Lineare Algebra und analytische Geometrie I (MIB)

Helmut Schwichtenberg

Mathematisches Institut der Ludwig-Maximilians-Universität München
Wintersemester 1997/98

Inhaltsverzeichnis

Vorwort

Dieses Skriptum gibt den Inhalt einer Vorlesung über lineare Algebra wieder, die ich im Wintersemester 1997/98 für Studienanfänger am Mathematischen Institut der Universität München gehalten habe. Quellen waren Vorlesungen meines Kollegen Hans-Jürgen Schneider sowie das Buch von Gerd Fischer [1] über lineare Algebra.

Bedanken möchte ich mich bei Ulrich Berger und Thomas Rudlof für ihre Unterstützung bei der Vorbereitung der Vorlesung, und bei Frau Gerlinde Bach für ihre Hilfe bei der Herstellung des Manuskripts. Für die gute Zusammenarbeit und viele hilfreiche Anmerkungen bedanke ich mich bei den Übungsgruppenleitern und Korrektoren Elke Fischer, Mathias Flurl, Georg Hoever, Stefanie Kuntze, Holger Wels und Christian Wittmann.

München, im Februar 1998
Helmut Schwichtenberg

Kapitel 1

Lineare Gleichungssysteme

Problemstellung: Gegeben seien Zahlen α_{ij} $(1 \leq i \leq m,\ 1 \leq j \leq n)$ und β_i $(1 \leq i \leq m)$. Gesucht sind Zahlen ξ_j $(1 \leq j \leq n)$ so daß

$$\begin{aligned} \alpha_{11}\xi_1 + \cdots + \alpha_{1n}\xi_n &= \beta_1 \\ &\vdots \\ \alpha_{m1}\xi_1 + \cdots + \alpha_{mn}\xi_n &= \beta_m \end{aligned} \tag{1.1}$$

(1.1) heißt *lineares Gleichungssystem*. Welche Eigenschaften von "Zahlen" sind für das Problem relevant? Sicherlich Addition, Subtraktion, Multiplikation, Division und Rechenregeln dafür, aber nicht z.B. die Quadratwurzel oder Grenzwertbildung. Deshalb wollen wir zunächst Zahlenbereiche festlegen, für die die Frage nach der Lösbarkeit von (1.1) sinnvoll ist, nämlich sogenannte *Körper*. Körper sind am einfachsten zu behandeln unter Verwendung von *Gruppen*, die wir deshalb zunächst einführen. Zu allererst aber einige einführende Betrachtungen über *Mengen*.

1.1 Mengen

Die Mengenlehre kann man als einen Rahmen auffassen, innerhalb dessen man Mathematik begründen und betreiben kann. Wir wollen die Mengenlehre hier informal (also nicht als eine formale Theorie innerhalb der mathematischen Logik) entwickeln. Zunächst ist es deshalb notwendig, sich eine inhaltliche Vorstellung des durch die Axiome zu beschreibenden Mengenbegriffs zu verschaffen.

1.1.1. Cantor gab 1895 die folgende Definition:

> Unter einer *Menge* verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die Elemente von M genannt werden) zu einem Ganzen.

Man kann versuchen, diese Definition wie folgt zu präzisieren. Sei V die Gesamtheit aller Objekte (unserer Anschauung oder unseres Denkens). Mit $A(x)$ bezeichnen wir Eigenschaften von Objekten x aus V. Dann kann man

$$\{\, x \mid A(x) \,\}$$

bilden, die Menge aller Objekte x aus V mit der Eigenschaft $A(x)$. Aufgrund von Cantors Definiton ist $\{\, x \mid A(x) \,\}$ wieder ein Objekt aus V. Beispiele für Eigenschaften:

1. x ist natürliche Zahl.
2. x ist Menge.
3. x ist Punkt, y ist Gerade und x liegt auf y.
4. y ist Menge und x ist Element von y, kurz: $\mathrm{Mg}(y) \wedge x \in y$.

Cantors Definiton ist jedoch in ihrer ursprünglichen Form nicht haltbar, da sie zu Widersprüchen führt. Am bekanntesten ist die sogenannte Russell*sche Antinomie*: Sei

$$x_0 := \{\, x \mid \mathrm{Mg}(x) \wedge x \notin x \,\}.$$

Dann gilt $x \in x_0 \equiv \mathrm{Mg}(x) \wedge x \notin x$ für alle Objekte x, also insbesondere

$$x_0 \in x_0 \equiv \mathrm{Mg}(x_0) \wedge x_0 \notin x_0 \equiv x_0 \notin x_0,$$

denn x_0 ist Menge. Der Grund für diesen Widerspruch liegt darin, daß man von der Vorstellung einer fertigen Gesamtheit aller Mengen ausgeht. Dies ist aber weder notwendig noch entspricht es dem Vorgehen in der Mathematik. Es reicht vollkommen aus, wenn man eine Menge nur dann bildet, wenn ihre Elemente bereits "zur Verfügung stehen". Dies führt auf eine Vorstellung einer stufenweisen Konstruktion von Mengen.

1.1.2. *Kumulative Typenstruktur.* Eine Möglichkeit, diese Vorstellung zu konkretisieren, bietet die sogenannte *kumulative Typenstruktur*: Man beginnt mit vorgegebenen Urelementen, die die Mengen der Stufe 0 bilden. Auf einer beliebigen Stufe kann man dann alle Mengen bilden, deren Elemente früheren Stufen angehören. Wählt man zum Beispiel als Urelemente die natürlichen Zahlen, so gehört $\{27, \{5\}\}$ zur Stufe 2.

Es stellen sich nun die folgenden natürlichen Fragen: 1. Welche Urelemente soll man wählen? 2. Wie weit reichen die Stufen?

Zu 1. Für die Zwecke der Mathematik ist es ausreichend, überhaupt keine Urelemente vorauszusetzen; man spricht in diesem Fall von *reinen Mengen*. Auf den ersten Stufen erhält man dann folgende Mengen:

$$\begin{array}{ll} \text{Stufe 0:} & - \\ \text{Stufe 1:} & \emptyset \\ \text{Stufe 2:} & \emptyset, \{\emptyset\} \\ \text{Stufe 3:} & \emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\} \\ & \text{und so weiter.} \end{array}$$

Zu 2. Shoenfield hat das folgende Prinzip formuliert.

> Man betrachte eine Gesamtheit $\mathcal{S}$ von Stufen. Kann man sich eine Situation vorstellen, in der alle Stufen aus $\mathcal{S}$ konstruiert sind, so soll es eine Stufe geben, die nach allen Stufen aus $\mathcal{S}$ kommt.

Aus diesem vagen Prinzip lassen sich exakte Folgerungen ziehen, die man durch Axiome fixieren kann. Dies würde jedoch innerhalb dieser Vorlesung zu weit führen; es sei deshalb auf Vorlesungen über mathematische Logik und über Mengenlehre verwiesen.

Unter einer *Menge* kann man also inhaltlich ein Objekt verstehen, das zu einer Stufe der kumulativen Typenstruktur gehört. Unter einer *Klasse* versteht man dann eine beliebige Gesamtheit von Mengen. Jede Menge ist also eine Klasse. Ferner gibt es Klassen, die keine Mengen sind, z.B. die Klasse V aller Mengen.

1.1.3. *Einfache Typen.* Für die meisten Zwecke in der Mathematik und insbesondere auch für diese Vorlesung reicht es aber aus, mit einem einfacheren und konkreteren Mengenbegriff zu arbeiten, nämlich mit Mengen von Objekten eines festen "einfachen" Typs τ. *Einfache Typen* sind aus dem Grundtyp ι induktiv gebildet mit den Operationen $\tau_1 \to \tau_2$ und $\tau_1 \times \tau_2$. Im folgenden lassen wir den Zusatz "einfach" weg und sprechen kurz von Typen. Unter *Objekten* eines Typs τ versteht man folgendes.

ι:	natürliche Zahlen,
$\tau_1 \to \tau_2$:	Funktionen f, die Objekten a des Typs τ_1 Objekte $f(a)$ des Typs τ_2 zuordnen,
$\tau_1 \times \tau_2$:	geordnete Paare (a, b) von Objekten a des Typs τ_1 und Objekten b des Typs τ_2.

Eine *Funktion* f des Typs $\tau_1 \to \tau_2$ fassen wir als eine Zuordnungsvorschrift auf, die zu jedem Objekt a des Typs τ_1 ein Objekt $f(a)$ des Typs τ_2 liefert. Zum Beispiel ist die Funktion $n \mapsto n^2$ ein Objekt des Typs $\iota \to \iota$, und das Paar $(27, 5)$ ein Objekt des Typs $\iota \times \iota$.

Eine *Menge* ist immer eine Gesamtheit von Objekten eines festen Typs τ. Meist ist der Typ aus dem Kontext klar und wird deshalb nicht extra erwähnt. Dies ist zum Beispiel bei der Menge $\mathbb{N} = \{0, 1, 2, \dots\}$ aller natürlichen Zahlen, bei $\{2, 4, 6, 8\}$ oder auch bei der Menge der Primzahlen der Fall. Sind M und N Mengen desselben Typs, so heißt M *Teilmenge* von N (geschrieben $M \subseteq N$), wenn jedes Element von M auch Element von N ist. Für $M \subseteq N$ schreibt man auch $N \supseteq M$.

Sind M und N Mengen desselben Typs, so definieren wir

$$\begin{aligned}
M \cap N &:= \{\, a \mid a \in M \text{ und } a \in N \,\} && \textit{Durchschnitt} \text{ von } M \text{ und } N, \\
M \cup N &:= \{\, a \mid a \in M \text{ oder } a \in N \,\} && \textit{Vereinigung} \text{ von } M \text{ und } N, \\
M \setminus N &:= \{\, a \mid a \in M \text{ und } a \notin N \,\} && \textit{Differenz} \text{ von } M \text{ und } N.
\end{aligned}$$

Aufgrund der bekannten Eigenschaften der logischen Verknüpfungen "und", "oder" und "nicht" gilt offenbar

$$\begin{aligned}
M \cap N &= N \cap M & M \cup N &= N \cup M, \\
(M \cap N) \cap K &= M \cap (N \cap K) & (M \cup N) \cup K &= M \cup (N \cup K), \\
M \cap (N \cup K) &= (M \cap N) \cup (M \cap K) & M \cup (N \cap K) &= (M \cup N) \cap (M \cup K).
\end{aligned}$$

sowie

$$\begin{aligned}
M \setminus (N \cap K) &= (M \setminus N) \cup (M \setminus K), \\
M \setminus (N \cup K) &= (M \setminus N) \cap (M \setminus K).
\end{aligned}$$

Ist M eine Menge von Objekten des Typs τ_1 und N eine Menge von Objekten des Typs τ_2, so ist

$$M \times N := \{\, (a, b) \mid a \in M \text{ und } b \in N \,\}$$

eine Menge von Objekten des Typs $\tau_1 \times \tau_2$. Man nennt $M \times N$ das *kartesische Produkt* der Mengen M und N. Unter einer *Relation* R zwischen Elementen von M und Elementen von N versteht man eine Teilmenge $R \subseteq M \times N$. Statt $(a, b) \in R$ schreibt man oft aRb.

Beispiel. Die Teilbarkeitsrelation $m|n$ auf $\mathbb{N}$ wird dargestellt durch die Relation $\{\, (m, n) \mid \exists k \in \mathbb{N}\, (m \cdot k = n) \,\} \subseteq \mathbb{N} \times \mathbb{N}$.

Unter einer *Abbildung* verstehen wir ein Tripel (M, f, N), wobei M eine Menge von Objekten des Typs τ_1, N eine Menge von Objekten des Typs τ_2 und f ein Objekt des Typs $\tau_1 \to \tau_2$ ist (also eine Funktion, die Objekten a des Typs τ_1 Objekte $f(a)$ des Typs τ_2 zuordnet), so daß für alle $a \in M$ gilt $f(a) \in N$. Man schreibt meist

$$f\colon M \to N.$$

Man beachte, daß f als Objekt des Typs $\tau_1 \to \tau_2$ auf *allen* Objekten a des Typs τ_1 definiert ist; die Werte von f für $a \notin M$ interessieren jedoch nicht. M heißt der *Definitionsbereich* und N der *Wertebereich* der Abbildung.

Ist $f\colon M \to N$ eine Abbildung und M' eine Teilmenge von M, so nennt man die durch

$$f\colon \begin{array}{l} M' \to N \\ a \mapsto f(a) \end{array}$$

definierte Abbildung die *Einschränkung* von f auf $M' \subseteq M$ und bezeichnet sie mit $f{\restriction}M'$.

Die Abbildung $f\colon M \to N$ heißt

1. *injektiv*, wenn aus der Gleichheit zweier Werte die Gleichheit der Argumente folgt, also wenn
$$\forall a_1, a_2{\in}M\, (f(a_1) = f(a_2) \to a_1 = a_2);$$
2. *surjektiv*, wenn jedes $b \in N$ als Funktionswert auftritt, also wenn
$$\forall b{\in}N\, \exists a{\in}M\; f(a) = b;$$
3. *bijektiv*, wenn sie injektiv und surjektiv ist.

Jede bijektive Abbildung $f\colon M \to N$ besitzt eine *Umkehrabbildung* $f^{-1}\colon N \to M$: sie ordnet jedem $b \in N$ das eindeutig bestimmte $a \in M$ mit $f(a) = b$ zu.

Sind $f\colon M \to N$ und $g\colon N \to K$ Abbildungen, so erklärt man die *Komposition* oder *Hintereinanderschaltung* $g \circ f$ (gelesen g *nach* f) als diejenige Abbildung $g \circ f\colon M \to K$, die durch die Zuordnungsvorschrift $(g \circ f)(a) := g(f(a))$ bestimmt ist. Ist noch $h\colon K \to L$, so gilt offenbar

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

man sagt hierfür, daß die Komposition von Abbildungen *assoziativ* ist. Man beachte, daß die Komposition von Abbildungen i.a. nicht kommutativ ist. Ist zum Beispiel

$$f\colon \begin{array}{l} \mathbb{N} \to \mathbb{N} \\ a \mapsto a+1 \end{array} \quad \text{und} \quad g\colon \begin{array}{l} \mathbb{N} \to \mathbb{N} \\ a \mapsto a^2, \end{array}$$

so gilt $(g \circ f)(a) = (a+1)^2$, aber $(f \circ g)(a) = a^2 + 1$, also $g \circ f \neq f \circ g$. – Ferner benötigen wir später eine sehr einfache Abbildung, nämlich die *Identität* $\mathrm{id}\colon M \to M$ auf einer gegebenen Menge M. Sie ist bestimmt durch die Zuordnungsvorschrift $\mathrm{id}(a) := a$. – Ist $f\colon M \to N$ bijektiv, so gilt offenbar $f^{-1} \circ f = \mathrm{id}$.

