})$ durch den Gaußschen Algorithmus vermöge solcher Operationen zur Form $(A'\overline{b'})$ von oben gebracht werden (s. Satz 3.5.4).

# 5 Determinanten

## 5.1 Determinantenformen

Der Begriff der Determinante entstand im Rahmen der Theorie linearer Gleichungssysteme. Die eindeutige Lösbarkeit haben wir bisher in der Sprache der Ränge formuliert. Für $n$ Gleichungen mit $n$ Unbekannten kann man dies auch in den Determinantenkalkül übersetzen. (Außerdem gestattet dieser eine explizite Darstellung der Lösung.) Im Laufe der Entwicklung wurde die Determinantenrechnung von Matrizen so verallgemeinert, daß auch andere Anwendungen möglich sind. Wir wählen hier diesen Zugang, der das Linearitätsprinzip betont.
In Abschnitt 3.2 wurde lineare Funktionale, die man auch Linearformen nennt, als lineare Abbildungen eines Vektorraumes $V$ in den zugehöringen Körper $K$ eingeführt. Wir betrachten jetzt ein mehrdimensionales Analogon.

**Definition.** Eine **$n$–fache Linearform (Multilinearform)** über den Vektorraum $V$ mit dem Körper $K$ ist eine Abbildung $L : V^n \to K$, die in jedem ihrer Argumente linear ist, d.h. für feste $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \in V$ ist $L(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)$ als Funktion in $v$ linear, $i = 1, \ldots, n$.

Für dieses Kapitel setzen wir nun voraus

$$\dim V = n$$

und interessieren uns für ganz bestimmte $n$–fache Linearformen.

**Definition.** Eine Abbildung $\Delta : V^n \to K$ heißt **Determinantenform**, falls folgende Bedingungen erfüllt sind:

(1) $\Delta$ ist Multilinearform.

(2) $\Delta$ ist nicht identisch Null.

(3) Für linear abhängige Vektoren $v_1, \ldots, v_n$ gilt $\Delta(v_1, \ldots, v_n) = 0$.

**Eigenschaften von Determinantenformen**

**5.1.1 Lemma.** $\Delta(\ldots, v_i, \ldots, v_k, \ldots) = \Delta(\ldots, v_i + \lambda v_k, \ldots, v_k, \ldots),\ \lambda \in K$

*Beweis.* Wegen (1) ist

$$\Delta(\ldots, v_i + \lambda v_k, \ldots, v_k, \ldots) = \Delta(v_1, \ldots, v_i, \ldots, v_k, \ldots) + \lambda\Delta(\ldots, v_k, \ldots, v_k, \ldots).$$

Da $v_1, \ldots, v_{i-1}, v_k, v_{i+1}, \ldots, v_k, \ldots, v_n$ linear abhängig sind, folgt mit (3), daß der letzte Summand verschwindet. □

**5.1.2 Lemma.**
$\Delta(v_{\pi(1)}, \ldots, v_{\pi(n)}) = \mathit{sgn}\,\pi\ \Delta(v_1, \ldots, v_n)$ *für jede Permutation* $\pi \in S_n$.



*Beweis.* Wir zeigen die Behauptung zuächst für die Vertauschung des $i$–ten mit dem $k$–ten Argument, indem wir nacheinander (3), (1) und wieder (3) anwenden:

$$\begin{aligned}
0 = \Delta(\ldots, v_i + v_k, \ldots, v_i + v_k, \ldots) &= \Delta(\ldots, v_i, \ldots, v_i, \ldots) \\
+ \Delta(\ldots, v_i, \ldots, v_k, \ldots) \qquad &+ \Delta(\ldots, v_k, \ldots, v_i, \ldots) \\
+ \Delta(\ldots, v_k, \ldots, v_k, \ldots) \qquad &= \Delta(\ldots, v_i, \ldots, v_k, \ldots) \\
+ \Delta(\ldots, v_k, \ldots, v_i, \ldots)\,, \qquad &
\end{aligned}$$

d.h.,

$$\Delta(\ldots, v_i, \ldots, v_k, \ldots) = -\Delta(\ldots, v_k, \ldots, v_i, \ldots)\,.$$

Für allgemeine $\pi \in S_n$ läßt sich $v_{\pi(1)}, \ldots, v_{\pi(n)}$ durch geeignete Transpositionen benachbarter Elemente in $v_1, \ldots, v_n$ umformen: Zunächst wird $v_1$ mit allen vorhergehenden $v_{\pi(i)}$ vertauscht, anschließend $v_2$ mit seinen Vorgängern bis auf $v_1$, usw. Die Anzahl dieser Transpositionen ist $\chi(\pi)$, die Charakteristik der Permutation. Die $\chi(\pi)$–fache Anwendung der obigen Gleichung ergibt

$$\Delta(v_{\pi(1)}, \ldots, v_{\pi(n)}) = (-1)^{\chi(\pi)}\Delta(v_1, \ldots, v_n)\,.$$

Mit $(-1)^{\chi(\pi)} = \mathrm{sgn}\,\pi$ folgt die Behauptung. □

Die Eigenschaften (1) und (2) implizieren bereits die Umkehrung von (3):

**5.1.3 Lemma.** *Aus* $\Delta(v_1, \ldots, v_n) = 0$ *folgt die lineare Abhängigkeit der* $v_1, \ldots, v_n$.

*Beweis.* Sei $b_1, \ldots, b_n$ eine beliebige Basis in $V$. Für $v_1, \ldots, v_n \in V$ benutzen wir die Koordinatendarstellung $v_i = \sum_{j=1}^{n} v_{ij} b_j$. Die Multilinearität von $\Delta$ führt zu

$$\Delta(v_1, \ldots, v_n) = \sum_{j_1=1}^{n} \ldots \sum_{j_n=1}^{n} v_{1j_1} \ldots v_{nj_n}\ \Delta(b_{j_1}, \ldots, b_{j_n})\,.$$

Dabei verschwinden alle Summanden, wo zwei gleiche Indizes vorkommen, d.h.

$$\Delta(v_1, \ldots, v_n) = \sum_{\pi \in S_n} v_{1\pi(1)} \ldots v_{n\pi(n)}\ \Delta(b_{\pi(1)}, \ldots, b_{\pi(n)})\,.$$

Wegen 5.1.2 können wir fortsetzen mit

$$\left( \sum_{\pi \in S_n} v_{1\pi(1)} \ldots v_{n\pi(n)}\ \mathrm{sgn}\,\pi \right)\ \Delta(b_1, \ldots, b_n)\,.$$

Falls nun $\Delta(b_1, \ldots b_n) = 0$ wäre, so würde $\Delta$ identisch Null sein, was einen Widerspruch zu (2) ergäbe. Demnach ist $\Delta(b_1, \ldots, b_n) \neq 0$ für jedes linear unabhängige Vektorsystem $\{b_1, \ldots, b_n\}$. □

Dieser Beweis liefert gleichzeitig eine weitere Eigenschaft von $\Delta$:

**5.1.4 Lemma.** $\Delta(v_1,\ldots,v_n) = \Delta(b_1,\ldots,b_n) \sum_{\pi\in S_n} sgn\,\pi\; v_{1\pi(1)}\ldots v_{n\pi(n)}$.

Auf dieser Grundlange können wir nun auch einen Existenznachweis für Determinantenformen führen:

**5.1.5 Lemma.** *Für jedes $d \in K \setminus \{0\}$ ist*

$$\Delta(v_1,\ldots,v_n) := d \sum_{\pi\in S_n} sgn\,\pi\; v_{1\pi(1)}\ldots v_{n\pi(n)}$$

*eine Determinantenform in $V$, wobei $v_i = \sum_{j=1}^{n} v_{ij}b_j$ bzgl. einer beliebigen Basis $b_1,\ldots,b_n$ gesetzt wurde.*

*Beweis.* **(1)** Aus der Linearität der Koordinatendarstellung der Vektoren folgt sofort die Linearität von $\Delta(v_1,\ldots,v_n)$ in jedem der Argumente $v_i$.

**(2)** Da die Basisvektoren $b_i$ die Koordinaten $\delta_{ij} = \left\{\begin{smallmatrix} 0\,, & i\neq j \\ 1\,, & i=j\end{smallmatrix}\right.$ besitzen, verschwinden in der Darstellung von $\Delta(b_1,\ldots,b_n)$ sämtliche Summanden mit Ausnahme des Falles $\pi = \mathrm{id}_{S_n}$, wo gilt $\mathrm{sgn}\,\pi = 1$ und $b_{1\pi(1)}\ldots b_{n\pi(n)} = b_{11}\ldots b_{nn} = 1$, d.h. $\Delta(b_1,\ldots .b_n) = d$, und damit ist $\Delta$ nicht identisch Null.

**(3)** Seien nun $v_1,\ldots,v_n$ linear abhängig: Wir nehmen o.E. an, daß $v_n = \sum_{i=1}^{n-1}\lambda_i v_i$. (Ansonsten vertauscht man die Indizes und erhält höchstens einen Vorzeichenwechsel bei $\Delta$.) Die Multilinearität ergibt dann

$$\Delta(v_1,\ldots,v_n) = \sum_{i=1}^{n-1}\lambda_i\,\Delta(v_1,\ldots,v_{n-1},v_i)\,.$$

Es genügt deshalb zu zeigen, daß

$$\Delta(v_1,\ldots,v_{n-1},v_i) = 0\,,\ i = 1,\ldots,n-1\,.$$

Nach Definition ist

$$\begin{aligned}\Delta(v_1,\ldots,v_{n-1},v_i) &= d\sum_{\pi\in S_n}\mathrm{sgn}\,\pi\; v_{1\pi(1)}\ldots v_{i\pi(i)}\ldots v_{n-1\,\pi(n-1)}v_{i\pi(n)}\\ &= d\sum_{\pi\in S_n^+} v_{1\pi(1)}\ldots v_{i\pi(i)}\ldots v_{n-1\,\pi(n-1)}v_{i\pi(n)}\\ &\quad + d\sum_{\sigma\in S_n^-}(-1)\,v_{1\sigma(1)}\ldots v_{i\sigma(i)}\ldots v_{n-1\sigma(n-1)}\,v_{i\sigma(n)},\end{aligned}$$



wobei $S_n^+ := \{\pi\in S_n : \mathrm{sgn}\,\pi = 1\}\,,\ S_n^- := \{\sigma\in S_n : \mathrm{sgn}\,\sigma = -1\}$.
Mit $\sigma = (i\,n)\circ\pi$ gilt aber $\pi\in S_n^+ \Leftrightarrow \sigma\in S_n^-$. Bei der Transposition $(i\,n)$ wird das Produkt $v_{1\pi(1)}\ldots v_{i\pi(i)}\ldots v_{n-1\,\pi(n-1)}\,v_{i\pi(n)}$ nicht verändert. Deshalb ist die obige Summe gleich

$$d\sum_{\pi\in S_n^+}(1-1)\,v_{1\pi(1)}\ldots v_{n-1\,\pi(n-1)}\,v_{i\pi(n)} = 0\,.$$

□

Als nächstes zeigen wir, daß Determinantenformen bis auf Vielfache eindeutig bestimmt sind. Wir benutzen im folgenden auch die Divisionsschreibweise für Körper:

$$\frac{\lambda}{\mu} := \lambda\mu^{-1}\,,\ \lambda,\mu\in K\,,\ \mu\neq 0\,.$$

**5.1.6 Lemma.** *Seien $\Delta\,,\ \Delta'$ Determinantenformen in $V$. Dann ist*

$$c := \frac{\Delta'(b_1,\ldots,b_n)}{\Delta(b_1,\ldots,b_n)}$$

*unabhängig von der Wahl der Basis $\{b_1,\ldots,b_n\}$, und es gilt die Abbildungsgleichung*

$$\Delta' = c\,\Delta\,.$$

*Beweis.* Nach Satz 5.1.4 ist

$$\begin{aligned}\Delta'(v_1,\ldots,v_n) &= \Delta'(b_1,\ldots,b_n)\sum_{\pi\in S_n}\mathrm{sgn}\,\pi\; v_{1\pi(1)}\ldots v_{n\pi(n)}\\ &= c\,\Delta(b_1,\ldots,b_n)\sum_{\pi\in S_n}\mathrm{sgn}\,\pi\; v_{1\pi(1)}\ldots v_{n\pi(n)}\\ &= c\,\Delta(v_1,\ldots,v_n)\end{aligned}$$

für alle $v_1,\ldots,v_n\in V$, d.h. $\Delta' = c\,\Delta$. Insbesondere ergibt sich für jede weitere Basis $\{b'_1,\ldots,b'_n\}$

$$\Delta'(b'_1,\ldots,b'_n) = c\,\Delta(b'_1,\ldots,b'_n)\,.$$

□

Wir führen jetzt Determinantenformen ein, die von Isomorphismen erzeugt werden:

**5.1.7 Lemma.** *Seien $\mathbf{A}\in L(V,W)$ ein Vektorraumisomorphismus und $\Delta$ eine Determinantenform in $W$. Dann ist*

$$\Delta_{\mathbf{A}}(v_1,\ldots,v_n) := \Delta(\mathbf{A}v_1,\ldots,\mathbf{A}v_n)$$

*eine Determinantenform in $V$.* •

## 5.2 Determinanten von Endomorphismen

Eine Abbildung $\mathbf{A} \in L(V,V)$ nennt man auch **Vektorraumendomorphismus**. Das Verhältnis

$$\frac{\Delta_{\mathbf{A}}(b_1,\dots,b_n)}{\Delta(b_1,\dots,b_n)}$$

ist wegen 5.1.7 und 5.1.6 unabhängig von der Wahl der Basis $\{b_1,\dots,b_n\}$. Es hängt jedoch auch nicht von der Wahl der Determinantenform $\Delta$ in $V$ ab: Wir betrachten für eine weiterre Determinantenform $\Delta'$

$$\frac{\Delta'_{\mathbf{A}}(b_1,\dots,b_n)}{\Delta'(b_1,\dots,b_n)} = \frac{\Delta'(\mathbf{A}b_1,\dots,\mathbf{A}b_n)}{\Delta'(b_1,\dots,b_n)} = \frac{c\,\Delta(\mathbf{A}b_1,\dots,\mathbf{A}b_n)}{c\,\Delta(b_1,\dots,b_n)} = \frac{\Delta_{\mathbf{A}}(b_1,\dots,b_n)}{\Delta(b_1,\dots,b_n)}$$

wobei wir 5.1.5 noch einmal benutzt haben.

**Definition.**

$$\det\mathbf{A} := \frac{\Delta_{\mathbf{A}}(b_1,\dots,b_n)}{\Delta(b_1,\dots,b_n)}$$

heißt **Determinante** von $\mathbf{A}$.
(Aufgrund der vorhergehenden Überlegungen ist die Definition korrekt.)

**5.2.1 Satz.** $\mathbf{A} \in L(V,V)$ *ist genau dann ein Vektorraumisomorphismus, wenn* $\det\mathbf{A} \neq 0$.

*Beweis.* Nach Satz 3.3.2 ist $\mathbf{A}$ genau dann Isomorphismus, wenn die Bilder $\mathbf{A}b_1,\dots,\mathbf{A}b_n$ einer Basis $b_1,\dots,b_n$ linear unabhängig sind. Dies gilt jedoch genau dann, wenn $\Delta(\mathbf{A}b_1,\dots,\mathbf{A}b_n) \neq 0$ für eine Determinantenform $\Delta$, d.h. wenn

$$\det\mathbf{A} = \frac{\Delta(\mathbf{A}b_1,\dots,\mathbf{A}b_n)}{\Delta(b_1,\dots,b_n)} \neq 0\,.$$

□

**5.2.2 Satz.**

(1) $\det(\mathbf{A}\circ\mathbf{B}) = \det\mathbf{A}\det\mathbf{B}$

(2) $\det\mathbf{I} = 1$

(3) $\det(\mathbf{A}^{-1}) = (\det\mathbf{A})^{-1}$, *falls* $\mathbf{A}$ *Isomorphismus ist.*

*Beweis.* **(1)** Mit $\mathbf{B}$ ist auch $\mathbf{A}\circ\mathbf{B}$ kein Isomorphismus.
In diesem Fall ergibt sich aus Satz 5.2.1

$$\det(\mathbf{A}\circ\mathbf{B}) = 0 = \det\mathbf{A}\cdot 0 = \det\mathbf{A}\det\mathbf{B}\,.$$



Falls $\mathbf{B}$ Isomorphismus ist, so ist $\{\mathbf{B}b_1,\dots,\mathbf{B}b_n\}$ für jede Basis $\{b_1,\dots,b_n\}$ ebenfalls eine Basis in $V$. Wir bekommen deshalb

$$\begin{aligned}\det(\mathbf{A}\circ\mathbf{B}) &= \frac{\Delta\big((\mathbf{A}\circ\mathbf{B})b_1,\dots,(\mathbf{A}\circ\mathbf{B})b_n\big)}{\Delta(b_1,\dots,b_n)}\\ &= \frac{\Delta\big(\mathbf{A}(\mathbf{B}b_1),\dots,\mathbf{A}(\mathbf{B}b_n)\big)}{\Delta(\mathbf{B}b_1,\dots,\mathbf{B}b_n)}\;\frac{\Delta(\mathbf{B}b_1),\dots,\mathbf{B}b_n)}{\Delta(b_1,\dots,b_n)}\\ &= \det\mathbf{A}\;\det\mathbf{B}\,.\end{aligned}$$

