# Lineare Algebra

E. W. Zink

# Inhaltsverzeichnis

# Kapitel 5

# Vektorräume

## 5.1 Definition und Eigenschaften

### 5.1.1 Definition

Sei $K$ ein fixierter Körper. Ein $K-$Vektorraum ist eine Menge $V$ mit folgenden Eigenschaften:

1. $V$ ist eine additiv geschriebene kommutative Gruppe, d.h. wir ordnen $v_1, v_2 \in V$ ein Element $v_1 + v_2 \in V$ zu, wir haben ein Nullelement $o$ und das Inverse von $v \in V$ ist $-v$. Es gelten das Assoziativgesetz und das Kommutativgesetz.

2. Es gibt eine Operation von $K$ auf $V$:

$$K \times V \longrightarrow V,$$

$$(\lambda, v) \mapsto \lambda v.$$

   und für alle $\lambda_i \in K,\ v, w \in V$ gilt

   (a) $\lambda_1(\lambda_2 v) = (\lambda_1 \lambda_2)v,$ Assoziativgesetz
   (b) $\lambda(v + w) = \lambda v + \lambda w,$ 1. Distributivgesetz
   (c) $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v,$ 2. Distributivgesetz
   (d) $1_K v = v.$

Die Elemente von $V$ heißen Vektoren , die Elemente von $K$ heißen Skalare, das Nullelement $o$ heißt Nullvektor.

Wenn speziell $K = R$ ist, spricht man von einem reellen, bei $K = C$ von einem komplexen Vektorraum.

Die Theorie der Vektorräume wurde von Herrmann G. Graßmann (1809 - 1877) in seinem Buch „Die lineale Ausdehnungslehre“ (1844, überarbeitet 1862) entwickelt. Graßmann war Gymnasiallehrer in Stettin, er betätigte sich als Mathematiker und Sanskritforscher.

### 5.1.2 Eigenschaften

Aus den Distributivgesetzen und den Assoziativgesetzen für die Addition in $V$ und $K$ folgt das allgemeine Distributvgesetz (Beweis durch Induktion):

$$(\sum_{i=1}^{m} \lambda_i)(\sum_{j=1}^{n} v_j) = \sum_{i=1}^{m}\sum_{j=1}^{n} \lambda_i v_j.$$

Weiter gilt $0 \cdot v = o$, denn $v = 1 \cdot v = (1+0) \cdot v = 1v + 0v = v + 0v$.
Es ist $\lambda \cdot o = o$, denn $\lambda \cdot v = (\lambda/v + o) = \lambda v + \lambda 0$.
Die Vorzeichenregel $-v = (-1)v$ folgt aus $0v = o$; daraus folgt weiter
$\lambda(-v) = (-\lambda)v = -(\lambda v)$, denn
$\lambda v + \lambda(-v) = \lambda(v - v) = \lambda o$, also $\lambda(-v) = -(\lambda v)$,
$\lambda v + (-\lambda)v = (\lambda - \lambda)v = 0v = o$, also $(-\lambda)v = -(\lambda v)$.
Weiterhin folgt $(-\lambda)(-v) = \lambda v$.
Wir haben folgende Kürzungsregel: Wenn $\lambda v = o$ ist, so folgt $\lambda = 0$ oder $v = o$, denn wenn $\lambda v = o$ und $\lambda \neq 0$ ist, so ist $v = 1v = (\lambda^{-1}\lambda)v = \lambda^{-1}(\lambda v) = \lambda^{-1} o = o$.

### 5.1.3 Beispiele

1. $V = K$ ist ein $K-$Vektorraum.

2. Funktionenräume: Sei $K$ ein Körper und $I$ eine Menge; wir betrachten alle Funktionen $f : I \longrightarrow K,\ x \mapsto f(x)$; die Menge aller derartiger Funktionen bezeichnen wir mit $K^I = \{f : I \longrightarrow K\}$. Wir führen hier eine Addition und eine $K-$Operation wie folgt ein: Seien $f, g : I \longrightarrow K$ zwei Funktionen und $\lambda \in K$, wir definieren
$$(f+g)(x) = f(x) + g(x),$$
$$(\lambda f)(x) = \lambda \cdot f(x)$$
und überprüfen z.B. ein Distributivgesetz:
$$[\lambda(f+g)](x) = \lambda(f+g)(x) = \lambda(f(x)+g(x)) = \lambda f(x) + \lambda g(x) = (\lambda f + \lambda g)(x),$$
für alle $x$, also $\lambda(f+g) = \lambda f + \lambda g$.

3. Die Menge $K[X]$ der Polynome mit Koeffizienten in $K$ ist ein $K-$Vektorraum.

4. Direktes Produkt von Vektorräumen: Seien $V_1, \ldots, V_n$ $K-$Vektorräume; wir betrachten das kartesische Produkt $V_1 \times \ldots \times V_n$ mit den Operationen
$$(v_1, \ldots, v_n) + (w_1, \ldots, w_n) = (v_1 + w_1, \ldots v_n + w_n),$$
$$\lambda(v_1, \ldots, v_n) = (\lambda v_1, \ldots, \lambda v_n).$$
Die Rechengesetze folgen aus deren Gültigkeit in den $V_i$.
Insbesondere ist $\underbrace{K \times \ldots \times K}_{n} =_{Df} K^n$ ein $K-$Vektorraum.

5. Sei $K$ ein Körper, dann ist $V = K^{m \times n}$ versehen mit der Matrixaddition und komponentenweiser $K-$Operation ein $K-$Vektorraum.

6. Das Grundbeispiel – der Ursprung des Vektorbegriffs:

   Ein Blick ins Konversationslexikon lehrt, daß es sich hier um eine ungeeignete Quelle handelt.

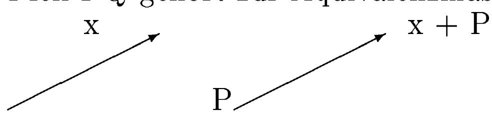
   1898: „Radiusvector: Bewegungslinie für einen sich vom Zentrum wegbewegenden Punkt."

   1932: „Vektor: Maßangabe für eine physikalische Größe, die eine bestimmte Richtung hat."

   (a) Sei $E$ der dreidimensionale Anschauungsraum und $P, Q \in E$ Punkte. Wir betrachten Pfeile $\vec{PQ}$ mit Anfangspunkt $P$ und Endpunkt$Q$. Zwei Pfeile $\overrightarrow{PQ}$ und $\overrightarrow{P'Q'}$ heißen äquivalent, wenn die Strecken $\overline{PQ}$ und $\overline{P'Q'}$ gleichlang sind und die Geraden durch $P, Q$ und $P', Q'$ parallel sind; wenn man diese Geraden durch Parallelverschiebung zur Deckung bringt, so sollen die Pfeile $\overrightarrow{PQ}$ und $\overrightarrow{P'Q'}$ dieselbe Orientierung haben.

   Ein Vektor $x$ ist dann eine Äquivalenzklasse von Pfeilen.

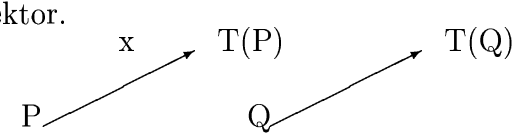
   (b) Ein Vektor induziert eine Abbildung $P \mapsto x + P = Q$ des Raumes $E$ auf sich, dabei ist $x + P = Q$ durch die Eigenschaft $\overrightarrow{PQ} \in x$ definiert, d.h. der Pfeil $\overrightarrow{PQ}$ gehört zur Äquivalenzklasse $x$.

   In der Äquivalenzklasse $x$ gibt es zu jedem Punkt $P$ genau einen Pfeil mit $P$ als Anfangspunkt, dann ist $x + P$ der Endpunkt dieses Pfeils.

   (c) Die Abbildung $P \mapsto x + P$ heißt die zum Vektor $x$ gehörige Translation $T_x$, dies ist eine Bijektion von $E$ auf $E$.

   (d) Umgekehrt: Eine Abbildung $T : E \longrightarrow E$ wird Translation genannt, falls alle Pfeile $\overrightarrow{PT(P)}$ eine Äquivalenzklasse $x$ bilden; $x$ heißt der zu $T$ gehörige Vektor.

   (e) Zwischen den Vektoren und den Translationen gibt es eine natürliche Bijektion:

   $$V(E) = \{Vektoren\} \leftrightarrow T(E) = \{Translationen\}$$

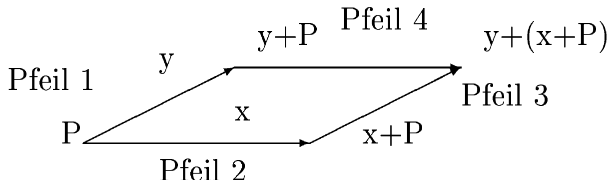
   Wir können nun eine Addition einführen: Seien $x, y$ Vektoren, wir definieren $x + y$ durch die Gleichung

   $$(x + y) + P = x + (y + P) \text{ für alle } P \in E,$$

d.h. $T_{x+y} = T_x \circ T_y$, die Summe der Vektoren entspricht der Hintereinanderausführung der entsprechenden Translationen.

Wir zeigen jetzt: Aus der Parallelogrammregel folgt $T_{x+y} = T_{y+x}$, d.h. die Addition ist kommutativ.

Zu zeigen ist $T_{x+y}(P) = T_{y+x}(P)$ für alle $P \in E$, d.h. $x+(y+P) = y+(x+P)$. Wie betrachten folgende Pfeile:




Nach Konstruktion sind Pfeil 1 und Pfeil 3 äquivalent, sie gehören zu $y$.

Die Parallelogrammregel besagt: Wenn Pfeil 1 und Pfeil 3 äquivalent sind (also: parallel, gleich lang, gleich gerichtet), dann müssen auch Pfeil 2 und Pfeil 4 äquivalent sein. Wenn also Pfeil 2 zu $x$ gehört, so auch Pfeil 4.

Der Pfeil 5 habe den Anfangspunkt $y + P$ und den Endpunkt $x + (y + P)$, dieser gehört zu $x$. Da Pfeil 4 und Pfeil 5 denselben Anfangspunkt haben und beide zum Vektor $x$ gehören, müssen sie auch denselben Endpunkt haben, also $y + (x + P) = x + (y + P)$. □

(f) Der Nullvektor $o$ ist nun die KLasse der Pfeile der Form $\overrightarrow{PP}$, die entsprechende Translation ist $T_o = id_E$: $T_o(P)$ ist der Endpunkt des Pfeils mit dem Anfangspunkt $P$, der zur Klass $o$ gehört, also $T_o(P) = P$.

(g) Multiplikation mit Skalaren: Sei $a \in R$ eine reelle Zahl und $v$ ein Vektor. Wenn $a = 0$ oder $v = o$ ist, setzen wir $av = o$.

Sei also $a \neq 0$ und $v \neq o$ und sei $\overrightarrow{PQ} \in V$ ein Pfeil. Wir betrachten den Zahlenstrahl mit Nullpunkt in $P$ und 1 in $Q$. Dann sei $Q'$ derjenige Punkt des Strahls, welcher zur Zahl $a$ gehört. Wir setzen fest: $av$ sei die Äquivalenzklasse, welche den Pfeil $\overrightarrow{PQ'}$ enthält.

Damit wird $V(E)$ zu einem $R-$Vektorraum; entsprechend wird $T(E)$ zu einem $R-$Vektorraum (durch $a \cdot T_v = T_{av}$). Die Zuordnung $v \mapsto T_v$ ist dann ein „Isomorphismus“ von $R-$Vektorräumen.

### 5.1.4 Beispiele

Neben den Grundbeispielen

- Zeilenvektoren,
- Spaltenvektoren,
- Vektoren im „anschaulichen Sinn“

lassen sich auch

- Funktionen, Polynome und
- Matrizen mit fixierten Format

als Vektoren auffassen.

Wir bemerken, daß die Definition des Vektorraumbegriffs so allgemein wie möglich gehalten ist, um ein Maximum an Anwendungsmöglichkeiten der Theorie zu sichern:

- Der Skalarkörper $K$ kann ein beliebiger Körper (vgl. 4.1) sein, nicht nur $R$ oder $C$.
- Es bleibt offen, was die Addition von Vektoren und die Multiplikation von Vektoren mit Skalaren genau sein soll, sondern es werden nur die „Spielregeln“ vorgegeben. In diesem Sinne versucht man auch, die Grundtatsachen der Theorie unabhängig von einem konkreten Beispiel zu entwickeln.

### 5.1.5 Unterräume

**Definition**

Sei $V$ ein $K$-Vektorraum und $U \subseteq V$ eine Teilmenge. Dann heißt $U$ Unterraum von $V$, falls folgendes gilt:

1. $(U, +)$ ist eine Untergruppe von $(V, +)$.
2. Für jedes $\lambda \in K$ und jedes $u \in U$ ist $\lambda u \in U$.

Aus der ersten Forderung folgt, daß $U$ mindestens das neutrale Element $o$ enthält. Die Teilmenge $\{o\}$ von $V$ ist bereits ein Unterraum ($\lambda o = o$).

**Kriterium**

Eine (somit nichtleere) Teilmenge $U \subseteq V$ ist ein Unterraum, wenn

1. für alle $u, v \in U$ auch $u + v \in U$ ist und
2. für alle $\lambda \in K$ und $u \in U$ auch $\lambda u \in U$ gilt.

Beweis: Die jeweils zweiten Bedingungen in der Definition und im Kriterium stimmen überein. Es ist nur die Gültigkeit der ersten Bedingung der Definition nachzuweisen. Da $U$ nicht leer ist, existiert ein $u \in U$. Wir wählen $\lambda = 0$. Dann folgt $0u = o \in U$, also enthält $U$ den Nullvektor. Weiter ist auch $-u = (-1) \cdot u \in U$ für alle $u \in U$, also ist $(U, +)$ eine Untergruppe von $(V, +)$. $\square$

**Beispiele für Unterräume**

1. Sei $V = R^{1\times 2} = \{(a,b) \mid a,b \in R\}$ der Vektorraum der Zeilenvektoren, sei $U$ die Teilmenge aller Vektoren der Form $(0,b)$, deren erste Koordinate null ist. Dies ist ein Unterraum.

2. Sei $V = V(E)$ der Vektorraum des 3-dimensionalen Anschauungsraums $E$. Wir betrachten im Raum $E$ irgendeine Ebene $F \subset E$. Sei $V(F)$ die Menge der Vektoren $x$, die Repräsentanten der Form $\overrightarrow{PQ}$ mit $P, Q \in F$ haben.

   Wir behaupten, daß diese Vektoren gerade den Translationen entsprechen, die die Ebene $F$ in sich überführen.

   **Beweis:** Sei also $x$ ein Vektor mit einem Repräsentanten $\overrightarrow{PQ} \in F$ und $P' \in F$ ein beliebiger Punkt. Dann gibt es genau einen Pfeil $\overrightarrow{P'Q'}$ mit dem Anfangspunkt $P'$, welcher zu $x$ gehört, und dessen Endpunkt $Q'$ liegt in $F$. Dies ist offensichtlich, wenn $P'$ auf der Geraden $\overline{PQ}$ liegt, denn die gesamte Gerade gehört zu $F$. Wenn $P' \notin \overline{PQ}$ ist, dann gehört das Dreieck $PQP'$ zur Ebene $F$. Wir betrachten den Vektor $y \in V(F)$ mit dem Repräsentanten $\overrightarrow{PP'}$. Dann ist $Q' = x + (y + P) = y + (x + P) = y + Q$ Eckpunkt des vom Dreieck $PQP'$ erzeugten Parallelogramms. Also ist $T_x(P') = Q' \in F$, d.h. $T_x$ führt die Ebene $F$ in sich über.

3. Sei $V = K[X]$ der Raum der Polynome mit Koeffizienten in $K$ und sei eine natürliche Zahl $n$ fixiert. Dann ist die Menge aller Polynome vom Grad $\leq n$ ein Unterraum.

### 5.1.6 Satz

*Sei $V$ ein $K$-Vektorraum und $\{U_i \mid i \in I\}$ eine Familie von Unterräumen. Dann ist der Durchschnitt $W = \bigcap_{i\in I} U_i$ ebenfalls ein Unterraum von $V$.*

Beweis: Wir verifizieren das Kriterium. Da $U_i$ ein Unterraum ist, gilt $o \in U_i$ für alle $U_i$. Also gehört $o$ auch zum Durchschnitt $W$, also ist $W$ nicht leer. Seien weiter $w_1, w_2 \in W$. Wir betrachten deren Summe. Nehmen wir ein beliebiges $U_i$; weil $w_1, w_2 \in U_i$ gilt und $U_i$ ein Unterraum ist, gilt $w_1 + w_2 \in U_i$ und damit liegt die Summe in $W$. Genauso sieht man, daß jedes skalare Vielfache eines Elements von $W$ wieder zu $W$ gehört, da es in jedem $U_i$ liegt. □

### 5.1.7 Folgerung / Definition

Sei $M$ eine Teilmenge eines Vektorraums $V$. Dann gibt es einen eindeutig bestimmten kleinsten Unterraum $W \subseteq V$, der alle Vektoren aus $M$ enthält. Man nennt $W$ den Spann von $M$; $W = Spann(M)$.

Beweis: Wir betrachten die Familie $\{U_i\}$ aller Unterräume von $V$, welche die Menge $M$ enthalten. Dann leistet $W = \bigcap U_i$ das Verlangte.

Beispiel: Sei $x \in V(E)$ und $\overrightarrow{PQ} \in x$ ein Pfeil mit $P \neq Q$, also $x \neq o$. Sei $G$ die Gerade durch $P, Q$ und $V(G)$ die Menge aller Vektoren, welche einen Repräsentanten $\overrightarrow{P'Q'}$ auf

$G$ haben (dies ist die Menge aller Translationen, die $G$ in sich überführen). Dann ist $V(G) = Spann(x)$.

Im Folgenden wollen wir eine direkte Beschreibung von $Spann(M)$ geben. Grundlegend ist der Begriff der Linearkombination.

### 5.1.8 Definition

Sei $M \subseteq V$ eine Teilmenge von Vektoren. Eine *Linearkombination* von $M$ ist ein Ausdruck der Form

$$u = \sum_{v \in M} \lambda_v v \in V,$$

wobei über skalare Vielfache der Vektoren aus $M$ summiert wird; dabei müssen fast alle (d.h. alle, bis auf endlich viele) Skalare $\lambda_v$ null sein. Der „Träger“ von $u$, also $\{v \mid \lambda_v \neq 0\}$ ist also eine *endliche* Teilmenge von $M$. (Wenn $M$ eine endliche Menge ist, gibt es also keine Einschränkung.)
Konvention: Wenn $M = \emptyset$, dann bezeichnet man den Nullvektor $o$ als einzige Linearkombination von $M$.

### 5.1.9 Satz

*Sei $M$ eine Teilmenge von $V$. Dann ist die Menge $W$ aller Linearkombinationen von $M$ ein Unterraum von $V$ und es gilt $W = Spann(M)$.*
Beweis: Für $M = \emptyset$ ist $W = \{o\}$.
Sei also $M \neq \emptyset$. Der Nullvektor ist immer eine Linearkombination von $M$, also ist $W \neq \emptyset$. Seien $u = \sum_{v \in S} \lambda_v v$ und $w = \sum_{v \in T} \mu_v v$ zwei Linearkombinationen mit den Trägern $S$ bzw. $T$; dann ist $S \cup T$ eine endliche Teilmenge von $M$ und

$$u = \sum_{v \in S \cup T} \lambda_v v,$$

und

$$w = \sum_{v \in S \cup T} \mu_v v,$$

also ist auch

$$u + w = \sum_{v \in S \cup T} (\lambda_v + \mu_v) v$$

eine Linearkombination von $M$.
Wenn $\lambda \in K$ ist, so ist auch

$$\lambda u = \sum_{v \in S} (\lambda \lambda_v) v$$

eine Linearkombination von $M$.
Somit ist $W$ ein Unterraum, welcher alle $v$ aus $M$ enthält, also $W \supseteq Spann(M)$. Andererseits ist $Spann(M)$ ein Vektorraum, der $M$ enthält, also enthält $Spann(M)$ alle Linearkombinationen aus $M$, also $W \subseteq Spann(M)$. □

Beispiel: Sei $V = R^{1 \times 3} = \{(a,b,c) \mid a,b,c \in R\}$ und $M = \{(1,0,0),(0,1,0)\}$, dann ist $Spann(M) = \{(a,b,0) \mid a,b \in R\}$.

### 5.1.10 Definition

Sei $V$ ein $K$-Vektorraum und $M$ eine Teilmenge. Man nennt $M$ ein Erzeugendensystem von $V$, falls $Spann(M) = V$ ist. Man nennt $V$ einen *endlich erzeugten* Vektorraum, falls es eine endliche Menge $M$ mit $V = Spann(M)$ gibt, andernfalls heißt $V$ *nicht endlich erzeugt*.

Beispiele: Der Polynomring $K[X]$ istein Beispiel für einen nicht endlich erzeugten $K$-Vektorraum. Die Polynome vom Grad $\leq n$ bilden darin eine endlich erzeugten Unterraum mit dem Erzeugendensystem $\{1, X, X^2, \ldots, X^n\}$.
Der Raum $R^{1,3}$ ist endlich erzeugt, er besitzt das Erzeugendensystem $\{(1,0,0),(0,1,0),(0,0,1)\}$.

Es gibt viele wichtige Beispiele von nicht endlich erzeugten Vektorräumen. Funktionsräume sind meist nicht endlich erzeugt. In dieser Vorlesung betrachten wir meistens nur endlich erzeugte Vektorräume, man kann sich diese oft geometrisch veranschaulichen.

Zunächst noch eine grundlegende

### 5.1.11 Definition

Sei $M$ eine Teilmenge von $V$. Die triviale Linearkombination von $M$ ist der Ausdruck $\sum_{v \in M} \lambda_v v$, wobei alle $\lambda_v$ gleich 0 sind. Das Ergebnis der trivialen Linearkombination ist offensichtlich der Nullvektor.
Wir nennen die Menge $M$ linear unabhängig, falls es nur eine Linearkombination von $M$ gibt, welche den Nullvektor darstellt, nämlich die triviale Lienarkombination. Andernfalls heißt $M$ linear abhängig.

Beispiel: Die Vektoren $e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1) \in R^{1,3}$ sind linear unabhängig.

### 5.1.12 Satz

1. *Sei die Menge $M \subset V$ linear unabhängig und sei $w \notin Spann(M)$. Dann ist auch $M \cup \{w\}$ linear unabhängig.*

2. *Umgekehrt: Ist $M$ linear unabhängig und $M \cup \{w\}$ linear abhängig, so folgt $w \in Spann(M)$.*

Beweis: Sei $M' = M \cup \{w\}$. Wir betrachten die Linearkombination

$$\sum_{v \in M'} \lambda_v v = o$$

und nehmen an, daß nicht alle $\lambda_v$ gleich 0 sind.
Wenn $\lambda_w \neq 0$, so multiplizieren wir $\sum_{v \in M} \lambda_v v = -\lambda_w w$ mit $(-\lambda_w)^{-1}$ und erhalten

$$w = -\sum_{v \in M} \frac{\lambda_v}{\lambda_w} v \in Spann(M),$$

ein Widerspruch. Also muß $\lambda_w = 0$ sein. Daher gilt

$$\sum_{v \in M'} \lambda_v v = \sum_{v \in M} \lambda_v v = o.$$

Da nach Voraussetzung $M$ linear unabhängig ist, ist $\lambda_v = 0$ für alle $v \in M$. Da wir schon sahen, daß $\lambda_w = 0$ ist, ist die betrachtete Linearkombination die triviale, also ist $M'$ linear unabhängig.

Die zweite Aussage folgt aus der ersten, denn wenn $w \notin Spann(M)$ wäre, so wäre $M \cup \{w\}$ linear unabhängig. □

Bemerkungen:

1. Einzelne Vektoren: Der Nullvektor $o$ ist linear abhängig, denn $1 \cdot o = 0$ und $1 \neq 0$. Jeder Vektor $v \neq o$ ist linear unabhängig, denn aus $\lambda v = o$ folgt $\lambda = 0$.

2. Der Satz gibt eine konstruktive Methode, um Mengen linear unabhängiger Vektoren aufzubauen. Wir beginnen mit $v_1 \neq o$, dieser Vektor ist linear unabhängig. Nun wählen wir $v_2 \notin Spann(v_1)$, dann sind $v_1, v_2$ linear unabhängig. Dann wählen wir $v_3 \notin Spann(v_1, v_2)$, dann sind $v_1, v_2, v_3$ linear unabhängig, usw. Das Verfahren bricht ab, sobald $Spann(v_1, \dots, v_n) = V$ ist, dann ist $V$ endlich erzeugt. Ist dagegen $V$ nicht endlich erzeugt, so kommt man nie zum Ende.

3. Der obige Satz wird auch als Abhängigkeitslemma bezeichnet.

Wir haben ein zweites Kriterium für die lineare Unabhängigkeit:

### 5.1.13 Satz (Eindeutigkeitslemma)

*Sei $M \neq \emptyset$ eine Teilmenge von $V$. Dann ist folgendes äquivalent:*

1. *$M$ ist linear unabhängig.*

2. *Jedes $u \in Spann(M)$ läßt sich auf genau eine Art und Weise als Linearkombination von $M$ realisieren.*

Beweis: (2) $\Rightarrow$ (1): Die triviale Linearkombination realisiert den Nullvektor $0 \in Spann(M)$. Nach Voraussetzung gibt es keine andere Linearkombination, die $o$ realisiert, also ist $M$ linear unabhängig.

(1) $\Rightarrow$(2): Sei $M$ linear unabhängig und $u \in Spann(M)$. Wir nehmen an, daß es zwei verschiedene Linearkombinationen von $M$ gibt, die $u$ realisieren:

$$u = \sum_{v \in S} \lambda_v v = \sum_{v \in T} \mu_v v.$$

Dann können wir in beiden Sumen auch über $S \cup T$ summieren. Wir bilden nun die Differenz:

$$\sum_{v \in S \cup T} (\lambda_v - \mu_v) v = o.$$

Da $M$ linear unabhängig ist, müssen für alle $v$ die Zahlen $\lambda_v - \mu_v = 0$ sein, also $\lambda_v = \mu_v$
□

### 5.1.14 Definition: Basis eines $K$-Vektorraums

Sei $V$ ein $K$-Vektorraum. Eine Teilmenge $B$ von $V$ heißt Basis, wenn folgende Eigenschaften erfüllt sind:

1. $B$ ist linear unabhängig.

2. $B$ erzeugt $V$, d.h. es gilt $V = Spann(B)$.

Bemerkungen:

1. Wenn $B$ eine Basis von $V$ ist, dann folgt aus dem Eindeutigkeitslemma, daß sich jeder Vektor $u \in V$ auf genau eine Art und Weise als Linearkombination von $B$ schreiben läßt.