Es ist oft notwendig, die Bilder und Urbilder unter einer Abbildung $f\colon M \to N$ nicht nur für Elemente von M bzw. N, sondern auch für Teilmengen von M bzw. N zu betrachten. Wir setzen für $A \subseteq M$ und $C \subseteq N$

$$f(A) := \{\, b \in N \mid \exists a{\in}A\; f(a) = b \,\},$$
$$f^{-1}(C) := \{\, a \in M \mid f(a) \in C \,\}.$$

Man beachte, daß wir in $f^{-1}(C)$ das Symbol f^{-1} verwenden, obwohl f nicht notwendig bijektiv ist, also i.a. keine Umkehrabbildung besitzt.

Bemerkung. Sei $f\colon M \to N$ eine Abbildung sowie $A, B \subseteq M$ und $C, D \subseteq N$. Dann gilt

$$\begin{aligned} f(A \cap B) &\subseteq f(A) \cap f(B), \\ f(A \cup B) &= f(A) \cup f(B), \\ f^{-1}(C \cap D) &= f^{-1}(C) \cap f^{-1}(D), \\ f^{-1}(C \cup D) &= f^{-1}(C) \cup f^{-1}(D). \end{aligned}$$

Durch Angabe eines Gegenbeispiels zeigt man leicht, daß i.a. $f(A \cap B) \neq f(A) \cap f(B)$ ist.

Ist I eine beliebige Menge (die wir als Indexmenge auffassen) und M eine weitere Menge, so nennt man eine Abbildung $f\colon I \to M$ auch eine mit I indizierte *Familie*. Man schreibt oft f_i anstelle von $f(i)$, und $(f_i)_{i \in I}$ anstelle von f. Ist speziell $I = \{1, \dots, n\}$, so hat man eine endliche Familie und schreibt $(f_1, \dots, f_n)$ anstelle von $(f_i)_{i \in \{1,\dots,n\}}$. Im Fall $I = \mathbb{N}$ nennt man die Familie $(f_i)_{i \in \mathbb{N}}$ auch eine *Folge*.

Ist $(M_i)_{i \in I}$ eine Familie von Mengen desselben Typs τ, so definieren wir

$$\begin{aligned} \bigcap_{i \in I} M_i &:= \{\, a \text{ vom Typ } \tau \mid \forall i{\in}I\; a \in M_i \,\}, \\ \bigcup_{i \in I} M_i &:= \{\, a \text{ vom Typ } \tau \mid \exists i{\in}I\; a \in M_i \,\}. \end{aligned}$$

Jeder Menge X von Objekten des Typs τ kann man eine 0-1-wertige Funktion c_X vom Typ $\tau \to \iota$ zuordnen:

$$c_X(a) := \begin{cases} 1 & \text{falls } a \in X \\ 0 & \text{falls } a \notin X. \end{cases}$$

c_X heißt *charakteristische Funktion* der Menge X. Offenbar enthält c_X dieselbe Information wie X; man kann deshalb die Menge X mit ihrer charakteristischen Funktion identifizieren. Auf diese Weise kann man die Menge aller Teilmengen einer gegebenen Menge M bilden, und zwar als Menge von Objekten des Typs $\tau \to \iota$:

$$\mathcal{P}(M) := \{\, X \mid X \subseteq M \,\} \quad (\textit{Potenzmenge} \text{ von } M).$$

Genauer müßte man $\mathcal{P}(M) = \{\, c_X \mid X \subseteq M \,\}$ schreiben.

Bemerkung. Für beliebige Mengen M, N gilt offenbar

$$\begin{aligned} \mathcal{P}(M \cap N) &= \mathcal{P}(M) \cap \mathcal{P}(N), \\ \mathcal{P}(M \cup N) &\supseteq \mathcal{P}(M) \cup \mathcal{P}(N), \\ \mathcal{P}(M \cup N) = \mathcal{P}(M) \cup \mathcal{P}(N) &\equiv M \subseteq N \text{ oder } N \subseteq M. \end{aligned}$$

1.2 Gruppen und Körper

Wir skizzieren einen präzisen Weg zur Einführung der ganzen, der rationalen, der reellen und der komplexen Zahlen. Dabei wird uns der Begriff einer Äquivalenzrelation gute Dienste leisten. Aus der Additionsstruktur der ganzen Zahlen abstrahieren wir den allgemeinen Gruppenbegriff, und entsprechend aus den Additions- und Multiplikationseigenschaften der rationalen Zahlen den Begriff eines Körpers.

1.2.1. *Ganze Zahlen* kann man darstellen als Paare (m, n) natürlicher Zahlen, also als Objekte vom Typ $\iota \times \iota$. Gemeint ist mit (m, n) die Differenz $m - n$. Man muß also zum Beispiel $(m + 1, n + 1)$ und (m, n) als gleich betrachten. Allgemein definieren wir eine Relation $\sim$ zwischen Paaren natürlicher Zahlen durch

$$(m, n) \sim (p, q) :\equiv m + q = p + n.$$

Es gilt dann:

Bemerkung. 1. $(m, n) \sim (m, n)$.
2. Wenn $(m, n) \sim (p, q)$, so ist $(p, q) \sim (m, n)$.
3. Wenn $(m, n) \sim (p, q)$ und $(p, q) \sim (k, l)$, so ist $(m, n) \sim (k, l)$.

Beweis. (1) und (2) sind klar. 3. Zu zeigen ist

$$m + l = k + n$$

und nach Annahme gilt $m + q = p + n$ und $p + l = k + q$. Man erhält

$$m + p + l = m + k + q = k + n + p,$$

also $m + l = k + n$. Qed

Allgemein nennt man eine Relation $\sim$ auf einer Menge M eine *Äquivalenzrelation*, wenn sie diese drei Eigenschaften besitzt, also wenn gilt

$a \sim a$	($\sim$ ist *reflexiv*),
wenn $a \sim b$, so ist $b \sim a$	($\sim$ ist *symmetrisch*),
wenn $a \sim b$ und $b \sim c$, so ist $a \sim c$	($\sim$ ist *transitiv*)

für alle $a, b, c \in M$. Eine Äquivalenzrelation kann man als eine Art verallgemeinerte Gleichheit auffassen.

Es ist möglich und oft üblich, das explizite Auftreten einer Äquivalenzrelation zu vermeiden und stattdessen mit der vertrauten Gleichheit zu arbeiten. Dazu faßt man äquivalente Objekte zusammen zu sogenannten Äquivalenzklassen (richtiger wäre "Äquivalenzmengen", aber der Name "Äquivalenzklasse" hat sich eingebürgert). Eine nicht leere Teilmenge α von M heißt *Äquivalenzklasse*, wenn gilt

Äk1 Wenn $a \in \alpha$ und $b \sim a$, so ist $b \in \alpha$ (α ist *abgeschlossen* unter $\sim$).

Äk2 Wenn $a, b \in \alpha$, so ist $a \sim b$.

Bemerkung. 1. Jedes $a \in M$ liegt in genau einer Äquivalenzklasse; sie wird mit $[a]$ bezeichnet.
2. Je zwei verschiedene Äquivalenzklassen sind disjunkt.

Beweis. 1. Existenz. Setze $[a] := \{ b \mid b \sim a \}$. Dann ist $a \in [a]$ wegen der Reflexivität von $\sim$. Zu zeigen bleibt, daß $[a]$ eine Äquivalenzklasse ist. Zu (Äk1). Sei $b \in [a]$ und $c \sim b$. Dann gilt $b \sim a$, also auch $c \sim a$ wegen der Transitivität von $\sim$, also $c \in [a]$. Zu (Äk2). Seien $b, c \in [a]$. Dann gilt $b \sim a$ und $c \sim a$. Aus $c \sim a$ folgt $a \sim c$ wegen der Symmetrie von $\sim$, also auch $b \sim c$ wegen der Transitivität von $\sim$.

Eindeutigkeit. Seien α, β Äquivalenzklassen mit $a \in \alpha$ und $a \in \beta$. Zu zeigen ist $\alpha = \beta$; aus Symmetriegründen genügt $\alpha \subseteq \beta$. Sei also $c \in \alpha$. Zu zeigen ist $c \in \beta$. Wegen $c, a \in \alpha$ folgt $c \sim a$ nach (Äk2) für α, also auch $c \in \beta$ nach (Äk1) für β.

2. Zu zeigen ist, daß aus $a \in \alpha, \beta$ folgt $\alpha = \beta$. Das ist aber gerade der Eindeutigkeitsteil der Aussage (1). QED

Eine Äquivalenzrelation zerlegt also die Trägermenge vollständig in disjunkte Äquivalenzklassen.

1.2.2. Die *Addition ganzer Zahlen* wird definiert durch

$$(m,n) + (p,q) := (m+p, n+q).$$

Diese Addition ist verträglich mit der Äquivalenzrelation $\sim$, d.h. aus $(m,n) \sim (m',n')$ und $(p,q) \sim (p',q')$ folgt $(m,n)+(p,q) \sim (m',n')+(p',q')$. Die ganze Zahl *Null* wird dargestellt durch $(0,0)$; sie hat die Eigenschaft $(0,0)+(p,q) \sim (p,q)$ (es gilt sogar die Gleichheit). Man beachte, daß für alle $p \in \mathbb{N}$ gilt $(0,0) \sim (p,p)$. Das *Negative* (oder *additive Inverse*) einer ganzen Zahl ist

$$-(m,n) := (n,m);$$

offenbar gilt $-(m,n)+(m,n) \sim (0,0)$. Das Bilden des Negativen ist verträglich mit $\sim$: aus $(m,n) \sim (m',n')$ folgt $-(m,n) \sim -(m',n')$.

Da also Addition und Inversenbildung mit der Äquivalenzrelation $\sim$ verträglich sind, kann man entsprechende Funktionen auch auf den Äquivalenzklassen definieren. Wir setzen für ganze Zahlen $a = (m,n)$ und $b = (p,q)$

$$\begin{aligned} [a] + [b] &:= [a+b], \\ -[a] &:= [-a]. \end{aligned}$$

Damit dies eine sinnvolle Definition ist, muß man die *Unabhängigkeit vom Repräsentanten* zeigen, d.h. man muß zeigen, daß

aus $[a] = [a']$ und $[b] = [b']$ folgt $[a+b] = [a'+b']$, und
aus $[a] = [a']$ folgt $[-a] = [-a']$.

Da aber $[c] = [c']$ gleichbedeutend ist mit $c \sim c'$, besagt dies gerade, daß

aus $a \sim a'$ und $b \sim b'$ folgt $a+b \sim a'+b'$, und
aus $a \sim a'$ folgt $-a \sim -a'$.

Das war aber die eben bemerkte Verträglichkeit der Äquivalenzrelation $\sim$ mit Addition und Inversenbildung.

Addition und Inversenbildung auf den Äquivalenzklassen haben die folgenden Eigenschaften:

1. $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ für alle Äquivalenzklassen α, β, γ (*Assoziativgesetz*).
2. Für das Nullelement $e := [(0,0)]$ gilt
   a. $e + \alpha = \alpha$ für alle Äquivalenzklassen α.
   b. $-\alpha + \alpha = e$ für alle Äquivalenzklassen α.

1.2.3. Diese Eigenschaften beinhalten in einem gewissen Sinne alles, was über die additive Struktur der ganzen Zahlen zu sagen ist. Wir führen jetzt einen Abstraktionsschritt durch und definieren den allgemeinen Gruppenbegriff.

DEFINITION. Sei G eine Menge und $\circ\colon G \times G \to G$ eine Abbildung. G (oder genauer das Paar $(G, \circ)$) heißt *Gruppe*, wenn gilt

1. $(\alpha \circ \beta) \circ \gamma = \alpha \circ (\beta \circ \gamma)$ für alle $\alpha, \beta, \gamma \in G$ (*Assoziativgesetz*).
2. Es gibt ein $e \in G$ (*neutrales Element* von G) mit folgenden Eigenschaften
   a. $e \circ \alpha = \alpha$ für alle $\alpha \in G$.
   b. Zu jedem $\alpha \in G$ gibt es ein $\alpha' \in G$ mit $\alpha' \circ \alpha = e$ (α' heißt *inverses Element* zu α).

G heißt *abelsch*, wenn außerdem noch gilt $\alpha \circ \beta = \beta \circ \alpha$ für alle $\alpha, \beta \in G$ (*Kommutativgesetz*).

BEISPIEL. In 1.2.2 haben wir die Gruppe $\mathbb{Z}$ oder genauer $(\mathbb{Z}, +)$ der *ganzen Zahlen* konstruiert.

BEMERKUNG. (Klammerung). Sei G eine Gruppe, ferner $n \geq 2$ und $\alpha_1, \ldots, \alpha_n \in G$. Dann hat $\alpha_1 \circ \alpha_2 \circ \ldots \circ \alpha_n$ bei jeder Klammerung denselben Wert.

BEWEIS. Gezeigt wird

$$\alpha_1 \circ \alpha_2 \circ \ldots \circ \alpha_{n-1} \circ \alpha_n = (\ldots(\alpha_1 \circ \alpha_2) \circ \ldots \circ \alpha_{n-1}) \circ \alpha_n$$

durch Induktion über n. *Basis* $n = 2$. Klar. *Schritt.* Wir nehmen an, daß die Behauptung für alle m mit $m \leq n$ schon bewiesen ist (Induktionsvoraussetzung). Zu zeigen ist die Behauptung für $n+1$. Gegeben sei also $(\alpha_1 \circ \alpha_2 \circ \ldots \circ \alpha_i) \circ (\alpha_{i+1} \circ \ldots \circ \alpha_n \circ \alpha_{n+1})$ mit $1 \leq i < n$ (dies kann man oBdA annehmen; andernfalls folgt die Behauptung sofort aus der Induktionsvoraussetzung). Dann gilt

$$\begin{aligned}
&(\alpha_1 \circ \alpha_2 \circ \ldots \circ \alpha_i) \circ (\alpha_{i+1} \circ \ldots \circ \alpha_n \circ \alpha_{n+1}) \\
&= (\alpha_1 \circ \alpha_2 \circ \ldots \circ \alpha_i) \circ ([\alpha_{i+1} \circ \ldots \circ \alpha_n] \circ \alpha_{n+1}) && \text{nach Induktionsvoraussetzung} \\
&= ((\alpha_1 \circ \alpha_2 \circ \ldots \circ \alpha_i) \circ [\alpha_{i+1} \circ \ldots \circ \alpha_n]) \circ \alpha_{n+1} && \text{nach dem Assoziativgesetz} \\
&= (\ldots(\alpha_1 \circ \alpha_2) \circ \ldots \circ \alpha_n) \circ \alpha_{n+1} && \text{nach Induktionsvoraussetzung. QED}
\end{aligned}$$

BEMERKUNGEN. Sei G eine Gruppe.

1. Sei $e \in G$ neutrales Element und $\alpha, \alpha' \in G$ so daß $\alpha' \circ \alpha = e$. Dann gilt auch $\alpha \circ \alpha' = e$.
2. Sei $e \in G$ neutrales Element. Dann gilt $\alpha \circ e = \alpha$ für alle $\alpha \in G$.
3. Es gibt genau ein neutrales Element $e \in G$.
4. Zu jedem $\alpha \in G$ gibt es genau ein inverses Element $\alpha' \in G$, bezeichnet mit α^{-1}.