**(2)** $\det\mathbf{I} = \dfrac{\Delta(\mathbf{I}b_1,\dots,\mathbf{I}b_n)}{\Delta(b_1,\dots,b_n)} = 1.$

**(3)** Wir setzen in (1) $\mathbf{B} := \mathbf{A}^{-1}$ und erhalten mit (2) die Behauptung. □

## 5.3 Determinanten von Matrizen

Jeder Abbildung $\mathbf{A} \in L(V,V)$ läßt sich ihre Matrix $A = (a_{ij})$ bzgl. ein und derselben Basis $\{b_1,\dots,b_n\}$ zuordnen. (Wir erinnern, daß diese Zuordnung einen Vektorraumisomorphismus von $L(V,V)$ auf $M(n\times n)$ darstellt.) Dabei sind $a_{ij},\ i = 1,\dots,n$, die Koordinaten des Vektors $\mathbf{A}b_j$ bzgl. der Basis $\{b_1,\dots,b_n\}$. Wir setzen in 5.1.4 $v_j := \mathbf{A}b_j$, d.h. $v_{ij} = a_{ji}$, und leiten folgendes ab:

$$\det\mathbf{A} = \frac{\Delta(\mathbf{A}b_1,\dots,\mathbf{A}b_n)}{\Delta(b_1,\dots,b_n)} = \sum_{\pi\in S_n} \operatorname{sgn}\pi\; a_{\pi(1)1}\dots a_{\pi(n)n}.$$

In dieser Darstellung ist noch eine Vertauschung in den Indexpaaren möglich, denn für $\sigma := \pi^{-1}$ gilt

$$a_{\pi(1)1}\dots a_{\pi(n)n} = a_{1\sigma(1)}\dots a_{n\sigma(n)}\ ,\ \operatorname{sgn}\pi = \operatorname{sgn}\sigma\,,$$

und mit $\pi$ durchläuft $\pi^{-1}$ alle Elemente von $S_n$.
Die obenstehende Summe ist deshalb gleich

$$\sum_{\sigma\in S_n} \operatorname{sgn}\sigma\; a_{1\sigma(1)}\dots a_{n\sigma(n)}.$$

**Definition.** Für eine $n\times n$–Matrix $A = (a_{ij})$ heißt

$$\det A := \sum_{\pi\in S_n} \operatorname{sgn}\pi\; a_{1\pi(1)}\dots a_{n\pi(n)}$$

**Determinante** von $A$. Man schreibt auch

$$\det A = \begin{vmatrix} a_{11} & \dots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \dots & a_{nn}\end{vmatrix}.$$

In dieser Sprache haben wir nun folgendes Resultat bewiesen:

**5.3.1 Satz.** *Für eine Abbildung* $\mathbf{A} \in L(V,V)$ *gilt*

$$\det \mathbf{A} = \det A$$

*für jede ihrer Matrizen* $A$.

Damit übertragen sich die Eigenschaften von Determinanten von Endomorphismen auf Determinanten von Matrizen. Satz 5.2.2 ergibt:

**5.3.2 Satz.**
(1) $\det(AB) = \det A \, \det B$

(2) $\det I = 1$

(3) *Falls* $\det A \neq 0$, *so existiert* $A^{-1}$ *und* $\det(A^{-1}) = (\det A)^{-1}$

(Für die Existenz von $A^{-1}$ in (3) benutzen wir Satz 5.2.1.)
Für die transponierte Matrix gilt:

**5.3.3 Satz.** $\det(A^T) = \det A$.

*Beweis.* Für $A^T = (a_{ij}^T)$ gilt $a_{ij}^T = a_{ji}$. Wir haben aber bereits oben gezeigt, daß

$$\sum_{\pi \in S_n} \operatorname{sgn} \pi \; a_{\pi(1)1} \dots a_{\pi(n)n} = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \; a_{1\sigma(1)} \dots a_{n\sigma(n)} .$$

□

**Weitere Eigenschaften von Determinanten**

Sei $\Delta$ die Determinantenform in $K^n$ mit $\Delta(e_1, \dots, e_n) = 1$ für die kanonische Basis $\{e_1, \dots, e_n\}$.
Wir fassen $A$ auch als Abbildung $\mathbf{A} \in L(K^n, K^n)$ auf und erinnern, daß $\mathbf{A}e_j = \bar{a}_j$ der $j$–te Spaltenvektor der Matrix $A$ ist. Dies führt zu

$$\det A = \Delta(\bar{a}_1, \dots, \bar{a}_n) .$$

Alle Eigenschaften der Determinantenformen $\Delta$ übertragen sich deshalb auf die Determinanten bzgl. ihrer Spaltenvektoren und wegen 5.3.3 auch bzgl. ihrer Zeilenvektoren. Determinanten von Matrizen sind insbesondere invariant gegenüber der Addition des Vielfachen einer Spalte oder Zeile zu einer anderen und ändern beim Vertauschen zweier Spalten oder Zeilen das Vorzeichen.
**Berechnung von Determinanten**

**Definition.** Die Determinante $M_{ij}$, die aus $\det A = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$ durch Streichen der $i$–ten Zeile und der $j$–ten Spalte entsteht heißt **Minor** des Elementes $a_{ij}$.



**Definition.** $A_{ij} := (-1)^{i+j} M_{ij}$ wird **Adjunkte** des Elementes $a_{ij}$ genannt.

Man kann Determinanten nach den Elementen einer Zeile entwickeln. Aus der Definition von $\det A$ und $M_{ij}$ ergibt sich sehr leicht:

**5.3.4 Satz.** $$\det A = \sum_{j=1}^{n} a_{ij} \, A_{ij} \, , \quad i = 1, \dots, n \, .$$ •

Für sogenannte **Blockmatrizen** der Struktur

$$A = \begin{pmatrix} B & 0 \\ D & C \end{pmatrix} \quad \text{für} \quad B \in M(m \times m) \quad \text{und} \quad C \in M(p \times p)$$

gilt folgende Rechenregel:

**5.3.5 Satz.** $\det A = \det B \, \det C$.

*Beweis.* Wir benutzen die vollständige Induktion nach $m$ und entwickeln $\det A$ gemäß 5.3.4 nach der ersten Zeile:

$$\det A = \sum_{j=1}^{m} b_{1j} \, A_{1j} = \sum_{j=1}^{m} (-1)^{1+j} \, b_{1j} \, M_{1j} \, .$$

Für $m = 1$ liefert dies schon die Behauptung. Allgemein hat $M_{ij}$ die Gestalt $\begin{vmatrix} B_j & 0 \\ D_j & C \end{vmatrix}$, wobei $B_j \in M\big((n-1) \times (n-1)\big)$. Im Sinne einer Induktionsvoraussetzung können wir annehmen, daß $M_{ij} = \det B_j \, \det C$. Dann ist die letzte Summe gleich

$$\left( \sum_{j=1}^{m} (-1)^{1+j} \, b_{1j} \, \det B_j \right) \det C \, .$$

Der erste der beiden Faktoren ist aber die Entwicklung von $\det B$ nach der ersten Zeile. □

**5.3.6 Folgerung.**

$$\begin{vmatrix} a_{11} & 0 & \dots & \dots & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & \dots & 0 \\ \vdots & & & & & \vdots \\ a_{n-11} & a_{n-12} & \dots & \dots & a_{n-1\,n-1} & 0 \\ a_{n1} & a_{n2} & \dots & \dots & \dots & a_{nn} \end{vmatrix} = a_{11} \, a_{22} \dots a_{nn} \, .$$

□

Damit ergibt sich aber folgende Berechnungsmöglichkeit füer beliebige Determinanten: Mit Hilfe des Gaußschen Algorithmus bringt man die Determinante auf obige Dreicksgestalt. Man muß dabei nur die entsprechende Änderung des Vorzeichen beachten.
Danach bildet man das Produkt der Elemente auf der (sogenannten) Hauptdiagonalen.
Abschließend wollen wir noch eine Anwendung auf die **Berechnung der inversen Matrix** angeben:

**5.3.7 Satz.** *Sei* $\det A \neq 0$. *Dann gilt*

$$A^{-1} = (\det A)^{-1} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} .$$

*Beweis.* Es genügt zu zeigen, daß

$$A \begin{pmatrix} A_{11} & \dots & A_{n1} \\ \vdots & & \vdots \\ A_{1n} & \dots & A_{nn} \end{pmatrix} = \det A \cdot I ,$$

was, elementeweise aufgeschrieben, bedeutet

$$\sum_{j=1}^{n} a_{ij} A_{kj} = \det A \ \delta_{ik} , \ \ i,k = 1, \dots, n .$$

Für $k = i$ sind die Formeln richtig, denn sie stellen dann die Entwicklung von $\det A$ nach der $i$–ten Zeile dar. Für $k \neq i$ benutzen wir

$$0 = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{k-11} & \dots & a_{k-1n} \\ a_{i1} & \dots & a_{in} \\ a_{k+11} & \dots & a_{k+1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \sum_{j=1}^{n} a_{ij} A_{kj} ,$$

d.h. die Entwicklung dieser wegen zwei gleichen Zeilen verschwindenden Determinante nach der $k$–ten Zeile. □

Insbesondere kann man dadurch eine **explizite Darstellung der Lösung eines linearen Gleichungssystems** der Gestalt $A\bar{x} = \bar{b}$ mit $\det A \neq 0$ angeben, die sogenannte **Cramersche Regel**:

**5.3.8 Satz.**

$$\bar{x} = A^{-1} b , \ \text{ d.h. } \ x_i = (\det A)^{-1} \sum_{j=1}^{n} A_{ji} \, b_j , \ \ i = 1, \dots, n .$$



# 6 Affine Räume und affine Abbildungen

## 6.1 Axiome der affinen Geometrie

In der (euklidischen) Schulgeometrie wurde ein synthetischer Zugang gewählt: Ausgangsobjekte waren Punkte, Geraden und Ebenen mit gewissen Inzidenzbeziehungen. Über die geometrische Vektorrechnung wurde ein Zusammenhang zur analytischen Geometrie hergestellt. Dort sind die Punkte und Vektoren Grundgegebenheiten, und Geraden und Ebenen werden daraus aufgebaut. In diesem Kapitel wählen wir nun den abstrakten analytischen Zugang zur Geometrie, der stark auf Methoden der linearen Algebra basiert.
Zunächst wollen wir aber auf solche Begriffe wir Abstand, Länge, Winkel, Orthogonalität verzichten und eine Geometrie kennenlernen, die das Prinzip der Parallelität in den Vordergrund stellt.

**Definition.** Eine **affine Geometrie** ist ein Tripel $(\mathcal{A}, V, f)$, wobei $\mathcal{A}$ und $V$ Mengen sind, deren Elemente Punkte bzw. Vektoren genannt werden, und $f$ eine Abbildung von $\mathcal{A} \times \mathcal{A}$ in $V$ (man schreibt auch $f(A,B) = \overrightarrow{AB}$) mit folgenden Eigenschaften:

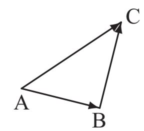
($\mathcal{A}$I) $V$ ist ein Vektorraum über einem Körper $K$.

($\mathcal{A}$II) Für jeden Punkt $A \in \mathcal{A}$ und jeden Vektor $v \in V$ gibt es einen Punkt $B \in \mathcal{A}$ mit $\overrightarrow{AB} = v$.
(Jeder Vektor läßt sich an jedem Punkt abtragen.)

($\mathcal{A}$III) Für zwei verschiedene Punkte $A, B$ gilt $\overrightarrow{AB} \neq \mathcal{O}$ .

($\mathcal{A}$IV) $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$ .

**Bemerkung.** Diejenigen geordneten Punktepaare $AB$, für die $\overrightarrow{AB} = v$ gilt, können als **Repräsentanten** des Vektors $v$ interpretiert werden. Diese Darstellung entspricht dem Begriff der Pfeilklassen aus der Schulgeometrie. Die Bedingungen $\mathcal{A}$IV liefert dann den Zusammenhang zwischen der Vektoraddition und der Repräsentantenaddition:




**Eigenschaften**

**6.1.1.** $\overrightarrow{AA} = \mathcal{O}$

*Beweis.* Nach ($\mathcal{A}$IV) ist $\overrightarrow{AA} + \overrightarrow{AA} = \overrightarrow{AA}$, so daß die Subtraktion von $\overrightarrow{AA}$ auf beiden Seiten zur Behauptung führt. □

**6.1.2.** Der Punkt $B$ in ($\mathcal{A}$II) ist eindeutig bestimmt.

*Beweis.* Falls $\overrightarrow{AB} = \overrightarrow{AB'} = \overrightarrow{AB} + \overrightarrow{BB'}$ wegen ($\mathcal{A}$IV,) so folgt $\overrightarrow{BB'} = \mathcal{O}$, was wegen ($\mathcal{A}$III) $B = B'$ bedeutet. □

**Definition.** $\dim \mathcal{A} := \dim V$ ist die **Dimension** des affinen Raumes $\mathcal{A}$.

**Beispiele.**

1) Wenn wir den synthetischen Zugang zur Schulgeometrie voraussetzen, passen die Ebene und der Raum zusammen mit ihren geometrischen Vektoren in diese Begriffsbildung.

2) Sei $A\bar{x} = \bar{b}$ ein beliebiges lineares Gleichungssystem wie in Kapitel 3. Wir setzen

$$\mathcal{A} := LM(A\bar{x} = \bar{b}) \subset K^n,\ V := LM(A\bar{x} = \bar{0}) \text{ und } f(\bar{x}, \bar{y}) := \bar{y} - \bar{x}\ .$$

Das Axiom ($\mathcal{A}$I) ist hier wegen Satz 4.2 erfüllt, und es gilt $\dim \mathcal{A} = \dim V = n-$Rang $A$ .

**Aufgabe.** Überprüfen Sie ($\mathcal{A}$II) – ($\mathcal{A}$IV). •

3) Für einen Vektorraum $V$ setzen wir $\mathcal{A} := V$ und $f : V \times V \to V\ ,\ f(u, v) := v-u$.

**Aufgabe.** Überprüfen Sie ($\mathcal{A}$II) – ($\mathcal{A}$IV). •

Demnach läßt sich jeder Vektorraum gleichzeitig auch als affiner Punktraum betrachten.

**Weitere Eigenschaften**

**6.1.3.** $\overrightarrow{AB} = -\overrightarrow{BA}$

*Beweis.* ($\mathcal{A}$IV) und 6.1.1 liefern $\overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \mathcal{O}$. □

**6.1.4.** $\overrightarrow{AB} = \overrightarrow{CD} \Rightarrow \overrightarrow{AC} = \overrightarrow{BD}$

(Dies entspricht der Parallelogrammmbildung in der Schulgeometrie.)

*Beweis.* $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BD} + \overrightarrow{DC} = \overrightarrow{AB} + \overrightarrow{BD} - \overrightarrow{CD} = \overrightarrow{BD}$ wegen ($\mathcal{A}$IV), 6.1.3 und der Voraussetzung. □

Als nächstes wollen wir den Punkten von $\mathcal{A}$ Ortsvektoren zuordnen. Wir zeichnen dafür in $\mathcal{A}$ einen festen Bezugspunkt $O$ aus, den wir auch **Ursprung** nennen.

**Definition.** $f_o(A) := f(O, A) = \overrightarrow{OA}$ heißt **Ortsvektor** des Punktes $A \in \mathcal{A}$.

**6.1.5.** Die Abbildung $f_o : \mathcal{A} \to V$ ist bijektiv.

*Beweis.* Die Surjektivität folgt aus ($\mathcal{A}$II) und die Injektivität aus 6.1.2. □

D.h. also, daß die Punkte von $\mathcal{A}$ mit ihren Ortsvektoren identifiziert werden können, die wiederum ganz $V$ ausschöpfen. Dies rechtfertigt auch den Dimensionsbegriff von oben.



## 6.2 Affine Unterräume (Ebenen)

Für eine Menge $\mathcal{M} \subset \mathcal{A}$ schreiben wir

$$V(\mathcal{M}) := \{\overrightarrow{AB} : A, B \in \mathcal{M}\}\ .$$

**Definition.** $\mathcal{E} \subset \mathcal{A}$ heißt **affiner Unterraum** oder auch **Ebene** in $\mathcal{A}$, falls folgende Bedingungen erfüllt sind:

($\mathcal{E}$I) $V(\mathcal{E})$ ist ein Unterraum von $V$.

($\mathcal{E}$II) $\forall\big(v \in V(\mathcal{E}),\ A \in \mathcal{E}\big)\ \exists B \in \mathcal{E}$ mit $\overrightarrow{AB} = v$.

**6.2.1 Lemma.** *$\mathcal{E}$, $V(\mathcal{E})$ und die Einschränkung von $f$ auf $\mathcal{E}$ liefern wieder eine affine Geometrie.*

*Beweis.* ($\mathcal{E}$I) entspricht ($\mathcal{A}$I), ($\mathcal{E}$II) dem Axiom ($\mathcal{A}$II), und ($\mathcal{A}$III) und ($\mathcal{A}$IV) gelten insbesondere für die Teilmenge $\mathcal{E}$. □

**Bemerkung.** Oft werden nur $k$–dimensionale affine Unterräume auch $k$–dimensionale Ebenen genannt ($k \in \mathbb{N}$), nicht aber unendlichdimensionale.

Sei nun $\mathfrak{A}$ ein beliebiges System von affinen Unterräumen von $\mathcal{A}$.