2. („Spitzfindigkeiten") Wir betrachten die leere Menge $M = \emptyset$. Es gibt keine nichttrivialen Linearkombinationen von $M$, welche den Nullvektor $o$ darstellen (denn eine nichttriviale Linearkombinationen müßte einen Träger haben (vgl 5.1.8)). In diesem Sinne ist die leere Menge linear unabhängig.

   $Spann(\emptyset)$ ist der kleinste Unterraum, welcher $\emptyset$ enthält. In diesem Sinne ist $Spann(\emptyset) = \{o\}$ der Raum, der nur aus dem Nullvektor besteht, denn die leere Menge ist Teilmenge jedes Unterraums. Deshalb wird die leere Menge als Basis des Vektorraums $\{o\}$ betrachtet. Der Nullvektor selbst kann ja keine Basis dieses Raums sein, da er linear abhängig ist ($1 \cdot o = o$ ist eine nichttriviale Linearkombination).

3. Die Vektoren $e_1, e_2, e_3 \in R^{1,3}$ bilden eine Basis dieses Vektorraums.

4. Wir betrachten den anschaulichen Vektorraum. Seien $x, y, z \in V(E)$ und $P \in E$ ein fixierter Punkt; $\overrightarrow{PQ}_1 \in x, \overrightarrow{PQ}_2 \in y, \overrightarrow{PQ}_3 \in z$ seien Repräsentanten, welche in $P$ beginnen. Dann sind $x, y, z$ genau dann eine Basis, wenn die in $P$ beginnenden Pfeile einen Körper (ein Parallelepiped) aufspannen.

### 5.1.15 Satz

*Sei $V$ ein $K$-Vektorraum. Seien $M \subseteq M'$ zwei Teilmengen von $V$ mit den Eigenschaften:*

1. *$M$ ist linear unabhängig.*

2. *$M'$ erzeugt $V$.*

*Dann gibt es eine Basis $B$ von $V$ mit $M \subseteq B \subseteq M'$.*

Beweis: Wir betrachten die Familie $F$ aller Mengen $S$, welche folgende Eigenschaften haben:

1. $S$ ist linear unabhängig.

2. $S$ liegt zwischen $M$ und $M'$: $M \subseteq S \subseteq M'$.

Wir kennen mindestens eine solche Menge, nämlich $S = M$.

**Hilfssatz:** *Sei $S$ ein maximales Element der Familie $F$, dann ist $S$ eine Basis von $V$.*

Beweis: Weil $S$ maximal ist, muß für alle $x \in M' \backslash S$ die Menge $S \cup \{x\}$ linear abhängig sein. Da $S$ linear unabhängig ist, folgt aus dem Abhängigkeitslemma $x \in Spann(S)$, also $M' \subseteq Spann(S)$. Daraus folgt weiter $V = Spann(M') \subseteq Spann(S)$, also wird $V$ von $S$ erzeugt. Andererseits ist $S$ linear unabhängig, also eine Basis von $V$. Damit ist der Hilfssatz bewiesen.

Um den Hilfssatz anwenden zu können, müssen wir zeigen, daß es in der Familie $F$ maximale Elemente gibt. Wir beschränken uns auf den Fall, daß die Menge $M'$ endlich ist; dann ist $V$ endlich erzeugt.
Wir beginnen mit $S = M$. Wenn $Spann(M) = V$ ist, dann ist $M$ eine Basis.
Wenn $Spann(M) \neq V$ ist, dann kann $M'$ nicht in $Spann(M)$ enthalten sein, weil $Spann(M') = V$ ist. Also exisitiert ein $x_1 \in M'$, $x_1 \notin Spann(M)$. Dann ist $M_1 = M \cup \{x_1\}$ linear unabhängig und $M \subseteq M_1 \subseteq M'$.
Wenn $Spann(M_1) = V$ ist, dann ist $M_1$ eine Basis von $V$. Andernfalls findet man wieder ein $x_2 \in M'$, $x_2 \notin Spann(M_1)$, also gehört $M_2 = M_1 \cup \{x_2\}$ zu $F$.
Weil die Menge $M'$ endlich ist, muß das Verfahren abbrechen, d.h. wir finden eine Basis $B = M_n$. □
Wenn die Menge $M'$ unendlich ist, dann muß das Verfahren nicht abbrechen, d.h. wir finden möglicherweise eine unendliche aufsteigende Folge

$$M \subset M_1 \subset M_2 \subset \ldots \subset M'$$

und alle $M_i$ gehören zu $F$. Dennoch kann man mit dem Zornschen Lemma zeigen, daß die Familie $F$ maximale Elemente haben muß.

### 5.1.16 Folgerung

1. *Jede linear unabhängige Teilmenge $M$ von $V$ kann zu einer Basis $B$ ergänzt werden.*
2. *Jedes Erzeugendensystem $M'$ von $V$ enthält eine Basis.*
3. *Jeder Vektorraum $V$ besitzt eine Basis.*

Beweis:

1. $M$ ist gegeben, wir wenden den Satz mit $M' = V$ an.
2. $M'$ ist gegeben, wir wenden den Satz mit $M = \emptyset$ an.
3. Wir wenden den Satz mit $M = \emptyset$ und $M' = V$ an. □

### 5.1.17 Beispiel: Der Zeilenraum einer Matrix

Sei $A \in K^{m\times n}$ eine Matrix vom Format $(m,n)$ mit Einträgen aus dem Körper $K$.
Als Zeilenraum $ZR(A)$ von $A$ bezeichnen wir den Unterraum von $K^{1\times n}$, der von den Zeilen von $A$ erzeugt wird.

**1.**
Sei $B \sim A$ eine zeilenäquivalente Matrix; dann gilt $ZR(B) = ZR(A)$.

Beweis: Die Matrix $B$ geht aus $A$ durch eine Folge elementarer Zeilenoperationen hervor, d.h. $B = CA$, wobei $C \in K^{m\times m}$ invertierbar ist. Also ist die $i$-te Zeile von $B$ das Produkt der $i$-ten Zeile von $C$ und $A$:

$$(b_{i1},\dots,b_{in}) = (c_{i1},\dots,c_{im})A = \sum_{j=1}^{m} c_{ij} Z_j(A), .$$

wobei die $j$-te Zeile $Z_j(A)$ mit $c_{ij}$ multipliziert wird. Also liegt die $i$-te Zeile von $B$ in $ZR(A)$, und das gilt für alle $i$, d.h. $ZR(B) \subseteq ZR(A)$. Weil $C$ invertierbar ist, gilt $A = C^{-1}B$, also auch $ZR(A) \subseteq ZR(B)$. □

**2.**
Sei $B \sim A$ und $B$ habe Zeilenstaffelung. Dann sind die von Null verschiedenen Zeilen von $B$ eine Basis des Zeilenraums von $A$.

Beweis: Die Zeilen von $B$ und die Zeilen von $A$ haben in $K^{1\times n}$ denselben Spann. Wir zeigen, daß die Zeilen von $B$ linear unabhängig sind:

$$B = \begin{pmatrix} b_{1i_1} & \dots & \dots & \dots & b_{1m} \\ & b_{2i_2} & \dots & & \\ & & b_{3i_3} & \dots & \\ & & \dots & & \end{pmatrix}$$

mit $i_1 < i_2 < i_3 < \dots$ und $b_{ki_k} \neq 0$.
Wir stellen nun die Null-Zeile als Linearkombination der Zeilen von $B$ dar:

$$\begin{array}{rl} \lambda_1 & (\, b_{1i_1} \quad \dots \quad \dots \quad \dots \quad b_{1m}\,) \\ +\lambda_2 & (\, \dots \quad b_{2i_2} \quad \dots \quad \dots \quad b_{2m}\,) \\ +\lambda_3 & (\, \dots \quad \dots \quad b_{3i_3} \quad \dots \quad b_{3m}\,) \\ & \qquad \dots \\ = & (\, 0 \quad 0 \quad \dots \quad \dots \quad \dots \quad 0\,)\,. \end{array}$$

Anhand der 1. Spalte sehen wir $\lambda_1 = 0$, also können wir die erste Zeile weglassen. Dann sehen wir $\lambda_2 = 0$, usw.
Also sind die Zeilenvektoren von $B$ linear unabhängig und erzeugen $ZR(A)$, bilden also eine Basis.

Beispiel:

$$A = \begin{pmatrix} 1 & 3 & 1 & -1 \\ -1 & -1 & -6 & 4 \\ -1 & -3 & -1 & 5 \\ 1 & 5 & -4 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 3 & 1 & -1 \\ 0 & 2 & -5 & 3 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Die von Null verschiedenen Zeilen von $B$ bilden eine Basis von $ZR(A)$.

## 5.2 Endlich erzeugte Vektorräume

In 5.1.9 hatten wir den Begriff des endlich erzeugten Vektorraums $V$, d.h. $V = Spann(M)$, wobei $M$ eine endliche Teilmenge von $V$ ist. Dann bekommen wir die Existenz einer Basis mit Hilfe von 5.1.15 *ohne* Zornsches Lemma. Wir nehmen dort $M = \emptyset$ und für $M'$ „unser“ $M$. Also:

### 5.2.1 Satz

*Jeder endlich erzeugte $K$-Vektorraum $V$ besitzt eine Basis $B$, welche aus endlich vielen Elementen besteht.*

Das Ziel dieses Abschnitts besteht darin zu zeigen, daß eine beliebige Basis $B'$ von $V$ nur endlich viele Elemente enthält, und zwar genauso viele, wie die schon gefundene Basis $B$.

### 5.2.2 Schrankenlemma

*Der Vektorraum $V$ besitze ein Erzeugendensystem aus $n$ Elementen. Dann müssen $n+1$ Elemente aus $V$ stets linear abhängig sein, d.h. linear unabhängige Teilmengen von $V$ haben höchstens $n$ Elemente.*

Beweis: Sei $S = \{v_1, \dots v_n\}$ ein Erzeugendensystem, d.h. $V = Spann(S)$. Seien $w_1, \dots, w_{n+1} \in V$ beliebige $n+1$ Vektoren. Dann existiert eine Matrix $A \in K^{n\times(n+1)}$, so daß

$$(1) \qquad (w_1, \dots, w_{n+1}) = (v_1, \dots, v_n) \cdot A,$$

d.h.

$$w_j = \sum_{i=1}^{n} v_i a_{ij} \in Spann(s).$$

**Hilfssatz F (= Fundamentallemma):** *Das System $AX = 0$ mit $n$ Gleichungen und $n+1$ Variablen hat mindestens eine nichttriviale Lösung $B \in K^{(n+1)\times 1}$.* (Der Beweis folgt gleich unter 5.2.3)

Wir multiplizieren (1) von rechts mit $B$ und erhalten

$$(w_1, \dots, w_{n+1})B = 0,$$

weil $AB = 0$ ist. Da $B$ nicht der Nullvektor ist, bedeutet das für $w_1, \dots, w_{n+1}$ eine lineare Abhängigkeit. □

### 5.2.3 Fundamentallemma

*Ein homogenes lineares Gleichungssystem $AX = 0$ mit mehr Variablen als Gleichungen, d.h. $A \in K^{m\times n}$ und $m < n$ hat immer eine nichttriviale Lösung. ($K$ ist ein beliebiger Körper.)*

Beweis: Wir betrachten die Matrix $A' \sim A$, die zu $A$ zeilenäquivalent ist und reduzierte Zeilenstufenform besitzt. Dann ist die Zahl der Pivots von $A'$ höchstens gleich der

Zeilenzahl $m$ von $A'$, also kleiner als die Zahl $n$ der Spalten von $A'$, also kleiner als die Zahl der Variablen. Wir wissen aus dem Abschnitt 1.1, daß die Zahl der frei wählbaren Variablen in der Lösung gleich der Differenz der Zahl der Variablen minus der Zahl der Pivots ist, also mindestens gleich $n - m > 0$. (Die Pivotspalten entsprechen den abhängigen Variablen, die restlichen Spalten den frei wählbaren Variablen.) Also ist mindestens eine Variable frei wählbar, d.h. es gibt von $X = 0$ verschiedene Lösungen des Systems $AX = 0$. □

### 5.2.4 Hauptsatz (Basissatz für endlich erzeugte Vektorräume)

*Sei $V \neq \{o\}$ ein endlich erzeugter $K$-Vektorraum, d.h. es gibt eine endliche Teilmenge $M \subset V$, so daß $V = Spann(M)$ ist. Dann gilt:*

1. *$V$ besitzt eine Basis $B$ mit $\#B \leq \#M$.*

2. *Je zwei Basen haben gleich viele Elemente. Diese Anzahl heißt die Dimension von $V$.*

3. **Ergänzungssatz:** *Jedes System $\{x_1, \ldots, x_r\}$ linear unabhängiger Elemente kann durch endlich viele Vektoren $x_{r+1}, \ldots, x_n$ zu einer Basis $B$ von $V$ ergänzt werden.*

4. **Austauschsatz von Steinitz:** *Ist bereits eine Basis $B = \{b_1, \ldots, b_n\}$ von $V$ gegeben, dann können die benötigten Elemente $x_{r+1}, \ldots, x_n$ innerhalb von $B$ gewählt werden.*

   *Andersherum: Man findet in $B$ ein Teilsystem von $r$ Vektoren, so daß die Eigenschaft, Basis zu sein, erhalten bleibt, wenn man diese Vektoren gegen $\{x_1, \ldots, x_r\}$ austauscht.*

Beweis: (1) Dies ist Satz 5.2.1.
(2) Seien $B_1, B_2$ Basen. Da $B_2$ den Raum $V$ erzeugt, folgt aus dem Schrankenlemma $\#B_1 \leq \#B_2$, da $B_1$ linear unabhängig ist. Aus denselben Gründen folgt $\#B_2 \leq \#B_1$.
(3, 4) Sei $M = \{x_1, \ldots x_r\}$ eine linear unabhängige Menge in $V$ und sei $B = \{b_1, \ldots, b_n\}$ irgendeine Basis. Wir betrachten $M' = M \cup B$, dies ist ein Erzeugendensystem. Wir wenden nun 5.1.14 an (wir benötigen nicht das Zornsche Lemma, weil die Mengen $M, M'$ endlich sind) und finden eine Basis $B'$ mit

$$M \subseteq B' \subseteq M'.$$

Nun gilt $\#B' = \#B = n$, also besteht $B'$ genau aus $M$ und $n - r$ Elementen aus $B$. □

### 5.2.5 Definition

Sei $V$ ein endlich erzeugter $K$-Vektorraum. Dann wird die eindeutig bestimmte Anzahl der Elemente aller Basen von $V$ als *Dimension* des $K$-Vektorraums bezeichnet, man schreibt $\dim_K(V)$ oder einfach $\dim(V)$. Die Dimension ist gleich der Maximalzahl linear unabhängiger Vektoren und gleich der Minimalzahl der Elemente eines Erzeugendensystems von $V$.

### 5.2.6 Folgerung

*Ist $V$ ein $K$-Vektorraum der Dimension $n$, dann muß jeder Unterraum $U \subseteq V$ eine Dimension $\leq n$ haben. Wenn $U$ dieselbe Dimension wie $V$ hat, so ist $U = V$. Andersherum: Wenn $U \neq V$ ist, dann hat $U$ eine echt kleinere Dimension als $V$.*

Beweis: Sei $B'$ eine Basis von $U$, dann sind deren Elemente linear unabhängige Vektoren von $V$. Also gilt $\#B' \leq \dim(V)$ nach dem Schrankenlemma.
Sei nun $U \neq V$, dann gibt es ein $x \in V$ mit $x \notin U = Spann(B')$. Also ist $B' \cup \{x\}$ auch linear unabhängig, also hat $U$ eine echt kleinere Dimension als $V$. □

### 5.2.7 Übersicht über die Basen eines Vektorraums (Basiswechselsatz)

*Es sei $B$ eine Basis von $V$ und $A = (\alpha_{ij}) \in K^{n \times n}$ eine Matrix. Wir setzen*

$$(b_1, \ldots, b_n) \cdot A := (\sum_i b_i \alpha_{i1}, \sum_i b_i \alpha_{i2}, \ldots) = (b'_1, b'_2, \ldots b'_n).$$

*Dann ist folgendes äquivalent:*

1. *$A \in GL_n(K)$ ist eine invertierbare Matrix.*
2. *$(b'_1, \ldots, b'_n)$ ist wieder eine Basis.*

Beweis: (1) $\Rightarrow$ (2) Wenn $A$ invertierbar ist, so existiert eine Matrix $B \in K^{n \times n}$ mit $AB = I_n$. Somit folgt

$$(b_1, \ldots, b_n) = (b'_1, \ldots, b'_n)B,$$

also können alle $b_i$ als Linearkombinationen der $b'_i$ angesehen werden. Also ist $b'_1, \ldots, b'_n$ ein Erzeugendensystem von $V$, weil ja $b_1, \ldots, b_n \in Spann(b'_1, \ldots, b'_n)$ gilt. Andererseits ist dieses Erzeugendensystem minimal, weil es genau $n = \dim(V)$ Elemente enthält. Also ist $(b'_1, \ldots, b'_n)$ eine Basis.
(2) $\Rightarrow$ (1) Umgekehrt sei $b'_1, \ldots, b'_n)$ auch eine Basis. Dann finden wir eine Matrix $A' \in K^{n \times n}$ mit

$$(b_1, \ldots, b_n) = (b'_1, \ldots, b'_n)A'.$$

Durch Substitution erhalten wir

$$(b_1, \ldots, b_n) = (b_1, \ldots, b_n)(AA'),$$

jedoch gilt

$$(b_1, \ldots, b_n) = (b_1, \ldots, b_n)I_n.$$

Da $b_1, \ldots, b_n$ linear unabhängig sind, können wir das Eindeutigkeitslemma anwenden, es folgt $AA' = I_n$, also ist die Matrix $A$ invertierbar. □

### 5.2.8 Konvention

Als Basis eines $n$-dimensionalen Vektorraums $V$ bezeichnen wir die *geordnete* Menge $B = (b_1, \ldots, b_n)$. Wenn $B' = (b'_1, \ldots, b'_n)$ ist, so sagen wir $B = B'$ genau dann, wenn $b_1 = b_1, b_2 = b'_2, \ldots, b_n = b'_n$ ist. Eine geänderte Reihenfolge ergibt eine andere Basis.

Sei $X$ die Menge aller Basen des Vektorraums $V$. Dann operiert $G = GL_n(K)$ von rechts auf der Menge $X$:

$$X \times G \longrightarrow X,\ (B, A) \mapsto BA$$

und diese Operation ist einfach transitiv, denn es gibt genau eine Bahn (von einer Basis $B$ ausgehend kommt man durch Anwendung von $G$ zu allen Basen) und wenn $BA = B$ ist, so muß $A = I$ sein (anders gesagt: die Anwendung von $A \neq I$ verändert die Basis). Wenn eine Basis $B$ fixiert wird, dann ist die Zuordnung

$$A \in G \mapsto BA \in X$$

eine Bijektion zwischen $G$ und $X$.

### 5.2.9 Folgerung

*Ist $K$ ein endlicher Körper mit $q$ Elementen, dann gilt*

$$\#GL_n(K) = (q^n - 1)(q^n - q) \cdots (q^n - q^{n-1})$$

*und dies ist gleich der Anzahl der Basen eines $n$-dimensionalen $K$-Vektorraums.*

Beweis:Sei $V = K^{1\times n}$, dies ist ein $n$-dimensionaler Vektorraum, also $\#GL_n(K) = \#X$ und $\#V = (\#K)^n = q^n$.
Alle von $o$ verschiedenen Vektoren sind linear unabhängig, also gibt es für die Auswahl des ersten Basisvektors $b_1$ genau $q^n - 1$ Möglichkeiten. Nun hat $Spann(b_1) = Kb_1$ genau $q$ Elemente, das zweite Basiselement $b_2$ muß in $V - Spann(b_1)$ liegen, dafür gibt es also $q^n - q$ Möglichkeiten. Weiter hat $Spann(b_1, b_2) = Kb_1 + Kb_2$ hat $q^2$ Elemente, also gibt es $q^n - q^2$ Möglichkeiten für den dritten Basisvektor, usw. Für den letzten Basisvektor bleiben $q^n - q^{n-1}$ Möglichkeiten. □

Als Beipiel betrachten wir den Körper mit 2 Elementen. Hier gibt es 16 $2 \times 2$-Matrizen, davon sind 6 invertierbar.

**Die Summe von Unterräumen**

Sei $V$ ein $K$-Vektorraum und seien $U, W$ zwei Unterräume.

**Definition:** Mit $U + W := Spann(U \cup W)$ bezeichnen wir die Summe dieser Unterräume; dies ist der kleinste Unterraum von $V$, der sowohl $U$ als auch $W$ enthält.

**Behauptung:** $Spann(U \cup W) = \{u + w \mid u \in U,\ w \in W\}$.

Beweis: ($\subseteq$) Mit $u \in U,\ w \in W$ muß $Spann(U, W)$ auch $u + w$ enthalten.
($\supseteq$) Umgekehrt sieht man leicht, daß $\{u + w \mid u \in U,\ w \in W\}$ ein Unterraum von $V$ ist, der $U$ und $W$ enthält, also ist er im kleinsten der diese Unterräume enthaltenden enthalten. □

Mit den Unterräumen $U, W \subseteq V$ erhalten wir also zwei neue Unterräume von $V$, nämlich $U \cap W$ und $U + W$; der erste ist in beiden enthalten, der zweite enthält beide.

### 5.2.10 Fakt

$$\dim_K(U \cap W) + \dim_K(U + W) = \dim_K(U) + \dim_K(W)$$

**Folgerung:** Sei $\dim_K(V) = n$. Dann ist $\dim_K(U + W) \leq n$, also folgt aus dem Fakt, daß $\dim_K(U \cap W) \geq \dim_K(U) + \dim_K(W) - n$ ist.
Wenn also $U, W$ Unterräume im $n$-dimensionalen Raum $V$ sind und $\dim_K(U)+\dim_K(W) > n$ ist, so ist $\dim_K(U\cap W) \geq 1$, d.h. $U\cap W$ kann nicht nur aus dem Nullvektor bestehen.

### 5.2.11 Beispiel

Sei wieder $V(E)$ der Vektorraum des Anschauungsraums und $P \in E$ ein fixierter Punkt. Die (echten) Unterräume von $V(E)$ entsprechen dann umkehrbar eindeutig einer der folgenden Teilmengen von $E$:

1. dem Punkt $P$,
2. den Geraden durch $P$,
3. den Ebenen durch $P$.

Denn wenn $U$ ein Unterraum von $V(E)$ ist, so betrachten wir die Menge $U+P$, d.h. wir wenden alle Vektoren aus $U$ auf den Punkt $P$ an und wir erhalten einen der angeführten Fälle.
Umgekehrt betrachten wir $P$ bzw. eine Gerade $g$ durch $P$ bzw. eine Ebene $F$ durch $P$ und bilden

$$U = \{x \in V \mid x + P \in \begin{cases} \{P\} \\ g \\ F \end{cases}$$

Im ersten Fall besteht $U$ nur aus dem Nullvektor, im zweiten Fall ist $U$ ein 1-dimensionaler Unterraum von $V(E)$, im dritten Fall ein zweidimensionaler Unterraum.
Wenn nun $U, W$ Unterräume mit $\dim U + \dim W \leq 3$ sind, also zwei Geraden durch $P$ oder eine Gerade und eine Ebene durch $P$, dann kann $U \cap W = \{o\}$ sein, d.h. $(U + P) \cap (W + P) = P$. Dies ist der Fall, wenn es sich um zwei verschiedene Geraden durch $P$ oder eine Ebene und eine Gerade durch $P$, die nicht in der Ebene liegt, handelt.
Wenn jedoch $\dim U + \dim W = 4$ ist, dann sind $U + P$ und $W + P$ zwei Ebenen durch $P$ und deren Durchschnitt hat mindestens die Dimension 1, enthält also (mindestens) eine Gerade.

### 5.2.12 Weitere Beispiele für Unterräume

Sei $A \in K^{m\times n}$ eine $m \times n$-Matrix. Der *Spaltenraum* $SR(A) \subseteq K^{m\times 1}$ von $A$ ist der Spann der Spaltenvektoren $s_1(A), \ldots ; s_n(A) \in K^{m\times 1}$.
Der *Nullraum* $NR(A) \subseteq K^{n\times 1}$ von $A$ ist der Raum aller Vektoren $X \in K^{n\times 1}$ mit $A \cdot X = o$.

### 5.2.13 Hauptsatz: Verfahren zur Bestimmung einer Basis von $NR(A)$

1. Sei $A' \sim A$ zeilenäquivalent mit reduzierter Zeilenstaffelung; dann ist $NR(A') = NR(A)$, denn die elementaren Zeilenoperationen verändern nicht die Lösungsmenge des zugehörigen linearen Gleichungssystems (in unserem Fall benötigen wir nicht die erweiterte Koeffizientenmatrix, da die rechte Seite von $AX = o$ null ist).

2. Nun bilden wir aus $A' \in K^{m\times n}$ die Matrix $A'' \in K^{n\times n}$ folgendermaßen:

   Die Zeilen von $A'$ werden (ggf. durch Einfügen von Nullzeilen) so auseinandergezogen, daß die Pivots alle auf der Hauptdiagonalen stehen. Das Format der MAtrix wird dann auf $n \times n$ geändert, indem man am Schluß noch Nullzeilen hinzufügt oder wegläßt.

3. Wir betrachten nun die Differenz $M = I_n - A''$; die von Null verschiedenen Spaltenvektoren dieser Matrix haben die Form

   $$s_j(M) = \begin{pmatrix} * \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

   dabei steht die 1 in der $j$-ten Zeile (oder diese Spalte ist null).

   Sei $r \leq n$ die Zahl der Pivots in $A'$ und $I \subset \{1, \ldots, n\}$ die Menge der Spaltennummern, in denen die Pivots von $A'$ stehen, weiter sei $J$ die Komplementärmenge von $I$. Dann ist $\#J = n - r$ und für $j \in J$ ist $s_j(M)$ kein Nullvektor. **Diese Vektoren bilden eine Basis von** $NR(A)$.

   Also: Wenn $X \in NR(A)$, so ist $X = \sum_{j\in J} s_j(M)x_j$ mit frei wählbaren $x_j \in K$ und die Vektoren $s_j(M)$ sind linear unabhängig. Insbesondere gilt $\dim_K(NR(A)) = \#J = n - r$, dies ist der Freiheitsgrad des zugehörigen Gleichungssystems.

### 5.2.14 Beispiel

Sei $A \in K^{2\times 4}$ die Matrix mit

$$A' = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{pmatrix},$$

dann ist

$$A'' = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \in K^{4\times 4}$$

und

$$M = I_4 - A'' = \begin{pmatrix} 0 & -2 & 0 & -3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Die Pivotspalten haben die Nummern 1, 3, also sind die gesuchten Basisvektoren diejenigen aus den Spalten 2 und 4.