BEWEIS. 1. Wähle α'' mit $\alpha'' \circ \alpha' = e$. Dann gilt

$$\alpha \circ \alpha' = e \circ \alpha \circ \alpha' = \alpha'' \circ \underbrace{\alpha' \circ \alpha}_{e} \circ \alpha' = \alpha'' \circ \alpha' = e.$$

2. $\alpha \circ e = \alpha \circ \alpha' \circ \alpha = e \circ \alpha = \alpha$ nach (1).
3. Seien e, e^* neutrale Elemente von G. Dann gilt $e^* = e \circ e^* = e$ nach (2).
4. Seien α', α^* inverse Elemente zu $\alpha \in G$. Dann gilt

$$\alpha^* = \alpha^* \circ e = \alpha^* \circ \alpha \circ \alpha' = e \circ \alpha = \alpha'. \quad \text{QED}$$

1.2.4. Lemma. *Sei G eine nicht leere Menge und $\circ\colon G \times G \to G$ eine Abbildung. G ist eine Gruppe genau dann, wenn gilt*

1. $(\alpha \circ \beta) \circ \gamma = \alpha \circ (\beta \circ \gamma)$ *für alle* $\alpha, \beta, \gamma \in G$.

2*. a. $\forall \alpha, \beta{\in}G \exists \gamma{\in}G \; \alpha \circ \gamma = \beta$.
 b. $\forall \alpha, \beta{\in}G \exists \delta{\in}G \; \delta \circ \alpha = \beta$.

Beweis. $\Longrightarrow$. Sei G eine Gruppe. Zu zeigen ist (2*). a. Setze $\gamma = \alpha^{-1} \circ \beta$. b. Setze $\delta = \beta \circ \alpha^{-1}$.

$\Longleftarrow$. G erfülle (1) und (2*). Zu zeigen ist (2). Wähle $\alpha_0 \in G$ fest (hier wird benötigt, daß G nicht leer ist). Wähle e so, daß $e \circ \alpha_0 = \alpha_0$ ist. a. Sei $\alpha \in G$ beliebig. Wähle $\gamma \in G$ mit $\alpha_0 \circ \gamma = \alpha$. Dann gilt

$$e \circ \alpha = e \circ \alpha_0 \circ \gamma = \alpha_0 \circ \gamma = \alpha.$$

Teil (b) folgt aus (2*b). Qed

Bemerkung. Sei G eine Gruppe und $\alpha, \beta \in G$. Dann gilt

$$(\alpha^{-1})^{-1} = \alpha \tag{1.2}$$

$$(\alpha \circ \beta)^{-1} = \beta^{-1} \circ \alpha^{-1} \tag{1.3}$$

Beweis. (1.2). Es gilt $(\alpha^{-1})^{-1} \circ \alpha^{-1} = e$ und auch $\alpha \circ \alpha^{-1} = e$. Die Behauptung folgt aus der Eindeutigkeit des Inversen.

(1.3). $\beta^{-1} \circ \alpha^{-1} \circ \alpha \circ \beta = \beta^{-1} \circ e \circ \beta = e$. Qed

Beispiel. Sei M eine nicht leere Menge und $S(M)$ die Menge der bijektiven Abbildungen von M auf sich selbst. Dann ist $(S(M), \circ)$, wobei $\circ$ für die Komposition von Abbildungen steht, eine Gruppe, die *symmetrische Gruppe* der Menge M. Neutrales Element ist die Identität $\mathrm{id}\colon M \to M$, und das inverse Element zu $f \in S(M)$ ist die Umkehrabbildung f^{-1}.

Man beachte, daß $S(M)$ i.a. nicht kommutativ ist. Ist zum Beispiel

$$f\colon \quad \{0,1,2\} \to \{0,1,2\} \qquad\qquad g\colon \quad \{0,1,2\} \to \{0,1,2\}$$

$$f(a) = \begin{cases} 0 & \text{falls } a = 0 \\ 2 & \text{falls } a = 1 \\ 1 & \text{falls } a = 2 \end{cases} \qquad \text{und} \qquad g(a) = \begin{cases} 1 & \text{falls } a = 0 \\ 0 & \text{falls } a = 1 \\ 2 & \text{falls } a = 2 \end{cases}$$

so gilt $(g \circ f)(0) = 1$, aber $(f \circ g)(0) = 2$, also $g \circ f \neq f \circ g$. – Im Spezialfall $M = \{1, \ldots, n\}$ schreibt man S_n statt $S(M)$. Jede Abbildung $\sigma \in S_n$ heißt *Permutation* der Zahlen $1, \ldots, n$.

1.2.5. *Rationale Zahlen* stellen wir dar als Paare $((m,n),k)$ einer ganzen Zahl (m,n) und einer von Null verschiedenen natürlichen Zahl k; gemeint ist der Quotient aus beiden, also $\frac{m-n}{k}$. Da wir zwei rationale Zahlen als gleich betrachten wollen, wenn sie denselben gekürzten Bruch darstellen, definieren wir eine Relation $\sim$ durch

$$((m,n),k) \sim ((p,q),l) :\equiv lm + kq = kp + ln.$$

Man sieht leicht, daß $\sim$ eine Äquivalenzrelation ist.

Die *Addition rationaler Zahlen* wird definiert durch

$$((m,n),k) + ((p,q),l) := ((lm + kp, ln + kq), kl).$$

Wieder ist diese Addition verträglich mit der Äquivalenzrelation $\sim$, d.h. aus $((m,n),k) \sim ((m',n'),k')$ und $((p,q),l) \sim ((p',q'),l')$ folgt $((m,n),k)+((p,q),l) \sim ((m',n'),k')+((p',q'),l')$. Die rationale Zahl *Null* wird dargestellt durch $((0,0),1)$; sie hat die Eigenschaft $((0,0),1) + ((p,q),l) \sim ((p,q),l)$ (es gilt sogar die Gleichheit). Das *Negative* (oder *additive Inverse*) einer rationalen Zahl ist

$$-((m,n),k) := ((n,m),k).$$

Auch die Bildung des Negativen ist wieder verträglich mit $\sim$: aus $((m,n),k) \sim ((m',n'),k')$ folgt $-((m,n),k) \sim -((m',n'),k')$. Wir können also entsprechende Funktionen auch auf den Äquivalenzklassen definieren und erhalten wie bei den ganzen Zahlen, daß die rationalen Zahlen zusammen mit der Addition eine abelsche Gruppe bilden.

Die *Multiplikation* rationaler Zahlen wird definiert durch

$$((m,n),k) \cdot ((p,q),l) := ((mp+nq, mq+np), kl).$$

Die rationale Zahl *Eins* wird dargestellt durch $((1,0),1)$; sie hat die Eigenschaft $((1,0),1) \cdot ((p,q),l) \sim ((p,q),l)$ (es gilt sogar die Gleichheit). Das *multiplikative Inverse* einer rationalen Zahl $((m,n),k) \not\sim ((0,0),1)$ ist

$$((m,n),k)^{-1} := \begin{cases} ((k,0), m-n) & \text{falls } m > n \\ ((0,k), n-m) & \text{falls } m < n. \end{cases}$$

Alle diese Operationen sind verträglich mit $\sim$. Wir können also wieder entsprechende Funktionen auch auf den Äquivalenzklassen definieren und erhalten wie bei den ganzen Zahlen, daß die rationalen Zahlen ohne die Null zusammen mit der Multiplikation eine abelsche Gruppe bilden. Ferner gilt das Distributivgesetz

$$((m,n),k) \cdot [((p,q),l) + ((p',q'),l')] = [((m,n),k) \cdot ((p,q),l)] + [((m,n),k) \cdot ((p',q'),l')].$$

1.2.6. Diese Eigenschaften beinhalten wieder alles das, was über die additive und multiplikative Struktur der rationalen Zahlen zu sagen ist. Wir führen deshalb einen Abstraktionsschritt durch und definieren den allgemeinen Körperbegriff.

Definition. Sei K eine Menge und $+\colon K \times K \to K$, $\cdot\colon K \times K \to K$ Abbildungen. Wir schreiben $\alpha + \beta$ für $+(\alpha,\beta)$ und $\alpha\beta$ oder $\alpha \cdot \beta$ für $\cdot(\alpha,\beta)$. Ferner soll $\cdot$ stärker binden als $+$. K (oder genauer das Tripel $(K,+,\cdot)$) heißt *Körper*, wenn

1. $(K,+)$ eine abelsche Gruppe ist,
2. $K^* := K \setminus \{0\}$ zusammen mit der auf K^* eingeschränkten Multiplikation eine abelsche Gruppe ist (insbesondere gilt $\alpha\beta \in K^*$ falls $\alpha, \beta \in K^*$), und
3. für alle $\alpha, \beta, \gamma \in K$ gilt

$$\alpha(\beta+\gamma) = \alpha\beta + \alpha\gamma \quad (\textit{Distributivgesetz}).$$

Beispiel. In 1.2.5 haben wir den Körper $\mathbb{Q}$ oder genauer $(\mathbb{Q},+,\cdot)$ der *rationalen Zahlen* konstruiert.

Schreibweise. 1. In der abelschen Gruppe $(K, +)$ wird das neutrale Element mit 0 und das zu $\alpha \in K$ inverse Element mit $-\alpha$ bezeichnet. Wir schreiben $\alpha - \beta$ für $\alpha + (-\beta)$.

2. In der abelschen Gruppe $(K^*, \cdot)$ wird das neutrale Element mit 1 und das zu $\alpha \in K^*$ inverse Element mit α^{-1} bezeichnet. Wir schreiben $\frac{\alpha}{\beta}$ für $\alpha \cdot \beta^{-1}$.

3. Wir schreiben

$$\begin{array}{lll} \sum_{i=1}^n \alpha_i & \text{für} & \alpha_1 + \alpha_2 + \ldots + \alpha_n, \\ \prod_{i=1}^n \alpha_i & \text{für} & \alpha_1 \cdot \alpha_2 \cdot \ldots \cdot \alpha_n, \\ n\alpha & \text{für} & \underbrace{\alpha + \ldots + \alpha}_{n\text{-mal}}, \\ \alpha^n & \text{für} & \underbrace{\alpha \cdot \ldots \cdot \alpha}_{n\text{-mal}}. \end{array}$$

Im folgenden werden wir oft Doppelsummen verwenden. Sei K ein Körper und $\alpha_{ij} \in K$ für $1 \leq i \leq m$ und $1 \leq j \leq n$. Dann gilt

$$\sum_{i=1}^{m}\Big(\sum_{j=1}^{n} \alpha_{ij}\Big) = \sum_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} \alpha_{ij} = \sum_{j=1}^{n}\Big(\sum_{i=1}^{m} \alpha_{ij}\Big).$$

Bemerkungen. Sei K ein Körper. Dann gilt für alle $\alpha, \beta \in K$

1. $0 \cdot \alpha = 0$.
2. Aus $\alpha \neq 0$ und $\beta \neq 0$ folgt $\alpha\beta \neq 0$.
3. $-\alpha = (-1)\alpha$.
4. $(-\alpha) \cdot (-\beta) = \alpha\beta$.

Beweis. 1. $0 \cdot \alpha = (0 + 0) \cdot \alpha = 0 \cdot \alpha + 0 \cdot \alpha$, also $0 \cdot \alpha = 0$.

2. Dies ist klar, da K^* eine Gruppe ist (bzgl. der Multiplikation).

3. Es ist $1 + (-1) = 0$, also $\alpha + (-1)\alpha = 0$ und damit $-\alpha = (-1)\alpha$.

4. Es ist $1 + (-1) = 0$, also $-1 + (-1) \cdot (-1) = 0$ und damit $(-1) \cdot (-1) = 1$. Es folgt $(-\alpha) \cdot (-\beta) = (-1)\alpha(-1)\beta = 1\alpha\beta = \alpha\beta$. Qed

Beispiele. 1. *Reelle Zahlen* kann man auffassen als Äquivalenzklassen von Cauchyfolgen rationaler Zahlen, also von Folgen $(\alpha_n)_{n\in\mathbb{N}}$ mit

$$\forall \varepsilon{>}0 \exists N{\in}\mathbb{N} \forall n, m{\geq}N \; |\alpha_n - \alpha_m| \leq \varepsilon.$$

Wir verwenden hier ε als Variable für rationale Zahlen. Zwei Cauchyfolgen rationaler Zahlen heißen äquivalent (in Zeichen $(\alpha_n)_{n\in\mathbb{N}} \sim (\beta_n)_{n\in\mathbb{N}}$), wenn gilt

$$\forall \varepsilon{>}0 \exists K{\in}\mathbb{N} \forall n{\geq}K \; |\alpha_n - \beta_n| \leq \varepsilon.$$

Addition und Multiplikation definiert man komponentenweise. Man kann leicht zeigen, daß Addition und Multiplikation mit der Äquivalenzrelation $\sim$ verträglich sind, und daß auf diese Weise ein Körper $(\mathbb{R}, +, \cdot)$ entsteht, den man den Körper der reellen Zahlen nennt. Für ein genaueres Studium dieses Körpers sei auf die Analysis-Vorlesung verwiesen.

2. Die Menge $\{0,1\}$ ist mit folgender Multiplikation und Addition ein Körper; er wird mit $\mathbb{F}_2$ bezeichnet.

$+$	0	1
0	0	1
1	1	0

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

Durch Unterscheiden der endlich vielen Fälle kann man leicht verifizieren, daß alle Körperaxiome erfüllt sind. In $\mathbb{F}_2$ gilt $1+1=0$. Man kann also $1+1\neq 0$ nicht aus den Körperaxiomen allein beweisen.

3. Sei K ein Körper und $\xi \in K$ kein Quadrat in K. Wir setzen

$$K(\sqrt{\xi}) := K \times K.$$

Stellt man sich (außerhalb der Legalität) (α,β) als $\alpha+\beta\sqrt{\xi}$ vor, so liegt die folgende Definition nahe:

$$(\alpha,\beta)+(\alpha',\beta') := (\alpha+\alpha',\beta+\beta'),$$
$$(\alpha,\beta)\cdot(\alpha',\beta') := (\alpha\alpha'+\beta\beta'\xi,\alpha\beta'+\beta\alpha').$$

Dadurch wird $K(\sqrt{\xi})$ zu einem Körper.

BEWEIS. 1. Daß $(K(\sqrt{\xi}),+)$ eine abelsche Gruppe ist, ergibt sich leicht aus der komponentenweisen Definition der Addition. Das neutrale Element ist $(0,0)$, und das inverse Element zu $(\alpha,\beta)\in K(\sqrt{\xi})$ ist $(-\alpha,-\beta)$.

2. Wir zeigen jetzt, daß $(K(\sqrt{\xi})^*,\cdot)$ eine abelsche Gruppe ist. Assoziativität und Kommutativität sind leicht nachzurechnen. Wir wollen zeigen, daß $(1,0)$ neutrales Element ist. a. Es gilt $(1,0)\cdot(\alpha,\beta)=(\alpha,\beta)$. b. Gegeben sei $(\alpha,\beta)\neq(0,0)$. Gesucht ist (α',β') mit

$$(\alpha',\beta')(\alpha,\beta)=(1,0) \tag{1.4}$$

Ansatz: Nehmen wir an, wir hätten schon (α',β') mit (1.4). Multiplikation beider Seiten mit $(\alpha,-\beta)$ ergibt

$$(\alpha',\beta')(\alpha^2-\beta^2\xi,0)=(\alpha,-\beta).$$

Nun gilt $\alpha^2-\beta^2\xi\neq 0$, da ξ kein Quadrat ist und $(\alpha,\beta)\neq(0,0)$. Also können wir definieren

$$\alpha' := \frac{\alpha}{\alpha^2-\beta^2\xi} \quad \text{und} \quad \beta' := \frac{-\beta}{\alpha^2-\beta^2\xi}.$$

Man verifiziert leicht (1.4).