**6.2.2.** Falls $\mathcal{A}' := \bigcap_{\mathcal{E} \in \mathfrak{A}} \mathcal{E} \neq \emptyset$, so ist $\mathcal{A}'$ ein affiner Unterraum, und es gilt

$$V(\mathcal{A}') = \bigcap_{\mathcal{A} \in \mathfrak{A}} V(\mathcal{E}).$$ •

**Definition.** Für $\mathcal{M} \subset \mathcal{A}$ heißt der kleinste affine Unterraum, der $\mathcal{M}$ enthält, d.h.

$$\bigcap\{\mathcal{E} : \mathcal{E} \text{ ist affiner Unterraum}, \mathcal{M} \subset \mathcal{E}\}$$

der von $\mathcal{M}$ **erzeugte Unterraum**.

**Definition.** Der von $\bigcup_{\mathcal{E} \in \mathfrak{A}} \mathcal{E}$ erzeugte affine Unterraum $\vee\{\mathcal{E} : \mathcal{E} \in \mathfrak{A}\}$ ist der **Verbindungsraum** des Ebenensystems $\mathfrak{A}$.

Im Fall $\mathfrak{A} = \{\mathcal{E}_1, \ldots, \mathcal{E}_m\}$ schreibt man $\mathcal{E}_1 \vee \ldots \vee \mathcal{E}_m$ und für Punkte $A_1, \ldots, A_m \in \mathcal{A}$ auch $A_1 \vee \ldots \vee A_m := \{A_1\} \vee \ldots \vee \{A_m\}$ (vgl. Beispiel (1) unten).

**Beispiele.**

(1) Jede einpunktige Menge $\{A\}$ ist ein nulldimensionaler affiner Unterraum.

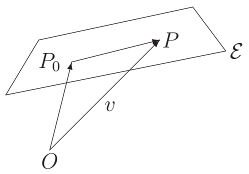
(2) Ein eindimensionaler affiner Unterraum $\mathcal{G}$ heißt **Gerade**. Für zwei verschiedene Punkte $A, B \in \mathcal{G}$ erhält man $\mathcal{G} = A \vee B$. (Offenbar ist $A \vee B \subset \mathcal{G}$, $V(A \vee B) = V(\mathcal{G})$, da beide Räume eindimensional sind, und die Zuordnungen $A \vee B \to V(A \vee B)$ sowie $\mathcal{G} \to V(\mathcal{G})$ gemäß $f_0$ sind bijektiv. □)

(3) $(n-1)$–dimensionale affine Unterräume eines $n$–dimensionalen affinen Raumes nennt man **Hyperebenen**.

Unser nächstes Ziel ist die Charakterisierung von Ebenen mit Hilfe der Menge $f_0(\mathcal{E})$ ihrer Ortsvektoren.

**6.2.3 Lemma.** *Für einen beliebigen Punkt $P_0$ einer Ebene $\mathcal{E}$ gilt $f_o(\mathcal{E}) = V(\mathcal{E}) + \overrightarrow{OP_0}$.*

*Beweis.* Für einen Ortsvektor $v \in f_o(\mathcal{E})$ sei $P \in \mathcal{E}$ der Endpunkt mit $v = \overrightarrow{OP}$. Wir haben $\overrightarrow{OP} = \overrightarrow{OP_0} + \overrightarrow{P_0P} \in V(\mathcal{E}) + \overrightarrow{OP_0}$, da $\overrightarrow{P_0P} \in V(\mathcal{E})$, d.h. $v \in V(\mathcal{E}) + \overrightarrow{OP_0}$.




Umgekehrt, läßt sich jedes $v \in V(\mathcal{E}) + \overrightarrow{OP_0}$ darstellen als $v = w + \overrightarrow{OP_0}$ mit $w \in V(\mathcal{E})$. Wir tragen $w$ an $P_0$ ab und gewinnen den Endpunkt $P \in \mathcal{E}$ mit $w = \overrightarrow{P_0P}$ Dies ergibt $v = \overrightarrow{P_0P} + \overrightarrow{OP_0} = \overrightarrow{OP}$, d.h. $v \in f_o(\mathcal{E})$. □

**6.2.4 Lemma.** *Für eine Menge $\mathcal{E} \subset \mathcal{A}$ sei $f_o(\mathcal{E}) = U + \overrightarrow{OP_0}$ mit einem Unterraum $U \subset V$. Dann folgt $P_0 \in \mathcal{E}$ und $U = V(\mathcal{E})$.*

*Beweis.* Wegen der Darstellung $\overrightarrow{OP_0} = \mathcal{O} + \overrightarrow{OP_0} \in U + \overrightarrow{OP_0}$ ist $\overrightarrow{OP_0}$ Ortsvektor eines Punktes aus $\mathcal{E}$, der wegen 6.1.2 mit $P_0$ zusammenfällt. Außerdem gilt $u \in U \Leftrightarrow u + \overrightarrow{OP_0} \in f_o(\mathcal{E}) \Leftrightarrow u + \overrightarrow{OP_0} = \overrightarrow{OP}$ für ein $P \in \mathcal{E} \Leftrightarrow u = \overrightarrow{P_0O} + \overrightarrow{OP} = \overrightarrow{P_0P}$ für ein $P \in \mathcal{E} \Leftrightarrow u \in V(\mathcal{E})$. □

Zusammenfassend läßt sich damit die Menge der Ortsvektoren einer Ebene wie folgt charakterisieren:

**6.2.5 Satz.** *$\mathcal{E} \subset \mathcal{A}$ ist genau dann eine Ebene, wenn sich $f_o(\mathcal{E})$ als Verschiebung eines Unterraumes darstellen läßt.*

*Beweis.* Die Notwendigkeit der Bedingung liefert 6.2.3. Sei nun $f_o(\mathcal{E}) = U + v$ für einen Unterraum $U$ und ein $v \in V$. Wegen 6.2.4 muß dann $V(\mathcal{E})$ mit $U$ zusammenfallen. $\mathcal{E}$ genügt also der Bedingung ($\mathcal{E}$I). Nach ($\mathcal{A}$I) gibt es für jeden Punkt $A \in \mathcal{E}$ und jeden Vektor $w \in V(\mathcal{E})$ einen Punkt $B \in \mathcal{A}$ mit $\overrightarrow{AB} = w$. Für ($\mathcal{E}$II) bleibt zu zeigen, daß $B$ in $\mathcal{E}$ liegt. Dies schließen wir aus $\overrightarrow{OA} = u + v$ für ein $u \in U$ und $\overrightarrow{OB} = \overrightarrow{OA} + \overrightarrow{AB} = u + v + \overrightarrow{AB} = (u + w) + v \in U + v = f_o(\mathcal{E})$. □

Wir betrachten nun Ebenen durch den Ursprung:



**6.2.6 Satz.** *Eine Menge $\mathcal{E} \subset \mathcal{A}$ ist genau dann eine Ebene mit $O \in \mathcal{E}$, wenn $f_o(V)$ einen Unterraum von $V$ bildet.*

*Beweis.* Für die Notwendigkeit setze man in 6.2.3 $P_0 = O$. Gilt andererseits für einen Unterraum $U$, daß $f_o(\mathcal{E}) = U = U + \overrightarrow{OO}$, so muß wegen 6.2.4 der Ursprung $O$ in $\mathcal{E}$ liegen, und 6.2.5 charakterisiert $\mathcal{E}$ als Ebene. □

**6.2.7 Satz.** *Für jeden Unterraum $U \subset V$ und jeden Punkt $O \in \mathcal{A}$ gibt es genau eine Ebene $\mathcal{E}$ mit $O \in \mathcal{E}$ und $V(\mathcal{E}) = U$.*

*Beweis.* Da $f_o$ bijektiv ist, können wir $\mathcal{E} := f_o^{-1}(U)$ setzen, d.h. $f_o(\mathcal{E}) = U$, und somit ist wegen Satz 6.2.6 die Menge $\mathcal{E}$ eine gesuchte Ebene. Falls für eine weitere Ebene $\mathcal{E}'$ ebenfalls $O \in \mathcal{E}'$ und $V(\mathcal{E}') = U$, so erhält man $V(\mathcal{E}) + \overrightarrow{OO} = V(\mathcal{E}') + \overrightarrow{OO}$ und wegen 6.2.3 $f_o(\mathcal{E}') = f_o(\mathcal{E})$, d.h. $\mathcal{E}' = \mathcal{E}$ aufgrund der Eineindeutigkeit von $f_o$. □

Aus der Schulgeometrie wissen wir, daß durch 3 nichtkollineare Punkte im Raum genau eine Ebene geht. Wir betrachten nun das höherdimensionale Analogon.

**Definition.** Die Punkte $A_0, A_1, \ldots, A_k \in \mathcal{A}$ heißen **linear unabhängig**, falls die Vektoren $\overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_k}$ lineare unabhängig sind.

**Aufgabe.** Zeigen Sie, daß diese Definition nicht von der Numerierung der Punkte abhängt. •

**6.2.8 Satz.** *Durch je $k+1$ linear unabhängige Punkte $A_0, \ldots, A_k$ geht genau eine $k$–dimensionale Ebene $\mathcal{E}$. Es gilt $\mathcal{E} = A_0 \vee \ldots \vee A_k$.*

*Beweis.* Der von der gesuchten Ebene erzeugte Vektorraum $V(\mathcal{E})$ muß $\{\overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_k}\}$ enthalten. Da seine Dimension gleich $k$ sein soll, muß gelten

$$V(\mathcal{E}) = L\Big(\{\overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_k}\}\Big).$$

Satz 6.2.7 für $O := A_0$ und $U := L\Big(\{\overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_k}\}\Big)$ sichert nun Existenz und Eindeutigkeit einer solchen Ebene, denn in diesem Fall ist $f_o(\mathcal{E}) = V(\mathcal{E})$ und damit $A_i \in \mathcal{E}$, für alle $i$. Da $A_0 \vee \ldots \vee A_k$ die kleinste Ebene ist, die $\{A_0, \ldots, A_k\}$ enthält, und ihre Dimension nicht kleiner als $k$ ist, muß sie mit $\mathcal{E}$ zusammenfallen. □

**6.2.9 Folgerung.** *Jede Ebene enthält mit je $l+1$ linear unabhängigen Punkten auch die durch diese Punkte gehende $l$–dimensionale Ebene.* □

## 6.3 Koordinaten im $n$–dimensionalen affinen Raum

### Affine Koordinatensysteme

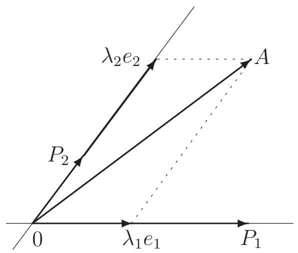
Sei jetzt $(\mathcal{A}, v, f)$ eine ***n*-dimensionale** affine Geometrie.

**Definition.** Jedes System $\{O,\ e_1, \ldots, e_n\}$ für einen Punkt $O \in \mathcal{A}$, der **Koordinatenursprung** genannt wird, und eine Basis $e_1, \ldots, e_n$ in $V$ heißt **affines Koordinatensystem**.

Wenn man den Basisvektor $e_i$ an $O$ abträgt, so geht durch den Endpunkt $P_i$ und den Ursprung genau eine Gerade, die man die ***i*–te Koordinatenachse** nennt. Durch die Basisvektoren ist wie früher eine Koordinatendarstellung in $V$ gegeben. Diese benutzen wir nun zur Bestimmung der Koordinaten eines Punktes $A \in \mathcal{A}$, indem wir zu seinem Ortsvektor $\overrightarrow{OA}$ übergehen:

$$\overrightarrow{OA} = \sum_{i=1}^{n} \lambda_i e_i$$

mit eindeutig bestimmten $\lambda_i \in K$.




**Definition.** Die $\lambda_1, \ldots, \lambda_n$ heißen die **Koordinaten** des Punktes $A$.

(Wie früher schreiben wir $\overline{\lambda}$ für den entsprechenden Spaltenvektor.)

**Parameterdarstellung von Ebenen und Koordinatengleichung**

Sei $\mathcal{E}_k$ eine $k$–dimensionale Ebene in $\mathcal{A}, k < n$. Wir wählen eine Basis $\{b_1, \ldots, b_k\}$ in $V(\mathcal{E}_k)$. Da sich die Menge der Ortsvektoren von $\mathcal{E}_k$ darstellen läßt als $f_o(\mathcal{E}_k) = V(\mathcal{E}_k) + \overrightarrow{OP_0}$ für beliebiges $P_0 \in \mathcal{E}_k$, liegt ein Punkt $P$ genau dann in $\mathcal{E}_k$, wenn für seinen Ortsvektor gilt

$$\boxed{\overrightarrow{OP} = \sum_{i=1}^{k} \mu_i b_i + \overrightarrow{OP_0}} \tag{4}$$

für gewisse Parameter $\mu_i \in K$. Außerdem liefert jede Wahl von $(\mu_1, \ldots, \mu_k) \in K^k$ in (1) einen Punkt $P$ der Ebene $\mathcal{E}_k$. Die $\mu_i$ heißen deshalb auch **freie Parameter**.

**Definition.** Die Gleichung (1) heißt **Parametergleichung** der Ebene $\mathcal{E}_k$.



Als nächstes werden wir (1) in eine Koordinatengleichung umformen. Ein affines Koordinatensystem $\{0, e_1, \ldots, e_n\}$ sei vorgegeben. Den Koordinatenvektor des Punktes $P$ wollen wir mit $\overline{x}$ bezeichnen, den von $P_0$ mit $\overline{y}$, und den des Vektors $b_i$ mit $\overline{b}_i, i = 1, \ldots, k$. Dann lautet (1) in Koordinatenschreibweise

$$\overline{x} = \sum_{i=1}^{k} \mu_i \overline{b}_i + \overline{y}\,,$$

und noch kompakter

$$\overline{x} = B\,\overline{\mu} + \overline{y}\,, \tag{1'}$$

wenn $B$ die $n \times k$–Matrix mit $\overline{b}_1, \ldots, \overline{b}_k$ als Spaltenvektoren bezeichnet. Als Hilfsmittel betrachten wir nun das homogene lineare Gleichungssystem $B^\top \overline{z} = \overline{0}$ mit unbekanntem $\overline{z}$, dessen Lösungsmenge ein Unterraum von $K^n$ ist mit der Dimension

$$n - \operatorname{Rang} B = n - k.$$

Es existieren deshalb $n - k$ linear unabhängige Lösungsvektoren $\overline{z}_1, \ldots, \overline{z}_k$, d.h. $B^\top \overline{z}_i = \overline{0}\ ,\ \forall i$. In Zeilenschreibweise bedeutet dies

$$\overline{z}_i^\top B = \overline{0}^\top\ , \quad i = 1, \ldots, n-k\,.$$

Mit den $\overline{z}_i^\top$ bilden wir die $(n-k) \times n$–Matrix

$$A := \begin{pmatrix} \overline{z}_1^\top \\ \vdots \\ \overline{z}_{n-k}^\top \end{pmatrix}$$

vom Rang $n - k$. Die Konstruktion ergibt

$$AB = \begin{pmatrix} \overline{z}_1^\top \\ \vdots \\ \overline{z}_{n-k}^\top \end{pmatrix} B = \begin{pmatrix} \overline{z}_1^\top B \\ \vdots \\ \overline{z}_{n-k}^\top B \end{pmatrix} = \begin{pmatrix} \overline{0}^\top \\ \vdots \\ \overline{0}^\top \end{pmatrix}\ ,$$

und deshalb ist wegen $(1')$

$$A\overline{x} = AB\overline{\mu} + A\overline{y} = A\overline{y}\,.$$

Wir setzen $\overline{b} := A\overline{y}$ und erhalten

$$\boxed{A\overline{x} = \overline{b}}\,. \tag{5}$$

Die Koordinaten $\overline{x}$ des Punktes $P$ genügen also dieser Gleichung. Aus dem nächsten Satz folgt, daß auch jede Lösung $\overline{x}$ von (2) Koordinatenvektor eines Punktes von $\mathcal{E}_k$ ist.

**6.3.1 Satz.** *Jede Parametergleichung der Form (1) mit linear unabhängigen Vektoren $b_1, \ldots, b_k$ bzw. jedes lineare Gleichungssystem der Form (2) mit* $\mathrm{Rang}(A, \overline{b}) = \mathrm{Rang} A = n - k$ *bestimmt eindeutig eine $k$–dimensionale Ebene.*

*Beweis.* Sei $\mathcal{E}$ die Menge der Punkte $P$, die (1) genügen. Dann gilt $f_o(\mathcal{E}) = L(\{b_1, \ldots, b_k\}) + \overrightarrow{OP_0}$, und wegen Satz 6.2.5 ist $\mathcal{E}$ eine Ebene.
Weiter wissen wir aus dem Kapitel über lineare Gleichungssysteme, daß $LM(A\overline{x} = \overline{b}) = U + \overline{v}$, wobei $U = LM(A\overline{x} = \overline{0})$ ein linearer Unterraum von $K^n$ der Dimension $k$ ist und $\overline{v}$ eine spezielle Lösung von (2). Jeder Lösungsvektor $\overline{x}$ von (2) läßt sich also darstellen als $\overline{x} = \sum_{i=1}^{k} \mu_i \overline{b}_i + \overline{v}$ für eine Basis $\{b_1, \ldots, b_k\}$ in $V$, und alle $\overline{\mu} \in K^k$ liefern Lösungsvektoren $\overline{x}$. Wir interpretieren nun $\overline{x}, \overline{v}$ und $\overline{b}_1, \ldots, \overline{b}_k$ als Koordinaten von Vektoren $x, v$ und $b_1, \ldots, b_k$ in $V$ bzgl. der Basis $\{e_1, \ldots, e_n\}$ und erhalten $x = \sum_{i=1}^{k} \mu_i b_i + v$. Schließlich definieren wir die Punkte $P$ und $P_0$ durch $\overrightarrow{OP} = x$ bzw. $\overrightarrow{OP_0} = v$ und kommen zu $\overrightarrow{OP} = \sum_{i=1}^{k} \mu_i b_i + \overrightarrow{OP_0}$. Dies ist wiederum die Parametergleichung einer $k$–dimensionalen Ebene. □

**Bemerkung.** Im Spezialfall der Hyperebene $\mathcal{H}$, wo $\dim \mathcal{H} = n - 1$ gilt, hat die Koordinatengleichung (2) die Form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b\ ,$$

mit mindestens einem von 0 verschiedenen $a_i$. Eine Gerade $\mathcal{G}$ ist durch $n-1$ Gleichungen dieser Art gegeben, wobei $\mathrm{Rang} A = \mathrm{Rang}(A, \overline{b}) = n - 1$.