Einige Erläuterungen zum Verfahren:

Wir bilden $A \mapsto A' \mapsto A''$.
Die Nullräume von $A$ und $A'$ stimmen überein (klar), die von $A'$ und $A''$ auch, da wir hier nur Nullzeilen hinzugefügt oder weggelassen haben.

**Fakt:** Die Matrix $A''$ hat die Eigenschaft $(A'')^2 = A''$, derartige Matrizen heißen idempotent (oder Projektoren). Rechen Sie es im obigen Beispiel nach! [1]

Aus dem Fakt folgt nun

$$NR(A) = SR(I - A''),$$

denn $I - A''$ ist der Projektor auf $NR(A'')$:
Sei $A''X = o$, dann ist $X = X - A''X = (I - A'')X$; mit den obigen Bezeichnungen ist also

$$M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = s_1(M)x_1 + \dots s_n(M)x_n$$

eine Linearkombination der Spalten von $M$, d.h

$$X = MX \in SR(M),$$

also

$$NR(A'') \subseteq SR(M).$$

Sei umgekehrt $X \in SR(M)$,, d.h.

$$X = s_1(M)y_1 + \dots + s_n(M)y_n = M \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (I - A'')Y.$$

Wegen der Idempotenz von $A''$ ist $A''(I - A'')$ die Nullmatrix, also

$$A''X = A''(I - A'')Y = o,$$

also ist $SR(M) \subseteq NR(A'')$, damit folgt die Gleichheit.

Wir bemerken, daß die $j$-te Spalte von $M$ entweder null ist oder eine 1 in der Zeile $j$ hat. Somit sind diese Spalten linear unabhängig. □

[1] Für Zahlen folgt aus $x^2 = x$, daß $x = 0$ oder $x = 1$ ist, weil man hier kürzen darf. In Matrixprodukten kann man nicht kürzen. Es gibt viele (nichttriviale) idempotente Matrizen.

### 5.2.15 Folgerung

*Für eine Matrix $A \in K^{m \times n}$ ist folgendes äquivalent:*

1. *Die Spaltenvektoren $s_1(A), \ldots, s_n(A) \in K^{m \times 1}$ sind linear unabhängig.*
2. *$NR(A) = \{o\}$.*
3. *Die Anzahl der Pivots von $A' \sim A$ ist gleich $n$.*
4. *Es ist $m \geq n$ und es gibt eine invertierbare Matrix $C \in GL_m(K)$, so daß $CA = \begin{pmatrix} I_n \\ R \end{pmatrix}$, wobei $R \in K^{(m-n) \times n}$ eine „Rest"-Matrix ist.*

(1) $\Rightarrow$ (2) : Es ist $AX = o$ genau dann, wenn

$$s_1(A)x_1 + \ldots + s_n(A)x_n = o,$$

wegen der linearen Unabhängigkeit der Spalten existiert also nur die triviale Lösung.
(2) $\Rightarrow$ (3) folgt aus $\dim(NR(A)) = n - r$.
(4) $\Rightarrow$ (3) : Sei

$$A \sim CA = \begin{pmatrix} I_n \\ R \end{pmatrix} \sim \begin{pmatrix} I_n \\ 0 \end{pmatrix},$$

damit ist die Anzahl der Pivots gleich $n$.
(3 $\Rightarrow$ 4) : Wenn die Zahl der Pivots gleich der Spaltenzahl ist, so steht in jeder Spalte von $A'$ eine 1. Damit ist $A'$ gleich $\begin{pmatrix} I_n \\ 0 \end{pmatrix}$, denn $A'$ hat reduzierte Zeilenstufenform, also existiert ein $C$ mit $CA = A' = \begin{pmatrix} I_n \\ 0 \end{pmatrix}$. □

## 5.3 Lineare Abbildungen

In diesem Abschnitt geben wir den Matrizen eine neue Interpretation. Matrizen sind lineare Abbildungen zwischen „Koordinatenräumen" der Form $K^{m \times 1}$ bzw. $K^{1 \times n}$. Sie ergeben sich durch Spezialisierung aus dem allgemeinen Begriff der linearen Abbildung.

### 5.3.1 Definition:

Eine Abbildung $f : V \longrightarrow W$ zwischen zwei $K$-Vektorräumen heißt linear, falls
(1) $f$ ist Homomorphismus der additiven Gruppen,, d.h.

$$f(v_1 +_V v_2) = f(v_1) +_W f(v_2),$$

(vgl. 2.8), es folgt $f(-v) = -f(v)$ und $f(o_v) = o_w$),
(2) $f$ ist $K$-homogen, d.h.

$$f(\lambda \cdot_V v) = \lambda \cdot_W f(v)$$

für alle $\lambda \in K$.

### 5.3.2 Grundeigenschaften (I):

Sei $f : V \longrightarrow W$ eine lineare Abbidung von $K$-Vektorräumen.
(1) Für jede Linearkombination $\sum_{v\in M} \lambda_v v$ einer Teilmenge $M \subset V$ gilt

$$f(\sum_{v\in M} \lambda_v) = \sum_{v\in M} \lambda_v f(v),$$

(2) Ist $M$ ein Erzeugendensystem von $V$, dann ist $f$ durch seine Werte $\{f(v); v \in M\}$ vollständig bestimmt.

Beweis: (1) Sei $\sum_{v\in M} \lambda_v v$ eine Linearkombination von $M$, diese hat nach Definition einen endlichen Träger, also nur endlich viele Summanden. Induktiv führt man (1) auf 5.3.1(1)/(2) zurück.
(2) $M$ sei ein Erzeugendensystem von $V$ und die Werte $f(v)$ für $v \in M$ seien bekannt. Dann ist $f(w)$ für jeden Vektor $w \in V$ vollständig bestimmt, denn $w \in Span(M)$, also $w = \sum \lambda_v v$, also $f(w) = f(\sum \lambda_v v) = \sum \lambda_v f(v)$. □

**Bemerkung:** Lineare Abbildungen zwischen Vektorräumen sind relativ starre Objekte. Sie sind durch wenige Ausgangsdaten schon vollständig bestimmt.

Anschaulich: $f(Kv) = Kf(v)$, d.h. die durch $v$ bestimmte Gerade durch 0 wird auf eine Gerade durch 0 abgebildet.

**Beispiel:** Wir betrachten $\mathbf{R}$ els 1-dimensionalen $\mathbf{R}$-Vektorraum, als Basis können wir 1 nehmen. Welche linearen Abbildungen von $\mathbf{R}$ in sich gibt es?
Wir haben $f(x) = f(x \cdot 1) = xf(1)$, d.h. der Graph der Funktion $f$ ist die Gerade duch den Nullpunkt und den Punkt $(1, f(1))$. Die Abbildung $f$ ist durch den Wert $f(1)$ voll bestimmt.

**Bezeichnung:** $Hom_K(V, W) = \{$ lineare Abbildungen von $V$ nach $W\}$.

### 5.3.3 Matrizen als Grundbeispiel für lineare Abbildungen

Sei $K^{n\times 1}$ der $K$-Vektorraum der Spaltenvektoren.
(1) Sei $A \in K^{m\times n}$ eine $m \times n$-Matrix. Dann ist

$$X \in K^{n\times 1} \mapsto A \cdot X \in K^{m\times 1}$$

eine $K$-lineare Abbildung.
(2) Andererseits: Sei $f : K^{n\times 1} \longrightarrow K^{m\times 1}$ eine beliebige $K$-lineare Abbildung und sei $S = \{e_1, \ldots, e_n\}$ die Standardbasis von $K^{n\times 1}$. Dann sind also $f(e_1), \ldots, f(e_n) \in K^{m\times 1}$. Wir bilden die Matrix $A \in K^{m\times n}$ mit den Spaltenvektoren $s_i(A) = f(e_i)$. Dann gilt $f(X) = AX$ für alle $X \in K^{n\times 1}$.

Beweis: (1) Daß $X \mapsto AX$ eine lineare Abbildung ist, folgt aus den Eigenschaften der Matrixmultiplikation.
(2) Sei $A$ die Matrix mit den Spaltenvektoren $f(e_1), \ldots, f(e_n)$ und $X \in K^{n\times 1}$. Dann ist

$$A \cdot X = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum a_{1i}x_i \\ \vdots \\ \sum a_{ni}x_i \end{pmatrix} = \sum_i x_i \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix} = \sum x_i s_i(A) = \sum x_i f(e_i) = f(\sum x_i e_i) = f(X).$$

Wir haben die Abbildung $f$ durch die Matrix $A$ realisiert. □

Das Grundbeispiel zeigt, daß die Zuordnung

$$\text{Matrix} \leftrightarrow \text{lineare Abbildung}$$

eine Bijektion

$$K^{m\times n} \leftrightarrow Hom_K(K^{n\times 1}, K^{m\times 1})$$

darstellt.

### 5.3.4 Grundeigenschaften (II)

Sei $f : V \longrightarrow W$ eine $K$-lineare Abbildung, dann gilt
(1) $Bild(f) = \{f(v); v \in V\}$ ist ein Unterraum von $W$, insbesondere gilt $dim(Bild(f)) \leq dim(W)$.
(2) $Ker(f) = \{v \in V; f(v) = o\}$ ist ein Unterraum von $V$.
(3) Eine lineare Abbildung ist genau dann injektiv, wenn $Ker(f) = \{o\}$.
(4) Injektivität hat zur Folge, daß linear unabhängige Vektoren auf linear unabhängige abgebildet werden. Surjektivität hat zur Folge, daß ein Erzeugendensystem auf ein Erzeugendensystem abgebildet wird.

Beweis: (1) $Bild(f) \neq \emptyset$, da $o_W = f(o_V) \in Bild(f)$ ist.
Seien $w_1, w_3 \in Bild(f)$, also $w_1 = f(v_1), w_2 = f(v_2)$, dann ist $w_1+w_2 = f(v_1)+f(v_2) = f(v_1 + v_2) \in Bild(f)$.
$\lambda w_1 = \lambda f(v_1) = f(\lambda v_1) \in Bild(f)$.
Sei $(b_1, \ldots, b_n)$ eine Basis von $V$, dann erzeugen $f(v_1), \ldots, f(v_n)$ den Raum $Bild(f)$, dessen Dimension ist also höchstens gleich $n$.
(2) $Ker(f) \neq \emptyset$, weil $o_v \in Ker(f)$.
Seien $v_1, v_2 \in Ker(f)$, also $f(v_1) = f(v_2) = o$. Dann ist $f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) = o$, also $\lambda_1 v_1 + \lambda_2 v_2 \in Ker(f)$.
(3) Wenn $f$ injektiv ist, so sind die Fasern der Abbildung einzelne Punkte, also ist $Ker(f) = \{o\}$.
Wenn $f(v) = f(w)$ ist, so ist $o = f(v) - f(w) = f(v - w)$, also $v - w \in Ker(f)$, d.h. $v = w$. □

### 5.3.5 Beispiel (Interpretation für Matrizen):

Wir betrachten zur Matrix $A \in K^{m\times n}$ die Abbildung $f_A \in Hom(K^{n\times 1}, K^{m\times 1})$ mit $f_A(X) = AX$. Dann ist
(1) $Bild(f_A)$ der Spaltenraum von $A$.
(2) $Ker(f_A)$ ist der Nullraum von $A$, d.h. die Lösungsmenge des homogenen Gleichungssystems $AX = 0$.

Beweis: Sei $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, dann ist $f_A(X) = x_1 s_1(A) + \cdots + s_n(A) \in Span(s_1(A), \ldots, s_n(A))$, also $Bild(f_A) \subseteq Span(s_1(A), \ldots, s_n(A))$. Andererseits ist $Bild(f_A)$ ein Unterraum und die $s_i(A) = f(e_i)$ sind darin enthalten, also gilt Gleichheit.

(2) $f_A(X) = 0$ gdw. $AX = 0$, d.h. $X$ ist Lösungsvektor des homogenen Gleichungssystems mit der Koeffizientenmatrix $A$ . □

### 5.3.6 Anschauliche Vorstellungen von linearen Abbildungen

Seien $v, w \in V$ zwei Vektoren, dann ist

$$\lambda v + (1-\lambda)w = w + \lambda(v - w),\ \lambda \in K$$

die Verbindungsgerade zwischen $v$ und $w$. Wenn $f : V \longrightarrow W$ linear ist, dann ist

$$f(\lambda v + (1-\lambda)w) = f(w) + \lambda(f(v) - f(w))$$

die Verbindungsgerade der Bildvektoren.

Man kann zeigen:
Sei $f : V \longrightarrow W$ irgendeine Abbildung zwischen$K$-Vektorräumen mit folgenden Eigenschaften (wir setzen $dim_K(V) > 1$ und $\#K > 2$ voraus):
(1) $f$ ist injektiv,
$f(o) = o$ und $Bild(f)$ ist ein Unterraum von $W$,
(3) jede Gerade in $V$ wird unter $f$ auf eine Gerade in $W$ abgebildet.
Dann ist $f$ „fast“ eine lineare Abbildung, d.h.

$$f(v_1 + v_2) = f(v_1) + f(v_2),\ f(\lambda v) = \sigma(\lambda) f(v).$$

Für $K = \mathbf{R}$ gilt immer $\sigma(\lambda) = \lambda$, d.h. $f$ ist tatsächlich linear.

## 5.4 Lineare Abbildungen im 2-dimensionalen Fall

Wir betrachten die reelle Zahlenebene $\mathbf{R}^{2\times 1}$.

### 5.4.1 Vorbemerkung:

Einen Vektor, z.B. $\begin{pmatrix} 5 \\ 1 \end{pmatrix}$ kann man sich in zweifachen Weise vorstellen:

a) $\begin{pmatrix} 5 \\ 1 \end{pmatrix}$ ist ein Ortsvektor, d.h. er repräsentiert die Äquivalenzklasse aller Pfeile $\vec{PQ}$ mit $x_Q - x_P = 5, y_Q - y_P = 1$.
Bemerkung: Wenn man einen Koordinatenursprung O festlegt, dann gibt es eine umkehrbar eindeutige Beziehung

$$\text{Vektor } x \longrightarrow \text{Punkt } P,$$

wobei $P$ der Endpunkt des zu $x$ gehörigen Ortsvektors ist, d.h. $P$ ist der Endpunkt des Pfeils aus der Äquivalenzklasse $x$, welcher in O ansetzt.
Umgekehrt Ordnen wir einem Punkt $P$ die Klasse $x$ der zu $\vec{OP}$ äquivalenten Pfeile zu.

b) Ebenso kann man $\begin{pmatrix} 5 \\ 1 \end{pmatrix}$ auch einfach als Punkt von $\mathbf{R}^{2\times 1}$ betrachten. Eine Addition $\begin{pmatrix} 5 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ hat zwei mögliche Interpretationen:
b1) Vektor + Vektor = Vektor,
b2) Punkt + Vektor = Punkt, dies ist das Ergebnis der entsprechenden Translation.
Bei b1) ist dann die Hintereinanderausführung der entsprechenen Translationen gemeint.

### 5.4.2

Eine lineare Abbildung $f : \mathbf{R}^{2\times 1} \longrightarrow \mathbf{R}^{2\times 1}$ muß folgende Eigenschaften haben:
A) Nullpunkt $\mapsto$ Nullpunkt,
$f(P+Q) = f(P) + f(Q)$,
$f(\lambda P) = \lambda f(P)$.
Wenn $f$ injektiv ist, dann bedeutet das:
B) Geraden $\mapsto$ Geraden,
(mit A):) Geraden durch O $\mapsto$ Geraden durch O,
C) Parallelen $\mapsto$ Parallenen,
D) Proportionalitäten bleiben erhalten.
Aus B) und C) folgt: Parallelogramme gehen in Parallelogramme über.

Beweis von C): Zwei Geraden $g_1, g_2$ sind genau dann parallel, wenn es einen Vektor $x$ gibt, der die eine in die andere verschiebt. Sei $f(g_i) = \{f(P); P \in g_i\}$, dann sind $f(g_1), f(g_2)$ wieder Geraden und $f(g_1) + f(x) = f(g_2)$.

### 5.4.3

$f : \mathbf{R}^{2\times 1} \longrightarrow \mathbf{R}^{2\times 1}$ hat immer die Form $f = f_A$ mit einer $2 \times 2$-Matrix $A$. Die Grunddaten sind $s_1(A) = f(e_1), s_2(A) = f(e_2)$.
Um das Bild von $P = \begin{pmatrix} x_P \\ y_P \end{pmatrix}$ zu erhalten, muß man die Koordinaten von $P' = f(P)$ bezüglich der Basis $e_1' = f(e_1), e_2' = f(e_2)$ finden:
Es gilt $f(P) = x_P f(e_1) + y_P f(e_2)$, und $x_P f(e_1)$ bzw. $y_P f(e_2)$ markiert denjenigen Punkt von $\mathbf{R}f(e_1)$ bzw. $\mathbf{R}f(e_2)$, für den

$$\frac{|O[x_P f(e_1)]|}{|Of(e_1)|} = \frac{|O[x_P e_1]|}{|Oe_1|}, \quad \frac{|O[y_P f(e_2)]|}{|Of(e_2)|} = \frac{|O[y_P e_2]|}{|Oe_2|}$$

ist. Das erhält man aus dem Strahlensatz.

## 5.4.4 Beispiele

a) Spiegelung an der $y$-Achse:

$$f(\begin{pmatrix} 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \ f(\begin{pmatrix} 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \ A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

b) Drehung um den Winkel $\theta$ entgegen dem Uhrzeigersinn:

$$f(\begin{pmatrix}1\\0\end{pmatrix}) = \begin{pmatrix}cos(\theta)\\sin(\theta)\end{pmatrix},\ f(\begin{pmatrix}0\\1\end{pmatrix}) = \begin{pmatrix}-sin(\theta)\\cos(\theta)\end{pmatrix},\ A = \begin{pmatrix}cos(\theta) & -sin(\theta)\\sin(\theta) & cos(\theta)\end{pmatrix}.$$

c) Spiegelung an der Achse mit dem Winkel $\theta/2$):
Der Vektor $\begin{pmatrix}1\\0\end{pmatrix}$ wird um $\theta$ gedreht, das Ergebnis ist $\begin{pmatrix}cos(\theta)\\sin(\theta)\end{pmatrix}$, der Vektor $\begin{pmatrix}0\\1\end{pmatrix}$ wird um $2(90^o - \theta/2)$ im Uhrzeigersinn gedreht, das Ergebnis ist dasselbe, als ob man $e_1$ um $90^o - \theta$ dreht, also $\begin{pmatrix}sin(\theta)\\-cos(\theta)\end{pmatrix}$, also $A = \begin{pmatrix}cos(\theta) & sin(\theta)\\sin(\theta) & -cos(\theta)\end{pmatrix}$.
d) Maßstabsverzerrungen: $e_1 \mapsto k_1 e_1, e_2 \mapsto k_2 e_2$, die alte $x$-Achse und die neue sind deckungsgleich, dasselbe gilt für die $y$-Achse, $A = \begin{pmatrix}k_1 & 0\\0 & k_2\end{pmatrix}$. Die Werte $k_1, k_2$ heißen die Verzerrungsfaktoren, sie können auch negativ sein.
Im Fall $k_1 = k_2 = k$ ist die Verzerrung in $x$- und $y$-Richtung dieselbe; jede Gerade durch den Nullpunkt wird in sich überführt und um den Faktor $k$ gestreckt, bei $k < 0$ wird die Orientierung geändert.

e) Scherung längs der $x$-Achse:
Der Einheitsvektor $e_1$ geht in sich über, der Vektor $e_2$ wird längs der $x$-Achse in $\begin{pmatrix}k\\1\end{pmatrix}$ verschoben: $A = \begin{pmatrix}1 & 0\\k & 1\end{pmatrix}$.
Die Scherung längs der $y$-Achse führt zur Matrix $A = \begin{pmatrix}1 & k\\0 & 1\end{pmatrix}$.

f) Spezialfall: Wenn $det(A) = 0$ ist, so ist der Nullraum von $A$ von $\{o\}$ verschieden, die Spaltenvektoren von $A$ sind linear abhängig und $Bild(f_A) = SR(A)$ ist nur eine Gerade (oder ein Punkt).

### 5.4.5 Der höherdimensionale Fall:

Die Fundamentalmasche $F$ ist hier ein $n$-dimensionaler Würfel und $F' = f(F)$ ist ein Parallelepiped. Wenn $f$ injektiv ist, so hat $F'$ dieselbe Dimension, sonst ist sie kleiner als $n$.
Der Raum $\mathbf{R}^{n\times 1}$ ist durch die Translationen von $F$ gepflastert, der Bildraum von $f$ ist durch die Translationen von $F'$ gepflastert.

### 5.4.6 Hintereinanderausführung linearer Abbildungen

**Satz:** Die Hintereinanderausführung linearer Abbildungen ist wieder eine lineare Abbildung. Bei der Entsprechung zwischen linearen Abbildungen und Matrizen korrespondiert die Hintereinanderausführung mit dem Matrixprodukt. Genauer:
(1) Seien $U, V, WK$-Vektorräume und $f : U \longrightarrow V, g : V \longrightarrow W$ linear, dann ist $g \circ f \in Hom(U, W)$ linear.
(2) Spezialfall: $U = K^{n\times 1}, V = K^{m\times 1}, W = K^{l\times 1}$. Zu $f$ gehöre die Matrix $A_f \in K^{m\times n}$, zu $g$ gehöre $A_g \in K^{l\times n}$. Dann gilt $A_{g\circ f} = A_g A_f$.

Beweis: (1)

$$(gf)(v_1 + v_2) = g(f(v_1 + v_2)) = g(f(v_1) + f(v_2)) = g(f(v_1)) + g(f(v_2)),$$

die Homogenität zeigt man analog.
(2) Sei $X \in K^{n\times 1}$, dann ist

$$A_{gf}X = g(f(X) = g(A_f X) = A_g A_f X$$

für alle $X$, also $A_{gf} = A_g A_f$. □

Bemerkung: Der Satz kann von zwei auf mehrere lineare Abbildungen bzw. Matrizen verallgemeinert werden.

### 5.4.7 Beispiele:

1. Zuerst eine Rotation um den Winkel $\theta$, dann eine Scherung in $x$-Richtung mit dem Faktor $k$:

$$A_{gf} = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix} = \begin{pmatrix} cos(\theta) + k \cdot sin(\theta) & -sin(\theta) + kcos(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix}.$$

2. a) Zuerst Scherung um den Faktor 2 in $x$-Richtung, dann Reflektion an der Achse $y = x$, Anstiegswinkel $\pi/4$.

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$$

b) Dieselben Abbildungen in umgekehrter Reihenfolge:

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$$

Die Anwendung der Abbildungen ist nicht vertauschbar!

### 5.4.8 Realisierung invertierbarer Abbildungen durch Abbildungsfolgen

Die Abbildung $f : \mathbf{R}^{2\times 1} \longrightarrow \mathbf{R}^{2\times 1}$ entstehe durch die Multiplikation mit einer Elementarmatrix.
Zeilenoperation 1: $\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$ : Maßstabsverzerrung
Zeilenoperation 2: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$: Spiegelung an der Achse $y = x$,
Zeilenoperation 3: $\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$ : Scherungen.

Die Matrix $A$ ist genau dann invertierbar, wenn sie Zeilenäquivalent zu $I_2$ ist, also wenn sie sich als Produkt von Elementarmatrizen darstellen läßt. Dann läßt sich $f_A$ als Hintereinanderausführung von Maßstabsverzerrungen, Spiegelungen und Scherungen realisieren.

### 5.4.9 Die inverse Abbildung

Wegen $AA^{-1} = A^{-1}A = I_2$ sind die Abbildungen $f_A$ und $f_{A^{-1}}$ zueinander invers, denn z.B.

$$f_A \circ f_{A^{-1}} = f_{AA^{-1}} = f_I = id$$

.

Wir wollen uns noch die inverse Abbildung veranschaulichen:
Die Abbildung $f$ ist durch das Bild $F'$ der Fundamentalmasche $F$ festgelegt, diese wird durch die Spaltenvektoren von $A$ aufgespannt. Um die Umkehrabbildung von $f$ zu beschreiben, müssen wir die Fundamentalmasche $F''$ finden, die auf $F$ abgebildet wird. Die Vektoren, die $F''$ aufspannen, sind die Spalten der Matrix $A^{-1}$.

## 5.5 Mehr über lineare Abbildungen und Matrizen

### 5.5.1 Hauptsatz:

Sei $f : V \longrightarrow W$ eine lineare Abbildung endlichdimensionaler $K$-Vektorräumen. Dann gilt die Dimensionsformel

$$dim(V) = dim(Ker(f)) + dim(Bild(f)).$$

Beweis: Sei $B' = (b'_1, \ldots, b'_r)$ eine Basis von $Ker(f)$ und $C = (c_1, \ldots, c_s)$ eine Basis von $Bild(f)$; wir wählen $c'_i \in V$ mit $f(c'_i) = c_i$.
1. Behauptung: Die Vektoren $b'_1, \ldots, b'_r, c'_1, \ldots, c'_s$ sind linear unabhängig.
Sei also

$$\sum \lambda_i b'_i + \sum \mu_i c'_i = o.$$

Wir wenden die Abbildung $f$ an und erhalten

$$\sum \mu_i f(c'_i) = \sum \mu_i c_i = o,$$

also müssen alle $\mu_i$ null sein. Dann ist aber $\sum \lambda_i b'_i = o$, woraus $\lambda_i = 0$ folgt.
2. Behauptung: Die Vektoren $b'_i, c'_j$ erzeugen $V$.
Sei $v \in V$ beliebig, dann ist $f(v) \in Bild(f)$, also

$$f(v) = \sum \mu_i c_i = \sum \mu_i f(c'_i) = f(\sum \mu_i c'_i).$$

Daraus folgt $f(v - \sum \mu_i c'_i) = o$, also $v - \sum \mu_i c'_i \in Ker(f)$, d.h. $v - \sum \mu_i c'_i = \sum \lambda_j b'_j$.
Damit ist

$$v = \sum \mu_i c'_i + \sum \lambda_j b'_j \in Span(b'_i, c'_j).$$

Aus beiden Behauptungen folgt, daß die $b'_i, c'_j$ eine Basis von $V$ bilden, also $dimV = r + s$. □

### 5.5.2 Anwendung auf Matrizen

Sei $A \in K^{m\times n}$, dann ist $ZR(A) \subset K^{1\times n}, SR(A) \subset K^{m\times 1}$. Wir wissen:
Zeilenrang$(A) = dimZR(A)$,
Spaltenrang$(A) = dimSR(A)$,
Defekt$(A) = dimNR(A)$.