3. Die Distributivität ergibt sich durch Nachrechnen. QED

Zwei interessante Spezialfälle wollen wir besonders betrachten. Sei $K=\mathbb{R}$ und $\xi=-1$; man beachte, daß -1 kein Quadrat in $\mathbb{R}$ ist (dies wird in der Analysis-Vorlesung unter Verwendung der Theorie der geordneten Körper bewiesen). Man schreibt $\mathbb{C} := \mathbb{R}(\sqrt{-1})$ und nennt $\mathbb{C}$ den Körper der *komplexen Zahlen*. Die Abbildung

$$\begin{array}{rcl} \mathbb{R} & \to & \mathbb{R}(\sqrt{-1}) \\ \alpha & \mapsto & (\alpha,0) \end{array}$$

ist injektiv. Da

$$\begin{aligned}(\alpha, 0) + (\beta, 0) &= (\alpha + \beta, 0)\\ (\alpha, 0) \cdot (\beta, 0) &= (\alpha \cdot \beta, 0),\end{aligned}$$

braucht man zwischen $\mathbb{R}$ und $\mathbb{R} \times \{0\} = \{ (\alpha, 0) \mid \alpha \in \mathbb{R} \}$ bzgl. Addition und Multiplikation nicht zu unterscheiden. Man kann also $\mathbb{R}$ mit $\mathbb{R} \times \{0\}$ identifizieren. In diesem Sinn ist $\mathbb{R} \subseteq \mathbb{C}$. Man setzt noch

$$i := (0, 1) \quad (\textit{imaginäre Einheit}).$$

Offenbar läßt sich jedes $\lambda \in \mathbb{C}$ eindeutig darstellen als $\lambda = \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$, denn es gilt $\alpha + i\beta = (\alpha, \beta)$. Die *konjugiert komplexe Zahl* zu $\lambda = \alpha + i\beta$ ist $\overline{\lambda} := \alpha - i\beta$. Man verifiziert leicht, daß gilt

$$\begin{aligned}\overline{\overline{\lambda}} &= \lambda,\\ \overline{\lambda + \mu} &= \overline{\lambda} + \overline{\mu},\\ \overline{\lambda\mu} &= \overline{\lambda}\,\overline{\mu},\\ \overline{0} &= 0,\\ \overline{1} &= 1.\end{aligned}$$

Ferner ist $\lambda\overline{\lambda} = \alpha^2 + \beta^2$ für $\lambda = \alpha + i\beta$. Aus der Analysis-Vorlesung ist bekannt, daß stets $\alpha^2 + \beta^2 \geq 0$ ist und man deshalb $\sqrt{\alpha^2 + \beta^2} \in \mathbb{R}$ bilden kann; man setzt $|\lambda| := \sqrt{\alpha^2 + \beta^2}$ und nennt dies den *Betrag* von $\lambda \in \mathbb{C}$. Für eine ausführlichere Diskussion des Körpers $\mathbb{C}$ der komplexen Zahlen sei wieder auf die Analysis-Vorlesung verwiesen.

Als weiteren Spezialfall betrachten wir $K = \mathbb{Q}$ und $\xi = 2$.

Bemerkung. 2 ist kein Quadrat in $\mathbb{Q}$.

Beweis. Andernfalls hätte man $2 = \frac{m^2}{n^2}$ mit teilerfremden $m, n \in \mathbb{N}$, also $2n^2 = m^2$, also m gerade, d.h. von der Form $2k$, also $n^2 = 2k^2$ und damit n^2 gerade, d.h. von der Form $2l$. Also wäre 2 doch ein gemeinsamer Teiler von m und n: Widerspruch. Qed

Also ist $\mathbb{Q}(\sqrt{2})$ ein Körper.

1.3 Matrizen

Unser Ausgangspunkt war es, das lineare Gleichungssystem (1.1) zu lösen. Wir wollen jetzt eine einfachere Schreibweise einführen, nämlich

$$A \cdot x = b \quad \text{mit} \quad A = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \quad b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}.$$

Dies wird die Lösung von (1.1) vereinfachen.

1.3.1. Definition. Sei K ein Körper, $m, n \in \mathbb{N}$ und $I = \{1, \dots, m\}$, $J = \{1, \dots, n\}$. Eine Abbildung $A\colon I \times J \to K$ heißt $(m \times n)$-*Matrix* mit Koeffizienten in K. Die $\alpha_{ij} := A(i, j) \in K$

$(1 \le i \le m,\ 1 \le j \le n)$ heißen *Koeffizienten* der Matrix A. Schreibweisen:

$$A = (\alpha_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}, \quad \text{kurz } A = (\alpha_{ij})$$

$$A = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}$$

Mit $K^{m \times n}$ bezeichnen wir die Menge aller $(m \times n)$-Matrizen mit Koeffizienten in K.

Definition. (Algebraische Operationen mit Matrizen). Sei K ein Körper, $m, n, p \in \mathbb{N}$, $A, B \in K^{m \times n}$, $C \in K^{n \times p}$ und $\alpha \in K$ mit $A = (\alpha_{ij})$, $B = (\beta_{ij})$ und $C = (\gamma_{jr})$. Dann setzen wir

$$\begin{aligned} A + B &:= (\eta_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \quad \text{mit } \eta_{ij} := \alpha_{ij} + \beta_{ij}, \\ A \cdot C &:= (\mu_{ir})_{\substack{1 \le i \le m \\ 1 \le r \le p}} \in K^{m \times p} \quad \text{mit } \mu_{ir} := \textstyle\sum_{j=1}^{n} \alpha_{ij}\gamma_{jr}, \\ \alpha A &:= (\alpha\alpha_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}. \end{aligned}$$

Wir schreiben AC für $A \cdot C$.

Beispiele. 1.

$$\begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \boxed{\alpha_{i1} \ \dots \ \alpha_{in}} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} \cdot \begin{pmatrix} \gamma_{11} & \dots & \boxed{\begin{matrix} \gamma_{1r} \\ \vdots \\ \gamma_{nr} \end{matrix}} & \dots & \gamma_{1p} \\ \vdots & & & & \vdots \\ \gamma_{n1} & \dots & & \dots & \gamma_{np} \end{pmatrix}$$

Der (i, r)-te Koeffizient des Matrizenprodukts ist $\sum_{j=1}^{n} \alpha_{ij}\gamma_{jr}$, also der einzige Koeffizient der Matrix

$$\begin{pmatrix} \alpha_{i1} & \dots & \alpha_{in} \end{pmatrix} \cdot \begin{pmatrix} \gamma_{1r} \\ \vdots \\ \gamma_{nr} \end{pmatrix}.$$

2. Sei $K = \mathbb{Q}$. Dann ist

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ -1 & 3 & -1 \end{pmatrix} = \begin{pmatrix} -2 & 8 & -1 \\ -1 & 3 & -1 \end{pmatrix}.$$

1.3.2. Bemerkungen. Wir geben jetzt einige *Rechenregeln* für Matrizen an.

1. $(K^{m \times n}, +)$ ist eine abelsche Gruppe; das neutrale Element ist

$$0 := \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix};$$

das inverse Element zu $(\alpha_{ij}) \in K^{m \times n}$ ist (β_{ij}) mit $\beta_{ij} := -\alpha_{ij}$. Die Beweise sind klar, da die Addition komponentenweise definiert ist,

2. Die Matrizenmultiplikation ist assoziativ, d.h. für beliebige $A \in K^{m\times n}$, $B \in K^{n\times p}$ und $C \in K^{p\times q}$ gilt

$$(AB)C = A(BC) \quad (\in K^{m\times q}).$$

Beweis. Sei $A = (\alpha_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$, $B = (\beta_{jr})_{\substack{1\le j\le n\\ 1\le r\le p}}$, $C = (\gamma_{rs})_{\substack{1\le r\le p\\ 1\le s\le q}}$, Setze

$$(AB)C = (\eta_{is})_{\substack{1\le i\le m\\ 1\le s\le q}} \quad \text{und} \quad A(BC) = (\zeta_{is})_{\substack{1\le i\le m\\ 1\le s\le q}}.$$

Seien i, s fest gewählt. Es gilt

$$\eta_{is} = \sum_{r=1}^{p}\Big(\sum_{j=1}^{n}\alpha_{ij}\beta_{jr}\Big)\gamma_{rs} \quad \text{und} \quad \zeta_{is} = \sum_{j=1}^{n}\alpha_{ij}\Big(\sum_{r=1}^{p}\beta_{jr}\gamma_{rs}\Big),$$

also

$$\eta_{is} = \sum_{r=1}^{p}\sum_{j=1}^{n}(\alpha_{ij}\beta_{jr}\gamma_{rs}) = \sum_{j=1}^{n}\sum_{r=1}^{p}(\alpha_{ij}\beta_{jr}\gamma_{rs}) = \zeta_{is}. \quad \text{QED}$$

3. *Neutrales Element* bei der Matrizenmultiplikation. Sei

$$E := (\delta_{ij})_{\substack{1\le i\le m\\ 1\le j\le m}} \quad \text{mit} \quad \delta_{ij} := \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \ne j. \end{cases}$$

Dann gilt $EA = A$ für alle $A \in K^{m\times n}$, und ebenso $BE = B$ für alle $B \in K^{n\times m}$.

Beweis. Sei $A = (\alpha_{jr})_{\substack{1\le j\le m\\ 1\le r\le n}}$ und $E = (\delta_{ij})_{1\le i,j\le m}$. Setze $EA := (\gamma_{ir})_{\substack{1\le j\le m\\ 1\le r\le n}}$. Dann gilt

$$\gamma_{ir} = \sum_{j=1}^{m}\delta_{ij}\alpha_{jr} = \alpha_{ir}.$$

$BE = B$ zeigt man analog. QED

4. *Distributivgesetze.* Für $A \in K^{m\times n}$ und $B, C \in K^{n\times p}$ gilt

$$A(B + C) = AB + AC.$$

Für $A', B' \in K^{m\times n}$ und $C' \in K^{n\times p}$ gilt

$$(A' + B')C' = A'C' + B'C'.$$

Die Beweise ergeben sich leicht aus dem Distributivgesetze für Körperelemente.

5. *Distributivgesetz* für die Multiplikation mit Körperelementen. Für $\alpha, \beta \in K$ und $A, B \in K^{m\times n}$ gilt

$$\alpha(A + B) = \alpha A + \alpha B \quad \text{und} \quad (\alpha + \beta)A = \alpha A + \beta B.$$

Die Beweise sind wieder klar, da die Operationen komponentenweise definiert sind.

6. Für $\alpha \in K$, $A \in K^{m\times n}$ und $B \in K^{n\times p}$ gilt

$$\alpha(AB) = (\alpha A)B = A(\alpha B).$$

Auch hier ist der Beweis aufgrund der Definitionen klar.

1.3.3. Bemerkungen. 1. Die Matrizenmultiplikation kann man als Komposition von Abbildungen deuten. Dazu sei $A \in K^{m\times n}$. A gibt Anlaß zu der Abbildung

$$\begin{aligned} f_A\colon K^{n\times 1} &\to K^{m\times 1} \\ x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} &\mapsto Ax. \end{aligned}$$

Ebenso gibt $B \in K^{n\times r}$ Anlaß zu $f_B\colon K^{r\times 1} \to K^{n\times 1}$. Es gilt

$$f_A \circ f_B = f_{AB} \quad : K^{r\times 1} \to K^{m\times 1}.$$

Beweis. Für beliebige $y \in K^{r\times 1}$ ist

$$(f_A \circ f_B)(y) = f_A(f_B(y)) = f_A(By) = A(By) = (AB)y = f_{AB}(y).$$

Hierbei haben wir das Assoziativgesetz der Matrizenmultiplikation verwendet. Qed

2. Die Matrizenmultiplikation ist i.a. *nicht* kommutativ. Beispiel:

$$A := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \in K^{2\times 2}$$

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{also } AB \neq BA.$$

3. Es gibt Matrizen $A \neq 0$, $B \neq 0$ mit $AB = 0$. Ein Beispiel dafür haben wir eben gesehen.

4. Das Gleichungssystem (1.1) läßt sich in Matrizenschreibweise darstellen als $Ax = b$ mit

$$A = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} \in K^{m\times n}, \quad x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^{n\times 1}, \quad b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \in K^{m\times 1}.$$

1.4 Lineare Gleichungssysteme

Das elementare Lösungsverfahren für lineare Gleichungssysteme ist gut bekannt. Betrachten wir zum Beispiel über dem Körper $\mathbb{Q}$ das Gleichungssystem

$$\begin{aligned} 2\xi_1 + 3\xi_2 &= 18 \\ 4\xi_1 + \xi_2 &= 11. \end{aligned}$$

Multipliziert man die erste Zeile mit 2 und zieht man das Ergebnis von der zweiten Zeile ab, so ergibt sich

$$\begin{aligned} 2\xi_1 + 3\xi_2 &= 18 \\ -5\xi_2 &= -25. \end{aligned}$$

und daraus

$$\begin{aligned} \xi_2 &= \frac{-25}{-5} = 5 \\ 2\xi_1 + 15 &= 18 \\ \xi_1 &= \frac{18-15}{2} = \frac{3}{2}. \end{aligned}$$

Unser Ziel ist es, dieses elementare Lösungsverfahrens auf m Gleichungen mit n Unbekannten zu verallgemeinern. Insbesondere will man entscheiden, ob ein Gleichungssystem lösbar ist, und eine Übersicht über alle Lösungen bekommen.

1.4.1. Definition. Sei K ein Körper. $A \in K^{m\times n}$ heißt *Zeilenstufenmatrix*, wenn $A = 0$ oder

$$A = \begin{pmatrix} 0 & \dots & 0 & \alpha_{1j_1} & \dots & . & \alpha_{1j_2} & \dots & . & \alpha_{1j_r} & . & \dots & . & \alpha_{1n} \\ 0 & \dots & 0 & 0 & \dots & 0 & \alpha_{2j_2} & \dots & . & \alpha_{2j_r} & . & \dots & . & \alpha_{2n} \\ \vdots & & & & & & & & & & & & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \alpha_{rj_r} & . & \dots & . & \alpha_{rn} \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & & & & & & & & & & & & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

wobei $1 \leq r \leq m, n$, $1 \leq j_1 < j_2 < \ldots < j_r \leq n$ und $\alpha_{ij_i} \neq 0$ für alle i mit $1 \leq i \leq r$; diese α_{ij_i} heißen *Pivots* (deutsch *Angelpunkte*) von A.

Beispiele.