## 6.4 Parallelität und Schnitte von Ebenen

Der Begriff der Parallelität spielt eine wichtige Rolle in der affinen Geometrie. Wir kommen darauf zurück, wenn wir die entsprechenden Abbildungen behandeln. Hier betrachten wir folgendes:

**Definition.** Zwei Ebenen eines affinen Raumes heißen **parallel**, falls einer der beiden erzeugten Vektorräume den anderen enthält.

**6.4.1 Satz.** *Die durch die Koordinatengleichungen $A\overline{x} = \overline{b}$ und $A'\overline{x} = \overline{b'}$ gegebenen Ebenen $\mathcal{E}$ bzw. $\mathcal{E}'$ sind genau dann parallel, wenn die Lösungsmenge eines der beiden zugehörigen homogenen Gleichungssysteme die des anderen enthält.*

*Beweis.* Aus dem Zusammenhang zwischen Parameterdarstellung und Koordinatengleichungen der Ebenen ergibt sich, daß $V(\mathcal{E})$ in Koordinaten zu $LM(A\overline{x} = \overline{0})$ gehört und $V(\mathcal{E}')$ zu $LM(A'\overline{x} = \overline{0})$. □

**Aufgabe.** Seien $\mathcal{E}$ und $\mathcal{E}'$ parallele Ebenen mit nichtleerem Durchschnitt und $\dim \mathcal{E} \leq \dim \mathcal{E}'$. Dann ist $\mathcal{E}$ in $\mathcal{E}'$ enthalten. •



Wir erinnern, daß der nichtleere Durchschnitt zweier Ebenen $\mathcal{E}$ und $\mathcal{E}'$ stets wieder eine Ebene bildet und $V(\mathcal{E} \cap \mathcal{E}') = V(\mathcal{E}) \cap V(\mathcal{E}')$. Für den Verbindungsraum $\mathcal{E} \vee \mathcal{E}'$ kann man folgendes zeigen.

**6.4.2 Satz.** $\quad \mathcal{E} \cap \mathcal{E}' \neq \emptyset \Leftrightarrow V(\mathcal{E} \vee \mathcal{E}') = V(\mathcal{E}) + V(\mathcal{E}')$.

*Beweis.* Unter der Voraussetzung $V(\mathcal{E} \vee \mathcal{E}') = V(\mathcal{E}) + V(\mathcal{E}')$ findet man für $A \in \mathcal{E}$ und $A' \in \mathcal{E}'$ die Darstellung $\overrightarrow{AA'} = v + v'$ mit $v \in V(\mathcal{E})$ und $v' \in V(\mathcal{E}')$. Der Punkt $O \in \mathcal{E}$ sei durch $\overrightarrow{AO} = v$ bestimmt. Aus $\overrightarrow{AA'} = \overrightarrow{AO} + \overrightarrow{OA'}$ folgt $\overrightarrow{OA'} = v'$, d.h. $\overrightarrow{A'O} = -v' \in V(\mathcal{E}')$. Somit gilt auch $O \in \mathcal{E}'$, so daß $O \in \mathcal{E} \cap \mathcal{E}'$.
Sei umgekehrt $O$ ein Punkt aus $\mathcal{E} \cap \mathcal{E}'$. Wir gehen in zwei Schritten vor:

1) Sei $w \in V(\mathcal{E}) + V(\mathcal{E}')$, d.h. $w = v + v'$ mit $v \in V(\mathcal{E}),\ v' \in V(\mathcal{E}')$. Die Punkte $A \in \mathcal{E}$ und $A' \in \mathcal{E}'$ bestimmen wir durch $-v = \overrightarrow{OA},\ v' = \overrightarrow{OA'}$ und bekommen $w = -\overrightarrow{OA} + \overrightarrow{OA'} = \overrightarrow{AO} + \overrightarrow{OA'} = \overrightarrow{AA'} \in V(\mathcal{E} \vee \mathcal{E}')$. Also haben wir

$$V(\mathcal{E}) + V(\mathcal{E}') \subset V(\mathcal{E} \vee \mathcal{E}')\ .$$

2) Wir betrachten die Ebene $\mathcal{E}'' := f_0^{-1}\big(V(\mathcal{E}) + V(\mathcal{E}')\big)$ mit $O \in \mathcal{E}''$ und $V(\mathcal{E}'') = V(\mathcal{E}) + V(\mathcal{E}')$. Offenbar gilt $\mathcal{E} \cup \mathcal{E}' \subset \mathcal{E}''$, demzufolge $\mathcal{E} \vee \mathcal{E}' \subset \mathcal{E}''$ und damit $V(\mathcal{E} \vee \mathcal{E}') \subset V(\mathcal{E}'')$, d.h.

$$V(\mathcal{E} \vee \mathcal{E}') \subset V(\mathcal{E}) + V(\mathcal{E}')\ .$$

1) und 2) ergeben $V(\mathcal{E} \vee \mathcal{E}') = V(\mathcal{E}) + V(\mathcal{E}')$ $\quad$ (und $\mathcal{E}'' = \mathcal{E} \vee \mathcal{E}'$). □

Unmittelbar daraus resultiert der folgende Dimensionssatz.

**6.4.3 Satz.** *Für zwei sich schneidende Ebenen $\mathcal{E}$ und $\mathcal{E}'$ gilt*

$$\dim(\mathcal{E} \vee \mathcal{E}') = \dim \mathcal{E} + \dim \mathcal{E}' - \dim(\mathcal{E} \cap \mathcal{E}').$$

*Beweis.* $\dim(\mathcal{E} \vee \mathcal{E}') = \dim V(\mathcal{E} \vee \mathcal{E}') = \dim\big(V(\mathcal{E}) + V(\mathcal{E}')\big) = \dim V(\mathcal{E}) + \dim V(\mathcal{E}') - \dim\big(V(\mathcal{E}) \cap V(\mathcal{E}')\big) = \dim \mathcal{E} + \dim \mathcal{E}' - \dim(\mathcal{E} \cap \mathcal{E}')$, wobei der Dimensionssatz für die Summe von Unterräumen benutzt wurde. □

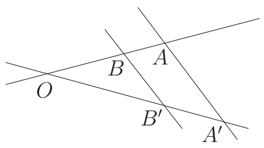
**Definition.** Zwei Ebenen heißen **windschief**, wenn sie nicht parallel sind und sich nicht schneiden.

**Aufgabe.**

1) Eine Ebene $\mathcal{E}$ und eine Hyperebene $\mathcal{H}$ sind nicht windschief.

2) Falls $\mathcal{E}$ und $\mathcal{H}$ nicht parallel sind, so gilt $\dim(\mathcal{E} \cap \mathcal{H}) = \dim \mathcal{E} - 1$. •

**Aufgabe.** Beweisen Sie den **Strahlensatz**:




$O, A, A'$ seien linear unabhängige Punkte, $B \in O \vee A,\ B' \in O \vee A'$ und $B \vee B'$ parallel zu $A \vee A'$. Dann gilt

$$\overrightarrow{OB} = \lambda \overrightarrow{OA} \qquad \text{und} \qquad \overrightarrow{OB'} = \lambda \overrightarrow{OA'}$$

für ein gemeinsames Teilverhältnis $\lambda \in K$. •

## 6.5 Affine Abbildungen

Wir suchen nun Abbildungen zwischen affinen Räumen, die mit linearen Abbildungen zwischen den erzeugten Vektorräumen korrespondieren. Dabei gehen wir wie folgt vor:

$(\mathcal{A}, V, f)$ und $\mathcal{A}', V', f')$ seien affine Geometrien über demselben Körper $K$.

**Definition.** Eine Abbildung $\mathcal{F} : \mathcal{A} \to \mathcal{A}'$ heißt **affin**, falls aus $\overrightarrow{AB} = \lambda \overrightarrow{CD}$ folgt $\overrightarrow{\mathcal{F}(A)\mathcal{F}(B)} = \lambda \overrightarrow{\mathcal{F}(C)\mathcal{F}(D)}$.
Eine bijektive affine Abbildung wird **Affinität** genannt.

Ehe wir Affinitäten genauer untersuchen, wollen wir einen wichtigen Spezialfall behandeln.

**Definition.** Eine Abbildung $\mathcal{T} : \mathcal{A} \to \mathcal{A}$ mit $\overrightarrow{\mathcal{T}(A)\mathcal{T}(B)} = \overrightarrow{AB}$ heißt **Translation** oder **Parallelverschiebung**.

**6.5.1 Lemma.** *Jede Translation $\mathcal{T}$ ist eine Affinität.*

*Beweis.* $\mathcal{T}$ ist offenbar affin.
Für die Injektivität setzen wir $\mathcal{T}(A) = \mathcal{T}(B)$, wählen ein beliebiges $O \in \mathcal{A}$ und erhalten $\overrightarrow{OA} = \overrightarrow{\mathcal{T}(O)\mathcal{T}(A)} = \overrightarrow{\mathcal{T}(O)\mathcal{T}(B)} = \overrightarrow{OB}$, d.h. $A = B$.
Um zu zeigen, daß $\mathcal{T}$ surjektiv ist tragen wir für ein beliebiges $A' \in \mathcal{A}$ den Vektor $\overrightarrow{\mathcal{T}(A')A'}$ an $A'$ ab. Der Endpunkt sei $A$. Dann ist $\overrightarrow{\mathcal{T}(A')A'} = \overrightarrow{A'A} = \overrightarrow{\mathcal{T}(A')\mathcal{T}(A)}$ und demnach $\mathcal{T}(A) = A'$. □

**Aufgabe.** Die Menge aller Translationen in $\mathcal{A}$ bildet eine Gruppe. •

Der folgende Satz beinhaltet, daß die Translationsgruppe mit $V$ identifiziert werden kann.



**6.5.2 Lemma.** *$\mathcal{T} : \mathcal{A} \to \mathcal{A}$ ist genau dann eine Translation, wenn es einen (eindeutig bestimmten) Vektor $v \in V$ gibt mit $\overrightarrow{O\mathcal{T}(A)} = \overrightarrow{OA} + v$, bei festem $O \in \mathcal{A}$.*

*Beweis.* Für $A, B \in \mathcal{A}$ gilt

$$\overrightarrow{AB} = \overrightarrow{\mathcal{T}(A)\mathcal{T}(B)} \Leftrightarrow \overrightarrow{AO} + \overrightarrow{OB} = \overrightarrow{\mathcal{T}(A)O} + \overrightarrow{O\mathcal{T}(B)} \Leftrightarrow$$
$$\overrightarrow{O\mathcal{T}(A)} - \overrightarrow{OA} = \overrightarrow{O\mathcal{T}(B)} - \overrightarrow{OB} =: v \Leftrightarrow \overrightarrow{O\mathcal{T}(A)} = \overrightarrow{OA} + v$$

für ein (eindeutig bestimmtes) $v \in V$. □

Jeder Affinität $\mathcal{F} : \mathcal{A} \to \mathcal{A}'$ ordnen wir nun die Abbildung $\mathbf{F} : V \to V'$ zu mit $\mathbf{F}v = v'$, falls $v = \overrightarrow{AB}$ und $v' = \overrightarrow{\mathcal{F}(A')\mathcal{F}(B')}$.

**Bemerkung.** Die Definition von $\mathbf{F}$ ist korrekt, denn aus $\overrightarrow{AB} = \overrightarrow{CD}$ folgt $\overrightarrow{\mathcal{F}(A)\mathcal{F}(B)} = \overrightarrow{\mathcal{F}(C)\mathcal{F}(D)}$.

**6.5.3 Satz.** (1) $\mathbf{F}$ *ist linear und mit $\mathcal{F}$ surjektiv bzw. injektiv, für Affinitäten $\mathcal{F}$ also ein Vektorraumisomorphismus.*

(2) *Wenn $\mathcal{F}_1$ eine weitere Affinität ist mit $\mathbf{F}_1 = \mathbf{F}$, so gibt es eine Translation $\mathcal{T}$ in $\mathcal{A}$ mit*

$$\mathcal{F}_1 = \mathcal{T} \circ \mathcal{F}\,.$$

*Beweis.* **(1) F** is *linear*: Für $v, w \in V$ sei $v = \overrightarrow{AB},\ w = \overrightarrow{BC}$ uns somit $v + w = \overrightarrow{AC}$.
Dann gilt $\mathbf{F}(v + w) = \overrightarrow{\mathcal{F}(A)\mathcal{F}(C)} = \overrightarrow{\mathcal{F}(A)\mathcal{F}(B)} + \overrightarrow{\mathcal{F}(B)\mathcal{F}(C)} = \mathbf{F}v + \mathbf{F}w$.
Aus $\lambda u = \overrightarrow{AB}$ und $u = \overrightarrow{CD}$ folgt nach Definition $\overrightarrow{\mathcal{F}(A)\mathcal{F}(B)} = \lambda \overrightarrow{\mathcal{F}(C)\mathcal{F}(D)}$, d.h. $\mathbf{F}(\lambda u) = \lambda \mathbf{F}u$.
$\mathbf{F}$ ist mit $\mathcal{F}$ *surjektiv*: Sei $\overrightarrow{A'B'} = v' \in V'$ beliebig. Wegen der Surjektivität vom $\mathcal{F}$ gibt es Punkte $A, B \in \mathcal{A}$ mit $\mathcal{F}(A) = A'$, $\mathcal{F}(B) = B'$. Wir setzen $v := \overrightarrow{AB}$ und erhalten $\mathbf{F}v = \overrightarrow{\mathcal{F}(A)\mathcal{F}(B)} = \overrightarrow{A'B'} = v'$.
$\mathbf{F}$ ist mit $\mathcal{F}$ *injektiv*: Sei $\mathbf{F}u = \mathbf{F}v$ und $u = \overrightarrow{AB}, v = \overrightarrow{AC}$. Dann ist $\overrightarrow{\mathcal{F}(A)\mathcal{F}(B)} = \overrightarrow{\mathcal{F}(A)\mathcal{F}(C)}$, d.h. $\mathcal{F}(B) = \mathcal{F}(C)$, und wegen der Injektivität von $\mathcal{F}$ folgt $B = C$, also auch $u = v$.

**(2)** Wir definieren die Translation $\mathcal{T} : \mathcal{A}' \to \mathcal{A}'$ durch

$$\overrightarrow{\mathcal{F}(O)\mathcal{T}(A')} = \overrightarrow{\mathcal{F}(O)A'} + \overrightarrow{\mathcal{F}(O)\mathcal{F}_1(O)}\,,\ A' \in \mathcal{A}' \ \text{ für ein } O \in \mathcal{A}.$$

Dann gilt

$$\begin{aligned}
\overrightarrow{\mathcal{F}(O)\mathcal{T}(\mathcal{F}(A))} &= \overrightarrow{\mathcal{F}(O)\mathcal{F}(A)} + \overrightarrow{\mathcal{F}(O)\mathcal{F}_1(O)} \\
&= \overrightarrow{\mathcal{F}_1(O)\mathcal{F}_1(A)} + \overrightarrow{\mathcal{F}(O)\mathcal{F}_1(O)} \\
&= \mathbf{F}\overrightarrow{OA} + \overrightarrow{\mathcal{F}(O)\mathcal{F}_1(O)} \\
&= \overrightarrow{\mathcal{F}_1(O)\mathcal{F}_1(A)} + \overrightarrow{\mathcal{F}(O)\mathcal{F}_1(O)} \\
&= \overrightarrow{\mathcal{F}(O)\mathcal{F}_1(A)}
\end{aligned}$$

und demnach

$$\mathcal{T}(\mathcal{F}(A)) = \mathcal{F}_1(A).$$

□

**Koordinatendarstellung von affinen Abbildungen zwischen endlichdimensionalen Räumen**

Seien $(O, \{e_1, \ldots, e_n\})$ und $(O', \{e'_1, \ldots, e'_{n'}\})$ affine Koordinatensysteme in $(\mathcal{A}, V)$ bzw. $(\mathcal{A}', V')$. Für eine Affinität $\mathcal{F} : \mathcal{A} \to \mathcal{A}'$ betrachten wir die Ortsvektoren

$$\overrightarrow{O'\mathcal{F}(A)} = \overrightarrow{O'\mathcal{F}(O)} + \overrightarrow{\mathcal{F}(O)\mathcal{F}(A)} = \mathbf{F}\overrightarrow{OA} + \overrightarrow{O'\mathcal{F}(O))} .$$

Wir benutzen folgende Bezeichnungen:

| | |
|---|---|
| $F \in M(n' \times n)$ | Matrix von $\mathbf{F}$. |
| $\overline{x} \in K^n$ | Koordinaten des Punktes $A$, |
| $\overline{y} \in K^{n'}$ | Koordinaten des Punktes $\mathcal{F}(A)$, |
| $\overline{c} \in K^{n'}$ | Koordinaten des Punktes $\mathcal{F}(O)$. |

Die obenstehende Beziehung liefert dann die **Koordinatendarstellung** der Abbildung $\mathcal{F}$:

$$\boxed{\overline{y} = F\overline{x} + \overline{c}} .$$

Für $n = n'$ und Affinitäten $\mathcal{F}$ gilt $\det F \neq 0$.