**Satz:** Für $A \in K^{m\times n}$ gilt Defekt$(A)$ + Spaltenrang$(A) = n$.

Beweis: Wir betrachten die zu $A$ gehörige lineare Abbildung

$$f_A : K^{n\times 1} \longrightarrow K^{m\times 1}.$$

Dann ist

$$dim(Ker(f_A)) + dim(Bild(f_A)) = n.$$

□

Bemerkung: Wir werden bald sehen, daß Zeilen- und Spaltenrang einer Matrix übereinstimmen, man spricht daher vom „Rang" der Matrix.

### 5.5.3 Isomorphismen, isomorphe Vektorräume

Eine lineare Abbildung $f : V \longrightarrow W$ heißt Isomorphismus, wenn die folgenden äquivalenten Eigenschaften erfüllt sind:
(1) Es existiert eine lineare Abbildung $g : W \longrightarrow V$, so daß

$$g \circ f = id_v, \ f \circ g = id_w.$$

(2) Die Abbildung $f$ ist bijektiv.

Beweis: (1) $\Rightarrow$ (2): Sei $w \in W$, dann ist $w = f(g(w))$, also ist $f$ surjektiv. Wenn $f(v_1) = f(v_2)$ gilt, so ist $v_1 = gf(v_1) = gf(v_2) = v_2$, also ist $f$ injektiv.
(2) $\Rightarrow$ (1): Wenn $f$ bijektiv ist, so haben wir zu jedem $w \in W$ einen eindeutig bestimmten Vektor $f^{-1}(w)$. Wir zeigen, daß $f^{-1}$ eine lineare Abildung ist:
Seien $w_1, w_2 \in W, f(v_1) = w_1, f(v_2) = w_2$, dann ist

$$f^{-1}(\lambda_1 w_1 + \lambda_2 w_2) = f^{-1}(\lambda_1 f(v_1) + \lambda_2 f(v_2)) = f^{-1}(f(\lambda_1 v_1 + \lambda_2 v_2)) = \lambda_1 v_1 + \lambda_2 v_2.$$

□

Wir nennen zwei $K$-Vektorräume $V, W$ isomorph, wenn ein Isomorphismus $f : V \longrightarrow W$ existiert.

Aus 5.5.1 folgt, daß isomorphe Vektorräume dieselbe Dimension haben. Wichtig ist, daß auch die Umkehrung gilt.

### 5.5.4 Beispiel:

Wir ordnen der Matrix $A \in K^{m\times n}$ die lineare Abbildung $f_A \in Hom(K^{n\times 1}, K^{m\times 1})$ zu. Dann ist $f_A$ genau dann ein Isomorphismus, wenn $m = n$ und $det(A) \neq 0$ ist. Die Umkehrabbildung ist durch die inverse Matrix gegeben: $(f_A)^{-1} = f_{A^{-1}}$.

### 5.5.5 Bemerkung: transponierte Matrix, Spaltenäquivalenz

1. Seien $A, B$ multiplizierbare Matrizen, dann gilt (vgl. 3.3.5): $(AB)^t = B^t A^t$.
2. Seien $s$ und $z$ elementare Spalten- bzw. Zeilenoprationen, die einander entsprechen, dann ist $s(A)^t = z(A^t)$.
3. Es ist $s(A) = A \cdot s(I_n)$.
4. Die Matrizen $A, B$ sind spaltenäquivalent, wenn $B$ aus $A$ durch elementare Spaltenoperationen hervorgeht, d.h wenn eine invertierbare Matrix $D$ mit $B = AD$ existiert.
5. Die Matrizen $A, B$ sind genau dann spaltenäquivalent, wenn $A^t, B^t$ zeilenäquivalent sind.

### 5.5.6 Satz:

Seien $A, B \in K^{m \times n}$.
1. Seien $B \sim A$ zeilenäquivalent, genauer $B = CA$ mit einer invertierbaren Matrix $C$. Dann gilt:
a) $NR(B) = NR(A)$,
b) $ZR(A) = ZR(B)$,
c) $SR(A) \cong SR(A), X \mapsto CX$.
2. Seien $B \sim A$ spaltenäquivalent, genauer $B = AD$ mit einer invertierbaren Matrix $D$. Dann gilt:
a) $SR(B) = SR(A)$,
b) $NR(B) \cong NR(A), X \mapsto DX$
c) $ZR(A) \cong ZR(B), Y \mapsto YD$.

Ausgewählte Beweise:
1. c) Wegen $B = CA$ gilt für die $i$-ten Spalten von $B$ und $A$

$$s_i(B) = Cs_i(A).$$

$SR(A)$ ist der Spann der $s_i(A)$, $SR(B)$ ist der Spann der $s_i(B)$. Wenn also $X \in SR(A)$ ist, so ist $CX \in SR(B)$. Es gibt eine Umkehrabbildung, weil $A = C^{-1}B$ ist.
2. b) Sei $X \in NR(B)$, also $BX = 0$. Wir substituieren $B = AD$, also gilt $ADX = 0$, also ist $DX \in NR(A)$. Wegen $A = BD^{-1}$ gibt es wieder eine Umkehrabbildung. □

### 5.5.7 Folgerung:

Bei elementaren Zeilen- und Spaltenoperationen bleiben der Defekt, Spaltenrang und Zeilenrang einer Matrix erhalten.

Beweis: Wir gehen von einer Matrix $A$ zu äquivalenten Matrizen $CA$ oder $AD$ über, dann sind die jeweiligen Zeilenräume, Spaltenräume und Nullräume entweder gleich oder isomorph. In jedem Fall bleibt die Dimension erhalten. □

### 5.5.8

Wir nennen $A, B \in K^{m\times n}$ schwach äquivalent, falls die folgenden gleichwertigen Eigenschaften erfüllt sind:
1. Es existieren invertierbare Matrizen $C \in GL_m(K),\ D \in GL_n(K)$, so daß $B = CAD$ ist.
2. $B$ entsteht aus $A$ durch eine Folge elementarer Zeilen- und Spaltenoperationen.
Dann gilt: Jede Matrix $A \in K^{m\times n}$ ist schwach äquivalent zu einer Matrix vom Typ

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

mit $p \leq \min(m, n)$. Dabei ist $r = \text{Zeilenrang}(A) = \text{Spaltenrang}(A)$.

Beweis: $A$ ist zeilenäquivalent zu

$$A' = \begin{pmatrix} 1 & * & \dots & * & 0 & \dots & 0 & \\ & & & & 1 & \dots & 0 & \\ & & & & & & 1 & \dots \end{pmatrix}$$

in reduzierter Zeilenstufenform. Bei schwacher Äquivalenz sind auch noch elementare Spaltenoperationen erlaubt, also

$$A \sim \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

und diese Matrix hat offensichtlich den Zeilen- und Spaltenrang $r$ und wegen 5.5.7 gilt dies auch für $A$. □

Im folgenden sprechen wir nur noch vom *Rang* einer Matrix.

### 5.5.9 Ein Rechenverfahren

Gegeben sei eine Menge $M = \{v_1, \dots v_k\}$ von Vektoren aus $K^{n\times 1}$, gesucht werden
a) ein Teilsystem $B = \{v_{j_1}, \dots v_{j_r}\}$ von $M$, welche eine Basis von $Span(M)$ ist.
b) eine explizite Darstellung der Vektoren aus $M$ als Linearkombinationen von $B$.

Zu a): Wir bilden die $n\times k$-Matrix $A$ mit den Spaltenvektoren $v_1, \dots, v_k$ und bestimmen eine zu $A$ zeilenäquivalente Matrix $B$ in Zeilenstufenform, die nicht reduziert sein muß. Wenn die Pivots von $B$ in den Spalten $j_1, \dots, j_r$ stehen, dann bilden die Vektoren $v_{j_1}, \dots v_{j_r}$ eine Basis von $Span(M)$ und es ist $r = Rang(A)$.

Zu b): Wir bilden nun die reduzierte Zeilenstufenform $A'$, die $j$-te Spalte von $A'$ sei

$$s_j(A') = \begin{pmatrix} \lambda_{1j} \\ \vdots \\ \lambda_{rj} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Dann ist

$$v_j = \sum_{i=1}^{r} \lambda_{ij} v_{j_i}.$$

Der Beweis beruht auf 5.5.6 1.c.

### 5.5.10 Satz von der Eindeutigkeit der reduzierten Zeilenstufenform

Sei $K^{m\times n}/\sim$ die Quotientenmenge von $K^{m\times n}$ bezüglich der Zeilenäquivalenz. Dann bilden die Matrizen mit reduzierter Zeilenstufenform einen Schnitt, d.h. in jeder Äquivalenzklasse gibt es genau eine solche Matrix. Das sehen wir wie folgt:
Sei $A \in K^{m\times n}$ gegeben. Dann ist $A' \sim A$, $A'$ in reduzierter Zeilenstufenform, folgendermaßen bestimmt:
Wir betrachten die Matrizen $A_j$, die aus $A$ durch Abschneiden nach der $j$-ten Spalte entstehen, und bilden die Flagge

$$V_0 = \{o\} \subseteq V_1 = SR(A_1) \subseteq V_2 = SR(A_2) \subseteq \ldots \subseteq V_n = SR(A).$$

Wir nennen den Index $j$ eine Sprungstelle, wenn $V_{j-1} \subset V_j$ echt enthalten ist. Seien $j_1 < j_2 < \ldots < j_r$ die Sprungstellen. Dann gilt:
1. die $j_1, \ldots, j_r$ sind genau die Nummern der Pivotsplaten von $A'$,
2. $B = \{s_{j_1}(A), \ldots, s_{j_r}(A)\}$ ist eine Basis von $SR(A)$,
3. Der $j$-teSpaltenvektor von $A'$ ist gleich dem Koordinatenvektor von $s_j(A)$ bezüglich der Basis $B$, ergänzt um $m-r$ Nullen am Ende der Spalte:

$$s_j(A) = \sum_{i=1}^{r} \lambda_{ij} s_{j_i}(A).$$

Beweis: Sei $A' = CA$, dann ist die Abbildung $X \in SR(A) \mapsto CX \in SR(A')$ ein Isomorphismus mit der Eigenschaft

$$s_i(A) \mapsto s_i(A') = Cs_i(A).$$

Deshalb wird der Unterraum $SR(A_j)$ nach $SR(A'_j)$ transportiert. Damit haben die Beiden Flaggen für $A$ und $A'$ dieselben Sprungstellen. Der Blick auf $A'$ zeigt, daß die Sprungstellen genau den Pivotspalten entsprechen.
Sei $s_j(A')$ eine Spalte.
Fall a) Dies ist eine Pivotspalte, etwa die $i$-te, dann ist $s_j(A) = e_i$.
Fall b) keine Pivotspalte, also

$$s_j(A') = \begin{pmatrix} \lambda_{1j} \\ \vdots \\ \lambda_{rj} \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

d.h.

$$s_j(A') = \sum e_i\lambda_{ij} = \sum s_{j_i}(A')\lambda_{ij}.$$

Wegen der obigen Isomorphie gilt die entsprechende Gleichung für die Spalten von $A$, d.h. die Koeffizienten $\lambda_{ij}$ sind durch $A$ eindeutig bestimmt.

### 5.5.11 Direkte Summen und Projektoren

**Definition:** Sei $V$ ein Vektorraum und $V_1, V_2$ Unterräume. Man nennt $V$ die direkte Summe von $V_1$ und $V_2$, fall sich jedes $v \in V$ auf genau eine Art und Weise in der Form

$$v = v_1 + v_2,\ v_i \in V_i$$

darstellen läßt, d.h. die Beiden Räume sind linear unabhängig.
Man schreibt dann $V = V_1 \oplus V_2$.
**Bemerkung:** $V$ ist genau dann direkte Summe von $V_1$ und $V_2$, wenn $V = V_1 + V_2$ und $V_1 \cap V_2 = \{o\}$ ist.

Beweis: Wenn $V$ direkte Summe ist, dann gilt $V = V_1 + V_2$. Sei $v \in V_1 \cap V_2$ und wir nehmen $v \neq o$ an. Dann hätte $v$ zwei verschiedene Darstellungen:

$$v = v + o = o + v.$$

Umgekehrt, wenn $V = V_1 + V_2$ und $V_1 \cap V_2 = \{o\}$ ist und zwei Darstellungen

$$v = v_1 + v_2 = w_1 + w_2$$

gegeben sind, dann ist

$$v_1 - w_1 = w_2 - v_2 \in V_1 \cap V_2,$$

also $v_1 = w_1, v_2 = w_2$. $\square$

### 5.5.12 Definition und Satz:

Eine lineare Abbildung $p : V \longrightarrow V$ heißt Projektor, falls folgende äquivalente Eigenschaften erfüllt sind:
1. $p^2 = p,$ d.h. $p(p(x)) = p(x)$ für alle $x \in V$,
2. $Bild(I - p) \subseteq Ker(p)$,
3. $Bild(I - p) = Ker(p)$,
4. $V = Ker(p) \oplus Bild(p)$,
und die Einschränkung von $p$ auf den Unterraum $Bild(p)$ ist die Identität.

Beweis: 1. $\Rightarrow$ 2. Sei $x \in Bild(I - p),\ x = (I - p)(y)$. Dann ist

$$p(x) = p((I - p)(y)) = (p - p^2)(y) = p(y) - p^2(y) = o,$$

also ist $x \in Ker(p)$.
2. $\Rightarrow$ 3. Es gilt immer $Ker(p) \subseteq Bild(I - p)$, denn $x = p(x) + (I - p)(x)$, wenn also $p(x) = o$ ist, so ist $x = (I - p)(x) \in Bild(I - p)$.

3. $\Rightarrow$ 2. ist klar.
2. $\Rightarrow$ 1. Es ist $p^2(x) - p(x) = p((I-p)(x))$, wenn also $Bild(I-p) \subseteq Ker(p)$ ist, so ist dies $o$.
1. $\Rightarrow$ 4. Es ist

$$x = p(x) + (I-p)(x) \in Bild(p) + Bild(I-p) = Bild(p) + Ker(p).$$

Sei nun $x \in Bild(p) \cap Ker(p),\ x = p(y)$, dann ist

$$p(x) = p^2(y) = p(y) = x,$$

andererseits ist $p(x) = o$, also $x = o$.
Also ist die Summe direkt; wegen $p^2 = p$ muß $p$ auf $Bild(p)$ die Identität sein.
4. $\Rightarrow$ 1. Sei $v = x + y$ mit $x \in Ker(p),\ y \in Bild(p)$. Dann ist

$$p(v) = p(x) + p(y) = p(y) = y,$$

weil $p$ auf $Bild(p)$ die Identität ist. Dann ist aber auch $p^2(v) = p(y) = y = p(v)$, also $p^2 = p$. □

## 5.5.13 Umkehrung

Sei $V = U \oplus W$ eine direkte Summenzerlegung, dann gibt es genau einen Projektor $p$ mit

$$Ker(p) = U,\ Bild(p) = W.$$

Die Fasern der Abbildung sind die zu $U$ parallelen affinen Unterräume $v + U$ und es gilt $p(v+U) = (v+U) \cap W$ ist ein einzelner Vektor.

Beweis: Jedes $v$ ist eindeutig als $v = x + y,\ x \in U,\ y \in W$ darstellbar. Wir definieren $p(v) = y$. Das ist ein Projektor:

$$p(p(v)) = p(y) = y = p(v),$$

da $y$ bereits in $W$ liegt, und es ist $Bild(p) = W,\ Ker(p) = U$.
Umgekehrt: Wenn $p$ ein Projektor mit $Ker(p) = U,\ Bild(p) = W$ ist, dann muß $p$ die oben beschriebene Abbildung sein.
Nun betrachten wir noch $p(v+U)$: Sei $v = x + y$, dann ist $v + U = y + x + U = y + U$, da $x \in U$ ist. Also $p(v+U) = p(y+U) = y$, da $U = Ker(p)$.
Andererseits ist $(v+U) \cap W = (y+U) \cap W = \{y\}$. □
**Bezeichnung:** Wir nennen $p$ die Parallelprojektion von $V$ auf $W$ längs $U$.

## 5.5.14 Bemerkung:

Projektoren treten immer paarweise auf. Wenn $p : V \longrightarrow V$ ein Projektor ist, so ist $I - p$ ebenfalls ein Projektor:

$$(I-p)^2 = I^2 - 2p + p^2 = I - p.$$

Bei beiden Projektoren vertauschen sich Kern und Bild.

### 5.5.15 Wiederholung(vgl 5.2.13):

Sei $A \in K^{m\times n}$ und $A'$ die zeilenäquivalente Matrix mit reduzierter Zeilenstaffelung. Sei weiter $A''$ die quadratische Matrix, die entsteht, indem man die Zeilen von $A'$ so auseinanderzieht, daß die Pivots auf der Hauptdiagonalen stehen. Dann gilt:
$A''$ ist ein Projektor mit $Ker(A'') = NR(A)$. Demzufolge hat der komplementäre Projektor $I - A''$ die Eigenschaft

$$SR(I - A'') = Bild(I - A'') = Ker(A'') = NR(A),$$

also ist $I - A''$ der Projektor auf den Nullraum von $A$.

### 5.5.16 Lineare Gleichungssysteme und Projektoren

A) Sei $A \in K^{m\times n}$. Wir betrachten das homogene Gleichungssystem $AX = 0$ mit dem Lösungsraum $NR(A) = NR(A')$. Sei weiter $\Delta_A = \Delta_{A'}$ die zugehörige Pivotmatrix und $\Delta_A^t$ deren transponierte, es ist $K^{n\times 1} = NR(A) \oplus SR(\Delta_A^t)$. Dann gilt:
(1) $A'' = \Delta_A^t A'$ ist der eindeutig bestimmte Projektor mit

$$NR(A'') = NR(A),\ SR(A'') = SR(\Delta_A^t),$$

(2) $I - A''$ ist der eindeutig bestimmte Projektor mit

$$NR(I - A'') = SR(\Delta_A^t),\ SR(I - A'') = NR(A).$$

Also hat $AX = 0$ die allgemeine Lösung $X = (I - A'')Y$ mit beliebigem $Y$.
(3) Die von Null verschiedenen Spalten der Matrix $I - A''$ bilden eine Basis von $NR(A)$.

Beweis: Man berechnet direkt $A'\Delta_A^t = I_m$, dann ist $(A'')^2 = A''$ ein Projektor mit der Eigenschaft (1): Es ist $NR(A'') = NR(A)$ und $SR(A'') \subseteq SR(\Delta_A^t)$. Wir betrachten die Dimensionen:
$dimSR(\Delta_A^t) =$ Zahl der Pivots $= r$,
$dimNR(A) = n - r$,
$dimSR(A'') = n - dimNR(A'') = n - (n - r) = r$.
Also haben oben beide Seiten dieselbe Dimension, sind also gleich.
Aus (1) folgt (2), weil $I - A''$ der komplementäre Projektor ist. Demzufolge ist $SR(I - A'') = NR(A)$, also $Rang(I - A'') = n - r$. Jedoch hat $I - A''$ genau $n - r$ von Null verschiedene Spalten, nämlich diejenigen, die komplementär zu den Pivotspalten sind.
□

B) Der inhomogene Fall $AX$ =B:
Wir betrachten die erweiterte Koeffizientenmatrix $(A, B)$, sei $(A', B')$ ihre reduzierte Zeilenstufenform. Eine Lösung existiert genau dann, wenn

$$B \in SR(A) \Leftrightarrow Rang(A, B) = Rang(A) \Leftrightarrow Rang(A', B') = Rang(A')$$

genau dann, wenn die letzte Spalte $B'$ keine Pivotspalte ist. Dann haben wir wieder die Pivotmatrix $\Delta_A$ und

$$\Delta_A^t(A', B') = (\Delta_A^t A', \Delta_A^t B') =: (A'', B'').$$

Es gilt: $B''$ ist eine spezielle Lösung von $AX = B$ und $X = B'' + (I - A'')Y$ ist die allgemeine Lösung.

Beweis: Es ist

$$AX = B \Leftrightarrow A'X = B' \Leftrightarrow A''X = B'',$$

(die beiden letzten sind dieselben Gleichungssysteme, es wurden nur Nullzeilen hinzugefügt), es ist also zu zeigen, daß $A''B'' = B''$ gilt. Nun hat aber $A''X = B''$ mindestens eine Lösung $X_0$, d.h. $B''$ liegt im Bild des Projektors $A''$. Dann ist

$$A''B'' = A''(A''X_0) = (A'')^2X_0 = A''X_0 = B''.$$

□

### 5.5.17 Lineare Gleichungssysteme und affine Räume

Sei $V$ ein $K$-Vektorraum. Ein affiner Unterraum ist eine Teilmenge der Form $v + U$, wobei $v$ ein „Stützvektor" und $U$ ein echter Unterraum von $V$ ist. Es gilt dann $v+U = U$ genau dann, wenn $v \in U$ und $v_1 + U = v_2 + U$ genau dann, wenn $v_1 - v_2 \in U$ (vgl. 2.5.4).

Im Spezialfall $V = K^{n\times 1}$ sind die affinen Unterräume genau die Lösungsmengen linearer Gleichungssysteme in $n$ Variablen, denn diese haben die Form $B'' + SR(I - A'') = B'' + NR(A)$.

Jeder affine Unterraum $v + U \subset K^{n\times 1}$ läßt sich als Lösungsmenge eines geeigneten Gleichungssystems darstellen: Sei $b_1, \ldots, b_s$ eine Basis von $U$ und $e_1, \ldots, e_n$ die Standardbasis. Dann kann $b_1, \ldots, b_s$ mit Hilfe geeigneter Elemente $e_{j_1}, \ldots, e_{j_r}, (r + s = n)$ zu einer Basis von $K^{n\times 1}$ ergänzt werden. Sei $A''$ der eindeutig bestimmte Projektor mit $NR(A'') = U, SR(A'') = Span(e_{j_1}, \ldots, e_{j_r})$. DAnn hat $A''X = A''v$ die spezielle Lösung $X = v$ und die allgemeine Lösung $v + U$. □

## 5.6 Koordinaten

Mit Vektoren kann man nur rechnen, wenn deren Koordinaten gegeben sind, diese beziehen sich stets auf eine fixierte Basis des Vektorraums. Wenn man eine andere Basis wählt, hat der gleiche Vektor andere Koordinaten. Für die Praxis ist es wichtig zu wissen, wie man zwischen verschiedenen Koordinatensystemen wechseln kann.

Als analytische Geometrie bezeichnet man die Methode, durch Einführung von Koordinaten geometrische Probleme rechnerisch anzugehen. Als Begründer gilt Rene Descartes = Renatus Cartesius (1596 - 1650).

### 5.6.1

Sei $V/K$ ein $n$-dimensionaler Vektorraum und $B = (b_1, \ldots, b_n)$ eine Basis. Dann bekommt man eine Koordinatenabbildung

$$x \in V \mapsto [x]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^{n \times 1}.$$

Der Zahlenvektor $[x]_B$ ist dabei durch die Eigenschaft

$$x = (b_1, \ldots, b_n) = b_1 x_1 + \ldots + b_n x_n$$

bestimmt. Wir nennen $[x]_B$ den Koordinatenvektor von $x$ zur Basis $B$.
Nach dem Eindeutigkeitslemma 5.1.12 läßt sich $x$ eindeutig als Linearkombination von $B$ schreiben.

### 5.6.2 Bemerkung:

Die Koordinatenabbildung ist eine lineare Abbildung, sogar ein Isomorphismus von Vektorräumen.

Beweis: Der Vektor $x + y$ hat die Koordinaten $[x + y]_B = [x]_B + [y]_B$ und für $\lambda \in K$ gilt $[\lambda x]_B = \lambda [x]_B$, also ist die Abbildung linear und besitzt offensichtlich eine Umkehrabbildung. □

Jeder $n$-dimensionale Vektorraum ist isomorph zu seinem Koordinatenraum $K^{n \times 1}$, jedoch ergeben sich für verschiedene Basen verschiedene Isomorphismen.

Im Spezialfall $V = K^{n \times 1}$ mit der Standardbasis $S = (e_1, \ldots, e_n)$ ist jeder Vektor sein eigener Koordinatenvektor.

### 5.6.3 Verfahren zur Koordinatenbestimmung

Sei $B = (b_1, \ldots, b_n)$ eine Basis des Koordinatenraums und $x \in K^{n \times 1}$. wir bilden die Matrix $(b_1, \ldots, b_n, x)$ und überführen sie in Zeilenstufenform, wir erhalten

$$\begin{pmatrix} 1 & & & \\ & \ddots & & y \\ & & 1 & \end{pmatrix},$$

dann ist $y = [x]_B$.

Wir wollen nun einer linearen Abbildung Koordinaten (eine Matrix) zuordnen. Im Spezialfall $f : K^{n \times 1}, \longrightarrow K^{m \times 1}$ entspricht der linearen Abbildung eine Matrix $A \in K^{m \times n}$, so daß $f(X) = AX$ ist. Jede lineare Abbildung hat also automatisch eine Koordinatenmatrix. Durch Zurckführung auf diesen Spezialfall werden wir jeder linearen Abbildung $f : V \longrightarrow W$ eine Koordinatenmatrix zuordnen.

Sei $B = (b_1, \ldots, b_n)$ eine Basis von $V$ und $C = (c_1, \ldots, c_m)$ eine Basis von $W$. Dann haben wir die Koordinatenabbildungen $b = [\ ]_B,\ c = [\ ]_C$ aus 5.6.1 und wir erhalten das folgende Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ b \downarrow & & \downarrow c \\ K^{n\times 1} & \xrightarrow{A} & K^{m\times 1} \end{array}$$

Dabei ergibt sich die Abbildung $A$ als Nacheinanderausfürung der anderen, wir fassen sie als Matrix auf und nennen diese die Koordinatenmatrix von $f$ bezüglich der Basen $B, C$; wir bezeichnen sie mit

$$A = [f]_{C,B}.$$

Beachten Sie, das die Indizes $B, C$ gegenläufig geschrieben sind.

### 5.6.4 Satz

Sei $f : V \longrightarrow W$ eine lineare Abbildung und seien $B, C$ Basen von $V$ bzw. $W$. Dann ist die Koordinatenmatrix $[f]_{C,B}$ eindeutig durch die Eigenschaft

$$[f]_{C,B}[x]_B = [f(x)]_C$$

bestimmt.

Beweis: Wir haben die Matrix $[f]_{C,B} = A$ gerade so definiert, daß dieses Eigenschaft gilt. Wir müssen noch zeigen, daß sie dadurch eindeutig bestimmt ist.
Wir nehmen $x = b_i$, dessen Koordinatenvektor ist $e_i$. Wir setzen ein: $[f]_{C,B}e_i = [f(b_i)]_C$.

Das bedeutet: Der $i$-te Spaltenvektor der Matrix $[f]_{C,B}$ ist genau der Koordinatenvektor von $f(b_i)$ bezüglich der Basis $C$.

### 5.6.5 Beipiele

Im trivialen Fall $f = f_A : K^{n\times 1} \longrightarrow K^{m\times 1}$ bei der Standardbasis $S$ ist $[f_A]_{S,S} = A$.