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

sind Zeilenstufenmatrizen, nicht aber

$$\begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

Definition. Sei K ein Körper, $A, A' \in K^{m\times n}$. A' entsteht aus A durch *elementare Zeilenumformung* (Spaltenumformung), wenn A aus A' hervorgeht durch

1. Vertauschen zweier Zeilen (Spalten), oder
2. Multiplikation einer Zeile (Spalte) mit einem $\lambda \in K$, $\lambda \neq 0$, oder
3. Addition einer Zeile (Spalte) zu einer anderen.

1.4.2. Satz. *(Lösungsverfahren für lineare Gleichungssysteme). Sei K ein Körper, $A \in K^{m\times n}$ und $b \in K^{m\times 1}$. Man betrachte die Gleichung (1.1), also $Ax = b$. Es wird ein Verfahren angegeben, das*

1. *entscheidet, ob* (1.1) *lösbar ist, und*
2. *falls es lösbar ist, eine Lösung $x_0 \in K^{n\times 1}$ liefert und weiter ein $r \in \mathbb{N}$ mit $0 \leq r \leq n$ und $x_1, \ldots, x_{n-r} \in K^{n\times 1}$ so, daß*

   a.
   $$X = \{\, x_0 + \lambda_1 x_1 + \ldots + \lambda_{n-r} x_{n-r} \mid \lambda_1, \ldots, \lambda_{n-r} \in K \,\}$$
   die Menge der Lösungen von (1.1) ist, und

   b. *die Darstellung der Lösungen in X eindeutig ist.*

Um den Bewies einfach führen zu können, stellen wir einige Hilfssätze bereit.

1.4.3. HILFSSATZ. *(Elementare Zeilenumformungen ändern nicht die Lösungsmenge).* $(A'\ b')$ *entstehe aus* $(A\ b)$ *durch elementare Zeilenumformungen. Dann gilt für alle* $x \in K^{n\times 1}$

$$Ax = b \equiv A'x = b'.$$

BEWEIS. OBdA entstehe $(A'\ b')$ aus $(A\ b)$ durch eine elementare Zeilenumformung.

Fall 1. Vertauschen zweier Zeilen von $(A\ b)$ verändert die Lösungsmenge nicht, da dieselben Bedingungen nur in anderer Reihenfolge aufgeschrieben werden.

Fall 2. Sei $\lambda \in K$, $\lambda \neq 0$. Multiplikation der i-ten Zeile von $(A\ b)$ mit λ ändert die Lösungsmenge nicht.

Fall 3. Sei $1 \leq s, t \leq m$, $s \neq t$. $(A'\ b')$ entstehe aus $(A\ b)$ durch Addition der s-ten Zeile von $(A\ b)$ zur t-ten. Die t-te Zeile hat dann folgende Gestalt:

$$\sum_{j=1}^{n} \alpha_{tj}\xi_j + \sum_{j=1}^{n} \alpha_{sj}\xi_j = \beta_t + \beta_s.$$

Alle anderen Zeilen von $A'x = b'$ bleiben unverändert. Zu zeigen ist

$$Ax = b \equiv A'x = b'.$$

$\Longrightarrow$. Gelte $Ax = b$. Dann ist

$$\begin{aligned} \sum_{j=1}^{n} \alpha_{tj}\xi_j &= \beta_t \qquad (t\text{-te Zeile}),\\ \sum_{j=1}^{n} \alpha_{sj}\xi_j &= \beta_s \qquad (s\text{-te Zeile}). \end{aligned}$$

Addition beider Zeilen ergibt

$$\sum_{j=1}^{n} \alpha_{tj}\xi_j + \sum_{j=1}^{n} \alpha_{sj}\xi_j = \beta_t + \beta_s.$$

$\Longleftarrow$. Gelte $A'x = b'$. Dann hat man

$$\begin{aligned} &\sum_{j=1}^{n} \alpha_{tj}\xi_j + \sum_{j=1}^{n} \alpha_{sj}\xi_j = \beta_t + \beta_s \quad && (t\text{-te Zeile}),\\ &\sum_{j=1}^{n} \alpha_{sj}\xi_j = \beta_s && (s\text{-te Zeile}). \end{aligned}$$

Subtraktion der s-ten Zeile von der t-ten Zeile ergibt

$$\sum_{j=1}^{n} \alpha_{tj}\xi_j = \beta_t. \quad \text{QED}$$

1.4.4. Der folgende Hilfssatz beschreibt das bekannte und schon im einleitenden Beispiel verwendete GAUSS*sche Eliminationsverfahren* in allgemeiner Form.

Hilfssatz. *(Reduktion auf Zeilenstufenmatrix). $(A\ b)$ läßt sich durch elementare Zeilenumformungen überführen in $(A'\ b')$, wobei A' Zeilenstufenmatrix ist.*

Beweis. OBdA sei $A \neq 0$. Wähle j_1 so, daß die ersten $j_1 - 1$ Spalten von A Null sind, aber die j_1-te nicht Null ist. Dann hat man mittels Zeilenvertauschung

$$(A \quad b) = \begin{pmatrix} 0 & \dots & 0 & \alpha_{1j_1} & \dots & \beta_1 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \alpha_{mj_m} & \dots & \beta_m \end{pmatrix} \mapsto \begin{pmatrix} 0 & \dots & 0 & \alpha'_{1j_1} & \dots & \beta'_1 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \alpha'_{mj_m} & \dots & \beta'_m \end{pmatrix},$$

wobei $\alpha'_{1j_1} \neq 0$. Multipliziert man die erste Zeile der neuen Matrix mit einem jeweils passenden $\lambda \in K$, $\lambda \neq 0$, und addiert das Ergebnis zur zweiten, dritten usw. Zeile, so erhält man eine neue Matrix der Gestalt

$$\begin{pmatrix} 0 & \dots & 0 & \alpha'_{1j_1} & \dots & \beta'_1 \\ 0 & \dots & 0 & 0 & & \\ \vdots & & \vdots & \vdots & \text{Rest} & \\ 0 & \dots & 0 & 0 & & \end{pmatrix}.$$

Man wende nun dasselbe Verfahren auf die Restmatrix an. Durch Iteration erhält man $(A'\ b')$, wobei A' Zeilenstufenmatrix ist. Qed

Im Satz können wir also oBdA annehmen, daß A Zeilenstufenmatrix ist. Wir haben deshalb

$$(A \quad b) = \begin{pmatrix} 0 & \dots & 0 & \alpha_{1j_1} & \dots & . & \alpha_{1j_2} & \dots & . & \alpha_{1j_r} & . & \dots & . & \alpha_{1n} & \beta_1 \\ 0 & \dots & 0 & 0 & \dots & 0 & \alpha_{2j_2} & \dots & . & \alpha_{2j_r} & . & \dots & . & \alpha_{2n} & \beta_2 \\ \vdots & & & & & & & & & & & & & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \alpha_{rj_r} & . & \dots & . & \alpha_{rn} & \beta_r \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \beta_{r+1} \\ \vdots & & & & & & & & & & & & & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \beta_m \end{pmatrix}.$$

Falls $\beta_i \neq 0$ für ein i mit $r < i \leq m$, so ist (1.1) unlösbar, da

$$\sum_{j=1}^{n} 0 \cdot \xi_j = \beta_j \neq 0$$

unmöglich ist. Im folgenden nehmen wir also $\beta_{r+1} = \dots = \beta_m = 0$ an.

1.4.5. Der folgende Hilfssatz sagt aus, daß die Komponenten der Lösung x, die im ausgeschriebene Gleichungssystem $Ax = b$ nicht an den Ecken der Stufenmatrix stehen, frei wählbar sind.

Hilfssatz. *Zu vorgegebenen ξ_j für $1 \leq j \leq n$, $j \notin \{j_1, \dots, j_r\}$ gibt es eindeutig bestimmte $\xi_{j_1}, \dots, \xi_{j_r}$ so daß*

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

die Gleichung $Ax = b$ erfüllt.

Beweis. Gegeben seien ξ_j, $1 \leq j \leq n$ mit $j \notin \{j_1, \ldots, j_r\}$. Wegen $\alpha_{rj_r} \neq 0$ (da A Zeilenstufenmatrix) ergibt sich ξ_{j_r} aus der r-ten Gleichung:

$$\xi_{j_r} = \Big(\beta_r - \sum_{j>j_r} \alpha_{rj}\xi_j\Big) \cdot \frac{1}{\alpha_{rj_r}}.$$

Ebenso ergibt sich $\xi_{j_{r-1}}$ aus der $(r-1)$-ten Gleichung und ξ_{j_r} und so weiter, schließlich also ξ_{j_1} aus der ersten Gleichung und $\xi_{j_2}, \ldots, \xi_{j_r}$. Qed

1.4.6. Der Beweis des Satzes wird nun in fünf Schritten geführt.

Schritt 1. (Konstruktion einer speziellen Lösung $x_0 \in K^{n\times 1}$). Setze $\xi_{0j} := 0$ für $1 \leq j \leq n$, $j \notin \{j_1, \ldots, j_r\}$ und konstruiere daraus

$$x_0 = \begin{pmatrix} \xi_{01} \\ \vdots \\ \xi_{0n} \end{pmatrix}$$

gemäß Hilfssatz 1.4.5 so daß $Ax_0 = b$.

Schritt 2. (Konstruktion von $x_1, \ldots, x_{n-r} \in K^{n\times 1}$ mit $Ax_j = 0$). Wir konstruieren dazu

$$x'_j = \begin{pmatrix} \xi'_{j1} \\ \vdots \\ \xi'_{jn} \end{pmatrix}$$

für $1 \leq j \leq n$, $j \notin \{j_1, \ldots, j_r\}$ mit $Ax'_j = 0$ und bilden daraus $x_1, \ldots, x_{n-r}$ durch Umnumerierung. Setze für $1 \leq \ell \leq n$, $\ell \notin \{j_1, \ldots, j_r\}$

$$\xi'_{j\ell} := \delta_{j\ell} := \begin{cases} 1 & \text{falls } j \neq \ell \\ 0 & \text{falls } j = \ell. \end{cases}$$

und konstruiere daraus x'_j gemäß Hilfssatz 1.4.5 so daß $Ax'_j = 0$.

Schritt 3. ($X \subseteq$ Lösungsmenge). Seien $\lambda_1, \ldots, \lambda_{n-r} \in K$, $x = x_0 + \sum_{j=1}^{n-r} \lambda_j x_j$. Dann ist x Lösung, denn

$$Ax = Ax_0 + \sum_{j=1}^{n-r} \lambda_j Ax_j = b + 0 = b.$$

Schritt 4. (Lösungsmenge $\subseteq X$). Sei $x \in K^{n\times 1}$ Lösung, also $Ax = b$. Zu zeigen ist, daß dann $\lambda'_1, \ldots, \lambda'_{n-r}$ existieren so daß

$$x = x_0 + \sum_{j=1}^{n-r} \lambda'_j x_j.$$

Setze

$$y = x - x_0 =: \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Dann gilt $Ay = Ax - Ax_0 = b - b = 0$. Es genügt also zu zeigen

$$y = \sum_{\substack{j=1,\dots,n\\ j\neq j_1,\dots,j_r}} \lambda_j x'_j =: z =: \begin{pmatrix}\mu_1\\ \vdots\\ \mu_n\end{pmatrix},$$

denn dann hat man y aus den x'_j, also aus den $x_1, \dots, x_{n-r}$ linear kombiniert. Es gilt

$$Az = \sum_{\substack{j=1,\dots,n\\ j\neq j_1,\dots,j_r}} \lambda_j \underbrace{Ax'_j}_{0} = 0.$$

Also haben wir $Ay = 0$ und $Az = 0$. Zum Beweis von $y = z$ genügt es nach dem Eindeutigkeitsteil von Hilfssatz 1.4.5 zu zeigen, daß und für $1 \leq \ell \leq n$ mit $\ell \notin \{j_1, \dots, j_r\}$ gilt $\lambda_\ell = \mu_\ell$. Für $1 \leq j, \ell \leq n$ mit $j, \ell \notin \{j_1, \dots, j_r\}$ gilt aber $\xi'_{j\ell} = \delta_{j\ell}$, also

$$\mu_\ell = \sum_{\substack{j=1,\dots,n\\ j\neq j_1,\dots,j_r}} \lambda_j \xi'_{j\ell} = \lambda_\ell.$$

Schritt 5. (Eindeutigkeit der Darstellung der Lösungen in X). Sei

$$x = x_0 + \sum_{\substack{j=1,\dots,n\\ j\neq j_1,\dots,j_r}} \lambda_j x'_j = x_0 + \sum_{\substack{j=1,\dots,n\\ j\neq j_1,\dots,j_r}} \mu_j x'_j$$

und $Ax = b$. Dann folgt

$$0 = \sum_{\substack{j=1,\dots,n\\ j\neq j_1,\dots,j_r}} (\lambda_j - \mu_j) x'_j.$$

Für $1 \leq j, \ell \leq n$ mit $j, \ell \notin \{j_1, \dots, j_r\}$ ergibt sich wegen $\xi'_{j\ell} = \delta_{j\ell}$

$$0 = \sum_{\substack{j=1,\dots,n\\ j\neq j_1,\dots,j_r}} (\lambda_j - \mu_j) \xi'_{j\ell} = \lambda_\ell - \mu_\ell. \quad \text{QED}$$

1.4.7. Beispiel. $K = \mathbb{Q}$ und

$$A = \begin{pmatrix} 0 & 2 & 3\\ 0 & 1 & 0\\ 0 & 2 & 0\end{pmatrix} \qquad b = \begin{pmatrix}1\\1\\2\end{pmatrix}.$$

Das Gleichungssystem $Ax = b$ lautet dann ausgeschrieben

$$\begin{aligned} 2\xi_2 + 3\xi_3 &= 1\\ \xi_2 &= 1\\ 2\xi_2 &= 2 \end{aligned}$$

Wir wollen das Gaussssche Lösungsverfahren hierauf anwenden. Zunächst bringen wir $(A\ b)$ durch elementare Zeilenumformungen in die Form einer Zeilenstufenmatrix:

$$\left(\begin{array}{ccc|c} 0 & 2 & 3 & 1\\ 0 & 1 & 0 & 1\\ 0 & 2 & 0 & 2\end{array}\right) \mapsto \begin{array}{c}(-2)\cdot \text{ zweite Zeile + erste Zeile}\\ (-1)\cdot \text{ dritte Zeile + erste Zeile}\end{array} \left(\begin{array}{ccc|c} 0 & 2 & 3 & 1\\ 0 & 0 & 3 & -1\\ 0 & 0 & 3 & -1\end{array}\right)$$

$$\mapsto (-1)\cdot \text{ dritte Zeile + zweite Zeile} \left(\begin{array}{ccc|c} 0 & 2 & 3 & 1\\ 0 & 0 & 3 & -1\\ 0 & 0 & 0 & 0\end{array}\right)$$

Also ist $r = 2$, $j_1 = 2$ und $j_2 = 3$. Wir erhalten deshalb

$$x_0 = \begin{pmatrix} 0 \\ 1 \\ -\frac{1}{3} \end{pmatrix}, \qquad x_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Die Lösungsmenge ist demnach

$$\{\, x \in \mathbb{Q}^3 \mid Ax = b \,\} = \{\, \begin{pmatrix} 0 \\ 1 \\ -\frac{1}{3} \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mid \lambda \in \mathbb{Q} \,\}.$$

1.4.8. FOLGERUNG. *Sei K ein Körper, $m < n$ und $A \in K^{m\times n}$. Dann existiert ein $x \in K^{n\times 1}$, $x \neq 0$, mit $Ax = 0$.*

BEWEIS. $Ax = 0$ ist lösbar, nämlich durch $x_0 = 0$. Nach dem Satz ist die Lösungsmenge

$$\{\, x_0 + \lambda_1 x_1 + \cdots + \lambda_{n-r} x_{n-r} \mid \lambda_1, \ldots, \lambda_{n-r} \in K \,\}.$$

Wegen $r \leq m, n$ und $m < n$ ist $n - r > m - r \geq 0$. Mit $x_1 \neq 0$ folgt die Behauptung. QED

1.4.9. Als weitere Folgerung geben wir jetzt ein entsprechendes Lösungsverfahren für die Matrizengleichung $AX = B$ an. Gegeben seien also $A \in K^{m\times n}$ und $B \in K^{m\times p}$. Gesucht ist ein $X \in K^{n\times p}$ mit

$$AX = B.$$

Seien $x_1, \ldots, x_p$ die Spalten von X und $b_1, \ldots, b_p$ die Spalten von B. Dann ist $AX = B$ gleichbedeutend mit

$$Ax_i = b_i \quad \text{für alle } i \text{ mit } 1 \leq i \leq p.$$

Man kann dann folgendes Lösungsverfahren verwenden. Durch elementare Zeilenumformungen bearbeite man $(A\ b_1\ \ldots\ b_p)$ solange bis A zu einer Zeilenstufenmatrix geworden ist. Die p entstehenden Gleichungssysteme löse man dann wie oben.