**6.5.4 Satz.** *Jede Gleichung der obigen Form kann als Koordinatengleichung einer affinen Abbildung interpretiert werden. (Für $n = n'$ und $\det F \neq 0$ ist diese eine Affinität.)*

*Beweis.* Die Matrix $F$ (mit $\det F \neq 0$) gehört zu einer linearen Abbildung (einem Isomorphismus) $\mathbf{F} : V \to V'$. Wir definieren die Abbildung $\mathcal{F} : \mathcal{A} \to \mathcal{A}'$ durch $\mathcal{F}(A) = A'$, falls $\overrightarrow{O'A'} = \mathbf{F}\overrightarrow{OA} + c$, wobei $c \in V'$ die Koordinaten $\overline{c}$ besitzt.

1) $\overrightarrow{\mathcal{F}(A)\mathcal{F}(B)} = \overrightarrow{A'B'} = -\overrightarrow{O'A'} + \overrightarrow{O'B'} = -\mathbf{F}\overrightarrow{OA} - c + \mathbf{F}\overrightarrow{OB} + c = \mathbf{F}\big(-\overrightarrow{OA} + \overrightarrow{OB}\big) = \mathbf{F}\overrightarrow{AB}$, d.h. $\overrightarrow{\mathcal{F}(A)\mathcal{F}(B)} = \mathbf{F}\overrightarrow{AB}$, insbesondere ist $\mathcal{F}$ dann **affin**. Nach Konstruktion hat $\mathcal{F}$ die gewünschte Koordinatendarstellung.



2) $\mathcal{F}$ ist mit $\mathbf{F}$ **injektiv**: Aus $\mathcal{F}(A) = \mathcal{F}(B)$ folgt $\mathbf{F}\overrightarrow{OA} = \mathbf{F}\overrightarrow{OB}$ und wegen der Injektivität von $\mathbf{F}$ auch $\overrightarrow{OA} = \overrightarrow{OB}$, d.h. $A = B$.

3) $\mathcal{F}$ ist mit $\mathbf{F}$ **surjektiv**: Da $\mathbf{F}$ surjektiv ist, gibt es für jedes $A' \in \mathcal{A}'$ ein $v \in V$ mit $\mathbf{F}v = \overrightarrow{O'A'} - c$. Wir definieren $A$ durch $\overrightarrow{OA} = v$ und schließen aus $\mathbf{F}\overrightarrow{OA} + c = \overrightarrow{O'A'}$ auf $\mathcal{F}(A) = A'$.

□

**Einige geometrische Eigenschaften von Affinitäten**

Die affine Geometrie als mathematische Theorie wird oft auch als die Lehre von den geometrischen Eigenschaften der Objekte des affinen Raumes, die bei Affinität erhalten bleiben, betrachtet. Einige davon wollen wir aufzählen.

**6.5.5 Satz.** *Affinitäten $\mathcal{F} : \mathcal{A} \to \mathcal{A}'$ sind **ebenentreu**, d.h. für jede Ebene $\mathcal{E} \subset \mathcal{A}$ und $\mathcal{E}' := \mathcal{F}(\mathcal{E})$ gilt*

(1) *$\mathcal{E}'$ ist eine Ebene in $\mathcal{A}'$,*

(2) *$V(\mathcal{E}') = \mathbf{F}(V(\mathcal{E}))$,*

(3) *$\dim \mathcal{E}' = \dim \mathcal{E}$.*

*Beweis.* Wir benutzen die Darstellungen aus Abschnitt 5.2. Für $O \in \mathcal{E}$ ist die Menge der Ortsvektoren von $\mathcal{E}$ gleich $f_o(\mathcal{E}) = V(\mathcal{E})$. Sei $O' := \mathcal{F}(O)$. In diesem Fall ist

$$\begin{aligned}
f'_{o'}(\mathcal{E}) &= \big\{\overrightarrow{O'A'} : A' \in \mathcal{E}'\big\} = \big\{\overrightarrow{\mathcal{F}(O)\mathcal{F}(A)} : A \in \mathcal{E}\big\} = \big\{\mathbf{F}\overrightarrow{OA} : A \in \mathcal{E}\big\} \\
&= \mathbf{F}\big(f_o(\mathcal{E})\big) = \mathbf{F}V(\mathcal{E}) .
\end{aligned}$$

Also bildet $f_{o'}(\mathcal{E})$ einen Unterraum von $V'$ und $O' \in \mathcal{E}'$. Deshalb ist $\mathcal{E}'$ eine Ebene und $f'_{O'}(\mathcal{E}') = V'(\mathcal{E}')$, was (1) und (2) beweist. Schließlich gilt $\dim \mathcal{E}' = \dim V'(\mathcal{E}') = \dim \mathbf{F}(V(\mathcal{E})) = \dim V(\mathcal{E}) = \dim \mathcal{E}$. □

Insbesondere sind Affinitäten geradentreu.
Als nächstes betrachten wir nun den Fall $K = \mathbb{R}$.
Seien $A, P, B$ Punkte auf einer Geraden.

**Definition.**

(1) Das **Teilverhältnis** $\tau = \tau(A, P, B) \in \mathbb{R}$ bestimmen wir durch $\overrightarrow{AP} = \tau\overrightarrow{PB}$.

(2) $P$ liegt **zwischen** $A$ und $B$, falls $\tau(A, P, B) > 0$.

(3) Die **Strecke** $AB$ ist die Menge aller Punkte der Geraden $A \vee B$, die zwischen $A$ und $B$ liegen, vereinigt mit den Endpunkten $A$ und $B$.

**6.5.6 Lemma.** *Wenn die Punkte $A, P$ und $B$ auf einer Geraden liegen und $P$ zwischen $A$ und $B$, so gilt das Gleiche für entsprechende Bildpunkte unter einer Affinität.*

*Beweis.* Jede Affinität ist geradentreu und erhält das Teilverhältnis. □

**6.5.7 Folgerung.** *Affinitäten sind* **streckentreu**. □

Zuletzt sei $K$ wieder beliebig.

**6.5.8 Lemma.** *Affinitäten sind* **parallelentreu**.

*Beweis.* Seien $\mathcal{E}_1$ und $\mathcal{E}_2$ parallele Ebenen in $\mathcal{A}$ mit $V(\mathcal{E}_1) \subset V(\mathcal{E}_2)$ und $\mathcal{F} : \mathcal{A} \to \mathcal{A}'$ eine Affinität. Für den zugehörigen Vektorraumisomorphismus $\mathbf{F}$ gilt dann $V\big(\mathcal{F}(\mathcal{E}_1)\big) = \mathbf{F}(V(\mathcal{E}_1)) \subset \mathbf{F}(V(\mathcal{E}_2)) = V\big(\mathcal{F}(\mathcal{E}_1)\big)$, d.h. $\mathcal{F}(\mathcal{E}_1)$ und $\mathcal{F}(\mathcal{E}_2)$ sind ebenfalls parallel. □



# 7 Euklidische Räume und ihre Abbildungen

## 7.1 Skalarprodukt Norm, Winkel, Metrik

Die aus der euklidischen Schulgeometrie bekannten Objekte wie Punkte, Geraden, Ebenen und Vektoren haben wir bereits verallgemeinert und ihre linearen und affinen Eigenschaften behandelt. Unser Ziel besteht nun darin, solche Begriffe wie Länge von Strecken und Vektoren, Winkel zwischen Geraden, Ebenen und Vektoren, das Prinzip der Orthogonalität, Volumen u.ä. mathematisch streng einzuführen. Wir brauchen dazu:

1) eine Metrik im Punktraum für den Abstand,

2) eine Norm im Vektorraum für die Länge,

3) ein Skalarprodukt im Vektorraum für die Winkel, das eng mit der Metrik und der Norm verbunden ist.

Es hat sich als günstig erwiesen, den Begriff des Skalarprodukts an den Anfang zu stellen: Ausgangspunkt ist eine reelle affine Geometrie $(\mathcal{A}, V, f)$ oder nur ein linearer Raum $V$ über dem Körper $\mathbb{R}$.

**Definition.** Ein **Skalarprodukt** in $V$ ist eine positiv definite symmetrische Bilinearform $B : V \times V \to \mathbb{R}$, d.h.

1) $B$ ist bilinear

2) $B(u, v) = B(v, u)$

3) $B(v, v) > 0$ für $v \neq 0$.

Üblich sind folgende Kurzschreibweisen: $B(u, v) = (u, v)$, $B(u, v) = \langle u, v\rangle$ oder $B(u, v) = u \cdot v$. Wir wollen hier die eckigen Klammern benutzen.

**Bemerkung.** Aus der Bilinearität folgt sofort $B(\mathcal{O}, v) = B(v, \mathcal{O}) = 0,\ \forall v \in V$.

**7.1.1 Satz.** *In jedem endlichdimensionalen Vektorraum über $\mathbb{R}$ existiert ein Skalarprodukt.*

*Beweis.* Wir wählen irgendeine Basis $b_1, \ \ldots, \ b_n$ in $V$ und bezeichnenen die Koordinaten von $u, v \in V$ mit $u_1, \ \ldots, \ u_n$ bzw. $v_1, \ \ldots, \ v_n$. Dann genügt

$$B(u, v) := \sum_{i=1}^{n} u_i\, v_i$$

allen Bedingungen an ein Skalarprodukt. □

**Bemerkung.** Aus dem obenstehenden Beweis ist ersichtlich, daß es in $V$ verschiedene Skalarprodukte gibt. Weiter unten werden wir diese beschreiben.

**Definition.** Ein endlichdimensionaler Vektorraum, versehen mit einem Skalarprodukt $\langle\cdot,\cdot\rangle$, heißt **euklidischer Vektorraum**. Im $n$–dimensionalen Fall schreibt man auch $\mathbb{E}_n$.

**Definition.** Ein **euklidischer Punktraum** ist ein affiner Raum, dessen erzeugter Vektorraum euklidisch ist.

Die folgende Eigenschaft eines Skalarproduktes in $V$ ist grundlegend für alles weitere.

**7.1.2 Satz** (Cauchy–Schwarzsche Ungleichung)**.**

(1) $\langle u,v\rangle^2 \le \langle u,u\rangle\,\langle v,v\rangle$

(2) *Die Gleichheit in* (1) *gilt genau dann, wenn $u$ und $v$ kollinear sind.*

*Beweis.* Falls $v=\mathcal{O}$ ist, so gilt die Gleichheit in (1), und $u$ und $v$ sind linear unabhängig. Sei nun $v\neq\mathcal{O}$. Wir setzen $\lambda := \frac{\langle u,v\rangle}{\langle v,v\rangle}$ und erhalten aus 1) – 3)

$$\begin{aligned}0 &\le \langle u-\lambda v,\ u-\lambda v\rangle = \langle u,u\rangle + \lambda^2\langle v,v\rangle - 2\lambda\langle u,v\rangle\\ &= \langle u,u\rangle + \frac{\langle u,v\rangle^2}{\langle v,v\rangle^2}\,\langle v,v\rangle - 2\frac{\langle u,v\rangle}{\langle v,v\rangle}\,\langle u,v\rangle = \langle u,u\rangle - \frac{\langle u,v\rangle^2}{\langle v,v\rangle}\end{aligned}$$

d.h. $\langle u,v\rangle^2 \le \langle u,u\rangle\langle v,v\rangle$. Wenn hier die Gleichheit gilt, dann folgt $\langle u-\lambda v,\ u-\lambda v\rangle = 0$, also $u-\lambda v=\mathcal{O}$. Ist umgekehrt $u=\mu v$ für ein $\mu\in\mathbb{R}$, so ergibt sich $\langle u,v\rangle^2 = \langle \mu v,v\rangle^2 = \mu^2\langle v,v\rangle^2 = \langle \mu v,\ \mu v\rangle\langle v,v\rangle = \langle u,u\rangle\langle v,v\rangle$, d.h. die gewünschte Gleichheit. □

Mit Hilfe des Skalarprodukts $\langle\cdot,\cdot\rangle$ in $V$ führen wir nun die Länge von Vektoren ein:

**Definition.** $\boxed{|v| := \sqrt{\langle v,v\rangle}}$ heißt **Betrag** oder **Länge** des Vektors $v\in V$.

**7.1.3 Satz.** *Die Länge von Vektoren bestimmt eine Norm im Vektorraum, d.h.*

(1) $|v|>0$ *für* $v\neq\mathcal{O}$ *(positive Definitheit)*

(2) $|\lambda v| = |\lambda|\,|v|$, $\lambda\in\mathbb{R}$, *(positive Linearität)*

(3) $|u+v| \le |u|+|v|$ *(Dreiecksungleichung)*

**Bemerkung.**

1) Aus (2) folgt $|\mathcal{O}|=0$.

2) Ein $v\in V$ mit $|v|=1$ heißt **normierter** oder **Einheitsvektor**.



*Beweis vom Satz 7.1.3.* **(1)** resultiert aus der positiven Definitheit des Skalarproduktes und **(2)** aus der entsprechenden Eigenschaft: $|\lambda v| = \sqrt{\langle\lambda v,\ \lambda v\rangle} = \sqrt{\lambda^2\langle v,v\rangle} = |\lambda|\,|v|$. Die Schwarzsche Ungleichung lautet in der Sprache der Beträge $|\langle u,v\rangle| \le |u|\,|v|$ und führt zu folgender Abschätzung:

$$\begin{aligned}|u+v|^2 &= \langle u+v,\ u+v\rangle = \langle u,u\rangle + \langle v,v\rangle + 2\langle u,v\rangle\\ &= |u|^2+|v|^2+2\langle u,v\rangle \le |u|^2+|v|^2+2|u|\,|v| = \big(|u|+|v|\big)^2.\end{aligned}$$

Dies ergibt aber die Dreiechsungleichung **(3)**. □

**Aufgabe.** In der Dreiecksungleichung gilt genau dann die Gleichheit, wenn $u$ und $v$ gleichgerichtet sind. •

Für den Winkelbegriff benutzen wir ebenfalls das Skalarprodukt. Eine weitere Konsequenz aus der Chauchy–Schwarzschen Ungleichung ist nämlich wegen

$$-1 \le \frac{\langle u,v\rangle}{|u|\,|v|} \le 1 \quad \text{für} \quad u,v\neq\mathcal{O}$$

die Existenz eines eindeutig bestimmten $\alpha\in[0,\pi]$ mit

$$\boxed{\cos\alpha = \frac{\langle u,v\rangle}{|u|\,|v|}}$$

**Definition.** $\alpha$ heißt **Winkel** zwischen den Vektoren $u$ und $v$. Man schreibt auch $\alpha = \sphericalangle(u,v)$.

**Bemerkung.** Der Begriff des Winkels zwischen Geraden läßt sich darauf zurückführen, wobei dann stets zwei Winkel auftreten.

Schließlich wollen wir noch den Abstand zwischen den Punkten eines affinen Raumes $\mathcal{A}$ mit Skalarprodukt $\langle\cdot,\cdot\rangle$ im erzeugten Vektorraum $V$ bestimmen:

**Definition.** $\varrho(A,B) := |\overrightarrow{AB}|$ heißt **Abstand** zwischen den Punkten $A,B\in\mathcal{A}$ oder auch **Länge der Stecke** $AB$.

**7.1.4 Satz.** *Durch die Abbildung $\varrho:\mathcal{A}\times\mathcal{A}\to\mathbb{R}$ ist eine Metrik in $\mathcal{A}$ gegeben, d.h.*

(1) $\varrho(A,B)\ge 0$, *und* $\varrho(A,B)=0 \Leftrightarrow A=B$ *(positive Definitheit)*

(2) $\varrho(A,B)=\varrho(B,A)$ *(Symmetrie)*

(3) $\varrho(A,C)\le\varrho(A,B)+\varrho(B,C)$ *(Dreiecksungleichung).*

*Beweis.* **(1)** Folgt aus (1) für die Norm und $|\mathcal{O}|=0$. Für **(2)** bemerken wir, daß $|\overrightarrow{AB}| = |-\overrightarrow{BA}| = |-1|\,|\overrightarrow{BA}| = |\overrightarrow{BA}|$, und **(3)** ergibt sich aus (3) für die Norm wegen $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$. □

## 7.2 Orthogonalität

Sei $V$ wieder ein Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$.

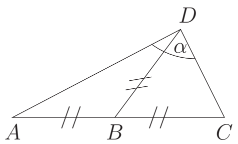
**Definition.** Die Vektoren $u, v \in V$ heißen **orthogonal**, falls $\langle u,v\rangle = 0$.