Sei $A = \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix} : K^{2\times 1} \longrightarrow K^{2\times 1}, B = \{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\}, C = \{\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\}$, dann erhalten wir nach einfacher Rechnung $[A]_{C,B} = \begin{pmatrix} 6 & 0 \\ -4 & -1 \end{pmatrix}$.

### 5.6.6

Wir betrachten die Menge $Hom_K(V, W)$ aller linearen Abbildungen von $V$ in $W$. Diese Menge ist ein Vektorraum:
Das Vielfache einer Abbildung $f$ ist die Abbildung $\lambda f$ mit $(\lambda f)(v) = \lambda f(v)$, die Summe von $f$ und $g$ ist durch $(f + g)(v) = f(v) + g(v)$ gegeben.

### 5.6.7 Satz:

Seien $B, C$ fixierte Basen von $V, W$. Dann ist

$$Hom_K(V,W) \longrightarrow K^{m\times n},\ f \mapsto [f]_{C,B}$$

ein Isomorphismus von $K$-Vektorräumen.

Eine lineare Abbildung ist eindeutig durch ihre Werte auf $B$ bestimmt. Diese Werte dürfen beliebig gewählt werden, deshalb hat man die Umkehrabbildung

$$K^{m\times n} \longrightarrow Hom_K(V,W),\ A \mapsto f,$$

wobei $f$ dadurch bestimmt ist, daß der Wert $f(b_i) \in W$ bezüglich der Basis $C$ den Koordinatenvektor $[f(b_i)]_C = s_i(A)$ haben soll.

Bemerkung: Wenn eine Teilmenge $M = (v_1, \ldots, v_r)$ von $V$ linear abhängig ist, so können die Werte von $f$ auf $M$ nicht beliebig gewählt werden

### 5.6.8 Hauptsatz:

Seien $f : U \longrightarrow V,\ g : V \longrightarrow W$ zwei lineare Abbildungen und es seien Basen $A, B, C$ von $U, V, W$ fixiert. Dann gilt für die Koordinatenmatrizen

$$[g]_{C,B}[f]_{B,A} = [gf]_{C,A}.$$

Beweis: Auf beiden Seiten der Gleichung stehen $l \times n$-Matrizen $L$ bzw. $R$, diese sind genau dann gleich, wenn $Lx = Rx$ für alle $x \in K^{n\times 1}$ gilt. Sei also $x \in K^{n\times 1}$ beliebig, dann ist $x = [u]_A$ für ein $u \in U$. Auf der linken Seite haben wir

$$[g]_{C,B}([f]_{B,A}[u]_A) = [g]_{C,B}[f(u)]_B = [g(f(u))]_C,$$

rechts haben wir

$$[gf]_{C,A}[u]_A = [(gf)(u)]_C$$

und diese Werte sind gleich. □

### 5.6.9 Basiswechsel

Wir betrachten im Vektorraum $V$ zwei Basen $B, C$. Wir bilden zur identischen Abbildung $id_V : V \longrightarrow V$ die Koordinatenmatrix $[id_V]_{C,B} = [1]_{C,B}$. Für jeden Vektor $x \in V$ gilt dann

$$[1]_{C,B}[x]_B = [x]_C,$$

d.h. die Matrix $[1]_{C,B}$ rechnet die $B$-Koordinaten in die $C$-Koordinaten um. Man nennt $[1]_{C,B}$ die Übergangsmatrix von $B$ zu $C$, diese ist quadratisch, weil zwei Basen desselben Raums gleichviele Elemente haben. Übergangsmatrizen sind auch invertierbar, denn

$$[1]_{C,B}[1]_{B,C} = [1]_{C,C} = I.$$

Wir geben $[1]_{C,B}$ explizit an: Die $i$-te Spalte ist gleich $[b_i]_C$.

## 5.6.10 Beispiel

Wir betrachten die Basis $B = \{\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\} \in K^{2\times 1}$, $S$ sei die Standardbasis. Wir wollen die Koordinaten von $x = [x]_S$ bezüglich der Basis $B$ ausrechnen. Dazu benötigen wir die Übergangsmatrix $[1]_{B,S}$, deren $i$-te Spalte ist gleich $[e_i]_B$, diese Koordinaten berechnen wir:

$$\begin{pmatrix}1 & 1 & 1 & 0\\ 2 & -1 & 0 & 1\end{pmatrix} \sim \begin{pmatrix}1 & 0 & 1/3 & 1/3\\ 0 & 1 & 2/3 & -1/3\end{pmatrix}.$$

Der rechte Teil ist $[1]_{B,S}$.

Dies ist genau die inverse Matrix zu $[1]_{S,B} = ([b_1]_S, [b_2]_S$.

Sei jetzt $x = \begin{pmatrix}3\\1\end{pmatrix}$, wir berechnen

$$[x]_B = [1]_{B,S}[x]_S = \begin{pmatrix}1/3 & 1/3\\ 2/3 & -1/3\end{pmatrix}\begin{pmatrix}3\\1\end{pmatrix} = \begin{pmatrix}4/3\\5/3\end{pmatrix}.$$

Allgemein: Sei in $K^{n\times 1}$ eine Basis $B$ gegeben, dann ist die Übergangsmatrix von $S$ zu $B$ die Inverse derjenigen Matrix, welche als Spaltenvektoren die Elemente von $B$ hat.

Mit den Übergangsmatrizen können wir nicht nur Koordinatenvektoren umrechnen, sondern auch Koordinatenmatrizen:

## 5.6.11 Satz:

Sei $f : V \longrightarrow W$ eine $K$-lineare Abbildung und seien $B, B'$ zwei Basen von $V$ und $C, C'$ zwei Basen von $W$. Dann gilt

$$[f]_{C',B'} = [1_W]_{C',C}[f]_{C,B}[1_V]_{B.B'}.$$

Der Beweis folgt sofort aus dem Hauptsatz, weil $id_W \circ f \circ id_V = f$ ist. □

## 5.6.12 Koordinatenmatrizen linearer Operatoren

Schließlich betrachten wir noch den Spezialfall $V = W, B = C$. Eine lineare Abbildung $f : V \longrightarrow V$ eines Vektorraums in sich nennt man einen linearen Operator. In 5.4.4 hatten wir Beispiele im 2-dimensionalen Fall betrachtet.

Hier schreiben wir kurz $[f]_B := [f]_{B,B}$.
Das Produkt (die Hintereinanderausführung) zweier linearer Operatoren ist wieder ein linearer Operator und es gilt $[gf]_B = [g]_B[f]_B$.
Die Menge $End_K(V) := Hom_K(V, V)$ ist ein $K$-Vektorraum und ein Ring, also eine $K$-Algebra.
Sei $dim(V) = n$, dann ist

$$End_K(V) \longrightarrow K^{n\times n},\ f \mapsto [f]_B$$

ein Isomorphismus von $K$-Algebren.

Wenn $C$ eine andere Basis von $V$ ist, so gilt

$$[f]_C = [1]_{C,B}[f]_B[1]_{B,C} = P^{-1}[f]_B P,$$

mit $P = [1]_{B,C}$.

### 5.6.13 Beispiel:

Sei $f : K^{2\times 1} \longrightarrow K^{2\times 1}$ durch $f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_1 \\ -2x_1 + 4x_2 \end{pmatrix}$ gegeben und sei $B = \{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}\}$. Dann ist $[f]_S = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$.
Die Matrix $[f]_B$ erhalten wir als $[1]_{B,S}[f]_S[1]_{S,B}$. Die Matrix $[1]_{S,B}$ hat als Spalten die Vektoren aus $B$. Deren Inverse ist $[1]_{B,S} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ und damit erhalten wir $[f]_B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$.
Dies ist eine Diagonalmatrix, d.h. bezüglich der Basis $B$ ist $f$ eine Maßstabänderung.

### 5.6.14 Ähnlichkeit von Matrizen

Quadratische Matrizen $A, B \in K^{n\times n}$ heißen ähnlich, falls es eine invertierbare Matrix $P \in K^{n\times n}$ gibt, so daß $B = P^{-1}AP$ gilt. Die Ähnlichkeit ist eine Äquivalenzrelation. Die Koordinatenmatrizen $[f]_B$ eines linearen Operators $f$ ergeben bei der Variation der Basis $B$ eine Äquivalenzklasse ähnlicher Matrizen.

### 5.6.15 Beispiel: Projektoren

Sei $p : V \longrightarrow V$ ein Projektor. Dann findet man immer eine Basis $B$ von $V$, so daß $[p]_B = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ ist, die Anzahl $r$ der Einsen ist der Rang. Bezüglich einer beliebigen Basis $C$ ist also $[p]_C$ ähnlich zu $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

Beweis: Es gilt $V = Bild(p) \oplus Ker(p)$. Wir wählen eine Basis $B$, die sich aus Basen beider Teilräume zusammensetzt. Dann ist $[p]_B = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ und mit $P = [1]_{B,C}$ ist $[p]_C = P^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}P$. □

# Kapitel 6

# Euklidische Vektorräume

In diesem Kapitel betrachten wir nur reelle Vektorräume, d.h. der Skalarkörper ist der Körper $R$.
wir betrachten Räume mit einer Zusatzstruktur: bisher erhielten wir als Summe zweier Vektoren einen neuen Vektor und als Produkt eines Skalars und eines Vektors einen Vektor. Zusätzlich führen wir nun ein sog. Skalarprodukt ein, es ordnet je zwei Vektoren einen Skalar zu. Die Haupteffekte dieser Konstruktion bestehen darin, daß wir jedem Vektor eine Länge zuordnen können und festlegen können, was der Winkel zwischen zwei Vektoren sein soll. Damit sind dann geometrische Aussagen möglich.

## 6.1 Definition und Grundlagen

### 6.1.1 Definition

Sei $V$ ein reeller Vektorraum. Ein Skalarprodukt auf $V$ ist eine Abbildung

$$V \times V \longrightarrow R,\ (v, w) \mapsto \langle v, w\rangle$$

mit den folgenden Eigenschaften:

1. Das Produkt ist bilinear , d.h.

$$\langle \lambda v + \lambda' v', w\rangle = \lambda\langle v, w\rangle + \lambda'\langle v', w\rangle,$$

$$\langle v, \mu w + \mu' w'\rangle = \mu\langle v, w\rangle + \mu'\langle v, w'\rangle,$$

   wobei $v, v', w, w' \in V,\ \lambda, \lambda', \mu, \mu' \in R$ beliebig sind.

2. Das Produkt ist symmetrisch , d.h.

$$\langle v, w\rangle = \langle w, v\rangle$$

   für alle $v, w \in V$.

3. Das Produkt ist positiv definit, d.h.

$$\langle v, v\rangle \geq 0 \text{ für alle } v \in V,$$

$$\langle v, v\rangle = 0 \Leftrightarrow v = o.$$

Ist ein Skalarprodukt gegeben, dann nennt man $\|v\| = \sqrt{\langle v, v\rangle}$ (die positive reelle Wurzel) die Norm oder die Länge des Vektors $v$.

Ein mit einem Skalarprodukt versehener reeller Vektorraum heißt euklidischer Vektorraum.

Das Standardbeispiel ist $V = R^{n\times 1}$:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},\ \langle x, y\rangle = x^t \cdot y = \sum x_i y_i.$$

### 6.1.2 Satz

Sei $V$ ein euklidischer Vektorraum, dann gilt für die Vektornorm:

1. $\|v\| = 0 \Leftrightarrow v = o$,
2. $\|\lambda v\| = |\lambda|\|v\|$ für alle $\lambda \in R$,
3. $\|v + w\| \leq \|v\| + \|w\|$ (Dreiecksungleichung, diese ist äquivalent zur Cauchy-Schwartzschen [1] Ungleichung:
4. $|\langle v, w\rangle| \leq \|v\| \cdot \|w\|$.

### 6.1.3 Bilinearformen und Gram-Matrizen

Sei $V$ ein $K$-Vektorraum, eine Abbildung $V \times V \longrightarrow K,\ (v, w) \mapsto \langle v, w\rangle$ heißt Bilinearform, wenn die Bedingungen 6.1.1 (1) erfüllt sind.

Ähnlich wie eine lineare Abbildung ist auch eine Bilinearform durch wenige Daten voll bestimmt. Sei nämlich $B = (b_1, \ldots, b_n)$ eine Basis von $V$. Wenn die Werte $\langle b_i, b_j\rangle$ bekannt sind, so ist die Bilinearform dadurch voll bestimmt. Wir nennen

$$G_B = (\langle b_i, b_j\rangle) \in K^{n\times n}$$

die Gram-Matrix von $\langle\ ,\ \rangle$ bezüglich der Basis $B$. Wenn nun $v, w \in V$ die Koordinatenvektoren $[v]_B,\ [w]_B \in K^{n\times 1}$ haben, dann gilt

$$\langle v, w\rangle = [v]_B^t \cdot G_B \cdot [w]_B \in K.$$

Beweis:

$$\langle v, w\rangle = \langle \sum \lambda b_i, \sum \mu_j b_j\rangle = \sum \lambda_i \mu_j \langle b_i, b_j\rangle.$$

Wir bezeichnen für einen Moment die Einträge von $G_B$ durch $g_{ij}$. Dann gilt

$$(\lambda_1, \ldots, \lambda_n) G_B \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \sum_j (\sum_i \lambda_i g_{ij}) \mu_j$$

und wir haben das gewüschte Ergebnis.

[1] Augustin Louis Cauchy, 1789 - 1857, Paris, ist der Begründer der modernen Analysis; Hermann Amandus Schwartz, 1843 - 1921, war Universitätslehrer in Berlin

### 6.1.4 Wichtiger Spezialfall

$V = K^{n\times 1}$ sei der Koordinatenraum und $S = (e_1, \ldots, e_n)$ die Standardbasis. Jedes $x \in V$ ist bezüglich $S$ sein eigener Koordinatenvektor. Ist nun $\langle , \rangle$ eine Bilinearform auf $V$, dann sei

$$A = G_S = (\langle e_i, e_j \rangle)$$

die Gram-Matrix bezüglich der Standardbasis. Dann folgt aus 6.1.3 mit $B = S$ und $a_{ij} = \langle e_i, e_j \rangle$:

$$\langle x, y \rangle = x^t A y = \sum a_{ij} x_i y_j.$$

**Satz:** Die Zuordnung, die einer Bilinearform auf $K^{n\times 1}$ ihre Gram-Matrix bezüglich der Standardbasis zuordnet, ist eine Bijektion zwischen der Menge der Bilinearformen und der Menge der $n \times n$-Matrizen.
Wir schreiben

$$\langle x, y \rangle_A := x^t A y.$$

Die Standardform gehört zur Einheitsmatrix $I$, d.h.

$$\langle x, y \rangle_I := x^t I y = x^t y = \sum x_i y_i.$$

Beweis: Einerseits findet man zu einer Bilinearform $\langle\ ,\ \rangle$ eindeutig die Matrix $A = G_S$. Andererseits kann man zu $A$ die Form $\langle x, y, \rangle_A$ bilden. Diese Abbildungen sind zueinander invers, denn

$$\langle\ ,\ \rangle \mapsto G_S \mapsto \langle\ ,\ \rangle_{G_S} = \langle\ ,\ \rangle,$$

$$A \mapsto \langle\ ,\ \rangle_A \mapsto G_S = A,$$

d.h. die Gram-Matrix $G_S$ der Form $\langle\ ,\ \rangle_A$ ist genau die Matrix $A$. □

### 6.1.5 Beispiele

A) $V = R^{2\times 1}$, $A = \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}$, dann ist $\langle x, y \rangle_A = x^t A y = 2x_1y_1 + 3x_1y_2 + x_2y_2$.

B) Das Standardprodukt auf dem 3-dimensionalen Anschauungsraum $E$

Zu $E$ gehört der Vektorraum $V(E)$ der Äquivalenzklassen von Pfeilen bzw. der Translationsraum $T(E)$ von $E$ (vgl. 5.1.3).
Wir fixieren in $E$ einen Ursprung $O$ und ein rechtwinkliges $(x_1, x_2, x_3)$-Koordinatensystem. Die Standardbasis $S$ von $V(E)$ besteht aus den Vektoren $e_1, e_2, e_3$, wobei $e_i$ die Translation um 1 in $x_i$-Richtung ist, d.h. die Äquivalenzklasse aller Pfeile der Länge 1 in $x_i$-Richtung $P \mapsto P + e_i$. Das Standardprodukt auf $V(E)$ bezüglich des fixierten Koordinatensystems hat die Gram-Matrix $G_S = I_3$, d.h. $\langle e_i, e_j \rangle_I = \delta_{ij}$ ist das Kronecker-symbol.

Wenn $[x]_S = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ ist, so ist $\langle x, y \rangle_I = \sum x_i y_i$.
Entsprechend führt man das Standardprodukt im 2-dimensionalen Anschauungsraum ein.

### 6.1.6 Symmetrische Bilinearformen

Wann ist eine Bilinearform $\langle , \rangle$ auf $K^{n\times 1}$ symmetrisch, d.h. $\langle x, y\rangle = \langle y, x\rangle$ für alle $x, y$ ?
Diese Frage entscheidetsich wieder an der Basis. Es genügt $\langle e_i, e_j\rangle = \langle e_j, e_i\rangle$ für alle Basisvektoren. Für die Gram-Matrix $A = (\langle e_i, e_j\rangle)$ bedeutet dies $A^t = A$.
Solche Matrizen nennt man symmetrisch.

Allgemeiner: Eine Bilinearform $\langle , \rangle$ auf dem $K$-Vektorraum $V$ ist genau dann symmetrisch, wenn für eine (für jede) Basis $B = (b_1, \ldots, b_n)$ die Gram-Matrix $G_B = (\langle b_i, b_j\rangle$ eine symmetrische Matrix ist.

### 6.1.7 Satz

Bei der Bijektion 6.1.4 entsprechen die symmetrischen Bilinearformen den symmetrischen Matrizen, d.h. eine Bilinearform ist genau dann symmetrisch, wenn ihre Gram-Matrix $A$ symmetrisch ist. In der ausgeschriebenen Form

$$\langle x, y\rangle = \sum a_{ij} x_i y_j$$

bedeutet das $a_{ij} = a_{ji}$; bei $x_i y_j$ und bei $x_j y_i$ steht derselbe Koeffizient.

### 6.1.8 Positiv definite symmetrische Bilinearformen

Dieser Begriff macht nur bei $K = R$ einen Sinn, denn wir fordern $\langle v, v\rangle \geq 0$ für alle $v \in V$ (allgemeiner macht es für „geordnete Körper“ Sinn).

Wir betrachten also $V = R^{n\times 1}$ und darauf die symmetrische Bilinearform

$$\langle x, y\rangle = x^t A y,$$

d.h. $A \in R^{n\times n}$ ist eine symmetrische Matrix. Man nennt die Matrix $A$ positiv definit, falls diese Form positiv definit ist, d.h. $x^t A x \geq 0$ und aus $x^t A x = 0$ folgt $x = o$.

Bemerkung: Die Standardform $\langle x, y\rangle = x^t y$ ist positiv definit, denn $\langle x, x\rangle = \sum x_i^2 \geq 0$ und wenn der Wert gleich 0 ist, so sind alle $x_i$ gleich 0.
Daraus folgt: Wenn eine Matrix $P$ mit $\det(P) \neq 0$ gegeben ist, dann ist die Matrix $A = P^t P$ positiv definit.

Beweis: $\langle x, x\rangle_A = x^t A x = x^t P^t P x = (Px)^t P x = \langle Px, Px\rangle_I \geq 0$ und $\langle x, x\rangle_A = 0$ bedeutet $\langle Px, Px\rangle_I = 0$, also $Px = o$ und da $P$ invertierbar ist, folgt $x = o$. □

Das einfachste Beispiel ist eine Matrix der Form

$$A = \begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_n \end{pmatrix} \text{ mit } w_i > 0,$$

wir nehmen einfach

$$P = \begin{pmatrix} \sqrt{w_1} & & \\ & \ddots & \\ & & \sqrt{w_n} \end{pmatrix}.$$

Wir zeigen später: Eine Matrix $A$ ist genau dann positiv definit und symmetrisch, wenn eine MAtrx $P \in GL_n(R)$ mit $A = P^t P$ existiert.

Damit haben wir eine Übersicht über alle Skalarprodukte auf dem Koordinatenraum. Es gibt sehr viele; trotzdem spielt das Standardprodukt mit $A = I_n$ eine besondere Rolle, wie wir sehen werden.

Es gibt einen praktischen Test für die positive Definitheit, wie wir später beweisen wollen:

Sei $A \in R^{n\times n}$ eine symmetrische Matrix. Für $m \leq n$ bezeichne $A[m] \in R^{m\times m}$ die Matrix, die aus $A$ durch Abschneiden nach der $m$-ten Zeile und Spalte entsteht. Dann gilt: $A$ ist genau dann positiv definit, wenn $\det(A[m]) > 0$ für alle $m = 1, \ldots, n$ gilt.

### 6.1.9 Ein weiteres Beipiel eines euklidischen Raums

Wir betrachten die Menge $V$ aller auf dem abgeschlossenen Intervall $[a, b] \subset R$ stetigen reellwertigen Funktionen. Dies ist ein Vektorraum: die Summe und die Multiplikation mit Skalaren ist argumentweise definiert, der Nullvektor ist die Nullfunktion. Wir führen in $V$ ein Skalarprodukt ein:

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx.$$

Dann gilt $\langle f, f\rangle = \int_a^b f(x)^2 dx \geq 0$ und wenn $f$ nicht die Nullfunktion ist, so ist $\langle f, f\rangle >$ 0.
Damit hat man als Norm $\|f\| = \sqrt{\int_a^b f(x)^2 dx}$ und die Cauchy-Schwartzsche Ungleichung lautet hier

$$\left[\int_a^b f(x)g(x)dx\right]^2 \leq \int_a^b f(x)^2 dx \cdot \int_a^b g(x)^2 dx.$$

Als nächstes wollen wir Abstände, Winkel und Orthogonalität im euklidischen Raum behandeln.

Sei $V$ ein euklidischer Vektorraum. Wir haben hier eine Längenfunktion $\|u\|$. Daraus ergibt sich eine

### 6.1.10 Abstandsfunktion

Es sei $d(u, v) := \|u - v\|$ der Abstand der Vektoren $u, v$.

Eigenschaften des Abstands: $d(u, v) = 0$ genau dann, wenn $u = v$.

$d(u, v) = d(v, u)$; der Abstand ist symmetrisch.

$d(\lambda u, \lambda v) = |\lambda| d(u, v)$, dies ist der Strahlensatz.

$d(u, v) \leq d(u, w) + d(w, v)$, die Dreiecksungleichung.

Im Spezialfall $V = R^{n\times 1}$ ist $d(u, v) = \sqrt{\sum(u_i - v_i)^2}$.

### 6.1.11 (Formale) Definition des Winkels zwischen zwei Vektoren

Wir betrachten die Cauchy-Schwartzsche Ungleichung

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|,$$

das bedeutet

$$-1 \leq \frac{\langle u, v\rangle}{\|u\| \cdot \|v\|} \leq +1.$$

**Lemma:** *Seien $u, v \in V$ vom Nullvektor verschiedene Vektoren. Dann gilt*

$$\frac{\langle u, v\rangle}{\|u\| \cdot \|v\|} = 1 \text{ gdw. } v = \lambda u \text{ mit } \lambda > 0,$$

$$\frac{\langle u, v\rangle}{\|u\| \cdot \|v\|} = -1 \text{ gdw. } v = \lambda u \text{ mit } \lambda < 0.$$

Beweis: Es ist

$$\frac{\langle u, v\rangle}{\|u\| \cdot \|v\|} = \langle \frac{u}{\|u\|}, \frac{v}{\|v\|}\rangle = \langle u_0, v_0\rangle,$$

wobei $u_0 = \frac{u}{\|u\|}, v_0 = \frac{v}{\|v\|}$ die Vektoren der Länge 1 in Richtung von $u, v$ sind. Es gilt

$$\|u_0 - v_0\|^2 = \langle u_0 - v_0, u_0 - v_0\rangle = \langle u_0, u_0\rangle - 2\langle u_0, v_0\rangle + \langle v_0, v_0\rangle = 2 \cdot (1 - \langle u_0, v_0\rangle).$$

Also $u_o = v_0$ gdw. $\|u_0 - v_0\| = 0$ gdw. $\langle u_0, v_0\rangle = 1$.
Entsprechend gilt $\|u_0 + v_0\|^2 = 2 \cdot (1 + \langle u_0, v_0\rangle)$, also $u_0 = -v_0$ gdw. $\langle u_0, v_0\rangle = -1$. □

Wir vergleichen das mit der Cosinus-Funktion im Intervall $[0, \pi]$. Wenn (anschaulich) der Winkel $\theta$ zwischen $u, v$ null ist, so ist der Wert des Quotienten $\frac{\langle u,v\rangle}{\|u\|\cdot\|v\|}$ gleich 1; wenn $\theta = \pi$ ist, so ist der Wert des Quotienten gleich $-1$.

### 6.1.12 Definition des Winkels

Der Winkel $\theta(u, v)$ zwischen zwei Vektoren ist der eindeutig bestimmte Wert $\theta(u, v) \in [0, \pi]$, so daß

$$\cos(\theta(u, v)) = \frac{\langle u, v\rangle}{\|u\| \cdot \|v\|}.$$

**Lemma:** $\theta(u, v) = 0$ *gdw.* $v = \lambda u$ *mit* $\lambda > 0$.
$\theta(u, v) = \pi$ *gdw.* $v = \lambda u$ *mit* $\lambda < 0$. □

Aus der Definition folgt

$$\theta(u, v) = \pi/2 \Leftrightarrow \langle u, v\rangle = 0,$$

in diesem Fall nennt man die Vektoren orthogonal .

### 6.1.13 Bemerkung

Im Fall $V = V(E) \equiv R^{3\times 1}$ mit dem Standardprodukt $\langle u, v\rangle = u^t v$ ist der Winkel genau das, was man sich darunter vorstellt. In der Schule wird das Skalarprodukt im Anschauungsraum durch

$$\langle u, v\rangle = \|u\| \cdot \|v\| \cdot \cos(\alpha)$$

definiert, wo $\alpha$ der eingeschlossene Winkel ist. Ansonsten hat man eine formale Analogie.

Bemerkung: Sei $\langle u, v\rangle_A = x^t A y$ ein Skalarprodukt auf $R^{n\times 1}$ und $A = P^t P$ mit invertierbarem $P$. Dann folgt für den Winkel bezüglich des Skalarprodukts $\langle u, v\rangle_A$:

$$\cos(\theta(u,v)_A) = \frac{\langle u, v\rangle_A}{\|u\|_A \cdot \|v\|_A} = \frac{\langle Pu, Pv\rangle_I}{\|Pu\|_I \cdot \|Pv\|_I} = \cos(\theta(Pu, Pv)_I).$$

Also ist der Winkel zwischen $u, v$ in der $A$-Metrik derselbe wie der Winkel zwischen $Pu, Pv$ in der Standardmetrik.