Interessant ist noch der folgende Spezialfall. Gegeben seien $A \in K^{n\times n}$ und $B \in K^{n\times n}$. Gesucht ist ein $X \in K^{n\times n}$ mit $AX = B$. Diesen Spezialfall werden wir gleich genauer betrachten. Zunächst benötigen wir aber noch eine Definition. Sei

$$A = (\alpha_{ij})_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} \in K^{m\times n}.$$

Dann nennt man

$$A^t := (\alpha_{ij})_{\substack{1\leq j\leq n \\ 1\leq i\leq m}} \in K^{n\times m}$$

die *transponierte Matrix* von A.

BEISPIEL.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 0 \\ 1 & 1 \end{pmatrix} \in \mathbb{Q}^{3\times 2} \qquad A^t = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & 1 \end{pmatrix} \in \mathbb{Q}^{2\times 3}$$

Bemerkungen. 1. Für $A, B \in K^{m\times n}$ und $\lambda \in K$ gilt

$$(A+B)^t = A^t + B^t,$$
$$(\lambda A)^t = \lambda A^t.$$

2. Für $A \in K^{m\times n}$ und $B \in K^{n\times p}$ gilt

$$(AB)^t = B^t A^t.$$

Beweis. (1) ist klar, da die Matrizenaddition und die Multiplikation einer Matrix mit einem Körperelement komponentenweise definiert sind. Zu (2) sei

$$A = (\alpha_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}, \qquad B = (\beta_{j\ell})_{\substack{1\le j\le n\\ 1\le \ell\le p}}.$$

Dann ist

$$A^t = (\alpha_{ij})_{\substack{1\le j\le n\\ 1\le i\le m}} \qquad B^t = (\beta_{j\ell})_{\substack{1\le \ell\le p\\ 1\le j\le n}}.$$

Es folgt

$$AB = \Big(\sum_{j=1}^{n} \alpha_{ij}\beta_{j\ell}\Big)_{\substack{1\le i\le m\\ 1\le \ell\le p}}, \qquad B^tA^t = \Big(\sum_{j=1}^{n} \beta_{j\ell}\alpha_{ij}\Big)_{\substack{1\le \ell\le p\\ 1\le i\le m}}.$$

Also ist $(AB)^t = B^tA^t$. Qed

1.4.10. Folgerung. *Für $A\in K^{n\times n}$ sind folgende Aussagen äquivalent.*

1. $\forall x\in K^{n\times 1}(Ax = 0 \to x = 0)$.
2. *A läßt sich durch elementare Zeilenumformungen in E überführen.*
3. $\exists X\in K^{n\times n}\ AX = E$.
4. $\forall y\in K^{1\times n}(yA = 0 \to y = 0)$.
5. *A läßt sich durch elementare Spaltenumformungen in E überführen.*
6. $\exists Y\in K^{n\times n}\ YA = E$.

In diesem Fall sind sowohl das X in (3) als auch das Y in (6) eindeutig bestimmt, und beide sind gleich. Das X in (3) läßt sich wie folgt berechnen:

$$(A\ E) \mapsto (E\ X) \qquad \text{durch elementare Zeilenumformungen.}$$

Hierbei sind $(A\ E)$ und $(E\ X)$ diejenigen $(n\times 2n)$-Matrizen, die man durch Nebeneinanderschreiben zweier $(n\times n)$-Matrizen erhält.

Beweis. 1 $\Longrightarrow$ 2. Nach Annahme ist $x = 0$ die einzige Lösung von $Ax = 0$. Das r im Lösungsverfahren des Satzes ist demnach $= n$. Daraus folgt (2).

2 $\Longrightarrow$ 3. Nach Annahme hat man

$$(A\ E) \mapsto (A'\ E') \qquad \text{durch elementare Zeilenumformungen}$$

mit $A' = E$. Seien $e_1, \dots, e_n$ die Spalten von E und $e'_1, \dots, e'_n$ die Spalten von E'. Dann gilt nach Hilfssatz 1.4.3 für $1 \le i \le n$

$$Ax_i = e_i \equiv \underbrace{A'x_i}_{x_i} = e'_i.$$

Mit $X := E'$ hat man also eine Lösung von $AX = E$.

$3 \implies 4$. Gelte $yA = 0$. Dann ist $y = yE = yAX = 0$.

$4 \implies 5$. Aus (4) folgt (1) mit A^t statt A, denn man hat

$$\begin{aligned} A^t x &= 0 \\ (A^t x)^t = x^t A &= 0 \\ x^t &= 0 \quad \text{nach } (4) \\ x &= 0. \end{aligned}$$

Da ($1 \implies 2$) schon bewiesen ist, folgt (5).

$5 \implies 6$. Aus (5) folgt (2) mit A^t statt A. Da ($2 \implies 3$) schon bewiesen ist, folgt $A^t X = E$, also $X^t A = E$.

$6 \implies 1$. Gelte $Ax = 0$. Dann hat man $x = Ex = YAx = 0$.

Im folgenden nehmen wir an, daß eine (und damit alle) der Bedingungen (1)-(6) zutrifft. Wir zeigen zunächst, daß das X in (3) eindeutig bestimmt ist:

$$\begin{aligned} AX_1 &= AX_2 = E \\ YAX_1 &= YAX_2 \\ X_1 &= X_2. \end{aligned}$$

Genauso zeigt man, daß das Y in (6) eindeutig bestimmt ist. Zu zeigen bleibt, daß beide gleich sind. Dies folgt aus

$$X = EX = YAX = YE = Y.$$

Die Richtigkeit des angegebenen Berechnungsverfahrens für X ergibt sich aus dem Beweis von ($2 \implies 3$). QED

DEFINITION. $A \in K^{n\times n}$ heißt *invertierbar*, wenn es ein $A^{-1} \in K^{n\times n}$ gibt mit

$$A^{-1}A = E.$$

Ist A invertierbar, so heißt das (eindeutig bestimmte) A^{-1} die *inverse Matrix* von A.

BEMERKUNGEN. 1. Gilt $A^{-1}A = E$, so ist nach obiger Folgerung auch $AA^{-1} = E$.

2. Sind A, B invertierbar, so auch AB.

BEWEIS. Setze $(AB)^{-1} := B^{-1}A^{-1}$. Dann ist

$$AB(AB)^{-1} = ABB^{-1}A^{-1} = E. \quad \text{QED}$$

3. Es gilt $(A^t)^{-1} = (A^{-1})^t$.

BEWEIS. $A^t(A^{-1})^t = (A^{-1}A)^t = E^t = E$. QED

Die Menge der invertierbaren $(n \times n)$-Matrizen über einem beliebigen Körper K bildet also bezüglich der Matrizenmultiplikation eine Gruppe. Man nennt sie die *allgemeine lineare Gruppe* und bezeichnet sie mit

$$\mathrm{GL}(n, K) := \{\, A \in K^{n\times n} \mid A \text{ invertierbar} \,\}.$$

(GL steht für "general linear").

Kapitel 2

Vektorräume

2.1 Vektorräume

Wir führen in diesem Abschnitt Vektorräume als Verallgemeinerung des $\mathbb{R}^n$ ein.

2.1.1. *Motivation.* Betrachten wir die anschaulich gegebene Ebene E_2 (sie wird also nicht mathematisch präzise definiert). In der Ebene können wir ein *Koordinatensystem* betrachten, in dem jeder Punkt P durch zwei Koordinaten $\xi_1(P)$ und $\xi_2(P)$ gegeben ist. Wir haben damit eine bijektive Abbildung

$$\begin{aligned} E_2 &\to \mathbb{R}^2 \quad (:= \mathbb{R}^{2\times 1}), \\ P &\mapsto \begin{pmatrix} \xi_1(P) \\ \xi_2(P) \end{pmatrix}. \end{aligned}$$

Ähnlich erhält man eine bijektive Abbildung von dem anschaulich gegeben Raum E_3 in den $\mathbb{R}^3$. Die Elemente von $\mathbb{R}^n$ für $n = 2, 3$ bezeichnet man oft als *Vektoren* und im Gegensatz dazu die Elemente von $\mathbb{R}$ als *Skalare.* Die *Addition* von Vektoren x und y kann man sich wie üblich graphisch als das Bilden der Diagonale im durch x und y aufgespannten Parallelogramm vorstellen. In der Koordinatendarstellung entspricht dies

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \vdots \\ \zeta_n \end{pmatrix} = \begin{pmatrix} \xi_1 + \zeta_1 \\ \vdots \\ \xi_n + \zeta_n \end{pmatrix}.$$

Ähnlich kann man sich die Skalarmultiplikation mit $\alpha \in \mathbb{R}$ als Streckung des Vektors x um den Faktor α vorstellen, also in Koordinatendarstellung

$$\alpha \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \alpha\xi_1 \\ \vdots \\ \alpha\xi_n \end{pmatrix}.$$

Folgende Eigenschaften sind dann offensichtlich.

1. $(\mathbb{R}^n, +)$ bildet eine abelsche Gruppe (sogar $(K^{m\times n}, +)$ ist eine). Das neutrale Element ist der Nullvektor

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

und

$$\begin{pmatrix} -\xi_1 \\ \vdots \\ -\xi_n \end{pmatrix} \text{ ist das inverse Element zu } \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}.$$

$x - y$ steht wie üblich für $x + (-y)$ und kann entsprechend graphisch konstruiert werden.

2. Es gilt

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

Für viele geometrische Untersuchungen kommt es nur auf diese Eigenschaften des $\mathbb{R}^n$ an. Sie geben Anlaß zur folgenden allgemeinen Definition eines Vektorraums.

2.1.2. Definition. Seien K ein Körper, V eine Menge und

$$\begin{aligned} +\colon V \times V &\to V \quad \text{(Vektor)addition,} \\ \cdot\colon K \times V &\to V \quad \text{Skalarmultiplikation} \end{aligned}$$

zwei Abbildungen. Wir schreiben $x + y$ für $+(x, y)$ und αx oder auch $\alpha \cdot x$ für $\cdot(\alpha, x)$. Ferner soll $\cdot$ stärker binden als $+$; zum Beispiel steht $\alpha x + \beta y$ für $(\alpha x) + (\beta y)$. V (oder genauer das Tripel $(V, +, \cdot)$) heißt *K-Vektorraum* (oder *Vektorraum über dem Körper K*), wenn

1. $(V, +)$ eine abelsche Gruppe ist, und
2. für alle $x, y \in V$ und alle $\alpha, \beta \in K$ gilt

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

Die Elemente $x, y, z \in V$ nennt man *Vektoren*, und die Elemente $\alpha, \beta, \gamma \in K$ heißen *Skalare*.

Wir schreiben wie üblich $x - y$ für $x + (-y)$ und nx für $x + x + \cdots + x$.

Beispiele. 1. K^n $(:= K^{n\times 1})$ mit folgender Addition und Skalarmultiplikation.

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \vdots \\ \zeta_n \end{pmatrix} := \begin{pmatrix} \xi_1 + \zeta_1 \\ \vdots \\ \xi_n + \zeta_n \end{pmatrix},$$

$$\alpha \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} := \begin{pmatrix} \alpha\xi_1 \\ \vdots \\ \alpha\xi_n \end{pmatrix}.$$

2. $\mathbb{R}[t]$, also die Menge aller Polynome in einer Variablen t mit Koeffizienten aus $\mathbb{R}$; Polynome werden hier als formale Ausdrücke (nicht als Funktionen) aufgefaßt. Addition und Skalarmultiplikation sind definiert durch

$$\begin{aligned}(\alpha_n t^n + \cdots + \alpha_1 t + \alpha_0) + (\beta_n t^n + \cdots + \beta_1 t + \beta_0) &= (\alpha_n+\beta_n)t^n + \cdots + (\alpha_1+\beta_1)t + (\alpha_0+\beta_0),\\ \alpha(\alpha_n t^n + \cdots + \alpha_1 t + \alpha_0) &= (\alpha\alpha_n t^n + \cdots + \alpha\alpha_1 t + \alpha\alpha_0).\end{aligned}$$

Das neutrale Elemente ist das Polynom 0, und das inverse Elemente zu $\alpha_n t^n + \cdots + \alpha_1 t + \alpha_0$ ist $-\alpha_n t^n - \cdots - \alpha_1 t - \alpha_0$.

3. Sei V ein K-Vektorraum und I eine nicht leere Indexmenge.

$$V^I := \{\, f \mid f\colon I \to V \text{ Abbildung} \,\}$$

wird ein K-Vektorraum mit

$$\begin{aligned}(f+g)(i) &:= f(i) + g(i),\\ (\alpha f)(i) &:= \alpha \cdot f(i).\end{aligned}$$

4.

$$\begin{aligned}C(\mathbb{R}) &:= \{\, f \mid f\colon \mathbb{R} \to \mathbb{R} \text{ stetig} \,\}\\ D(\mathbb{R}) &:= \{\, f \mid f\colon \mathbb{R} \to \mathbb{R} \text{ differenzierbar} \,\}\end{aligned}$$

sind $\mathbb{R}$-Vektorräume, wenn man Addition und Skalarmultiplikation komponentenweise definiert.

5. $\mathbb{R}$ ist ein $\mathbb{Q}$-Vektorraum.