Insbesondere ist der Nullvektor orthogonal zu jedem anderen. Für $u, v \neq \mathcal{O}$ haben wir $\langle u,v\rangle = |u|\,|v| \cos \sphericalangle(u,v)$. Also sind $u$ und $v$ genau dann orthogonal, wenn gilt

$$\boxed{\sphericalangle(u,v) = \frac{\pi}{2}}.$$

Man spricht von einem **rechten Winkel** und schreibt auch $u \perp v$.

**Aufgabe.** Beweisen Sie den Thales–Satz:




Wenn die Strecken $AB,\ BC$ und $BD$ gleichlang sind, so ist der Winkel $\alpha$ ein rechter.
Hinweis: Benutzen Sie die Vektordarstellung !

**Definition.** Ein Vektorsystem $S \subset V$ heißt **orthogonal**, falls alle Vektoren aus $S$ paarweise orthogonal sind.

**Definition.** $S \subset V$ heißt **orthonormiertes** System, falls es orthogonal ist und nur aus normierten Vektoren besteht.

**7.2.1 Satz.** *Jedes orthogonale System $S$ mit $\mathcal{O} \notin S$ ist linear unabhängig.*

*Beweis.* Sei $\sum\limits_{i=1}^{k} \lambda_i v_i = \mathcal{O}$ mit $v_i \in S,\ \lambda_i \in \mathbb{R}$.

Dann ist für $j = 1,\dots,k$, $\quad 0 = \big\langle v_j,\ \sum\limits_{i=1}^{k} \lambda_i v_i\big\rangle = \sum\limits_{i=1}^{k} \lambda_i \,\langle v_j, v_i\rangle = \lambda_j \big\langle v_j, v_j\big\rangle$ aufgrund der paarweisen Orthogonalität. Somit ist $\lambda_j = 0$, denn $v_j = \mathcal{O}$ wurde ausgeschlossen. □

**Koordinatendarstellung des Skalarproduktes in euklidischen Vektorräumen**

Sei $\dim V = n$ und $b_1,\dots,b_n$ eine Basis. Wir benutzen die Koordinatendarstellung $u = \sum\limits_{i=1}^{n} u_i b_i$ und $v = \sum\limits_{j=1}^{n} v_j b_j$ und die Bilinearität des Skalarproduktes für $\langle u,v\rangle = \sum\limits_{i,j=1}^{n} u_i v_j \,\langle b_i, b_j\rangle$. In der Bezeichnung

$$g_{ij} := \langle b_i, b_j\rangle$$



ergibt sich

$$\boxed{\langle u,v\rangle = \sum_{i.j=1}^{n} g_{ij}\, u_i v_j}$$

**7.2.2 Satz.** *Eine Basis $\{e_1,\dots,e_n\}$ in $E_n$ ist genau dann orthonormiert, wenn für die Koordinatendarstellung gilt*

$$\langle u,v\rangle = \sum_{i=1}^{n} u_i v_i$$

*Beweis.* Die Bedingung ist offenbar notwendig, denn im orthonormierten Fall haben wir $g_{ij} = \delta_{ij}$.
Setzen wir $\langle u,v\rangle = \sum u_i v_i$ voraus, so erhalten wir insbesondere $\langle e_i, e_j\rangle = \delta_{ij}$, denn für einen Basisvektor $e_k$ ist die $k$–te Koordinate gleich 1, und alle anderen Koordinaten sind 0. □

**Das Schmidtsche Orthonormierungsverfahren**

**7.2.3 Satz.** *Seien $\{a_1, a_2, \dots\}$ linear unabhängige Vektoren eines Vektorraumes $V$ mit Skalarprodukt. Dann gibt es genau ein Orthonormalsystem $\{e_1, e_2, \dots\}$ in $V$, so daß für alle $k \in \mathbb{N}$*

(1) *$\{a_1,\dots,a_k\}$ und $\{e_1,\dots,e_k\}$ denselben Unterraum $U_k$ von $V$ aufspannen, und*

(2) *die Determinante der Koordinatentransformation $\{a_1,\dots,a_k\} \to \{e_1,\dots,e_k\}$ in $U_k$ positiv ist.*

*Beweis.* (Induktion nach $k$)
Für $k = 1$ erhalten wir mit $e_1 := |a_1|^{-1} a_1$ das Gewünschte.
Die Aussage sei für $k \le n$ richtig, und wir betrachten den Fall $k = n+1$. Der Vektor

(3) $\qquad b_{n+1} := a_{n+1} - \sum\limits_{i=1}^{n} \langle a_{n+1}, e_i\rangle\, e_i$

ist orthogonal zu $\{e_1,\dots,e_n\}$, denn es gilt

$$\langle b_{n+1}, e_j\rangle = \langle a_{n+1}, e_j\rangle - \sum_{i=1}^{n} \langle a_{n+1}, e_i\rangle \,\langle e_i, e_j\rangle = \langle a_{n+1}, e_j\rangle - \langle a_{n+1}, e_j\rangle \cdot 1 = 0\,.$$

Außerdem ist nach Konstruktion (3) $b_{n+1} \neq 0$, da die Vektoren $e_1,\dots,e_n,\ a_{n+1}$ linear unabhängig sind. (Anderenfalls ergäbe $a_{n+1} \in L\big(\{e_1,\dots,e_n\}\big) = L\big(\{a_1,\dots,a_n\}\big)$ einen Widerspruch zur linearen Unabhängigkeit der $a_1,\dots,a_{n+1}$.) Wir können deshalb setzen

(4) $\qquad e_{n+1} := |b_{n+1}|^{-1}\, b_{n+1}\,,$

so daß $\{e_1, \ldots, e_{n+1}\}$ ein Orthonormalsystem bildet.
Aus $e_{n+1} \in L\big(\{e_1, \ldots, e_n, e_{n+1}\}\big)$, $a_{n+1} \in L\big(\{e_1, \ldots, e_n, e_{n+1}\}\big)$ und der Induktionsvoraussetzung folgt $U_{n+1} = L\big(\{a_1, \ldots, a_{n+1}\}\big) = L\big(\{e_1, \ldots, e_{n+1}\}\big)$, d.h. (1). Die Transformationsmatrix beim Übergang von $\{a_1, \ldots, a_{n+1}\}$ zu $\{e_1, \ldots, e_{n+1}\}$ hat die Gestalt

$$C_{n+1} = \begin{pmatrix} c_{11} & 0 & 0 & \ldots & & 0 \\ c_{12} & c_{22} & 0 & \ldots & & 0 \\ \vdots & & & & & \vdots \\ c_{n1} & \ldots & \ldots & & c_{nn} & 0 \\ c_{n+11} & \ldots & \ldots & & c_{n+1n} & |b_{n+1}|^{-1} \end{pmatrix}$$

so daß $\det C_{n+1} = |b_{n+1}^{-1} \det C_n$. Aufgrund der Induktionsvoraussetung ist aber $\det C_n > 0$, also ist (2) auch für $k = n+1$ erfüllt.
Zum Nachweis der Eindeutigkeit von $e_{n+1}$ nehmen wir an, ein zweiter Vektor $e'_{n+1}$ genüge allen Bedingungen. Insbesondere ist dann $a_{n+1} = \sum\limits_{i=1}^{n} \lambda_i e_i - \lambda_{n+1} e'_{n+1}$ für gewisse $\lambda_i$, die folgendermaßen bestimmt werden können:
$\langle a_{n+1}, e_j \rangle = \sum\limits_{i=1}^{n} \lambda_i \langle e_i, e_j \rangle - \lambda_{n+1} \langle e'_{n+1}, e_j \rangle = \lambda_j$. Es folgt

$$\lambda_{n+1} e'_{n+1} = a_{n+1} - \sum_{i=1}^{n} \langle a_{n+1}, e_i \rangle e_i \,.$$

Der Vergleich mit (3) und (4) und die Bedingung (2) ergeben

$$e'_{n+1} = e_{n+1} \,.$$

□

**Bemerkung.** Die Gleichungen (3) und (4) liefern gleichzeitig ein Konstruktionsverfahren für das Orthonormalsystem, das so von E. Schmidt entwickelt wurde.

**7.2.4 Folgerung.** *$V$ besitze eine abzählbare Basis. Dann läßt sich jedes Orthogonalsystem $\{e_1, \ldots, e_k\}$ zu einer orthonormierten Basis in $V$ ergänzen.*

*Beweis.* Nach dem Austauschsatz gibt es Vektoren $b_{k+1}, b_{k+2}, \ldots$ in $V$ derart, daß $\{e_1, \ldots, e_k,\ b_{k+1}, b_{k+2}, \ldots\}$ eine Basis bildet. Die Anwendung des Verfahrens von Schmidt auf diese Basisvektoren führt nun zu einer Orthonormalbasis. □

Insbesondere gibt es in einem solchen $V$ überhaupt eine Orthonormalbasis.

**Orthogonales Komplement und orthogonale Projektion:**

**Definition.** Zwei Mengen $S_2$ und $S_2$ von $V$ heißen **orthogonal** ($S_1 \perp S_2$), falls $\langle v_1, v_2 \rangle = 0,\ v_1 \in S_1,\ v_2 \in S_2$.



**Aufgabe.** $S_1 \perp S_2 \Leftrightarrow S_1 \perp L(S_2) \Leftrightarrow L(S_1) \perp L(S_2)$. •

**Definition.** Für einen Unterraum $U \subset V$ heißt

$$U^\perp := \{v \in V : \langle u, v \rangle = 0,\ u \in U\}$$

**orthogonales Komplement** zu $U$ in $V$.

**Aufgabe.** (1) $U^\perp$ ist wieder ein Unterraum.
(2) $(U^\perp)^\perp \supset U$ •

**7.2.5 Satz.** $\dim U < \infty \Rightarrow$

$$\boxed{V = U \oplus U^\perp.}$$

*Beweis.* Wir wählen eine Orthonormalbasis $\{e_1, \ldots, e_n\}$ in $U$ und setzen für $v \in V$

$$u := \sum_{i=1}^{n} \langle v, e_i \rangle e_i \,, \quad u^\perp := v - u$$

so daß

$$v = u + u^\perp \,.$$

Offenbar ist $u \in U$. Für die Beziehung $u^\perp \in U^\perp$ genügt es zu zeigen, daß $u^\perp$ orthogonal zu allen $e_j$ ist:

$$\langle u^\perp, e_j \rangle = \langle v, e_j \rangle - \langle u, e_j \rangle = \langle v, e_j \rangle - \sum \langle v, e_i \rangle \langle e_i, e_j \rangle = \langle v, e_j \rangle - \langle v, e_j \rangle = 0 \,.$$

Aus diesen Überlegungen folgt $V = U + U^\perp$. Außerdem ist aber $U \cap U^\perp = \{\mathcal{O}\}$, denn für $w \in U \cap U^\perp$ gilt $\langle w, w \rangle = 0$. Satz 1.5.1 besagt nun, daß die Summe direkt ist. □

**Definition.** Für die Zerlegung $v = u + u^\perp$ mit

$$\boxed{u = \sum_{i=1}^{n} \langle v, e_i \rangle \, e_i}$$

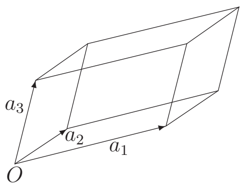
heißt $u$ **orthogonale Projektion** von $v$ auf $U$ und $u^\perp$ **Lot** von $v$ auf $U$.

## 7.3 Volumen

Mit Hilfe des Skalarproduktes läßt sich der aus der Schulgeometrie bekannte Flächen– bzw. Volumenbegriff abstrakt fassen und auf höhere Dimensionen übertragen. Wir wählen dabei einen geometrisch anschaulichen Zugang, den wir dann in eine algebraische Sprache übersetzen. Ausgangspunkt ist ein affiner Raum $(\mathcal{A}, V)$ mit Skalarprodukt, der auch unendlichdimensional sein darf. Für jeden Punkt $O \in \mathcal{A}$ und beliebiege Dektoren $a_1, \ldots, a_k \in V$ heißt die Punktmenge

$$\{P : \overrightarrow{OP} = \sum_{i=1}^{k} \lambda_i a_i\,, \quad 0 \leq \lambda_i \leq 1\}$$

das von den Vektoren $a_1, \ldots, a_k$ aufgespannte **Parallelotop** mit $O$ als Eckpunkt. Sind die $a_i$ linear unabhängig, so nennt man das Parallelotop **k–dimensional**.




Für orthogonale Vektoren $\{a_1, \ldots, a_k\}$ entsteht das mehrdimensionale Analogon von Rechtecken und Quadern, insbesonders ein $k$–dimensionaler Würfel im Falle von Vektoren gleicher Länge. Anstelle des Punktraumes kann man auch nur den Vektorraum $V$ betrachten, wobei man bei der Definition des Parallelotops auf die Interpretation als Ortsvektoren verzichtet. Für jedes $k$ und alle $a_1, \ldots, a_k$ soll nun dem aufgespannten Parallelotop eine nichtnegative Zahl $\mathcal{V}(a_1, \ldots, a_k)$ zugeordnet werden, die die Rolle eines Volumens spielen soll. Folgende Axiome sollen dabei erfüllt sein.

($\mathcal{V}$I) $\mathcal{V}(a_1, \ldots, a_{k-1}, \lambda a_k) = |\lambda|\, \mathcal{V}(a_1, \ldots, a_{k-1}, a_k),\ \lambda \in \mathbb{R}$

($\mathcal{V}$II) $b \in L\big(\{a_1, \ldots, a_{k-1}\}\big) \Rightarrow \mathcal{V}(a_1, \ldots, a_{k-1}, a_k + b) = \mathcal{V}(a_1, \ldots, a_{k-1}, a_k)$

($\mathcal{V}$III) $a_k \perp \{a_1, \ldots, a_{k-1}\} \Rightarrow \mathcal{V}(a_1, \ldots, a_{k-1}, a_k) = \mathcal{V}(a_1, \ldots, a_{k-1})\, \mathcal{V}(a_1)$

($\mathcal{V}$IV) $|a| = 1 \Rightarrow \mathcal{V}(a) = 1\,.$

**Bemerkung.** Für $k = 2$ besagt Axiom ($\mathcal{V}$II), daß Parallelogramme mit gleichen Grundseiten und gleichen “Höhen”, denselben Flächeninhalt haben, und ($\mathcal{V}$III) entspricht der Flächenformel für Rechtecke, wenn man noch folgendes berücksichtigt: Wegen ($\mathcal{V}$I) und ($\mathcal{V}$III) ist für jedes $a \in \mathcal{V}$

$$\mathcal{V}(a) = |a|\, \mathcal{V}\left(\frac{1}{|a|}\, a\right) = |a|\,.$$

**7.3.1 Lemma.** *Sei $a_k^\perp$ das Lot des Vektors $a_k$ auf $L\big(\{a_1, \ldots, a_{k-1}\}\big)$. Dann gilt*

$$\mathcal{V}(a_1, \ldots, a_k) = \mathcal{V}(a_1, \ldots, a_{k-1})\ |a_k^\perp|\,.$$



*Beweis.* Wir zerlegen $a_k = a_k' + a_k^\perp$ und benutzen ($\mathcal{V}$II), ($\mathcal{V}$III) und $\mathcal{V}(a) = |a|$:

$$\begin{aligned}\mathcal{V}(a_1, \ldots, a_k) &= \mathcal{V}(a_1, \ldots, a_{k-1}, a_k' + a_k^\perp) = \mathcal{V}(a_1, \ldots, a_{k-1}, a_k^\perp)\\ &= \mathcal{V}(a_1, \ldots, a_{k-1})\, \mathcal{V}(a_k^\perp) = \mathcal{V}(a_1, \ldots, a_{k-1})\, |a_k^\perp|\,.\end{aligned}$$

□

Durch sukzessives Anwenden dieser Formel ergibt sich
$\mathcal{V}(a_1, \ldots, a_k) = |a_1|\, |a_2^\perp| \ldots a_k^\perp|$ und somit:

**7.3.2 Folgerung.** *Die Volumenfunktion ist eindeutig bestimmt.*

Bisher wissen wir jedoch nicht, ob es überhaupt eine Funktion gibt, die den Forderungen ($\mathcal{V}$I) – ($\mathcal{V}$IV) genügen. Eine Lösung des Problems und gleichzeitig eine Berechnungsmöglichkeit liefert der folgende Satz. Wir benötigen dazu

**Definition.**

$$\Gamma(a_1, \ldots, a_k) := \begin{vmatrix} \langle a_1, a_1\rangle & \ldots & \langle a_1, a_k\rangle \\ \ldots & & \\ \langle a_k, a_1\rangle & \ldots & \langle a_k, a_k\rangle \end{vmatrix}$$

heißt **Gramsche Determinante** der Vektoren $a_1, \ldots, a_k$.