Natürlich hängen Normen und Abstände von dem jeweiligen Skalarprodukt ab, z.B. sind Kreise in einer Metrik in einer anderen evtl. Ellipsen.

## 6.2 Orthogonale Projektoren und das Gram-Schmidtsche Orthogonalisierungsverfahren

Wir betrachten einen euklidischen Vektorraum $V/\mathbf{R}, \langle\ ,\ \rangle$. Aufgrund der Definition des Winkels (6.1.12) heißen zwei Vektoren $u, v$ orthogonal, falls $\langle u, v\rangle = 0$ gilt. Dafür schreiben wir auch $u \perp v$.

### 6.2.1 Definition:

a) Zwei Teilmengen $M, N \subset V$ heißen orthogonal, falls $\langle u, v\rangle = 0$ für alle Paare von Vektoren $u \in M, v \in N$. Wir schreiben dafür $M \perp N$.

b) Das volle Orthogonal $M^\perp$ einer Menge $M$ ist definiert als $M^\perp = \{v \in V \mid v \perp M\}$.

### 6.2.2 Bemerkung

a) $M^\perp$ ist stets ein Unterraum von $V$.

b) Wie üblich sei $Span(M)$ der von $M$ erzeugte Teilraum. Dann ist $M^\perp = Span(M)^\perp$.

Der Beweis folgt aus der Bilinearität des Skalarprodukts.

In 5.5.12 ff hatten wir *Projektoren* auf einem $K$-Vektorraum $V$ betrachtet. Charakteristisch war die Eigenschaft $p^2 = p$, und daraus folgte, daß

$$V = Ker(p) \oplus Bild(p) \qquad (1)$$

die direkte Summe von Kern und Bild von $p$ ist. Jetzt ist $V$ ein euklidischer Raum. Deswegen macht folgende Definition Sinn:

### 6.2.3 Definition / Satz:

a) Wir nennen einen Projektor orthogonal, falls Kern und Bildaufeinander senkrecht stehen, d.h.

$$Ker(p) \perp Bild(p).$$

b) Wenn $p$ ein orthogonaler Projektor ist, dann bestimmen sich Kern und Bild gegenseitig, nämlich

$$Ker(p) = Bild(p)^{\perp}, \quad Bild(p) = Ker(p)^{\perp}.$$

Beweis von b): Aus der Definition $Ker(p) \perp Bild(p)$ folgt offensichtlich $Ker(p) \subseteq Bild(p)^{\perp}$ sowie $Bild(p) \subseteq Ker(p)^{\perp}$. Wir zeigen, daß diese Inklusionen Gleichheiten sind. Dabei beschränken wir uns auf den ersten Fall. Wir nehmen an, daß $x \in Bild(p)^{\perp}, x \notin Ker(p)$ ist. Aus der Eigenschaft (1) von Projektoren folgt, daß wir $x = a + b$ mit $a \in Ker(p), b \in Bild(p)$ schreiben können. Sei $v = x - a = b$. Offensichtlich ist $x - a \in Bild(p)^{\perp}, b \in Bild(p) \subseteq Ker(p)^{\perp}$. Also steht $v$ sowohl auf $Bild(p)$ als auch $Ker(p)$ senkrecht. Wieder wegen (1) folgt $v \perp V$. Jedoch ist $\langle v, v \rangle = 0$ nur für $v = o$ möglich, also ist $x = a, b = o$. □

**Bemerkungen:** a) Ist $W \subseteq V$ ein Unterraum, dann kann es nach 6.2.3 b) höchstens einen orthogonalen Projektor $p$ mit $Bild(p) = W$ geben, denn es folgt notwendig $Ker(p) = W^{\perp}$. Im weiteren sehen wir, daß $p$ wirklich existiert.

b) Mit $p$ ist auch der komplementäre Projektor $I - p$ orthogonal, denn hier wird nur die Rolle von Kern und Bild des Projektors vertauscht. Die Tatsache, daß Kern und Bild aufeinander senkrecht stehen, bleibt dabei unverändert.

c) Aus den Gleichungen 6.2.3 b) folgt direkt

$$Ker(p)^{\perp\perp} = Ker(p), \; Bild(p)^{\perp\perp} = Bild(p).$$

Wegen $W \perp W^{\perp}$ ist die Inklusion $W \subseteq W^{\perp\perp}$ immer offensichtlich.

### 6.2.4 Definition:

Sei $V, \langle \ , \ \rangle$ ein euklidischer Raum und $W \subseteq V$ ein Unterraum. Eine Basis $b_1, \ldots, b_r$ von $W$ heißt Orthogonalbasis, falls $b_i \perp b_j$ für alle $i \neq j$ gilt. Man spricht von einer Orthonormalbasis, falls

$$\langle b_i, b_j \rangle = \delta_{ij}, \; i, j = 1, \ldots, r$$

gilt.

Anschaulich heißt das, daßs alle Vektoren $b_i$ die Länge 1 haben und paarweise senkrecht aufeinander stehen.

**Bemerkung:** Ein System paarweise orthogonaler Vektoren ist stets linear unabhängig:

$$\sum \lambda_i b_i = o \Rightarrow \lambda_k \langle b_k, b_k \rangle = 0 \Rightarrow \lambda_k = 0.$$

### 6.2.5 Satz:

Sei $W \subseteq V$ ein Unterraum und $b_1, \ldots, b_r$ eine Orthonormalbasis von $W$, dann ist die Abbildung

$$v \in V \mapsto p(v) = \sum_{i=1}^{r} \langle v, b_i \rangle b_i$$

ein orthogonaler Projektor von $V$ auf $W$. Außerdem ist $I-p$ ein orthogonaler Projektor von $V$ auf $W^\perp$.

a) $p$ ist eine lineare Abbildung, da das Skalarprodukt bilinear ist, und das Bild von $p$ ist offensichtlich in $W$ enthalten.

b) Für alle Basisvektoren $b_k$ haben wir $p(b_k) = \sum \langle k_k, k_i \rangle b_i = b_k$, weil es sich um eine Orthonormalbasis handelt. Da $p$ linear ist und die Basisvektoren von $W$ auf sich abbildet, gilt $p(w) = w$ für alle $w \in W$. Somit ist $p(p(v)) = p(v)$, d.h. $p$ ist Projektor und $Bild(p) = W$.

c) $p$ ist orthogonaler Projektor:
Aus der Definition von $p$ folgt direkt $p(v) = o$ genau dann, wenn $\langle v, b_i \rangle = 0$ für alle $i = 1, \ldots, r$ gilt, also wenn $v \perp W = Spann(b_1, \ldots, b_r)$ ist. Damit haben wir $Ker(p) \perp Bild(p)$.

Schließlich betrachten wir noch den komplementären Projektor $I - p$. Dann wissen wir

$$Bild(I-p) = Ker(p) = Bild(p)^\perp, \quad Ker(I-p) = Bild(p).$$

Diese Räume stehen aufeinander senkrecht, also ist $I - p$ ebenfalls ein orthogonaler Projektor. □

**Variante:** Seien $b_1, \ldots, b_r$ paarweise orthogonal, aber von beliebiger Länge ($\neq 0$), dann leistet $p(v) = \sum_{i=1}^{r} \frac{\langle v, b_i \rangle}{||b_i||^2} \cdot b_i$ ebenfalls das Verlangte.

### 6.2.6 Satz:

Sei $V/\mathbf{R}, \langle\ ,\ \rangle$ ein endlichdimensionaler euklidischer Raum und sei $B = (b_1, \ldots, b_n)$ eine beliebige Basis. Dann kann man ausgehend von $B$ eine Orthonormalbasis von $V$ konstruieren.

**Gram-Schmidt-Verfahren**[2] **:** Wir betrachten in $V$ die Flagge

$$W_1 = Spann(b_1) \subset W_2 = Spann(b_1, b_2) \subset \ldots \subset W_n = Spann(b_1, \ldots, b_n) = V$$

und machen induktiv aus $b_1, \ldots, b_r$ eine Orthonormalbasis von $W_r$.
Induktionsanfang: $dim W_1 = 1$, also ist $c_1 = \frac{b_1}{||b_1||}$ eine Orthonormalbasis von $W_1$.

[2] Jörgen Pederson Gram (1850 - 1916), dänischer Mathematiker, beschäftigte sich hauptsächlich mit Versicherungsstatistik; Erhardt Schmidt (1876 - 1959), studierte in Göttingen bei David Hilbert, Promotion 1905, ab 1917 Professor an der Berliner Universität, er hat den Begriff des Hilbertraums klar herausgearbeitet

Induktionsschritt: Wir haben bereits eine Orthonormalbasis $c_1, \ldots, c_r$ für $W_r$ konstruiert. Dann bekommen wir daraus den orthogonalen Projektor $P_r$ von $V$ auf $W_r$ (vgl. 6.2.8):

$$p_r(v) = \sum_{i=1}^{r} \langle v, c_i \rangle c_i.$$

Wir betrachten nun $p_r$ als orthogonalen Projektor von $W_{r+1}$ auf $W_r$, dann ist $W_{r+1} = Ker(p_r) \oplus W_r$ und diese Räume sind orthogonal. Für den komplementären Projektor $I - p_r : W_{r+1} \longrightarrow W_{r+1}$ gilt dann $Ker(p_r) = (I - p_r)(W_{r+1})$. Das heißt, das Bild des komplementären Projektors steht senkrecht auf $W_r$. Wir nehmen also

$$c_{r+1} = \frac{(I - p_r)(b_{r+1})}{\|(I - p_r)(b_{r+1})\|}.$$

Nach Voraussetzung ist $b_{r+1} \notin W_r = Ker(I - p_r)$, also ist $(i - p_r)(b_{r+1}) \neq o$ und steht senkrecht auf $W_r$. Deshalb ist $c_1, \ldots, c_{r+1}$ eine Orthonormalbasis von $W_{r+e}$. □

Explizit hat man folgenden Algorithmus:
1. Bilde $c'_1 = b_1$ und $c_1 = \frac{c'_1}{\|c'_1\|}$.
2. Bilde $c'_2 = (I - p_1)(b_2) = b_2 - \langle b_2, c_1 \rangle c_1$ und $c_2 = \frac{c'_2}{\|c'_2\|}$
3. Bilde $c'_3 = (I - p_2)(b_3) = b_3 - \langle b_3, c_1 \rangle c_1 - \langle b_3, c_2 \rangle c_3$ und $c_3 = \frac{c'_3}{\|c'_3\|}$
...
$(r+1)$. Bilde $c'_{r+1} = (I - p_r)(b_{r+1}) = b_{r+1} - \langle b_{r+1}, c_1 \rangle c_1 - \ldots - \langle b_{+1}, c_r \rangle c_r$ und $c_{r+1} = \frac{c'_{r+1}}{\|c'_{r+1}\|}$

Bemerkung: Man kann auch die Variante von 6.2.5 enwenden, um eine Orthonormalbasis zu konstruieren.

### 6.2.7 Folgerung

Sei $V/\mathbf{R}, \langle\ ,\ \rangle$ ein endlichdimensionaler euklidischer Vektorraum und sei $B = (b_1, \ldots, b_n)$ eine Orthonormalbasis von $V$, dann gilt:
1. Zu $v \in V$ läßt sichder Koordinatenvektor $[v]_B$ leicht ausrechnen, es gilt $[v]_B = \begin{pmatrix} \langle b_1, v \rangle \\ \vdots \\ \langle b_n, v \rangle \end{pmatrix}$.
2. Die Koordinatenabbildung $v \in V \mapsto [v]_B \in \mathbf{R}^{n \times 1}$ ist eine Isometrie zwischen $V, \langle\ ,\ \rangle$ und dem Standardraum $\mathbf{R}^{n \times 1}, \langle\ ,\ \rangle_I$, d.h. es gilt

$$\langle v, w \rangle = \langle [v]_B, [w]_B \rangle_I = [v]_B^t [w]_B = \sum \langle b_i, v \rangle \langle b_i, w \rangle.$$

3. Insbesondere ergibt sich aus 1. und 2.

$$\|v\|^2 = \sum \langle b_i, v \rangle^2.$$

Beweis. Sei $[v]_B = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$, d.h. $v = \lambda_1 b_1 + \ldots + \lambda_n b_n$, also $\langle v, b_i \rangle = \lambda_i$. Wie wir wissen, ist die Koordinatenabbildung $v \mapsto [v]_B$ ein Isomorphismus zwischen $V$ und

$\mathbf{R}^{n\times 1}$. Außerdem gilt für jede Basis $B$ von $V$ (vgl. 6.1.3) $\langle v,w\rangle = [v]_B^t G_B [w]_B$, wobei $G_B = (\langle b_i, b_j\rangle$ die Gram-Matrix ist. Da $B$ eine Orthonormalbasis ist, gilt $G_B = I_n$. Der Rest ist klar. □

Wir merken uns: Man kann jeden euklidischen Raum auf den Standardraum zurückführen, indem man bezüglich einer Orthonormalbasis zum Koordinatenraum übergeht.

### 6.2.8 Folgerung:

Ist $V, \langle\ ,\ \rangle$ ein euklidischer Raum und $W \subseteq V$ ein endlichdimensionaler Unterraum, dann gilt immer $V = W \oplus W^\perp$. Man nennt $W^\perp$ das orthogonale Komplement von $W$. Es gilt $W^{\perp\perp} = W$.

Beweis: Wenn wir eine Orthonormalbasis von $W$ wählen, könen wir wie oben einen orthogonalen Projektor $p : V \longrightarrow V$ mit $Bild(p) = W$ konstruieren. Aus der Projektoreigenschaft folgt $V = Ker(p) \oplus Bild(p)$. Nun ist $p$ orthogonal, also $Ker(p) = Bild(p)^\perp$, damit ist $V = Bild(p)^\perp \oplus Bild(p) = W^\perp \oplus W$. □

### 6.2.9 Eine rechnerische Anwendung

Sei $A \in \mathbf{R}^{n\times n}$ positiv definit und symmetrisch, d.h. $\langle x,y\rangle_A = x^t A y$ ist ein Skalarprodukt auf $\mathbf{R}^{n\times 1}$. Wir konstruieren mit Gram-Schmidt eine Orthonormalbasis $B = (b_1, \ldots, b_n)$, dann gilt

$$x^t A y = [x]_B^t [y]_B.$$

Andererseits ist $[x]_B = [1]_{B,S}[x]_S = Px$, wobei $P = [1]_{B,S}, x = [x]_S$ ist. Daraus ergibt sich

$$x^t A y = (Px)^t (Py) = x^t P^t P y$$

für alle $x, y \in \mathbf{R}^{n\times 1}$, also

$$A = P^t P.$$

Wir verfahren also wie folgt: Zur positiv definiten symmetrischen Matrix $A$ konstruieren wir eine Orthonormalbasis $B = (b_1, \ldots, b_n)$ für das Skalarprodukt $\langle x,y\rangle_A$ auf $\mathbf{R}^{n\times 1}$. Nun bilden wir die Matrix $Q = (b_1, \ldots, b_n)$ mit den $b_i$ als Spalten. Dann ist $Q = [1]_{S,B}$, also $P = Q^{-1}$ und $A = P^t P = (Q^{-1})^t Q^{-1}$.
Wegen $(Q^{-1})^t = (Q^t)^{-1}$ folgt $A^{-1} = QQ^t$.

**Beispiel:** $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ ist positiv definit und symmetrisch. Wir beginen mit der Standardbasis $e_1, e_2$, setzen $d_1 = e_1$ und

$$d_2 = (I - p_1)(e_2) = e_2 - \frac{\langle e_2, e_1\rangle_A}{\langle e_1, e_1\rangle_A} e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ 1 \end{pmatrix}.$$

Wir normieren diese Vektoren, es ist $\|d_1\|^2 = 2, \|d_2\|^2 = \frac{3}{2}$ und erhalten

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & \sqrt{\frac{2}{3}} \end{pmatrix},\ det(Q) = \frac{1}{\sqrt{3}},\ P = Q^{-1} = \begin{pmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{3}{2}} \end{pmatrix}.$$

Nun haben wir $x^tAy = (Px)^t(Py) = \langle Px, Py\rangle$.

**Mehr über orthogonale Projektoren**

Sei $W$ ein endlichdimensionaler Teilraum im euklidischen Raum $V, \langle\ ,\ \rangle$. Dann wissen wir, daß ein eindeutig bestimmter orthogonaler Projektor $p_W$ mit $Bild(p_W) = W$ existiert und daß $V = W \oplus W^\perp = Bild(p_W) \oplus Ker(p_W)$ ist.

### 6.2.10 Satz vom Minimalabstand

Sei $p_W$ der orthogonale Projektor von $V$ auf den Unterraum $W$, dann gilt: 1. Für jedes $x \in V$ ist $p_W(x) \in W$ das eindeutig bestimmte Element von $W$ mit minimalem Abstand zu $x$, also

$$d(x, p_W(x)) = min\{d(x,y); y \in W\} =_{def} d(x,W).$$

Wir bezeichnen diesen Wert als den Abstand von $x$ zu $W$.

2. Explizit gilt $d(x,W) = \sqrt{\|x\|^2 - \|p_W(x)\|^2}$.

Beweis: Sei $y \in W$ und $y \neq p_W(x)$. Wir zeigen $\|x-y\| > \|x - p_W(x)\|$. Es ist nämlich $x - p_W(x) \in Ker(p_W)$, $p_W(x) - y \in W = Bild(p_W)$. Daher ist $(x - p_w(x)) \perp (p_W(x) - y)$, also

$$\|x-y\|^2 = \|(x - p_W(x) + (p_W(x) - y\|^2 = \|x - p_W(x)\|^2 + \|p_W(x) - y\|^2$$

und der letzte Summand ist unglich Null, also gilt $\|x-y\| > \|x - p_W(x)\|$.

Wenn wir in dieser Formel $y = o$ einsetzen, erhalten wir $\|x|^2 = \|x - p_W(x)\|^2 + \|p_W(x)\|^2$, daraus ergibt sich die Abstandsformel. □

### 6.2.11 Berechnung von $d(x,W)$ mit Hilfe einer Orthonormalbasis von $W$

Sei $p_W$ wie in 6.2.10 und sei $B = (b_1, \ldots, b_n)$ eine Orthonormalbasis von $W$, dann ist

$$p_W(x) = \sum_{i=1}^{n} \langle x, b_i\rangle b_i,$$

$$\|p_W(x)\|^2 = \sum_{i=1}^{r} \langle x, b_i\rangle^2 \leq \|x\|^2,$$

(Besselsche Ungleichung; Gleichheit gilt genau dann wenn $x \in W$)

$$d(x,W) = \sqrt{\|x\|^2 - \sum_{i=1}^{r} \langle x, b_i\rangle^2}$$

Bemerkung: $p_W(x)$ ist durch die Eigenschaften $p_W(x) \in W, x - p_W(x) \in W^\perp$ eindeutig bestimt, weil $V = W^\perp \oplus W$ ist.

### 6.2.12 Orthogonale Projektionen auf 1-dimensionale Unterräume

Sei $W = \mathbf{R}w$, dann ist

$$p_W(x) = \|x\| cos(\theta(x,w)) \cdot \frac{w}{\|w\|},$$
$$\|p_W(x) = \|x\| \left| cos(\theta(x,w)) \right|,$$
$$d(x,W) = \|x\| sin(\theta(x,w)).$$

Wir bemerken, daß die Sinusfunktion im Intervall $[0,\pi]$ nur nichtnegative Werte annimmt.

Wenn $dimV = n$ ist, so ist $w^\perp$ eine $(n-1)$-dimensionale Hyperebene und $p_{w^\perp} = I - p_W$ ist der komplementäre Projektor. Aus 6.2.12 folgen damit Formeln für die orthogonale Projektion auf eine Hyperebene.

### 6.2.13 Berechnung der orthogonalen Projektion und des Abstandes ohne Orthonormalbasen

Sei $B = (b_1, \dots b_r)$ eine beliebige Basis von $W$ und $G_B = (\langle b_i, b_j \rangle)$ die zugehörige Gram-Matrix. Sei $x \in V$, wir bilden den Spaltenvektor $s_B(x) = \begin{pmatrix} \langle b_1, x\rangle \\ \vdots \\ \langle b_r, x\rangle \end{pmatrix} \in \mathbf{R}^{r\times 1}$.
Für den Koordinatenvektor $[p_W(x)]_B \in \mathbf{R}^{r\times 1}$ der Projektion von $x$ auf $W$ gilt dann

$$G_B[p_W(x)]_B = s_B(x), \text{ d.h. } [p_W(x)]_B = G_B^{-1} s_B(x).$$

Mit der Cramerschen Regel erhält man

$$p_W(x) = \sum_{i=1}^r \alpha_i b_i \text{ mit } \alpha_i = \frac{det(G_B^i)}{det(G_B)},$$

wobei $G_B^i$ entsteht, wenn man die $i$-te Spalte von $G_B$ durch $s_B(x)$ ersetzt.
Für den Abstand gilt

$$d(x,W)^2 = \|x\|^2 - \|p_W(x)\|^2 = \|x\|^2 - s_B(x)^t G_B^{-1} s_B(x).$$

Beweis: Sei $p_W(x) = \alpha_1 b_1 + \ldots + \alpha_r b_r$. Es ist $x - p_W(x) \in W^\perp$, also $\langle b_i, x - p_W(x)\rangle$ für alle $i$, also

$$\langle b_i, x\rangle = \langle b_i, p_W(x)\rangle = \sum_j \langle b_i, b_j\rangle \alpha_j = z_i(G_B) \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix}.$$

Insgesamt erhalten wir

$$s_B(x) = G_B \cdot [p_W(x)]_B.$$

Weiter gilt

$$\|p_W(x)\|^2 = \langle p_W(x), p_W(x)\rangle = [p_W(x)]_B^t G_B [p_W(x)]_B = (G_B^{-1} s_B(x))^t s_B(x) = s_B(x)^t G_B^{-1} s_B(x),$$

weil $G_B$ und damit auch $G_B^{-1}$ symmetrisch ist.
Nach Pytharoras ist $d(x,W)^2 = \|x\|^2 - \|p_W(x)\|^2$, daraus folgt die Behauptung. □

Bemerkung: Wenn $B$ eine Orthonormalbasis ist, so ist $G_B = I_r$, also $s_B(x) = [p_W(x)]_B$ und wir erhalten die alten Formeln.

Nachtrag: Wieso ist $G_B$ invertierbar?
Wenn $G_B$ nicht vollen Rang hat, so gibt es ein $y \neq o$ mit $G_B[y]_B = o$, dann wäre $\langle x, y\rangle = 0$ für alle $x$, insbsondere für $x = y$, ein Widerspruch zur positiven Definitheit.

## 6.3 Anwendungen der orthogonalen Projektion

**Thema 1: Ausgleichsrechnung**

Es sei $P_s \subset \mathbf{R}[X]$ die Menge der Polynome vom Grad $< s$, dies ist ein $s$-dimensionaler Vektorraum mit der Basis $1, X, \ldots, X^{s-1}$.

### 6.3.1 Satz

Sind $t_1, \ldots, t_m \in \mathbf{R}$ $m$ verschiedene Argumente und ist $m \geq s$, dann ist die Evaluationsabbildung

$$(1) \qquad p(X) \in P_s \mapsto ev(p) = \begin{pmatrix} p(t_1) \\ \vdots \\ p(t_m) \end{pmatrix} \in \mathbf{R}^{m\times 1}$$

eine injektive lineare Abbildung von $\mathbf{R}$-Vektorräumen.

Beweis: Die Linearität ist offensichtlich.
Zur Injektivität: Sei $ev(p) = o$ der Nullvektor, d.h. $p(t_1) = \ldots = p(t_m) = 0$. Dann hat das Polynom $p$ $m$ verschiedene Nullstellen. Weil $Grad(p) < s \leq m$ ist, muß $p$ das Nullpolynom sein (vgl. Identitässatz für Polynome (4.4.9)). □

Wir setzen $W_s$ = Bild der Evaluationsabbildung (1). Dies ist ein $s$-dimensionaler Teilraum von $\mathbf{R}^{m\times 1}$. Wir betrachten auf $\mathbf{R}^{m\times 1}$ das Standardprodukt $\langle x, y\rangle = x^t y$.
Zu gegebenem $x \in \mathbf{R}^{m\times 1}$ sei $y = p_{W_s}(x)$ die orthogonale Projektion von $x$ auf $W_s$. Dann gilt:

1. $y = ev(p_0) = \begin{pmatrix} p_0(t_1) \\ \vdots \\ p_0(t_m) \end{pmatrix}$ für ein eindeutig bestimmtes Polynom $p_0$ vom Grad $< s$.
2. $\|x-y\|^2 = \|x - ev(p_0)\|^2 = \sum_{i=1}^m \mid x_i - p_0(t_i) \mid^2$ ist das Minimum aller Abstände $\|x - ev(p)\|^2$ für $p \in P_s$..

### 6.3.2 Kommentar:

Gegeben ist ein „Argumentvektor“ $t \in \mathbf{R}^{m\times 1}$ und ein „Wertevektor“ $x \in \mathbf{R}^{m\times 1}$. Wir betrachten für laufendes $p \in P_s$ die Summe der Fehlerquadrate

$$\delta_p = \sum_{i=1}^{m} \mid x_i - p(t_i) \mid^2 .$$

Dann ist $p = p_0$ mit $ev(p_0) = p_{W_s}$ das eindeutig bestimmte Polynom, für welches die Summe der Fehlerquadrate minimal wird. Wenn zufällig $x \in W_s \subset \mathbf{R}^{m\times 1}$ ist, dann ist $p = p_0$ das eindeutig bestimmte Polynom mit $p_0(t_i) = x_i$ für $i = 1, \ldots, m$. (Weil $m \geq s$ kann es höchstens ein solches $p \in P_s$ geben.)