Bemerkung. Sei V ein K-Vektorraum. Dann gilt für alle $\alpha \in K$, $x \in V$:

$$\begin{aligned}0x &= 0.\\ \alpha 0 &= 0.\\ \alpha x &\neq 0, \quad \text{falls } \alpha \neq 0 \text{ und } x \neq 0\\ (-1)x &= -x.\end{aligned}$$

Beweis. Aus den Axiomen für Vektorräume erhält man

$$\begin{aligned}0x &= (0+0)x = 0x + 0x.\\ \alpha 0 &= \alpha(0+0) = \alpha 0 + \alpha 0.\end{aligned}$$

Seien $\alpha \neq 0$ und $x \neq 0$. Annahme: $\alpha x = 0$. Dann hat man

$$x = 1x = (\alpha^{-1}\alpha)x = \alpha^{-1}(\alpha x) = \alpha^{-1}0 = 0,$$

also einen Widerspruch. Für die letzte Behauptung ergibt sich

$$x + (-1)x = 1x + (-1)x = (1 + (-1))x = 0x = 0. \quad \text{Qed}$$

2.2 Lineare Unabhängigkeit, Basis, Dimension

2.2.1. *Motivation.* Wir betrachten den Spezialfall des $\mathbb{R}^3$. Zwei Vektoren $x, y \in \mathbb{R}^3$ heißen linear unabhängig, wenn für alle $\alpha, \beta \in \mathbb{R}$

$$\text{aus } \alpha x + \beta y = 0 \text{ folgt } \alpha = \beta = 0.$$

$x, y \in \mathbb{R}^3$ heißen linear abhängig, wenn das nicht der Fall ist, also wenn es $\alpha, \beta \in \mathbb{R}$ gibt, die nicht beide verschwinden, so daß

$$\alpha x + \beta y = 0.$$

Wir wollen uns überlegen, daß für $x, y \in \mathbb{R}^3$ die folgenden Aussagen äquivalent sind.

1. x, y sind linear abhängig.
2. $\exists \beta \in \mathbb{R}\,(x = \beta y)$ oder $\exists \alpha \in \mathbb{R}\,(y = \alpha x)$.

Beweis. $1 \Longrightarrow 2$. Gelte $\alpha x + \beta y = 0$, wobei α und β nicht beide verschwinden. Im Fall $\alpha \neq 0$ hat man dann $x = -\frac{\beta}{\alpha} y$, und im Fall $\beta \neq 0$ ist $y = -\frac{\alpha}{\beta} x$.

$2 \Longrightarrow 1$. Sei etwa $x = \beta y$. Dann ist $1x - \beta y = 0$. Qed

Bemerkung. x, y sind linear unabhängig genau dann, wenn $\mathbb{R}x + \mathbb{R}y := \{\, \alpha x + \beta y \mid \alpha, \beta \in \mathbb{R} \,\}$ eine Ebene durch 0 ist, und x, y sind linear abhängig genau dann, wenn $\{\, \alpha x + \beta y \mid \alpha, \beta \in \mathbb{R} \,\}$ eine Gerade durch 0 oder $= \{0\}$ ist.

Entsprechende Überlegungen kann man auch für drei Vektoren anstellen. $x, y, z \in \mathbb{R}^3$ heißen linear unabhängig, wenn für alle $\alpha, \beta, \gamma \in \mathbb{R}$

$$\text{aus } \alpha x + \beta y + \gamma z = 0 \text{ folgt } \alpha = \beta = \gamma = 0.$$

$x, y, z \in \mathbb{R}^3$ heißen linear abhängig, wenn das Gegenteil der Fall ist, also wenn es $\alpha, \beta, \gamma \in \mathbb{R}$ gibt, die nicht alle verschwinden, so daß

$$\alpha x + \beta y + \gamma z = 0.$$

Wie eben beweist man, daß für $x, y, z \in \mathbb{R}^3$ die folgenden Aussagen äquivalent sind.

1. x, y, z sind linear abhängig.
2. $\exists \beta, \gamma \in \mathbb{R}\,(x = \beta y + \gamma z)$ oder $\exists \alpha, \gamma \in \mathbb{R}\,(y = \alpha x + \gamma z)$ oder $\exists \alpha, \beta \in \mathbb{R}\,(z = \alpha x + \beta y)$.

2.2.2. Wir definieren jetzt allgemein die Begriffe der linearen Unabhängigkeit und der linearen Abhängigkeit für Familien von Vektoren.

Definition. Sei V ein K-Vektorraum und $x_1, \ldots, x_n \in V$. Die Familie $(x_1, \ldots, x_n)$ heißt *linear unabhängig*, wenn sich der Nullvektor nur auf die triviale Weise aus $x_1, \ldots, x_n$ linear kombinieren läßt, d.h. wenn für alle $\alpha_1, \ldots, \alpha_n \in K$

$$\text{aus } \alpha_1 x_1 + \cdots + \alpha_n x_n = 0 \text{ folgt } \alpha_1 = \cdots = \alpha_n = 0.$$

Die Familie $(x_1, \ldots, x_n)$ heißt *linear abhängig*, wenn sie nicht linear unabhängig ist, also wenn sich der Nullvektor auf nicht triviale Weise aus $x_1, \ldots, x_n$ linear kombinieren läßt, d.h. wenn es $\alpha_1, \ldots, \alpha_n \in K$ gibt, die nicht alle verschwinden, so daß

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = 0.$$

Schreibweise. Man schreibt oft einfach, daß $x_1, \dots, x_n$ linear (un)abhängig sind, wenn man meint, daß die Familie $(x_1, \dots, x_n)$ linear (un)abhängig ist.

Definition. Sei V ein K-Vektorraum, I eine Indexmenge und $(x_i)_{i \in I}$ eine Familie von Vektoren aus V. Ist I unendlich, so heißt $(x_i)_{i \in I}$ *linear unabhängig*, wenn jede endliche Teilfamilie linear unabhängig ist, und *linear abhängig*, wenn es eine linear abhängige endliche Teilfamilie gibt. Ist $I = \emptyset$, so nennt man die leere Familie $(x_i)_{i \in \emptyset}$ *linear unabhängig.*

Beispiele. 1. Sei $e_i \in \mathbb{R}^n$ derjenige Vektor, der in der i-ten Zeile eine 1 und sonst 0 als Einträge hat. e_i heißt der i-te *Einheitsvektor* im $\mathbb{R}^n$. Dann sind im $\mathbb{R}^n$ die Vektoren $e_1, \dots, e_n$ linear unabhängig, denn aus

$$\alpha_1 e_1 + \dots + \alpha_n e_n = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = 0$$

folgt offenbar $\alpha_1 = \dots = \alpha_n = 0$.

2. In $\mathbb{R}[t]$ ist die Familie

$$(1, t, t^2, t^3, \dots)$$

linear unabhängig. Zum Beweis genügt es zu zeigen, daß für jedes n die Familie $(1, t, \dots, t^n)$ linear unabhängig ist. Gelte also $\alpha_0 + \alpha_1 t + \dots + \alpha_n t^n = 0$. Dann ist offenbar $\alpha_0 = \alpha_1 = \dots = \alpha_n = 0$.

Bemerkung. Die lineare Unabhängigkeit im K^n läßt sich wie folgt mittels Matrizen beschreiben. Seien $A \in K^{m \times n}$ und $a_1, \dots, a_n \in K^{m \times 1}$ die Spalten von A. Dann sind

$$a_1, \dots, a_n \text{ linear unabhängig } \equiv \forall x \in K^{n \times 1} \, (Ax = 0 \to x = 0).$$

Beweis.

$$\begin{aligned} Ax = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} &= \begin{pmatrix} \alpha_{11}\xi_1 + \dots + \alpha_{1n}\xi_n \\ \vdots \\ \alpha_{m1}\xi_1 + \dots + \alpha_{mn}\xi_n \end{pmatrix} \\ &= \xi_1 \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{pmatrix} + \dots + \xi_n \begin{pmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{mn} \end{pmatrix} \\ &= \xi_1 a_1 + \dots + \xi_n a_n. \end{aligned}$$

Daraus folgt die Behauptung aufgrund der Definitionen.

Definition. Sei V ein K-Vektorraum und $x_1, \dots, x_n \in V$. Ein $x \in V$ heißt *Linearkombination* der Familie $(x_1, \dots, x_n)$, wenn es $\alpha_1, \dots, \alpha_n \in K$ gibt so daß $x = \alpha_1 x_1 + \dots + \alpha_n x_n$.

Bemerkung. Aus der leeren Familie läßt sich 0, aber kein anderer Vektor linear kombinieren.

2.2.3. LEMMA. *Sei V ein K-Vektorraum und $x_1, \dots, x_n \in V$.*

1. *$x_1, \dots, x_n$ sind linear abhängig genau dann, wenn sich ein x_i aus den übrigen linear kombinieren läßt.*

2. *Seien $y_1, \dots, y_m \in V$ linear unabhängig. Dann sind $y_1, \dots, y_m, x_1, \dots, x_n$ linear abhängig genau dann, wenn sich ein x_i aus $y_1, \dots, y_m, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ linear kombinieren läßt.*

BEWEIS. (1) ist ein Spezialfall von (2), da die leere Familie linear unabhängig ist. Es genügt also, (2) zu beweisen.

$\Longrightarrow$. Gelte $\beta_1 y_1 + \dots + \beta_m y_m + \alpha_1 x_1 + \dots + \alpha_n x_n = 0$, wobei $\beta_1, \dots, \beta_m, \alpha_1, \dots, \alpha_n$ nicht alle verschwinden. Dann muß ein $\alpha_i \neq 0$ sein, denn $y_1, \dots, y_m$ sind linear unabhängig. Sei etwa $\alpha_1 \neq 0$. Man erhält

$$x_1 = -\frac{\beta_1}{\alpha_1} y_1 - \dots - \frac{\beta_m}{\alpha_1} y_m - \frac{\alpha_2}{\alpha_1} x_2 - \dots - \frac{\alpha_n}{\alpha_1} x_n.$$

$\Longleftarrow$. Sei etwa $x_1 = \beta_1 y_1 + \dots + \beta_m y_m + \alpha_2 x_2 + \dots + \alpha_n x_n$. Dann ist

$$x_1 - \beta_1 y_1 - \dots - \beta_m y_m - \alpha_2 x_2 - \dots - \alpha_n x_n = 0.$$

QED

2.2.4. DEFINITION. Sei V ein K-Vektorraum, I eine Indexmenge und $(x_i)_{i \in I}$ eine Familie von Vektoren aus V. $(x_i)_{i \in I}$ heißt *Basis* von V, wenn

a. $(x_i)_{i \in I}$ linear unabhängig ist und

b. $(x_i)_{i \in I}$ ein *Erzeugendensystem* von V ist, d.h. jedes $x \in V$ sich aus einer endlichen Teilfamilie von $(x_i)_{i \in I}$ linear kombinieren läßt.

V heißt *endlichdimensional*, wenn V eine endliche Basis besitzt.

BEISPIELE. 1. $e_1, \dots, e_n$ ist eine Basis des K^n.

BEWEIS. Daß $e_1, \dots, e_n$ linear unabhängig sind, wurde bereits oben bemerkt. Zu zeigen bleibt, daß $e_1, \dots, e_n$ ein Erzeugendensystem ist. Sei also

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^n.$$

Dann gilt

$$x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \vdots \\ \xi_n \end{pmatrix} = \xi_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \xi_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + \xi_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \xi_1 e_1 + \xi_2 e_2 + \dots + \xi_n e_n.$$

2. $(t_n)_{n \in \mathbb{N}}$ ist Basis von $\mathbb{R}[t]$.

BEWEIS. Daß $(t_n)_{n \in \mathbb{N}}$ linear unabhängig sind, wurde bereits oben bemerkt. Ferner ist jedes Polynom $\alpha_0 + \alpha_1 t + \dots + \alpha_n t^n \in \mathbb{R}[t]$ Linearkombination von endlich vielen t^m.

3. $\emptyset$ ist Basis von 0, dem nur aus dem Nullelement bestehenden Vektorraum.

2.2.5. Wir geben jetzt einige nützliche Charakterisierungen des Basisbegriffs.

SATZ. *Sei V ein K-Vektorraum und $\mathcal{B} = (x_1, \ldots, x_n)$ eine Familie von Vektoren aus V. Dann sind folgende Aussagen äquivalent.*

1. *$\mathcal{B}$ ist Basis, d.h. linear unabhängiges Erzeugendensystem von V.*
2. *$\mathcal{B}$ ist maximal linear unabhängig, d.h. $\mathcal{B}$ ist linear unabhängig und $(x_1, \ldots, x_n, y)$ ist linear abhängig für jedes $y \in V$.*
3. *$\mathcal{B}$ ist eindeutiges Erzeugendensystem von V, d.h. $\mathcal{B}$ ist Erzeugendensystem und für jedes $x \in V$ ist die Darstellung $x = \alpha_1 x_1 + \cdots + \alpha_n x_n$ eindeutig.*
4. *$\mathcal{B}$ ist minimales Erzeugendensystem von V, d.h. $\mathcal{B}$ ist Erzeugendensystem und kein x_i ist aus den übrigen linear kombinierbar.*

BEWEIS. $1 \Longrightarrow 2$. Sei $x_1, \ldots, x_n$ Basis und $y \in V$. Dann ist y Linearkombination von $x_1, \ldots, x_n$, also nach Lemma 2.2.3(1) $x_1, \ldots, x_n, y$ linear abhängig.

$2 \Longrightarrow 3$. Wir zeigen zunächst, daß $x_1, \ldots, x_n$ Erzeugendensystem von V ist. Sei also $y \in V$. Dann ist nach Annahme $x_1, \ldots, x_n, y$ linear abhängig, also nach Lemma 2.2.3(2) wegen der linearen Unabhängigkeit von $x_1, \ldots, x_n$ der Vektor y Linearkombination von $x_1, \ldots, x_n$. Wir zeigen jetzt, daß für jedes $x \in V$ die Darstellung $x = \alpha_1 x_1 + \cdots + \alpha_n x_n$ eindeutig ist. Gelte also $\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta_1 x_1 + \cdots + \beta_n x_n$. Dann ist $(\alpha_1 - \beta_1)x_1 + \cdots + (\alpha_n - \beta_n)x_n = 0$, also $\alpha_1 - \beta_1 = \cdots = \alpha_n - \beta_n = 0$.

$3 \Longrightarrow 4$. Wäre etwa x_1 aus den übrigen Vektoren linear kombinierbar, so hätte man $x_1 = \alpha_2 x_2 + \cdots + \alpha_n x_n$, also denselben Vektor auf zwei verschiedene Weisen aus $x_1, \ldots, x_n$ linear kombiniert.

$4 \Longrightarrow 1$. Die Behauptung folgt unmittelbar aus Lemma 2.2.3(1). QED

2.2.6. Wir wollen jetzt die Längen verschiedener Basen miteinander vergleichen. Dazu werden uns unsere Kenntnisse über lineare Gleichungssysteme nützlich sein.