Unten werden wir zeigen, dass diese Determinante stets nichtnegativ ist, zunächst aber folgendes:

**7.3.3 Satz.** $\mathcal{V}(a_1, \ldots, a_k) := \sqrt{|\Gamma(a_1, \ldots, a_k)|}$,
*genügt den Axiomen ($\mathcal{V}$I) – ($\mathcal{V}$IV) und ist symmetrisch in $a_1, \ldots, a_k$, $k \in \mathbb{N}$.*

*Beweis.*
Die Symmetrie folgt aus der Definition von $\Gamma$.

$$(\mathcal{V}\mathrm{I}):\quad \Gamma(a_1, \ldots, a_{k-1}, \lambda a_k) = \begin{vmatrix} \langle a_1, a_1\rangle & \ldots & \langle a_1, \lambda a_k\rangle \\ \ldots & & \\ \langle \lambda a_k, a_1\rangle & \ldots & \langle \lambda a_k, \lambda a_k\rangle \end{vmatrix}$$

$$= \lambda^2 \begin{vmatrix} \langle a_1, a_1\rangle & \ldots & \langle a_1, a_k\rangle \\ \ldots & & \\ \langle a_k, a_1\rangle & \ldots & \langle a_k, a_k\rangle \end{vmatrix} = \lambda^2\, \Gamma(a_1, \ldots, a_{k-1}, a_k)$$

$$(\mathcal{V}\mathrm{II}):\quad \begin{vmatrix} \langle a_1, a_1\rangle & \ldots & \langle a_1, a_k + \sum\limits_{j=1}^{k-1} \lambda_j a_j\rangle \\ \ldots & & \\ \langle a_k + \sum\limits_{i=1}^{k-1} \lambda_i a_i,\ a_1\rangle & \ldots & \langle a_k + \sum\limits_{i=1}^{k-1} \lambda_i a_i,\ a_k + \sum\limits_{j=1}^{k-1} \lambda_j a_j\rangle \end{vmatrix}$$

$$= \left| \begin{matrix} \langle a_1,a_1\rangle & \dots & \langle a_1,a_k\rangle + \sum\limits_{j=1}^{k-1}\lambda_j\langle a_1,a_j\rangle \\ \dots & & \\ \langle a_k+\sum\limits_{i=1}^{k-1}\lambda_i a_i,\ a_1\rangle & \dots & \langle a_k+\sum\limits_{j=1}^{k-1}\lambda_i a_i,\ a_k\rangle + \sum\limits_{j=1}^{k-1}\lambda_j\langle a_k+\sum\limits_{i=1}^{k-1}\lambda_i a_i, a_j\rangle \end{matrix} \right| .$$

Der zweite Summand der letzten Spalte ist eine Linearkombination der ersten $k-1$ Spalten und kann deshalb weggelassen werden. Es ergibt sich

$$\left| \begin{matrix} \langle a_1,a_1\rangle & \dots & \langle a_1,a_k\rangle \\ \dots & & \\ \langle a_k+\sum\limits_{i=1}^{k-1}\lambda_i a_i,\ a_1\rangle & \dots & \langle a_k+\sum\limits_{i=1}^{k-1}\lambda_i a_i,\ a_k\rangle \end{matrix} \right|$$

$$= \left| \begin{matrix} \langle a_1,a_1\rangle & \dots & \langle a_1,a_k\rangle \\ \dots & & \\ \langle a_k,a_1\rangle + \sum\limits_{i=1}^{k-1}\lambda_i\langle a_1,a_1\rangle & \dots & \langle a_k,a_k\rangle + \sum\limits_{i=1}^{k-1}\lambda_i\langle a_i,a_k\rangle \end{matrix} \right| .$$

Hier ist wiederum der zweite Summand der letzten Zeile Linearkombination der vorhergehenden Zeilen, so daß die Determinante gleich ist zu

$$\left| \begin{matrix} \langle a_1,a_1\rangle & \dots & \langle a_1,a_k\rangle \\ \dots & & \\ \langle a_k,a_1\rangle & \dots & \langle a_k,a_k\rangle \end{matrix} \right| .$$

$(\mathcal{V}\mathrm{III})$: $\quad a_k \perp \{a_1,\dots,a_{k-1}\}$ bedeutet $\langle a_k,a_i\rangle = 0\,,\ i<k\,.$

Dann ist wegen Satz 5.3.3 und 5.1.5

$$\Gamma(a_1,\dots,a_k) = \left| \begin{matrix} & & & \vdots \\ \dots & \dots & \dots & \\ 0 & \dots & 0 & \langle a_k,a_k\rangle \end{matrix} \right| = \left| \begin{matrix} \langle a_1,a_1\rangle & \dots & \langle a_1,a_{k-1}\rangle \\ \dots & & \\ \langle a_{k-1},a_1\rangle & \dots & \langle a_{k-1},a_{k-1}\rangle \end{matrix} \right| \langle a_k,a_k\rangle$$

$$= \Gamma(a_1,\dots,a_{k-1})\,\Gamma(a_k)\,.$$

$(\mathcal{V}\mathrm{IV})$: $\quad \sqrt{|\Gamma(a)|} = \sqrt{\langle a,a\rangle} = |a|\,.$ $\square$



**Koordinatendarstellung im euklidischen Fall**

Seien $\begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1k} \\ \vdots \\ a_{nk} \end{pmatrix}$ die Koordinatenvektoren von $a_1,\dots,a_k \in \mathbb{E}_n$ bzgl. einer beliebigen Orthonormalbasis in $\mathbb{E}_n$. Für die Matrizen $A = (a_{ij})$ und $C = (c_{ij})$ mit $c_{ij} := \langle a_i,a_j\rangle = \sum\limits_{k=1}^{n} a_{ki}a_{kj}$ gilt offenbar $C = A^\top A$ und damit $\det C = \det(A^\top A)$.
Andererseits ist $\det C = \Gamma(a_1,\dots,a_k)$, also

$$\Gamma(a_1,\dots,a_k) = \det(A^\top A)\,.$$

Im Fall $k=n$ erhält man

$$\det(A^\top A) = \det A^\top \det A = (\det A)^2 \geq 0$$

und demnach $\sqrt{\Gamma(a_1,\dots,a_n)} = |\det A|$.
Dies liefert eine geometrische Interpretation des Betrags der Determinante einer $n\times n$–Matrix.
Für $k\leq n$ können wir Ähnliches herleiten: Wenn man die $a_1,\dots,a_k \in V$ ($V$ beliebig) in Koordinaten bzgl. einer Orthonormalbasis in $L(\{a_1,\dots,a_k\})$ darstellt, so bleiben die obigen Überlegungen für $k=n$ richtig. Sei jetzt $A'$ anstelle von $A$ die zugehörige Matrix. Dann gilt

$$\Gamma(a_1,\dots,a_k) = (\det A')^2 \geq 0\,.$$

Zusammengefaßt erhalten wir folgendes:

**7.3.4 Satz.** *(i) Die Gramsche Determinante in $V$ ist stets nichtnegativ.*

*(ii) $\mathcal{V}(a_1,\dots,a_k) = \sqrt{\det(A^\top A)}$, insbesondere $\mathcal{V}(a_1,\dots,a_n) = |\det A|$ für $k=n$, wobei die Spalten der Matrix $A$ aus den Koordinatenvektoren der $a_1,\dots,a_k$ bzgl. einer beliebigen Orthonormalbasis in $\mathbb{E}_n$ gebildet werden.*

## 7.4 Orientierung und Vektorprodukt in euklidischen Räumen

**Definition.** Zwei Basen im $\mathbb{E}_n$ heißen **gleich orientiert**, falls die Determinante der zugehörigen Koordinatentransformation positiv ist, und ansonsten entgegengesetzt orientiert.

Wenn $C_B^{B'}$ die Transormationsmatrix beim Übergang von der Basis $B$ zur Basis $B'$ bezeichnet, so gilt $C_{B'}^{B''}\,C_B^{B'} = C_B^{B''}$ und damit $\det C_{B'}^{B''}\cdot\det C_B^{B'} = \det C_B^{B''}$. Wenn also $B$ und $B'$ gleich orientiert sind und $B'$ und $B''$ ebenfalls, so sind auch $B$ und $B''$ gleich orientiert. Die Menge aller Basen von $\mathbb{E}_n$ zerfällt deshalb in zwei Äquivalenzklassen bzgl. ihrer Orientierung.

**Definition.** Eine **Orientierung** in $\mathbb{E}_n$ ist die Wahl einer der beiden obigen Äquivalenzklassen als positiv orientiert. (Die Basen aus der anderen Äquivalenzklasse heißen dann negativ orientiert.)

Aus 7.3.5. und der Definition der Orientierung ergibt sich folgendes: Für linear unabhängige Vektoren $A_1, \ldots, a_n \in \mathbb{E}_n$ kann man $\det A = \det(a_{ij})$ als **orientiertes Volumen** des Parallelotops mit diesen Erzeugungsvektoren interpretieren.

Seien nun $a_1, \ldots, a_{n-1}$ linear unabhängige Vektoren eines $n$–dimensionalen orientierten euklidischen Raumes $\mathbb{E}_n$.

**7.4.1 Lemma.** *Es gibt genau einen Vektor $\mathfrak{n} \in \mathbb{E}_n$ mit folgenden Eigenschaften:*

(1) $\langle \mathfrak{n}, a_i \rangle = 0\ ,\ i = 1, \ldots, n-1$

(2) $|\mathfrak{n}| = \mathcal{V}(a_1, \ldots, a_{n-1})$

(3) $\{a_1, \ldots, a_{n-1}, \mathfrak{n}\}$ *ist positiv orientiert.*

*Beweis.* Wegen (1) muß der gesuchte Vektor in $L\big(\{a_1, \ldots, a_{n-1}\}\big)^\perp$ liegen. Dies ist ein eindimensionaler Unterraum von $\mathbb{E}_n$. Durch die Länge $\mathcal{V}(a_1, \ldots, a_{n-1})$ ist der Vektor $\mathfrak{n}$ dann bis auf sein Vorzeichen bestimmt, welches sich aus der Orientierungsbedingung ergibt. □

**Definition.**

$$\boxed{\mathfrak{n} =: a_1 \times \ldots \times a_{n-1}}$$

wird **Vektorprodukt** der linear unabhängigen Vektoren $a_1, \ldots, a_{n-1}$ genannt.

**7.4.2 Aufgabe.** $\langle a_1 \times \ldots \times a_{n-1}, v \rangle = \det(\bar{a}_1 \ldots \bar{a}_{n-1} \bar{v})\ ,\ v \in \mathbb{E}_n$, wobei $\bar{a}_i, \bar{v}$ die Koordinatenvektoren bzgl. einer Orthonormalbasis sind.
(Hinweis: Zerlegen Sie $v$ nach der Basis $\{a_1, \ldots, a_{n-1},\ a_1 \times \ldots \times a_{n-1}\}$ und benutzen Sie $|\det(\bar{a}_1, \ldots, \bar{a}_{n-1}, \bar{a}_n)| = \mathcal{V}(a_1, \ldots, a_n)$.) •

Wir wenden diese Formel an, um die Koordinatendarstellung des Vektorprodukts zu finden: Wenn $\{e_1, \ldots, e_n\}$ die zugrundeliegende $ONB$ ist und $\bar{\mathfrak{n}} = \begin{pmatrix} \mathfrak{n}_1 \\ \vdots \\ \mathfrak{n}_n \end{pmatrix}$, so gilt ja $\mathfrak{n}_i = \langle \mathfrak{n}, e_i \rangle$ und deshalb

**7.4.3 Lemma.** $\boxed{\mathfrak{n}_i = \det(\bar{a}_1, \ldots, \bar{a}_{n-1}, \bar{e}_i)}\ ,\ i = 1, \ldots, n\,.$

Vektoren, die senkrecht auf einem gegebenen Unterraum $U$ stehen, werden auch **Normalen** zu $U$ genannt. Unter einer Normalen zu einer Hyperebene in einem euklidischen Punktraum versteht man eine Normale zum erzeugten Vektorraum. Mit Hilfe einer Basis $\{a_1, \ldots, a_{n-1}\}$ in diesem Raum und deren Vektorprodukt kann man dann alle Normalen bestimmen, da diese sich von $\mathfrak{n} = a_1 \times \ldots \times a_{n-1}$ nur durch ein



Vielfaches unterscheiden.

Umgekehrt gibt es zu einem beliebig vorgegebenen Vektor $\mathfrak{n} \in \mathbb{E}_n$ mit $|\mathfrak{n}| = 1$ und einem Punkt $P \in \mathcal{A}_n$ genau eine Hyperebene $\mathcal{H}$ durch $P_0$ Normalenvektor $\mathfrak{n}$. Ihre Gleichung lautet in der Sprache der Ortsvektoren der Punkte $P$:

**7.4.4 Lemma.** $\langle \overrightarrow{OP}, \mathfrak{n} \rangle = \langle \overrightarrow{OP_0}, \mathfrak{n} \rangle$ ***(Hessesche Normalform)***

*Beweis:.* Wir ergänzen $a_n := \mathfrak{n}$ zu einer Orthonormalbasis $a_1, \ldots, a_{n-1}, a_n$ in $E_n$. Die Parametergleichung

$$\overrightarrow{OP} = \overrightarrow{OP_0} + \sum_{i=1}^{n-1} \lambda_i\, a_i$$

von $\mathcal{H}$ ist dann wegen $\overrightarrow{OP} - \overrightarrow{OP_0} = \overrightarrow{P_0P}$ äquivalent zu $\langle \overrightarrow{P_0P},\, a_i \rangle = \lambda_i,\ i = 1, \ldots, n-1$, was wegen der beliebigen Wahl der $\lambda_i \in \mathbb{R}$ keine Festlegung bedeutet, und $\langle \overrightarrow{P_0P},\, a_n \rangle = 0$ . □

## 7.5 Orthogonale Abbildungen und Isometrien

### Orthogonale Abbildungen zwischen Vektorräumen mit Skalarprodukt

**Definition.** Eine Abbildung $\mathbf{A} : V \to V'$ zwischen zwei Vektorräumen mit Skalarprodukt heißt **orthogonal**, falls

$$\langle \mathbf{A}u, \mathbf{A}v \rangle = \langle u, v \rangle\ ,\ u, v \in V\,.$$

$\mathbf{A}$ sei nun eine solche Abbildung.

**Eigenschaften:**

**7.5.1.** $\mathbf{A}$ *ist normtreu, d.h.* $|\mathbf{A}v| = |v|$. □

**7.5.2.** $\mathbf{A}$ *ist linear und injektiv.*

*Beweis.* $\langle \mathbf{A}(u+v) - \mathbf{A}u - \mathbf{A}v,\ \mathbf{A}(u+v) - \mathbf{A}u - \mathbf{A}v \rangle$
$= |\mathbf{A}(u+v)|^2 + |\mathbf{A}u|^2 + |\mathbf{A}v|^2 - 2\langle \mathbf{A}(u+v),\ \mathbf{A}u \rangle - 2\langle \mathbf{A}(u+v),\ \mathbf{A}v \rangle + 2\langle \mathbf{A}u, \mathbf{A}v \rangle$
$= |u+v|^2 + |u|^2 + |v|^2 - 2\langle u+v, u \rangle - 2\langle u+v, v \rangle + 2\langle u, v \rangle$
$= \langle u+v-u-v,\ u+v-u-v \rangle = 0$, d.h. $\mathbf{A}(u+v) = \mathbf{A}u + \mathbf{A}v$.
Die Homogenität von $\mathbf{A}$ beweist man analog.
Gilt nun $\mathbf{A}u = \mathbf{A}v$, so folgt $0 = |\mathbf{A}(u-v)| = |u-v|$, also $u = v$, so daß $\mathbf{A}$ injektiv ist. □

**7.5.3.** $\mathbf{A}$ *ist winkeltreu, d.h.* $\sphericalangle(\mathbf{A}u, \mathbf{A}v) = \sphericalangle(u, v)$.

*Beweis.* $\cos \sphericalangle(\mathbf{A}u, \mathbf{A}v) = \dfrac{\langle \mathbf{A}u, \mathbf{A}v\rangle}{|\mathbf{A}u|\,|\mathbf{A}v|} = \dfrac{\langle u, v\rangle}{|u|\,|v|} = \cos \sphericalangle(u, v)$. $\square$

**7.5.4.** $\mathbf{A}$ *ist volumentreu, d.h.* $\mathcal{V}(\mathbf{A}a_1, \dots, \mathbf{A}a_k) = \mathcal{V}(a_1, \dots, a_k)$.

*Beweis.* $\mathcal{V}(\mathbf{A}a_1, \dots, \mathbf{A}a_k) = \sqrt{\det(\langle \mathbf{A}a_i, \mathbf{A}a_j\rangle)} = \sqrt{\det(\langle a_i, a_j\rangle)} = \mathcal{V}(a_1, \dots, a_k)$. $\square$

**Definition.** Eine orthogonale Bijektion zwischen zwei Vektorräumen mit Skalarprodukt heißt **Isomorphismus**.

**Bemerkung.** Zusätzlich zum Isomorphiebegriff allgemeiner linearer Räume wird hier also auch noch das Skalarprodukt übertragen.