### 6.3.3 Zur Berechnung von $p_{W_s}(x) = y = ev(p_0)$

Wir konstruieren eine Basis $B$ des Raumes $W_s$ und wenden 6.2.13 an.
Zunächst bilden $1, X, X^2, \ldots, X^{s-1}$ eine Basis von $P_s$. Folglich sind die zugehörigen Evaluationsvektoren eine Basis $B = (b_1, \ldots, b_s)$ von $W_s$:

$$b_1 = ev(1) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \ b_2 = ev(X) = \begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix}, \ \ldots, b_s = ev(x^{s-1}) = \begin{pmatrix} t_1^{s-1} \\ \vdots \\ t_m^{s-1} \end{pmatrix}.$$

Sei $x \in \mathbf{R}^{m\times 1}$ gegeben; die Koordinaten von $p_{W_s}(x)$ bezüglich der Basis $B$ sind dann nach 6.2.13

$$[p_{W_s}(x)]_B = G_B^{-1} \begin{pmatrix} \langle b_1, x\rangle \\ \ldots \\ \langle b_s, x\rangle \end{pmatrix},$$

wobei $G_B = (\langle b_i, b_j\rangle)$ die Gram-Matrix ist.
Wir bilden jetzt

$$M = (b_1, \ldots, b_s) = \begin{pmatrix} 1 & t_1 & \ldots & t_1^{s-1} \\ \vdots & \vdots & & \vdots \\ 1 & t_m & \ldots & t_m^{s-1} \end{pmatrix}$$

die $m \times s$-Matrix mit den Spaltenvektoren $b_1, \ldots, b_s$. Dann ist $G_B = M^t M$ die Gram-Matrix bezüglich des Standardprodukts.
Ebenso ist

$$s_B(x) = \begin{pmatrix} \langle b_1, x\rangle \\ \ldots \\ \langle b_s, x\rangle \end{pmatrix} = M^t x,$$

also $[p_{W_s}(x)]_B = (M^t M)^{-1}(M^t x)$. Andererseits bedeutet

$$[p_{W_s}(x)]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_s \end{pmatrix},$$

daß $p_{W_s}(x) = \sum \alpha_i b_i = \sum \alpha_i ev(X^{i-1}) = ev(\sum \alpha_i X^{i-1})$ ist. Also ist $p_0(X) = \sum \alpha_i X^{i-1}$ das gesuchte Polynom, denn $ev(p_0) = p_{W_s}(x)$ und somit ist $\delta_{p_0} = \sum(x_i - p_0(t_i))^2$ minimal. Man nennt $p_0(X)$ das ausgleichende Polynom zum gegebenen Argument- und Wertevektor. Dessen Koeffizienten erhält man aus der Gleichung

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_s \end{pmatrix} = (M^t M)^{-1}(M^t x).$$

Beispiel: $s = 2, m = 4$. Wir betrachten die Punkte (0,1), (1,3), (2,4), (3,4) und suchen die ausgleichende Gerade $\alpha_1 + \alpha_2 X$, welche zu den Argumenten $t = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}$ die Werte $x = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 4 \end{pmatrix}$ am besten approximiert. Zur Matrix $M = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$ erhalten wir $M^tM = \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}$, $(M^tM)^{-1} = \frac{1}{10}\begin{pmatrix} 7 & -3 \\ -3 & 2 \end{pmatrix}$ und $M^tx = \begin{pmatrix} 12 \\ 23 \end{pmatrix}$ und damit $\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} 1,5 \\ 1 \end{pmatrix}$, die ausgleichende Gerade ist also $p_0(X) = 1,5 + X$.

### 6.3.4 Variante:

Gegeben sei eine reelle Funktion $f(t)$ mit reellem Argument $t$. Wir wählen verschiedene $t_1, \ldots t_m$ im Definitionsbereich und bilden $x = ev(f) = \begin{pmatrix} f(t_1) \\ \vdots \\ f(t_m) \end{pmatrix}$. Sei $s \leq m$
Wir suchen das ausgleichende Polynom $p_0 \in P_s$ für diesen Argumentenvektor. Dies geschieht mit dem angegebenen Verfahren. Das Polynom $p_0$ ist also dasjenige unter den Polynomen $p \in P_s$, für welches $\delta_p = \|ev(f) - ev(p)\|^2$ minimal ist; $p_0$ ist die beste Approximation für $f$.

Wenn speziell $s = m$ ist, so ist die Evaluationsabbildung ein Isomorphismus, also gibt es für vorgegebene $t_i, x_i, i = 1, \ldots, s$ genau ein Polynom $p \in P_s$ mit $p(t_i) = x_i$ (Lagrange-Interpolation). Wenn $s > m$ ist, so ist die Evaluationsabbildung surjektiv, es gibt also unendlich viele Lösungen.

**Thema 2: Entwicklung in eine Fourierreihe**

Wir betrachten ein abgeschlossenes Intervall $I = [a, b] \subseteq \mathbf{R}$ und denVektorraum $V_I$ der stetigen $\mathbf{R}$-wertigen Funktionen auf $I$. Ein Skalarprodukt in $V_I$ ist durch

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx$$

gegeben, dies ist offensichtlich bilinear, symmetrisch und positiv definit (Achtung: hierfür ist die Stetigkeit der Funktionen notwendig). Der zugehörige Abstand ist

$$d(f, g) = \|f - g\| = \sqrt{\int_a^b (f(x) - g(x))^2 dx}.$$

Der Wert $\|f - g\|^2$ heißt die mittlere quadratische Abweichung von $f$ und $g$.
Wir nehmen $I = [0, 2\pi]$ und betrachten in $V_I$ den Unterraum

$$W_{2n+1} = Spann(1, \cos(x), \cos(2x), \ldots, \cos(nx), \sin(x), \ldots, \sin(nx)),$$

man nennt ihn den Raum der trigonometrischen Polynome der Ordnung $\leq n$.

### 6.3.5 Satz:

Sei $f \in V_I$ gegeben und $p_n(f)$ die orthogonale Projektion von $f$ auf $W_{2n+1}$, dies ist das eindeutig bestimmte trigonometrische Polynom vom Grad $\leq n$, welches minimale quadratische Abweichung von $f$ hat.

Beweis: $W_{2n+1}$ ist endlichdimensional, daher existiert die orthogonale Projetion. Die Minimalität des Abstands folgt aus 6.2.10. □

Im Gegensatz zum Thema 1 haben wir hier sogar eine Orthonormalbasis:

### 6.3.6 Satz:

Die Funktionen $g_0(x) = 1/\sqrt{2\pi}, g_k(x) = 1/\sqrt{\pi}\cos(kx), g_{n+k} = 1/\sqrt{\pi}\sin(kx), k = 1, \ldots, n$ bilden eine Orthonormalbasis vo $W_{2n+1}$.

Beweis: Seien $a \neq \pm b$ ganze Zahlen. Aus den Additionstheoremen folgt, daß sich die Produkte von Sinus- und Kosinusfunktionen durch Kosinusfunktionen darstellen lassen, weiter ist

$$\int_0^{2\pi} \sin(ax)\cos(bx)dx = -\frac{1}{2(a+b)}[\cos((a+b)x)]_{x=0}^{x=2\pi} - \frac{1}{2(a-b)}[\cos((a-b)x)]_{x=0}^{x=2\pi} = 0,$$

unsere Funktionen sind also orthogonal. Jetzt muß man nur noch normieren und erhält die Behauptung. □

### 6.3.7 Folgerung:

Bezüglich der Basis $B = (g_0, \ldots, g_{2n})$ gilt

$$p_n(f) = \sum_{k=0}^{2n} \langle f, g_k \rangle g_k = a_0(f)/2 + \sum_{k=1}^{n} a_k(f)\cos(kx) + \sum_{k=1}^{n} b_k(f)\sin(kx),$$

wobei

$$a_0(f) = \frac{1}{\pi}\langle f, 1\rangle,\ a_k(f) = \frac{1}{\pi}\langle f, \cos(kx)\rangle,\ b_k(f) = \frac{1}{\pi}\langle f, \sin(kx)\rangle.$$

Der mittlere quadratische Fehler ist

$$\|f - p_n(f)\| = \|f\|^2 - (\pi/2)a_0(f)^2 + \pi\sum_{k=1}^{n}(a_k(f)^2 + b_k(f)^2).$$

□

Die Koeffizienten $a_0(f), a_k(f), b_k(f)$ heißen die Fourierkoeffizienten von $f$.

### 6.3.8 Fakt:

$$\lim_{n\to\infty} \|f - p_n(f)\| = 0.$$

Wenn man ein bißchen rechnet, erhält man für die Funktion $f(x) = x$ die Annäherung

$$x \sim \pi - 2(\sin(x) + 1/2\sin(2x) + \ldots 1/n\sin(nx)).$$

Bemerkung: Der Fakt 6.3.8 bedeutet nur, daß $f$ bis auf eine Menge vom Maß 0 approximiert wird, die Grenzfunktion $p_\infty(f)$ kann sich von $f$ auf einer Menge vom Maß 0 unterscheiden: bei $f(x) = x$ ist dies in den Punkten 0 und $2\pi$ der Fall, die Grenzfunktion hat dort den Wert $\pi$.

## 6.4 Volumenberechnung im euklidischen Raum

Weiterhin sei $V|\mathbf{R}, < , >$ ein euklidischer Vektorraum.

### 6.4.1 Definition:

Seien $b_1, \dots, b_r \in V$ linear unabhängige Vektoren, dann nennen wir

$P(b_1, \dots, b_n) = \{\sum \lambda_i b_i,\ 0 \leq \lambda_i \leq 1\}$ das von $b_1, \dots b_r$ aufgespannte Parallelepiped (Parallelotop). Wenn diese Vektoren eine Basis $B$ bilden, so spricht man auch von der zu $B$ gehörigen Fundamentalmasche.
Allgemeiner wird auch jede Verschiebung $x + P(b_1, \dots b_r)$ als Parallelotop bezeichnet. Für $r = 2$ ist das ein Parallelogramm, das $m+1$-dimensionale Parallelotop besteht aus Schichten $m$-dimensionaler Parallelotope, die bis auf Verschiebungen einander gleich sind:

$$P(b_1, \dots, b_{r+1}) = \bigcup_{0 \leq \lambda \leq 1} \lambda b_{r+1} + P(b_1, \dots, b_r).$$

Zur Volumenberechnung nutzen wir folgende Prinzipien:

### 6.4.2 Prinzipien der Volumenberechnung:

A) Cavalierisches Prinzip (Bonaventura Cavalieri, 1591 - 1647, Bologna): Gebilde der Ebene bzw. des Raums sind inhaltsgleich, wenn in gleicher Höhe geführte Schnitte gleiche Strecken bzw. gleiche Flächen ergeben.
B) Induktiv soll gelten: $vol(P(b_1)) = \|b_1\|$, und wenn $b_{r+1} \perp Spann(b_1, \dots, b_r)$, dann ist $vol(P(b_1, \dots, b_{r+1})) = \|b_{r+1}\| vol(P(b_1, \dots b_r))$.
Insbesondere folgt hier, daß das von einer Orthonormalbasis aufgespannte Parallelepiped, also der Würfel, das Volumen 1 hat.
C) Das Volumen ist translationsinvariant, d.h. $vol(x+P(b_1, \dots, b_r)) = vol(P(b_1, \dots, b_r))$.

Bemerkung: Sei $B$ eine Orthonormalbasis von $V$, dann ist die Koordinatenabbildung $x \mapsto [x]_B \in \mathbf{R}^{n \times 1}$ eine Isometrie auf den Standardraum, wobei die Basis $B$ in die Standardbasis $S$ übergeht. Wir fordern, daß diese Abbildung das Volumen invariant läßt.

Aufgrund der Prinzipien erhalten wir den

### 6.4.3 Satz:

Sei $U = Spann(b_1, \dots, b_r)$ und sei $b'_{r+1} = p_{U^\perp}(b_{r+1})$ die orthogonale Projektion von $b_{r+1}$ auf $U^\perp$. Dann gilt:

1. $b_{r+1} + U = b'_{r+1} + U$ ist derselbe affine Unterraum von $V$,
2. $\|b'_{r+1}\| = d(b_{r+1}, U)$,
3. $vol(P(b_1, \ldots, b_{r+1})) = vol(P(b_1, \ldots, b'_{r+1})) = \|b'_{r+1}\| vol(P(b_1, \ldots, b_r))$.

Beweis: Wir wissen $V = U \oplus U^\perp$, also

$$b_{r+1} = p_U(b_{r+1}) + p_{U^\perp}(b_{r+1}).$$

Wegen $p_U(b_{r+1}) \in U$ folgt (2.5.4)

$$b_{r+1} + U = b'_{r+1} + U$$

und auch

$$\lambda b_{r+1} + U = \lambda b'_{r+1} + U.$$

Für ein $\mu \neq \lambda$ ist dann

$$(\lambda b_{r+1} + U) \cap (\mu b_{r+1} + U) = \emptyset,$$

weil $(\lambda - \mu) b_{r+1} \notin U$ ist.
Dann ist $(x + U) \cap P(b_1, \ldots, b_{r+1}) = \lambda b_{r+1} + P(b_1, \ldots b_r)$, falls $x + U = \lambda b_{r+1} + U$ ist, mit einem $\lambda \in [0, 1]$. Ansonsten ist der Durchschnitt leer. Entsprechendes gilt für $b'_{r+1}$.
Wegen $\lambda b_{r+1} + U = \lambda b'_{r+1} + U$ folgt

$$(x+U) \cap P(b_1, \ldots, b_{r+1}) = \lambda b_{r+1} + P(b_1, \ldots b_r) \ \Leftrightarrow\ (x+U) \cap P(b_1, \ldots, b'_{r+1}) = \lambda b'_{r+1} + P(b_1, \ldots b_r),$$

die beiden rechten Seiten haben wegen C) dasselbe Volumen. Wir können also A) und B) anwenden:

$$vol(P(b_1, \ldots, b_{r+1})) = vol(P(b_1, \ldots b'_{r+1})) = \|b'_{r+1}\| vol(P(b_1, \ldots, b_r)).$$

Damit ist 3. bewiesen.
Schliesslich gilt

$$\|b'_{r+1}\| = \|p_{U^\perp}(b_{r+1})\| = \|b_{r+1} - p_U(b_{r+1})\|.$$

□

Daraus ergibt sich nun der

### 6.4.4 Hauptsatz:

Seien $b_1, \ldots, b_r$ linear unabhängige Vektoren im euklidischen Vektorraum $V$. Dann ist

$$vol(P(b_1, \ldots, b_r)) = (det(G_{b_1,\ldots,b_r})^{1/2},$$

wobei $G_{b_1,\ldots,b_r} = (\langle b_i, b_j \rangle)$ die „eingeschränkte“ Gram-Matrix bezeichnet (deren Determinante ist immer positiv). Wenn insbesondere $B$ eine Basis von $V$ ist, so ist $vol(P(B)) = (det(G_B))^{1/2}$ das Volumen der zugehörigen Fundamentalmasche.

Bemerkung: 1. Achtung: Die Zahlenwerte beziehen sich immer auf den Einheitswürfel mit dem Volumen $= 1$.

2. Das Volumen ist immer im Rahmen der betrachteten Dimension zu verstehen: eine Fläche von $5m^2$ hat als Volumen $0m^3$.

Beweis: Induktion über $r$.
Für $r = 1$ ist $G_{b_1} = (\langle b_1, b_1\rangle) = \|b_1\|^2$ positiv und $\sqrt{G_{b_1}} = \|b_1\| = vol(P(b_1))$ gemäß b).
Induktionsschritt $r > 1$: Nach 6.4.3 wissen wir schon

$$vol(P(b_1, \dots b_{r+1})) = vol(P(b_1, \dots, b_r))d(b_{r+1}, U),$$

wobei $U = Spann(b_1, \dots, b_r)$ ist. Nach der Abstandsformel aus 6.2.13 ist dann

$$\star \qquad d(b_{r+1}, U)^2 = \|b_{r+1}\|^2 - s_{B_r}(b_{r+1})^t G_{B_r}^{-1} s_{B_r}(b_{r+1})$$

mit $s_{B_r}(b_{r+1}) = \begin{pmatrix} \langle b_1, b_{r+1}\rangle \\ \vdots \\ \langle b_r, b_{r+1} \end{pmatrix} =: q.$

Wir betrachten nun die Matrix $A = \begin{pmatrix} I_r & G_{B_r}^{-1}q \\ 0 & 1 \end{pmatrix}$, dann ist (unter Beachtung, daß $G_{B_r}$ symmetrisch ist)

$$A^t \begin{pmatrix} G_{B_r} & 0 \\ 0 & d^2 \end{pmatrix} A = \begin{pmatrix} G_{B_r} & q \\ q & q^t G_{B_r}^{-1} q + d^2 \end{pmatrix} = G_{B_{r+1}},$$

wegen $\star$.
Also ist

$$A^t \begin{pmatrix} G_{B_r} & 0 \\ 0 & d^2 \end{pmatrix} A = G_{B_{r+1}}.$$

Wir berechnen die Determinante: Wegen $det(A) = 1$ ist

$$det(G_{B_{r+1}}) = det(G_{B_r})d^2,$$

aus der Induktionsvorauusetzung folgt, daß diese Zahl positiv ist, und

$$\sqrt{det(G_{B_{r+1})}} = \sqrt{det(G_{B_r})}d = vol(P(b_1, \dots, b_r))d(b_{r+1}, U) = vol(P(b_1, \dots, b_{r+1})).$$

□

### 6.4.5 Folgerung:

Es seien $B, C$ zwei Systeme linear unabhängiger Vektoren in $V$ und $Spann(B) = Spann(C) = W$. Dann gibt es zwischen ihnen eine Übergangsmatrix $[1]_{C,B}$ und für die Paralleotope gilt

$$vol(P(B)) = vol(P(C))) \mid det[1]_{C,B} \mid .$$

Wenn also $C$ eine Orthonormalbasis von $W$ ist, so gilt $vol(P(B))) =\mid det[1]_{C,B} \mid$ .

Zum Beweis brauchen wir die

### 6.4.6 Transformationsformel für Gram-Matrizen (vgl. 6.1.3)

$$G_B = [1]_{C,B}^t G_C [1]_{C,B}$$

Beweis: Es ist

$$\langle x, y\rangle = [x]_B^t G_B [y]_B = [x]_C^t G_C [y]_C,$$

andererseits ist $[y]_C = [1]_{C,B}[y]_B$ und dasselbe gilt für $x$. Also ist

$$[x]_C^t G_C [y]_C = [x]_B^t([1]_{C,B}^t G_C [1]_{C,B})[y]_B$$

für alle $x, y$. Daraus folgt die Behauptung. □

Wir beweisen nun 6.4.5: Wenn $B, C$ zwei Basen von $W \subseteq V$ mit der Übergangsmatrix $Q$ sind,so ist $G_B = Q^t G_C Q$, also $\sqrt{det(G_B)} = \sqrt{det(G_C)} \mid det(Q) \mid$. □

### 6.4.7 Folgerung:

Sei $B$ eine Basis von $V$ und $f : V \longrightarrow V$ eine lineare Abbildung. Dann gilt
1.$f(P(B)) = P(f(B))$,
2. $vol(f(P(B))) = vol(P(B)) \mid det(f) \mid$.

### 6.4.8 Erklärung: Die Determinante einer linearen Abbildung

Man nehme irgendeine Basis $B$ von $V$ und deren Koordinatenmatrix $[f]_B$; wir definieren

$$det(f) = det([f]_B),$$

dies ist unabhängig von der Wahl von $B$, denn für eine andere Basis $C$ gilt

$$[f]_C = [1]_{C,B}[f]_B[1]_{B,C} = P^{-1}[f]_B P,$$

also $det[f]_C = det[f]_B$.
Wir beweisen nun 6.4.7:
1. folgt aus der Linearität von $f$.
2. Aus 6.4.5 folgt

$$vol(f(P(B))) = vol(P(f(B))) = vol(P(B)) \mid det[1]_{B,f(B)} \mid.$$

Es ist nun aber $[1]_{B,f(B)} = [f]_B$, denn die $i$-te Spalte von $[1]_{B,C}$ ist $[c_i]_B$ und die $i$-te Spalte von $[f]_B = [f(b_i)]_B$. Jetzt nehmen wir $C = f(B)$, d.h. $c_i = f(b_i)$. daraus folgt die Behauptung. □

Bemerkung: 6.4.7 ist auch richtig, wenn $f$ einen Defekt hat. Dann folgt $f(P(B)) = P(f(B))$ mit kleineren Dimensionen. Die Fundamentalmasche $P(B)$ wird unter $f$ „plattgedrückt“, also $vol(f(P(B))) = 0$ im $n$-dimensionalen Sinn. Aber hier ist $det(f) = 0$ und die Formel stimmt.

**Anwendung von 6.4.7 in der Integralrechnung**

In $V$ betrachten wir jetzt Parallelotope der Form $v+P(B)$, wobei $B$ eine Basis und $v$ ein Translationsvektor ist. Solche Parallelotope heißen nicht-überlappend, wenn sie keine gemeinsamen inneren Punkte haben. Eine Intervallsumme $I\sum = \bigcup P_i$ ist eine endliche Vereinigung nichtüberlappender Parallelotope. Sie hat das Volumen $\mid I\sum \mid = \sum \mid P_i \mid$. Eine Menge $M$ heißt quadrierbar, falls ihr inneres und äußeres Maß übereinstimmen, d.h.

$$\mid M \mid_i = sup\{\mid I\sum \mid; I\sum \subseteq M\} = \mid M \mid_a = inf\{\mid I\sum \mid; I\sum \supseteq M\}.$$

Ohne Schwierigkeiten übertragen wir nun 6.4.7 auf Intervallsummen und quadrierbare Mengen:

### 6.4.9 Satz:

Ist $f : V \longrightarrow V$ eine lineare Abbildung und $P \subset V$ ein beliebiges Parallelotop, dann gilt $\mid f(P) \mid = \mid P \mid\mid det(f) \mid$. Ebenso gilt für jede quadrierbare Menge $\mid f(M) \mid = \mid M \mid\mid det(f) \mid$.

Beweis: Wenn $Q = v + P(B)$ ist, so ist $f(Q) = f(v) + f(P(B))$, daher $\mid f(Q) \mid = \mid f(P(B)) \mid = \mid P(B) \mid\mid det(f) \mid = \mid v + P(B) \mid\mid det(f) \mid = \mid P \mid\mid det(f) \mid$.
Für die Intevallsumme $I\sum = \bigcup P_i$ folgt $f(I\sum) = \bigcup f(P_i)$. Wenn $f$ keinen Defekt hat, dann ist $f(I\sum)$ wieder eine Intervallsumme und

$$\mid f(I\sum) \mid = \sum \mid f(P_i) \mid = (\sum \mid P_i \mid) \mid det(f) \mid = \mid I\sum \mid\mid det(f) \mid .$$

Durch Limesbildung überträgt sich dies von Intervallsummen auf quadrierbare Mengen.

Ab jetzt ist $V = \mathbf{R}^n$ der euklidische Standardraum.

### 6.4.10 Variante; Variablentransformation in Integralen

Ist $(x_1, \ldots, x_n) = (\phi_1(u_1, \ldots, u_n), \ldots, \phi_n(u_1, \ldots, u_n))$ eine differenzierbare Variablentransformation des $\mathbf{R}^n$, dann gilt für die integrierbare Funktion $f : \mathbf{R}^n \longrightarrow \mathbf{R}$ und eine quadrierbare Menge $M \subseteq \mathbf{R}^n$ die Transformationsgleichung

$$\int_{x\in\phi(M)} f(x)dx = \int_{u\in M} f(\phi(u)) \mid det(\frac{\partial\phi_i}{\partial u_i})du.$$

Beweis: Differenzierbare Funktionen sind „im Kleinen" linear. Es ist

$$\int_{M'} f(x)dx = \lim \sum_{P\in I\sum} f(x_p)vol(P),$$

wobei die Intervallsumme $I\sum$ die quadrierbare Menge $M'$ approximiert und $x_p \in P$ irgendein Vertreter in dem Volumenelement $P$ ist. Weil $x = \phi(u)$ eine differenzierbare Varialblentransformation ist, so gilt in der Nähe irgendeines ausgewählen Punkts $x^* = \phi(u^*)$ die Matrixgleichung

$$x - x^* \sim A_{u^*}(u - u^*),$$

wobei $A = \left(\frac{\partial\phi_i}{\partial u_j}\mid_{u=u^*}\right)$ die Jacobische Matrix von $\phi$ im Punkt $u^*$ ist. Wir bezeichnen mit $\phi_{u^*}$ die lineare Abbildung $x \mapsto A_{u^*}x$.
In einer kleinen Umgebung von $u^*$ verhält sich die differenzierbare Abbildung $\phi$ wie die lineare Abbildung $\phi_{u^*}$ und für ein kleines Parallelotop $u^* \in P$ gilt $\phi(P) \sim \phi_{u^*}(P)$. Für das Volumen folgt

$$\mid \phi(P) \mid\sim\mid \phi_{u^*}(P) \mid=\mid det(\phi_{u^*}) \mid\mid P \mid .$$

Jetzt betrachten wir $P = du$ als infinitesimales Volumenelement, das $u^*$ enthält und entsprechend $\phi(P) = dx$ als infinitesimales Volumenelement, das $x^* = \phi(u^*)$ enthält. Dann folgt $dx =\mid det(A_{u^*}) \mid du$. Jetzt lassen wir den Punkt $x^*$ wandern und beachten, daß $x, x^*$ bzw. $u, u^*$ stets infinitesimal benachbart sind. Dann gilt

$$dx = d\phi(u) =\mid det(\frac{\partial\phi_i}{\partial u_j}) \mid du,$$

wobei die Determinante jetzt ebenfalls als Funktion von $u_1, \ldots, u_n$ aufzufassen ist (die approximierende lineare Abbildung ändert sich von Punkt zu Punkt). Damit folgt schließlich

$$\int_{x\in\phi(M)} f(x)dx = \int_{u\in M} f(\phi(u))d\phi(u) = \int_{u\in M} f(\phi(u)) \mid \frac{\partial\phi_i}{\partial u_j} \mid du.$$

□

Bemerkung: Umgekehrt kommt man von 6.4.10 zurück zu 6.4.9, wenn man eine lineare Abbildung $\phi$ betrachtet, die durch eine Matrix $A$ gegeben ist. Dann ist $\left(\frac{\partial\phi_i}{\partial u_j}\right) = A$, da die $\phi_i$ linear in den $u_j$ sind.
Für $f$ nimmt man dann die charakteristische Funktion von $M$ mit dem Wert 1 in $M$ und 0 außerhalb von $M$. Aus der Integralformel wird dann einfach die Formel $\mid \phi(M) \mid=\mid M \mid\mid det(\phi) \mid$ .

## 6.5 Linearformen

Zunächst sei $V$ ein beliebiger $K$-Vektorraum.

### 6.5.1 Definition:

Eine Linearform auf dem $K$-Vektorraum $V$ ist eine lineare Abbildung $\phi : V \longrightarrow K$.

### 6.5.2 Ausrechnen einer Linearform für endlichdimensionale Vektorräume

Sei $\phi$ eine Linearform auf $V$ und $B = (b_1, \ldots, b_n)$ eine Basis von $V$. Für einen Vektor $v = \sum x_i b_i$ gilt dann

$$\phi(v) = \sum \phi(b_i)x_i = (\phi(b_1), \ldots, \phi(b_n))[v]_B,$$

der Zeilenvektor ist einfach die Koordinatenmatrix der Abbildung $\phi$ bezüglich der Basen $B$ von $V$ und 1 von $K$.