**7.5.5.** *Im Fall* $\dim V = \dim V' < \infty$ *ist* $\mathbf{A}$ *ein Isomorphismus.* $\square$

**7.5.6.** *Alle* $n$*–dimensionalen euklidischen Vektorräume sind isomorph (und können mit dem Koordinatenraum* $\mathbb{R}^n$ *mit dem kanonischen Skalarprodukt identifiziert werden).*

*Beweis.* Seine $V$ und $V'$ zwei solcher Räume. Mit Hilfe zweier Orthonormalbasen $\{e_1, \dots, e_n\}$ bzw. $\{e'_1, \dots, e'_n\}$ definieren wir eine Abbildung $\mathbf{A} \in L(V, V')$ über $\mathbf{A}e_i = e'_i$, $i = 1, \dots, n$. Diese ist orthogonal:
$\langle \mathbf{A}u, \mathbf{A}v\rangle = \langle (\sum\limits_i u_i e_i),\ \mathbf{A}(\sum\limits_j v_j e_j)\rangle = \sum_{i,j} u_i v_j \langle \mathbf{A}e_i, \mathbf{A}e_j\rangle$
$= \sum_{i,j} u_i v_i \langle e'_i, e'_j\rangle = \sum\limits_{i,j} u_i v_j \delta_{ij} = \sum\limits_i u_i v_i = \langle u, v\rangle$,
wobei $u_i$ und $v_i$ die Koordinaten von $u$ bzw. $v$ sind. $\square$

**7.5.7.** *Ist* $\mathbf{A}$ *ein Isomorphismus, so gilt dies auch für* $\mathbf{A}^{-1}$.

*Beweis.* $\langle \mathbf{A}^{-1}u',\ \mathbf{A}^{-1}v'\rangle = \langle \mathbf{A}(\mathbf{A}^{-1}u'),\ \mathbf{A}(\mathbf{A}^{-1}v')\rangle = \langle u', v'\rangle$. $\square$

**7.5.8 Satz.** *Wenn* $V$ *und* $V'$ *abzählbare Basen besitzen, so sind folgende Aussagen äquivalent:*

(1) $\mathbf{A} : V \to V'$ *ist orthogonal.*

(2) $\mathbf{A}$ *ist linear und überführt irgendeine Orthonormalbasis, und damit jede, wieder in eine Orthonormalbasis.*

*Setzen wir zusätzlich voraus* $\dim V = \dim V' < \infty$*, so ergibt sich eine weitere äquivalents Aussage:*

(3) $A^\top = A^{-1}$*, wobei* $A$ *die Matrix von* $\mathbf{A}$ *bzgl. zweier Orthonormalbasen bezeichnet.*

*Beweis.* Jede orthogonale Abbildung überführt Orthonormalbasen wieder in Orthonormalbasen. Ist umgekehrt $\mathbf{A} \in L(V, V')$ und das Bild $\{\mathbf{A}e_1, \mathbf{A}e_2, \dots\}$ irgendeiner Orthonormalbasis $\{e_1, e_2, \dots\}$ in $V$ eine Orthonormalbasis in $V'$, so schließt man wie beim Beweis von 7.5.6. auf die Orthogonalität von $\mathbf{A}$. (1) und (2) sind also äquivalent.



Sei nun $\dim V = \dim V' = n$ und $A$ die Matrix von $\mathbf{A}$ bzgl. der Orthonormalbasen $\{e_1, \dots, e_n\}$ bzw. $\{e'_1, \dots, e'_n\}$. Weil die Spaltenvektoren $\bar{a}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$ von $A$ die Koordinatenvektoren von $\mathbf{A}e_j$ bzgl $\{e'_1, \dots, e'_n\}$ bilden, schließen wir:
$\{\mathbf{A}e_1, \dots, \mathbf{A}e_n\}$ ist Orthonormalbasis $\Leftrightarrow$ $\langle \mathbf{A}e_i, \mathbf{A}e_j\rangle = \delta_{ij} \Leftrightarrow \sum\limits_{k=1}^{n} a_{ki}\, a_{kj} = \delta_{ij} \Leftrightarrow$
$A^\top A = I \Leftrightarrow A^\top = A^{-1}$. $\square$

Die letzte Aussage legt folgende Begriffsbildung nahe.

**Definition.** Eine $n \times n$–Matrix $A$ über $\mathbb{R}$ heißt **orthogonal** falls $A^\top = A^{-1}$.

**Aufgabe.** Die orthogonalen Abbildungen $\mathbb{E}_n \to \mathbb{E}_n$ und die othogonalen $n \times n$–Matrizen bilden zwei isomorphe Gruppen. (Für erstere schreibt man auch $O(n)$.) •

**Isometrien und Bewegungen in affinen Räumen mit Skalarprodukt im erzeugten Vektorraum**

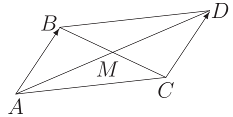
Ähnlich wir beim Zusammenhang zwischen linearen Abbildungen und affinen Abbildungen wollen wir nun Abbildungen zwischen Punkträumen einführen, die in enger Beziehung zu orthogonalen Abbildungen stehen.
Seien $(\mathcal{A}, V, f)$ , $(\mathcal{A}', V', f')$ affine Geometrien über $\mathbb{R}$ mit Skalarprodukten.

**Definition.** Eine Abbildung $\mathcal{F} : \mathcal{A} \to \mathcal{A}'$ heißt **Isometrie**, falls

$$|\overrightarrow{\mathcal{F}(A)\mathcal{F}(B)}| = |\overrightarrow{AB}|\,,\ A, B \in \mathcal{A}.$$

(Dabei bleiben also der Abstand zwischen den Punkten erhalten.) Im Fall $\mathcal{A} = \mathcal{A}'$ wird die Isometrie $\mathcal{F}$ auch **Bewegung** in $\mathcal{A}$ genannt.

Jeder Isometrie $\mathcal{F}$ läßt sich genau eine Vektorabbildung $\mathbf{F} : V \to V'$ zuordnen mit $\mathbf{F}\overrightarrow{AB} = \overrightarrow{\mathcal{F}(A)\mathcal{F}(B)}$. Um zu zeigen, daß die Definition von $\mathbf{F}$ korrekt ist, d.h. nicht von der Wahl der Repräsentanten der Vektoren abhängt, benutzen wir das folgende Mittelpunktprinzip:
Sei $\overrightarrow{AB} = \overrightarrow{CD}$. Dann teilt der Schnittpunkt $M$ der Diagonalen im Parallelogramm mit den Eckpunkten $A, B, C, D$ die Diagonalen im Verhältnis 1:1, d.h. $|\overrightarrow{AM}| = |\overrightarrow{MD}| = \frac{1}{2}|\overrightarrow{AD}|$ und $|\overrightarrow{CM}| = |\overrightarrow{MB}| = \frac{1}{2}|\overrightarrow{CB}|$, denn es gilt auch $\overrightarrow{AC} = \overrightarrow{BD}$.




Aus der Isometrie ergibt sich $|\overrightarrow{\mathcal{F}(A)\mathcal{F}(M)}| = |\overrightarrow{\mathcal{F}(M)\mathcal{F}(D)}| = \frac{1}{2}|\overrightarrow{\mathcal{F}(A)\mathcal{F}(D)}|$ und $|\overrightarrow{\mathcal{F}(C)\mathcal{F}(M)}| = |\overrightarrow{\mathcal{F}(M)\mathcal{F}(B)}| = \frac{1}{2}|\overrightarrow{\mathcal{F}(C)\mathcal{F}(B)}|$, was wegen der Dreicksungleichung für die Norm wiederum nur möglich ist, wenn $M$ gemeinsamer Mittelpunkt der Strecken $\mathcal{F}(A)\mathcal{F}(D)$ und $CF(C)\mathcal{F}(B)$ ist.
Damit folgt aber die gewünschte Gleichheit $|\overrightarrow{\mathcal{F}(A)\mathcal{F}(B)}| = |\overrightarrow{\mathcal{F}(C)\mathcal{F}(D)}|$. $\square$

**7.5.9 Lemma.** *Die zu einer Isometrie $\mathcal{F}$ gehörende Vektorabbildung* $\mathbf{F}$ *ist orthogonal.*

*Beweis.* Nach Konstruktion ist $\mathbf{F}$ normtreu. Außerdem ist $\mathbf{F}$ additiv:
Sei $u = \overrightarrow{AB}$, $v = \overrightarrow{BC}$. Dann gilt $u + v = \overrightarrow{AC}$ und
$\mathbf{F}(u+v) = \overrightarrow{\mathbf{F}(A)\mathbf{F}(C)} = \overrightarrow{\mathbf{F}(A)\mathbf{F}(B)} + \overrightarrow{\mathbf{F}(B)\mathbf{F}(C)} = \mathbf{F}\overrightarrow{AB} + \mathbf{F}\overrightarrow{BC} = \mathbf{F}u + \mathbf{F}v$.
Wir können nun zeigen, daß $\mathbf{F}$ das Skalarprodukt erhält, indem wir dieses über die Norm darstellen:

$$\begin{aligned}\langle \mathbf{F}u, \mathbf{F}v\rangle &= \frac{1}{2}\left(|\mathbf{F}u + \mathbf{F}v|^2 - |\mathbf{F}u|^2 - |\mathbf{F}v|^2\right) = \frac{1}{2}\left(|\mathbf{F}(u+v)|^2 - |\mathbf{F}u|^2 - |\mathbf{F}v|^2\right) \\ &= \frac{1}{2}\left(|u+v|^2 - |u|^2 - |v|^2\right) = \langle u, v\rangle .\end{aligned}$$

$\square$

**7.5.10 Folgerung.** *Jede Isometrie ist affin. Isometrien zwischen Räumen gleicher Dimension sind spezielle Affinitäten.*

*Beweis.* Da das zu einer Isometrie $\mathcal{F}$ gehörende $\mathbf{F}$ eine orthogonale Abbildung und somit linear ist, folgt unmittelbar, daß $\mathcal{F}$ affin sein muß. Bei gleicher Dimension von $V$ und $V'$ ist $\mathbf{F}$ außerdem bijektiv. Wie beim Beweis von Satz 6.5.4, 2) und 3) schließt man nun, daß auch $\mathcal{F}$ bijektiv ist. $\square$

## 7.6 Adjungierte Abbildungen

Sei $(V, \mathbb{R}, \langle\cdot,\cdot\rangle)$ jetzt wieder ein beliebiger reeller Vektorraum mit Skalarprodukt. Für jede lineare Selbstabbildung $\mathbf{A} \in L(V, V)$ soll eine zugehörige lineare Abbildung $\mathbf{A}^*$ eingeführt werden, die in der Matrixsprache der transponierten Matrix der darstellenden Matrix von $\mathbf{A}$ entspricht.

**Definition.** $\mathbf{A}^* : V \to V$ heißt zu $\mathbf{A}$ **adjungiert**, falls

$$\langle \mathbf{A}u, v\rangle = \langle u, \mathbf{A}^*v\rangle , \quad u, v \in V.$$

**Bemerkung.** Für die Untersuchung von Eigenschaften solcher Abbildungen und auch unabhängig davon ist folgende Überlegung nützlich: Für zwei Vektoren $v, w \in V$ gilt

$$v = w \Leftrightarrow \langle u, v\rangle = \langle u, w\rangle, \ \forall u \in V$$

$\big($*Beweis:* $v - w = 0 \Rightarrow \langle u, v - w\rangle = 0, \ \forall u \in V; \quad \langle u, v - w\rangle = 0$ für $u := v - w \Rightarrow v - w = 0 \quad \square\big)$

**7.6.1 Lemma.** $\mathbf{A}^*$ *ist eindeutig bestimmt.*

*Beweis.* Wenn $\mathbf{B}$ eine weitere solche Abbildungen bezeichnet, so gilt $\langle u, \mathbf{A}^*v\rangle = \langle u, \mathbf{B}v\rangle$, d.h., $\langle u, (\mathbf{A}^* - \mathbf{B})v\rangle = 0, \ \forall u \in U$, und deshalb wegen der obigen Bemerkung $(\mathbf{A}^* - \mathbf{B})v = 0, \ \forall v \in V$. Folglich ist $\mathbf{A}^* - \mathbf{B} = \mathbf{O}$. $\square$



**7.6.2 Lemma.** $\mathbf{A}^*$ *ist linear.*

*Beweis.* Wir benutzen wieder die obige Bemerkung.
Für jedes $u \in V$ gilt

$$\begin{aligned}\langle u, \mathbf{A}^*(v + \lambda w)\rangle &= \langle \mathbf{A}u, v + \lambda w\rangle \\ &= \langle \mathbf{A}u, v\rangle + \lambda\langle \mathbf{A}u, w\rangle \\ &= \langle u, \mathbf{A}^*v\rangle + \lambda\langle u, \mathbf{A}^*w\rangle \\ &= \langle u, \mathbf{A}^*v\rangle + \langle u, \lambda\mathbf{A}^*w\rangle = \langle u, \mathbf{A}^*v + \lambda\mathbf{A}^*w\rangle\end{aligned}$$

und deshalb $\mathbf{A}^*(v + \lambda w) = \mathbf{A}^*v + \lambda\mathbf{A}^*w$. $\square$

Die Adjungiertenbildung ist folgendermaßen mit den Abbildungsoperationen verknüpft. (Existenzaussagen haben wir bisher nicht formuliert.)

**7.6.3 Satz.**

(1) $(\mathbf{A}^*)^* = \mathbf{A}$

(2) $(\mathbf{A} + \mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*$

(3) $(\mathbf{A} \circ \mathbf{B})^* = \mathbf{A}^* \circ \mathbf{B}^*$

(4) $(\mathbf{A}^{-1})^* = (\mathbf{A}^*)^{-1}$

*wobei die Existenz der einen Seite jeweils die der anderen impliziert.*

(5) $\mathbf{A}^*$ *existiert* $\Rightarrow$ ($\mathbf{A}^* \circ \mathbf{A} = \mathbf{I} \Leftrightarrow \mathbf{A}$ *ist orthogonal)*

(6) $\mathbf{A}^*$ *ist bijektiv* $\Rightarrow$ ($\mathbf{A}^* = \mathbf{A}^{-1} \Leftrightarrow \mathbf{A}$ *ist orthogonal).*

*Beweis.* (1) – (4) stellen wir als Aufgabe. •

(5): Da $\langle u, \mathbf{A}^* \circ \mathbf{A}v\rangle = \langle \mathbf{A}u, \mathbf{A}v\rangle$,so folgt

$$\begin{aligned}\mathbf{A}^* \circ \mathbf{A} = \mathbf{I} &\Leftrightarrow \langle u, \mathbf{A}^* \circ \mathbf{A}v\rangle = \langle u, v\rangle \\ &\Leftrightarrow \langle \mathbf{A}u, \mathbf{A}v\rangle = \langle u, v\rangle \\ &\Leftrightarrow: \mathbf{A} \text{ orthogonal.}\end{aligned}$$

(6): Unter Voraussetzung der Bijektivität gilt $\mathbf{A}^* = \mathbf{A}^{-1} \Leftrightarrow \mathbf{A}^* \circ \mathbf{A} = \mathbf{I}$. Wegen (5) ist dieser aber äquivalent zur Orthogonalität von $\mathbf{A}$.
$\square$

In unendlichdimensionalen Räumen muß nicht jede lineare Abbildung eine Adjungierte besitzen. Für euklidische Vektorräume gilt aber folgendes.

**7.6.4 Satz.** *Sei* $\mathbf{A} \in L(\mathbb{E}_n, \mathbb{E}_n)$. *Dann existiert* $\mathbf{A}^*$, *und für die darstellende Matrix* $A$ *von* $\mathbf{A}$ *bzgl. einer Orthonormalbasis ist die Transponierte* $A^T$ *die darstellende Matrix von* $\mathbf{A}^*$ *bzgl. dieser Basis.*

*Beweis.* Seien $\{e_i, \ldots, e_n\}$ eine Orthonormalbasis in $\mathbb{E}_n$ und

$$u = \sum_{i=1}^{n} u_i e_i \, , \quad v = \sum_{i=1}^{n} v_i e_i \, , \quad A = (a_{ij})$$

die zugehörigen Koordinatendarstellungen.
Für das kanonische Skalarprodukt in $\mathbb{R}^n$ schreiben wir hier

$$\bar{u} \bullet \bar{v} := \sum_{i=1}^{n} u_i v_i \, .$$

Dann ist

$$\begin{aligned} \langle \mathbf{A}u, v \rangle &= \overline{\mathbf{A}u} \bullet \bar{v} = A\bar{u} \bullet \bar{v} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} u_j v_i = \sum_{j=1}^{n} u_j \sum_{i=1}^{n} a_{ij} v_i \\ &= \bar{u} \bullet A^T \bar{v} = \langle u, \mathbf{A}^T v \rangle, \end{aligned}$$

falls $\mathbf{A}^T$ die zur Matrix $A^T$ gehörende lineare Abbildung bezeichnet. Wir erhalten somit

$$\langle \mathbf{A}u, v \rangle = \langle u, \mathbf{A}^T v \rangle$$

und nach Definition von $\mathbf{A}^*$ muss $\mathbf{A}^T$ mit $\mathbf{A}^*$ zusammenfallen. Insbesonders ist dann $A^T$ die darstellende Matrix von $\mathbf{A}^*$. □

**Definition.** Für einen endlichdimensionalen reellen Vektorraum $V$ mit Skalarprodukt heißt $\mathbf{A} \in L(V, V)$ **symmetrisch** (oder **selbstadjungiert** im allgemeinen Fall), falls $\mathbf{A}^* = \mathbf{A}$, und **schiefsymmetrisch**, falls $\mathbf{A}^* = -\mathbf{A}$.

**7.6.5 Folgerung.** *Sei* $\mathbf{A} \in L(\mathbb{E}_n, \mathbb{E}_n)$ *und* $A$ *die darstellende Matrix bzgl. einer Orthonormalbasis. Dann gilt:*

(1) $\mathbf{A}$ *ist symmetrisch* $\Leftrightarrow A^T = A$.

(2) $\mathbf{A}$ *ist schiefsymmetrisch* $\Leftrightarrow A^T = -A$ .

(3) $\mathbf{A} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^*) + \frac{1}{2}(\mathbf{A} - \mathbf{A}^*)$ *ist die eindeutige Zerlegung von* $\mathbf{A}$ *in einen symmetrischen und einen schiefsymmetrischen Anteil.*

*Beweis.* (1) und (2) folgen aus Satz 7.6.4 und (3) aus 7.6.3. □