### 6.5.3 Satz:

1. Die Linearformen auf $V$ bilden einen Vektorrau $V^* = Hom_K(V, k)$, die Nullabbildung ist der Nullvektor.
2. Wenn $V$ endlichdimensional und $B = (b_1, \ldots, b_n)$ eine Basis ist, so ist

$$\phi \in V^* \mapsto [\phi]_{1,B} \in K^{1\times n}$$

ein Isomorphismus, insbesondere ist $dim(V^*) = n = dim(V)$.
3. Wir betrachten die Linearformen $b_i^*$ mit $b_i^*(b_j) = \delta_{ij}$, d.h. $[b_i^*]_{1,B} = e_i$ ist der Einheitsvektor. Dann ist $B^* = (b_1^*, \ldots b_n^*)$ eine Basis von $V^*$, sie heißt die zu $B$ duale Basis.

4. Sei $\phi \in V^*, \phi \neq 0$, dann ist $Ker(\phi) \subset V$ ein Unterraum der Kodimension 1. Für zwei Linearformen $\phi,\ \psi$ gilt $Ker(\phi) = Ker(\psi)$ genau dann, wenn $\phi = \lambda\psi$ mit $\lambda \in K,\ \lambda \neq 0$. Jeder Unterraum $W \subset V$ der Kodimension 1 läßt sich als Kern einer Linearform realisieren und die Zuordnung
{Unterräume der Dimension 1 von $V^*$} $\leftrightarrow$ {Unterräume der Dimension n-1 von $V$}
ist bijektiv.

Beweis: 1. ist klar.
2. folgt aus 5.6.7.
3. ist klar.
4. Sei $\phi$ eine nichtriviale Linearform, diese ist dann surjektiv, denn ihr Bild ist ein nichttrivialer Unterraum von $K$. Aus 5.5.1 folgt $dim(Ker(\phi)) = dim(V) - 1$. Offensichtlich ist $Ker(\lambda\phi) = Ker(\phi)$.
Umgekehrt: Sei $Ker(\phi) = Ker(\psi) = W$. Dann findet sich ein Vektor $b$ mit $V = W \oplus Kb$ und $\phi(b)$ sowie $\psi(b)$ sind beide nicht null. Also ist $\lambda = \frac{\phi(b)}{\psi(b)}$.
Sei schließlich $W$ gegeben und $V = W \oplus Kb$. Wir definieren $\phi(w + \lambda b) = \lambda$, dies ist eine Linearform mit $Ker(\phi) = W$. $\square$

Beispiel: Wenn $V = K^{n\times 1}$ der $n$-dimensionale Spaltenraum ist, so kann man $V^*$ mit dem $n$-dimensionalen Zeilenraum $K^{1\times n}$ identifizieren.

### 6.5.4 Der Gradient einer Linearform

Jetzt sei $V$ ein euklidischer Vektorraum, dann hat man eine Abbildung

$$V \longrightarrow V^*,\ v \mapsto \phi_v = \{x \mapsto \langle x, v\rangle\}.$$

Diese Abbildung ist linear, injektiv und auch surjektiv, wenn $V$ endlichdimensional ist. Wir setzen nun $dim(V) < \infty$ voraus, die obige Abbildung ist dann ein Isomorphismus. Es gibt also zu jeder Linearform $\phi \in V^*$ einen eindeutig bestimmten Vektor $v \in V$ mit $\phi = \phi_v$. Man nennt diesen Vektor $v = grad(\phi)$ den Gradienten von $\phi$. Es gilt

$$grad(\phi) \perp Ker(\phi),\ \|grad(\phi)\|^2 = \phi(grad(\phi)).$$

Beweis: Die Linearität der Abbildung $v \mapsto \phi_v$ ist klar, die Injektivität folgt aus der positiven Definitheit des Skalarprodukts und die Surjektivität aus Dimensionsgründen.

Aus $\phi(x) = \langle grad(\phi), x\rangle$ folgt $\phi(x) = 0$ gdw. $grad(\phi) \perp x$ und schließlich ist $\phi(grad(\phi)) = \langle grad(\phi), grad(\phi)\rangle$. □

Bemerkung: Auf dem zu $Ker(\phi)$ parallelen affinen Unteraum $grad(\phi) + Ker(\phi)$ nimmt $\phi$ den positiven Wert $\|grad(\phi)\|^2$ an.
Auf den zu $Ker(\phi)$ parallenen affinen Räumen $x + Ker(\phi)$ hat $\phi$ den konstanten Wert $\phi(x)$. Diese Unterräume heißen die Niveauebenen von $\phi$.

Der Vektor $v = grad(\phi)$ kann als der Vektor charakterisiert werden, für den gilt:

$$v \perp Ker(\phi),\ \phi(v) = \|v\|^2,\ v \neq o.$$

Beispiel: Sei $V = \mathbf{R}^{n\times 1}$ mit der Standardmetrik versehen. Ein Zeilenvektor, aufgefaßt als Koordinatenmatrix einer Linearform, hat als Gradienten den entsprechenden Spaltenvektor.

## 6.6 Selbstadjungierte und orthogonale Transformationen eines euklidischen Raums

Lineare Operatoren auf einem Vektorraum $V$ sind lineare Abbildungen $f : V \longrightarrow V$. Andere Namen sind „lineare Transformationen", „lineare Selbstabbildungen" oder „Endomorphismen" . Die Menge $End_{\mathbf{R}}V$ aller linearen Abbildungen von $V$ in sich ist ein $\mathbf{R}$-Vektorraum.

Sei nun $V/\mathbf{R}, \langle\ ,\ \rangle$ ein euklidischer Raum.

### 6.6.1 Satz:

A) Zu jedem linearen Operator $f \in End(V)$ gibt es einen eindeutig bestimmten Operator $f^* \in End(V)$, welcher durch die folgenden äquivalenten Eigenschaften bestimmt ist:

$$(1) \qquad \langle x, f(y)\rangle = \langle f^*(x), y\rangle$$

$$(2) \qquad \langle f(x), y\rangle = \langle x, f^*(y)\rangle$$

für alle $x, y \in V$.
Der Operator $f^*$ heißt der zu $f$ adjungierte Operator.
Zu fixiertem $x \in V$ ist $f^*(x)$ als Gradient der Linearform $y \mapsto \langle x, f(y)\rangle$ bestimmt.

B) Wenn $B$ eine Orthonormalbasis von $V$ ist, dann gilt für die Koordinatenmatrizen von $f$ unf $f^*$ die Beziehung:

$$[f^*]_B = ([f]_B)^t.$$

C) Ist $V = \mathbf{R}^{n\times 1}$ der euklidische Raum mit dem Standardprodukt $\langle x, y\rangle_I = x^t y$ und ist $A \in \mathbf{R}^{n\times n}$ eine Matrix, die wir als linearen Operator auf $\mathbf{R}^{n\times 1}$ auffassen, dann gilt $A^* = A^t$, der adjungierte Operator ist in diesem Fall dasselbe wie die transponierte Matrix.

Beweis: A) Es sei $f$ gegeben, $x \in V$ sei fixiert. Dann ist

$$\phi : y \mapsto \langle x, f(y) \rangle$$

eine Linearform auf $V$, denn

$$\lambda y_1 + \mu y_2 \mapsto \langle x, f(\lambda y_1 + \mu y_2) \rangle = \lambda \langle x, f(y_1) \rangle + \mu \langle x, f(y_2) \rangle = \lambda \phi(y_1) + \mu \phi(y_2).$$

Wir setzen $f^*(x) = grad(\phi)$, dann ist $x \mapsto f^*(x)$ eine lineare Abbildung mit den Eigenschaften (1), (2).

B) Sei $B$ eine Orthonormalbasis von $V$, dann haben wir bereits gesehen (6.2.7):

$$\langle x, y \rangle = \langle [x]_B, [y]_B \rangle_I,$$

dem Skalarprodukt auf $V$ entspricht das Standardprodukt auf $\mathbf{R}^{n \times 1}$. Deshalb folgt aus 5.6.4

$$\langle x, f(y) \rangle = \langle [x]_B, [f(y)]_B \rangle_I = \langle [x]_B, [f]_B \cdot [y]_B \rangle_I,$$

$$\langle f^*(x), y \rangle = \langle [f^*(x)]_B, y]_B \rangle_I = \langle [f^*]_B \cdot [x]_B, [y]_B \rangle_I.$$

Die linken Seiten sind gleich, also auch die rechten, folglich gilt nach Definition des Standardprodukts

$$[x]_B^t [f]_B [y]_B = ([f^*]_B [x]_B)^t [y] B = [x]_B^t ([f^*]_B)^t [y]_B.$$

Da diese Gleichung für alle $[x]_B, [y]_B$ gilt, folgt $[f]_B = ([f^*]_B)^t$.

C) Sei $\mathbf{R}^{n \times 1}, \langle \ , \ \rangle$ der Standardraum, dann gilt

$$\langle x, Ay \rangle = x^t A y = x^t (A^t)^t y = (A^t x)^t y = \langle A^t x, y \rangle.$$

Wenn wir $A$ und $A^t$ als lineare Operatoren auf dem Spaltenraum auffassen, dann ist $A^t$ zu $A$ adjungiert. □

Merke: Die Bildung des adjungierten Operators ist eine Verallgemeinerung der Bildung der transponierten Matrix. Wenn man bezüglich einer Orthonormalbasis vom euklidischen Raum zum Standardraum übergeht, dann gehen beide Bildungen ineinander über.

Vorsicht: Früher (im Kapitel 3) hatten wir den Begiff $A^\# = C(A)^t$, dies ist die Transponierte der Kofaktormatrix, diese wird manchmal auch als die zu $A$ adjungierte Matrix (oder Adjunktenmatrix[3] von $A$) bezeichnet; das ist aber etwas anderes.

[3] Ein Adjunkt ist im Behördendeutsch des 19. Jh. ein Stellvertreter oder Begleiter eines Beamten. Das englische Verb *to adjoin* heißt anstoßen, was ja hier auch Sinn macht (s.u)

### 6.6.2 Definition

Ein linearer Operator $f \in End_{\mathbf{R}}(V)$ heißt
a) selbstadjungiert, wenn $f^* = f$ ist, d.h. $\langle x, f(x)\rangle = \langle f(x), y\rangle$,
b) orthogonale Transformation, falls $f^* \circ f = id_V$ ist, d.h. $\langle f(x), f(y)\rangle = \langle x, y\rangle$.
Es folgt, daß $f^*$ surjektiv und $f$ injektiv ist. Weil $dimV < \infty$ ist, sind beide bijektiv, also gilt $f^* f f^* = f^*$ und damit auch $f \circ f^* = id_V$.

Für den Standardraum ist ein selbstadjungierter Operator dasselbe wie eine symmetrische Matrix. Eine orthogonale Transformation ist dasselbe wie eine orthogonale Matrix, d.h. $AA^t = I_n = A^t A$.

Wir zeigen die Äquivalenz der unter b) angegebenen Bedingungen:
Wir setzen in $\langle x, f(y)\rangle = \langle f^*(x), y\rangle$ speziell $x = f(z)$ und erhalten $\langle f(z), f(y)\rangle = \langle f^*(f(z)), y\rangle = \langle z, y\rangle$.
Sei umgekehrt $\langle f(x), f(y)\rangle = \langle x, y\rangle$, dann folgt $\langle f(x), f(y)\rangle = \langle x, f^*(f(y))\rangle$, also müssen die rechten Seiten für alle $x, y$ gleich sein, also $f^* f = id_V$.

### 6.6.3 Anschauliche Vorstellungen über selbstadjungierte bzw. orthogonale Operatoren

(Wir nehmen hier einige Ergebnisse vorweg.)
Sei $f \in End(V)$ und $P(B)$ eine Fundamentalmasche zu einer Orthonormalbasis $B$ von $V$; d.h. $P(B)$ ist ein Würfel, die Seiten haben die Länge 1 und die Kanten an einer Ecke stehen paarweise senkrecht aufeinander. Die lineare Abbildung $f$ ist dadurch charakterisiert, wie sie auf $P(B)$ wirkt.

a) Wenn $f$ selbstadjungiert ist, so existiert eine Orthonormalbasis $B = (b_1, \ldots, b_n)$, wo $f$ eine Veränderung der Seitenlängen bewirkt, d.h.$B$ geht in $(\lambda_1 b_1, \ldots, \lambda_n b_n)$ über. Dabei können einige der Faktoren Null sein, d.h. die entsprechende Dimension wird zusammengedrückt.

Ein typisches Beispiel ist eine orthogonale Projektion $p$, dann ist $V = Ker(p) \oplus Bild(p)$ eine orthogonale Zerlegung. Wir finden also Orthonormalbasen $b_1, \ldots, b_r$ von $Bild(p)$ und $b_{r+1}, \ldots b_n$ von $Ker(p)$. Wegen $Ker(p) \perp Bild(p)$ ist $b_1, \ldots b_n$ eine Orthonormalbasis von $V$ und $p(b_1) = b_1, \ldots p(b_r) = b_r, p(b_{r+1}) = \ldots = p(b_n) = o$. Hier sind also $\lambda_1 = \ldots = \lambda_r = 1, \lambda_{r+1} = \ldots \lambda_n = 0$.

Außerdem bilden die selbstadjungierten Operatoren auf $V$ einen Vektorraum, d.h. man kann Linearkombinationen selbstadjungierter Operatoren bilden und erhält wieder einen solchen.

b) Wenn $f$ ein orthogonaler Operator ist, dann erhält die Anwendung von $f$ alle Längen (alle Abstände) und alle Winkel. Wenn also $P(B)$ ein Würfel ist, so ist $f(P(B)) = P(f(B))$ wieder ein Würfel, d.h. $f(B)$ ist wieder eine Orthonormalbasis. Anschaulich ist deshalb $f$ eine Spiegelung an einer Hyperebene oder eine Drehung um eine Achse durch den Nullpunkt oder eine Kombination solcher Abbildungen.

Die orthogonalen Operatoren bilden keinen Vektorraum, aber bezüglich der Hintereinanderausführung von Abbildungen eine Gruppe: wenn $f, g$ orthogonal sind, so sind $fg$ und $f^{-1}$ ebenfalls orthogonal.

### 6.6.4 Satz:

(1) Die Abbildung $f \mapsto f^*$, die einem linearen Operator die dazu adjungierte Abbildung zuordnet, ist linear.
(2) Die selbstadjungierten Operatoren bilden einen Vektorraum.
(3) Jeder orthogonale Projektor ist selbstadjungiert, also ist auch jede Linearkombination orthogonaler Projektoren selbstadjungiert.

Beweis: Zu (1) wollen wir nur bemerken, daß für quadratische Matrizen die Zuordnung $A \mapsto A^t$ linear ist.
(2) Die selbstadjungierten Operatoren sind die Fixpunkte der linearen Abbildung $f \mapsto f^*$, diese bilden also einen Unterraum von $End(V)$.
(3) Sei $p : V \longrightarrow V$ ein Ortogonaler Projektor mit der orthogonalen Zerlegung

$$V = Ker(p) \oplus Bild(p),$$

$$x = x_1 + x_2,$$

$$y = y_1 + y_2.$$

Dann gilt

$$\langle p(x), y\rangle = \langle x_2, y\rangle = \underbrace{\langle x_2, y_1\rangle}_{=0} + \langle x_2, y_2\rangle = \langle p(x), p(y)\rangle,$$

$$\langle x, p(y)\rangle = \langle x, y_2\rangle = \underbrace{\langle x_1, y_2\rangle}_{=0} + \langle x_2, y_2\rangle = \langle p(x), p(y)\rangle.$$

Insgesamt haben wir also

$$\langle p(x), y\rangle = \langle p(x), p(y)\rangle = \langle x, p(y)\rangle.$$

□Achtung: Ein orthogonaler Projektor ist kein orthogonaler Operator! Denn $p^*p = pp = p$ ist nicht die Identität.

Auf Matrizen übertragen heißt 6.6.5(1): Die symmetrischen Matrizen bilden einen Vektorraum.

Später werden wir sehen, daß jeder selbstadjungierte Operator eine Linearkombination orthogonaler Projektoren ist (Spektralsatz).

### 6.6.5 Satz:

(1) Orthogonale Operatoren erhalten Längen, Abstände und Winkel.
(2) Die orthogonalen Operatoren bilden bezüglich der Hintereinanderausführung eine Gruppe.
(3) Die Determinante eines orthogonalen Operators ist $\pm 1$.

Beweis: (1) Wegen 6.6.6.b folgt

$$\|f(x)\|^2 = \langle f(x), f(x)\rangle = \langle x, x\rangle = \|x\|^2$$

sofort $\|f(x)\| = \|x\|$. Dann ist auch

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

also stimmen beide Winkel überein. Insbesondere folgt auch

$$d(f(x), f(y)) = \|f(x) - f(y)\| = \|f(x - y)\| = \|x - y\| = d(x, y).$$

(2) Ein orthogonaler Operator erhält die Abstände, ist also injektiv. Nch 5.5.1 hat dann $Bild(f)$ dieselbe Dimension wie $V$, also ist $f$ bijektiv. Also hat $f$ eine Inverse, die ebenfalls linear ist. Wir substituieren in

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

$x = f^{-1}(x'), y = f^{-1}y'$ und erhalten

$$\langle x', y')\rangle = \langle f^{-1}(x'), f^{-1}(y')\rangle,$$

also ist auch $f^{-1}$ orthogonal.
Wenn $f, g$ orthogonal sind so auch $fg$ und die Identität ist ebenfalls orthogonal, also bilden diese Abbildungen eine Gruppe.

(3) Es sei $P(B)$ der Einheitswürfel, dann ist $f(P(B)) = P(f(B))$ wieder ein Einheitswürfel, also $1 = vol(f(P(B))) = vol(P(B))\cdot \mid det(f) \mid$, also $\mid det(f) \mid= 1$ und da $[f]_B$ reell ist folgt $det(f) = \pm 1$. □

### 6.6.6 Orthogonale Transformationen im Standardfall

Jetzt betrachten wir den euklidischen Standardraum $\mathbf{R}^{n\times 1}$.
Eine Matrix$A \in \mathbf{R}^{n\times n}$, aufgefaßt als lineare Abbildung, ist eine orthogonale Transformation, wenn die folgenden äquivalenten Bedingungen erfüllt sind:

$$(1) \qquad A^tA = I,$$

$$(2) \qquad \langle Ax, Ay\rangle = \langle x, y\rangle,$$

(3) die Spaltenvektoren von $A$ bilden eine Orthonormalbasis von $\mathbf{R}^{n\times 1}$,

(4) die Zeilenvektoren bilden eine Orthonormalbasis von $\mathbf{R}^{1\times n}$.

Die orthogonalen Matrizen bilden eine Untergruppe $O_n(\mathbf{R}) \subset GL_N(\mathbf{R})$.

Beweis: Die Äquivalenz von (1) und (2) ist schon klar.
Außerdem wissen wir: $A^tA$ ist die Gram-Matrix der Spaltenvektoren von $A$ bezüglich des Standardprodukts; $AA^t$ ist die Gram-Matrix der Zeilenvektoren von $A$. Daher ist

$A^tA = I$ genau dann, wenn die Spaltenvektoren eine Orthonormalbasis bilden, und $AA^t = I$ genau dann, wenn die Zeilenvektoren eine Orthonormalbasis bilden. □

**Der Fall** $n = 2$

Alle $2 \times 2$-Matrizen mit diesen Eigenschaften sind von der Form

$$A = \begin{pmatrix} cos(\alpha) & -sin(\alpha) \\ sin(\alpha) & cos(\alpha) \end{pmatrix} \text{ oder } B = \begin{pmatrix} cos(\alpha) & sin(\alpha) \\ sin(\alpha) & -cos(\alpha) \end{pmatrix}.$$

Die zur Matrix $A$ gehörige Abbildung ist die Drehung der Ebene um den Winkel $\alpha$ entgegen dem Uhrzeigersinn.
Zur Matrix $B$ gehört die Spiegelung an der Achse mit dem Anstiegswinkel $\alpha/2$.
Nach Abschnitt 5.4 muss man dazu nur prüfen, wie die Abbildungen auf die Standardbasis wirken. Diese Bilder sind gerade die Spalten der entsprechenden Matrix. Man mache eine Zeichnung.
Die Drehungen zeichnen sich durch $det(A) = 1$ aus, für die Spiegelungen gilt $det(B) = -1$. Die orthogonale Gruppe $O_2(\mathbf{R})$ besteht aus Drehungen und Spiegelungen (sie heißt auch stetige Diedergruppe).
Also: $SO_2(\mathbf{R})$ ist die Drehgruppe und $O_2(\mathbf{R}) = SO_2(\mathbf{R}) \sqcup SO_2(\mathbf{R})B$, wobei $B$ eine beliebige Spiegelung ist.

### 6.6.7 Spiegelungen

Sei $v \neq o$ ein Vektor aus $V$. Dann zerlegt sich der Raum:

$$V = \mathbf{R}v \oplus v^\perp,$$

also gibt es für jedes $x \in V$ eine eindeutige Zerlegung $x = x_1 + x_2$ mit $x_1 = \lambda v,\ x_2 \perp v$.

**Definition:** Die Abbildung

$$(*) \qquad x = x_1 + x_2 \mapsto s_v(x) = -x_1 + x_2$$

heißt (orthogonale) Spiegelung an der Hyperebene $v^\perp$.

Offensichtlich ist die Spiegelung $s_v$ eine lineare Abbildung von $V$ in sich mit der Eigenschaft $s_v^2 = id_V$. Man nennt das eine Involution des Vektorraums; $s_v$ ist zu sich selbst invers.

Wenn $\mu v,\ \mu \neq 0$ ein Vielfaches von $v$ ist, so ist $\mathbf{R}v = \mathbf{R}(\mu v),\ v^\perp = (\mu v)^\perp$ und deshalb $s_{\mu v} = s_v$.

**Satz:**
(1) Die Spiegelungen $s_v$ sind orthogonale Transformationen und es ist $det(s_v) = -1$.
(2) Die Spiegelungen $s_v$ kann man beschreiben, ohne die Zerlegung (*) auszurechnen:

$$s_v(x) = x - 2\frac{\langle v, x\rangle}{\langle v, v\rangle}v.$$

Beweis: (1)

$\langle s_v(x), s_v(y)\rangle = \langle -x_1+x_2, -y_1+y_2\rangle = \langle -x_1, -y_1\rangle + \langle x_2, y_2\rangle = \langle x_1, y_1\rangle + \langle x_2, y_2\rangle = \langle x, y\rangle.$

Wir wählen eine Basis $B$ von $V$: $b_1 = v,\ b_2, \ldots, b_n$ sei eine Basis von $v^\perp$. Dann ist

$$[s_v]_B = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix},$$

deren Determinante ist gleich -1.

Daraus folgt auch, daß Spiegelungen selbstadjungiert sind:

$$\langle s_v(x), y\rangle = \langle s_v(x), s_v^2(y)\rangle = \langle x, s_v(y)\rangle$$

wegen (1). (2) Die Spiegelung $s_v$ ist durch folgende Eigenschaften charakterisiert:
a) $s_v$ ist eine lineare Abbildung,
b) $s_v(v) = -v$,
c) $s_v(x) = x$ für alle $x \perp v$.
Diese Eigenschaften hat aber gerade $s_v(x) = x - 2\frac{\langle v,x\rangle}{\langle v,v\rangle} v$. □
Die Koordinatenmatrix von $s_v$ bezüglich einer Orthonormalbasis $B$ von $V$ ist

$$[s_v]_B = I_n - \frac{2}{\|v\|^2}([v]_B [v]_B^t),$$

dabei hat die letzte Matrix den Rang 1.

Da $s_v$ sowohl orthogonal als auch selbstadjungiert ist, ist $[s_v]_B$ eine symmetrische orthogonale Matrix.

### 6.6.8 Drehungen

Seien $s_v, s_w$ zwei Spiegelungen. Wenn $v, w$ linear abhängig sind, so ist $s_v = s_w$, also $s_v s_w = id_V$.
Sind dagegen $v, w$ linear unabhängig, so spannen sie eine Ebene $E = \mathbf{R}v + \mathbf{R}w$ auf, und es gilt:

(1) $s_v s_w$ ist eine orthogonale Transformation, welche $E^\perp$ elementweise fest läßt. Die Determinante ist gleich 1.

(2) $E$ wird von $s_v s_w$ in sich überführt. Die Ebene wird um den Winkel $2\theta(v, w)$ gedreht, und zwar in der Richtung von $w$ nach $v$.

Beweis: Die ersten Aussagen sind klar.
(2) Es ist $E^\perp = (\mathbf{R}v)^\perp \cap (\mathbf{R}w)^\perp$. Wenn also $x \in E^\perp$ ist, so ist $s_v(x) = x$ und $s_w(x) = x$, somit $s_v s_w(x) = x$.
.................................. ???

### 6.6.9 Satz:

Jede orthogonale Transformation eines $n$-dimensionalen Raums ist das Produkt von höchstens $n$ Spiegelungen.

(ohne Beweis)

### 6.6.10 Charakterisierung der orthogonalen Matrizen als Übergangsmatrizen

(1) Sind $B, C$ zwei Orthonormalbasen des euklidischen Raums $V$, dann ist die Übergangsmatrix $[1]_{B,C}$ orthogonal.
(2) Umgekehrt: Wenn $[1]_{B,C}$ eine orthogonale Matrix ist und eine der Basen orthonormal ist, dann ist es auch die andere.

Beweis: Für die Gram-Matrizen gilt

$$(*) \qquad G_C = [1]_{B,C}^t G_B [1]_{B,C}.$$

(1) Wenn beides Orthonormalbasen sind, also $G_B = G_C = I$, so ist

$$I = [1]_{B,C}^t [1]_{B,C},$$

d.h. $[1]_{B,C}$ ist orthogonal.

(2) Sei $B$ eine Orthonormalbasis und $[1]_{B,C}$ eine orthogonale Matrix. Dann folgt aus (*) $G_C = I$, also ist $C$ orthonormal. □

### 6.6.11 Orientierung eines reellen Vektorraums

Es sei daran erinnert, daß eine Basis $B = (b_1, \dots, b_n)$ als *geordnete* Menge zu verstehen ist.
Wir nennen Basen $B, C$ gleichorientiert, wenn die Übergangsmatrix $[1]_{B,C}$ eine positive Determinante hat. Dies ist eine Äquivalenzrelation.
Wenn $B, C$ zwei Basen mit $det([1]_{B,C}) < 0$, also nicht gleichorientiert sind, dann ist jede weitere Basis $D$ entweder zu $B$ oder zu $C$ gleichorientiert. Es gibt also zwei Äquivalenzklassen gleichorientierter Basen.
Eine Orientierung im Vektorraum bedeutet also die Auszeichnung einer der Klassen als „positiv“.

Im Anschauungsraum $E$ mit dem rechtwinkligen Koordinatensystem $(x, y, z)$ nennt man das System positiv orientiert, wenn eine Rechtsdrehung mit einem in $z$-Richtung angesetzten Schraubenzieher die $x$-Achse auf die $y$-Achse hinbewegt.

# Index