SATZ. *Seien $x_1, \ldots, x_n$ linear unabhängig und $y_1, \ldots, y_m$ ein Erzeugendensystem von V. Dann ist $n \leq m$.*

BEWEIS. Wir nehmen wir $m < n$ an und leiten daraus einen Widerspruch her. Für $1 \leq j \leq n$ ist

$$x_j = \sum_{i=1}^{m} \alpha_{ij} y_i$$

für eine Matrix $A = (\alpha_{ij}) \in K^{m \times n}$. Wegen der Annahme $m < n$ hat das durch A bestimmte homogene lineare Gleichungssystem mehr Unbekannte als Gleichungen. Nach 1.4.8 existieren deshalb $\beta_1, \ldots, \beta_n \in K$, die nicht alle verschwinden, mit

$$A \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = 0, \quad \text{also} \quad \sum_{j=1}^{n} \alpha_{ij} \beta_j = 0.$$

Also ist

$$\sum_{j=1}^{n} \beta_j x_j = \sum_{j=1}^{n} \beta_j \sum_{i=1}^{m} \alpha_{ij} y_i = \sum_{i=1}^{m} \Bigl(\sum_{j=1}^{n} \alpha_{ij} \beta_j \Bigr) y_i = 0.$$

Damit haben wir einen Widerspruch zur linearen Unabhängigkeit von $x_1, \ldots, x_n$. QED

FOLGERUNG. *In einem endlichdimensionalen K-Vektorraum haben je zwei Basen dieselbe endliche Länge.*

Wir können deshalb die *Dimension* eines K-Vektorraums V ist wie folgt definieren.

$$\dim(V) := \begin{cases} 0 & \text{falls } V = 0, \\ n & \text{falls eine Basis der Länge } n \text{ von } V \text{ existiert,} \\ \infty & \text{sonst.} \end{cases}$$

2.2.7. Wir zeigen jetzt den *Basisergänzungssatz*. Er sagt aus, daß sich in einem endlichdimensionalen Vektorraum jede linear unabhängige Familie aus einem gegebenen Erzeugendensystem heraus zu einer Basis ergänzen läßt.

SATZ. *Sei V ein endlichdimensionaler K-Vektorraum, ferner $x_1, \dots, x_n$ linear unabhängig und $y_1, \dots, y_m$ ein Erzeugendensystem. Dann gibt es $i_1, \dots, i_k$ so daß $x_1, \dots, x_n, y_{i_1}, \dots, y_{i_k}$ eine Basis von V ist.*

BEWEIS. Wir zeigen die folgende Aussage, aus der offensichtlich die Behauptung folgt: Sind $x_1, \dots, x_n$ linear unabhängig und ist $x_1, \dots, x_n, y_1, \dots, y_m$ ein Erzeugendensystem, so gibt es $i_1, \dots, i_k$ derart daß $x_1, \dots, x_n, y_{i_1}, \dots, y_{i_k}$ eine Basis von V ist.

Den Beweis führen wir durch Induktion über m. Im Fall $m = 0$ ist nichts zu zeigen. Sei also $m > 0$. Sind $x_1, \dots, x_n, y_1, \dots, y_m$ linear unabhängig, so ist wieder nichts zu zeigen. Andernfalls ist nach Lemma 2.2.3(2) ein y_i Linearkombination der übrigen Vektoren, etwa y_m. Dann ist $x_1, \dots, x_n, y_1, \dots, y_{m-1}$ ein Erzeugendensystem, und die Induktionsvoraussetzung liefert die Behauptung. QED

Insbesondere haben wir (mit $n = 0$) damit gezeigt, daß jeder endlich erzeugte Vektorraum eine endliche Basis besitzt.

2.2.8. Wir wollen jetzt unter Verwendung mengentheoretischer Hilfsmittel – dem sogenannten ZORNschen Lemma – zeigen, daß der Basisergänzungssatz auch ohne die Voraussetzung der Endlichdimensionalität von V richtig ist. Insbesondere haben wir dann gezeigt, daß jeder Vektorraum eine Basis besitzt.

Zunächst formulieren wir das ZORNsche Lemma; dazu sind einige Vorbetrachtungen über partiell geordnete Mengen erforderlich. Eine Relation $\leq$ auf einer nicht leeren Menge M heißt *partielle Ordnung*, wenn für alle $a, b, c \in M$ gilt

1. $a \leq a$ (*Reflexivität*).
2. Wenn $a \leq b$ und $b \leq a$, so ist $a = b$ (*Antisymmetrie*).
3. Wenn $a \leq b$ und $b \leq c$, so ist $a \leq c$ (*Transitivität*).

Gilt zusätzlich $a \leq b$ oder $b \leq a$ für alle $a, b \in M$, so heißt $\leq$ eine *lineare Ordnung* von M (man spricht auch von einer *totalen* oder *vollständigen* Ordnung).

BEISPIEL. Auf der Potenzmenge $\mathcal{P}(M)$ einer beliebigen Menge M ist die Teilmengenrelation $\subseteq$ eine partielle, aber im allgemeinen keine lineare Ordnung.

Sei $\leq$ eine partielle Ordnung auf M. Ein Element $b \in M$ heißt *maximales Element* von M, wenn es kein von b verschiedenes $a \in M$ gibt mit $b \leq a$. Sei noch $X \subseteq M$. Ein Element $b \in M$ heißt *obere Schranke* von X, wenn für alle $a \in X$ gilt $a \leq b$. Weiter heißt $b \in M$ *kleinste obere Schranke* von X (oder *Supremum* von X), wenn

1. b obere Schranke von X ist und

2. für jede andere obere Schranke c von X gilt $b \leq c$.

Man beachte, daß es (wegen der Antisymmetrie) höchstens eine kleinste obere Schranke von X geben kann.

LEMMA. *(ZORNsches Lemma). Sei $\leq$ eine partielle Ordnung auf M. Jede durch Einschränkung von $\leq$ linear geordnete Teilmenge $X \subseteq M$ besitze ein Supremum in M. Dann gibt es ein maximales Element in M.*

BEWEIS. Man benötigt das *Auswahlsaxiom* (es läß sich sogar zeigen, daß auf der Basis der übrigen Axiome der Mengenlehre das ZORNsche Lemma äquivalent zum Auswahlsaxiom ist); für den genauen Beweis muß auf Vorlesungen über Mathematische Logik oder Mengenlehre verwiesen werden. QED

Um das ZORNsche Lemma auf unser Problem anwenden zu können, ist es bequem, auch für Mengen (statt Familien) die Begriffe der linearen (Un)abhängigkeit und eines Erzeugendensystems zu verwenden. Offenbar übertragen sich die Definitionen unverändert, und es gilt

1. Eine Familie $(x_i)_{i\in I}$ aus *verschiedenen* Vektoren x_i ist genau dann linear unabhängig, wenn es die Menge $\{\, x_i \mid i \in I \,\}$ ist.

2. Eine Familie $(x_i)_{i\in I}$ ist ein Erzeugendensystem genau dann, wenn es die Menge $\{\, x_i \mid i \in I \,\}$ ist.

SATZ. *Sei V ein K-Vektorraum, ferner $X \subseteq V$ linear unabhängig und $Y \subseteq V$ ein Erzeugendensystem. Dann gibt es eine Teilmenge $Y_0 \subseteq Y$ so daß $X \cup Y_0$ eine Basis von V ist.*

BEWEIS. Wir wenden das ZORNsche Lemma an auf das Mengensystem

$$\mathcal{S} := \{\, Z \mid X \subseteq Z \subseteq X \cup Y \text{ und } Z \text{ linear unabhängig} \,\}$$

mit der Teilmengenrelation als partieller Ordnung. Dazu ist die Bedingung zu überprüfen, daß jedes linear geordnete Teilsystem $\mathcal{T} \subseteq \mathcal{S}$ ein Supremum in $\mathcal{S}$ besitzt. Sei also $\mathcal{T} \subseteq \mathcal{S}$ linear geordnet. Wir zeigen, daß dann $\bar{Z} := \bigcup_{Z\in\mathcal{T}} Z$ ein Supremum von $\mathcal{T}$ in $\mathcal{S}$ ist. Offenbar gilt

$$X \subseteq \bar{Z} \subseteq X \cup Y,$$

da dies für jedes einzelne Z der Vereinigung erfüllt ist. Zu zeigen bleibt die lineare Unabhängigkeit von $\bar{Z}$. Seien also $z_1, \ldots, z_n \in \bar{Z}$ und nehmen wir an, daß $\alpha_1 z_1 + \cdots + \alpha_n z_n = 0$. Dann gibt es $Z_1, \ldots, Z_n \in \mathcal{T}$ mit $z_i \in Z_i$ für $i = 1, \ldots, n$. Da $\mathcal{T}$ linear geordnet ist, muß eines dieser Z_i das größte sein, etwa Z_1. Dann gilt also $z_1, \ldots, z_n \in Z_1$ und damit $\alpha_1 = \cdots = \alpha_n = 0$, denn Z_1 ist linear unabhängig. – Die Bedingung des ZORNschen Lemmas ist also erfüllt und wir erhalten, daß es ein maximales $Z_0 \in \mathcal{S}$ gibt.

Zu zeigen ist, daß Z_0 Erzeugendensystem ist. Sei also $x \in V$. Da Y Erzeugendensystem ist, läßt sich x schreiben in der Form $x = \beta_1 y_1 + \cdots + \beta_n y_n$ mit $y_j \in Y$. Es genügt also zu zeigen, daß alle y_j aus Z_0 linear kombinierbar sind. Wäre ein y_j *nicht* aus Z_0 linear kombinierbar, so wäre $Z_0 \cup \{y_j\}$ linear unabhängig und wir hätten $Z_0 \subsetneq Z_0 \cup \{y_0\} \subseteq X \cup Y$. Dies widerspricht aber der Maximalität von Z_0. QED

2.2.9. SATZ. *Sei V ein n-dimensionaler K-Vektorraum.*

1. *Je $n+1$ Vektoren aus V sind linear abhängig.*
2. *Seien $x_1, \ldots, x_n \in V$. Folgende Aussagen sind äquivalent.*
   a. *$x_1, \ldots, x_n$ ist Basis von V.*
   b. *$x_1, \ldots, x_n$ ist linear unabhängig.*
   c. *Jedes $x \in V$ ist aus $x_1, \ldots, x_n$ linear kombinierbar.*

BEWEIS. 1. Wären $x_1, \ldots, x_{n+1}$ linear unabhängig, so müßten sie zu einer Basis von V ergänzbar sein, was nicht der Fall sein kann.

2. a $\Longrightarrow$ b ist klar.

b $\Longrightarrow$ c. Seien $x_1, \ldots, x_n$ linear unabhängig und $x \in V$. Dann sind $x_1, \ldots, x_n, x$ linear abhängig nach Teil (1), und der Koeffizient von x ist von 0 verschieden.

c $\Longrightarrow$ a. Jedes $x \in V$ sei aus $x_1, \ldots, x_n$ linear kombinierbar. Zu zeigen ist, daß $x_1, \ldots, x_n$ linear unabhängig sind. Gelte

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = 0.$$

Zu zeigen ist $\alpha_1 = \cdots = \alpha_n = 0$. Sei etwa $\alpha_n \neq 0$. Dann ist

$$x_n = -\frac{\alpha_1}{\alpha_n} x_1 - \cdots - \frac{\alpha_{n-1}}{\alpha_n} x_{n-1}.$$

Also wäre auch $x_1, \ldots, x_{n-1}$ ein Erzeugendensystem. Dies widerspricht aber dem oben bewiesenen Satz 2.2.6. QED

2.2.10. AUFGABE. Man beweise den folgenden *Austauschsatz*: Es seien V ein K-Vektorraum, $x_1, \ldots, x_n$ eine Basis von V und $y_1, \ldots, y_m$ linear unabhängig. Dann ist $m \leq n$, und es gibt verschiedene $i_1, \ldots, i_m$ so daß nach Austausch von x_{i_1} durch y_1, x_{i_2} durch y_2 und schließlich x_{i_m} durch y_m wieder eine Basis entsteht. QED

BEWEIS. $m \leq n$ gilt nach dem Längenvergleichsatz 2.2.6. Die restliche Aussage folgt aus dem Basisergänzungssatz 2.2.7. Man ergänze $y_1, \ldots, y_m$ aus dem Erzeugendensystem $x_1, \ldots, x_n$ heraus zu einer Basis. Deren Länge muß n sein. Wir haben also eine Basis, die aus $x_1, \ldots, x_n$ durch Austausch von x_{i_1} durch y_1, x_{i_2} durch y_2 und schließlich x_{i_m} durch y_m entsteht.

BEMERKUNG. Man kann die Anwendung des Basisergänzungssatzes auch vermeiden und den Beweis (der restlichen Aussage) durch Induktion nach m führen. *Basis* $m = 0$. Klar. *Schritt* $m-1 \Longrightarrow m$. Anwendung der Induktionsvoraussetzung auf $y_1, \ldots, y_{m-1}$ liefert oBdA eine Basis $y_1, \ldots, y_{m-1}, x_m, \ldots, x_n$. Dann hat man

$$y_m = \beta_1 y_1 + \cdots + \beta_{m-1} y_{m-1} + \alpha_m x_m + \cdots + \alpha_n x_n. \tag{2.1}$$

Wegen der linearen Unabhängigkeit von $y_1, \ldots, y_m$ muß ein $\alpha_i \neq 0$ sein; etwa $\alpha_m \neq 0$. Also ist $y_1, \ldots, y_{m-1}, y_m, x_{m+1}, \ldots, x_n$ ein Erzeugendensystem. Zu zeigen bleibt die lineare Unabhängigkeit. Gelte also

$$\mu_1 y_1 + \cdots + \mu_{m-1} y_{m-1} + \mu_m y_m + \lambda_{m+1} x_{m+1} + \cdots + \lambda_n x_n = 0.$$

Einsetzen von (2.1) ergibt

$$\begin{aligned}(\mu_1 + \mu_m \beta_1) y_1 + \cdots + (\mu_{m-1} + \mu_m \beta_{m-1}) y_{m-1} +\\ \mu_m \alpha_m x_m + (\lambda_{m+1} + \mu_m \alpha_{m+1}) x_{m+1} + \cdots + (\lambda_n + \mu_m \alpha_n) x_n = 0.\end{aligned}$$

Aus der linearen Unabhängigkeit von $y_1, \dots, y_{m-1}, x_m, \dots, x_n$ folgt zunächst $\mu_m \alpha_m = 0$, also wegen $\alpha_m \neq 0$ auch $\mu_m = 0$, und weiter durch Einsetzen von $\mu_m = 0$ in die restlichen Koeffizienten auch $\mu_1 = \cdots = \mu_{m-1} = \lambda_{m+1} = \cdots = \lambda_n = 0$.

2.2.11. Das folgende *Austauschlemma* kann etwa in der Graphentheorie benutzt werden zum Beweis, daß der *Kruskal-Algorithmus* in einem endlichen zusammenhängenden bewerteten Graphen einen minimalen aufspannenden Baum bestimmt.

AUFGABE. In einem endlichdimensionalen K-Vektorraum seien $x_1, \dots, x_n$ linear unabhängig und auch $y_1, \dots, y_m$ linear unabhängig, mit $n < m$. Dann gibt es ein j so daß $x_1, \dots, x_n, y_j$ linear unabhängig sind.

BEWEIS. Man ergänze $y_1, \dots, y_m$ zu einer Basis $y_1, \dots, y_m, y_{m+1}, \dots, y_k$. Dann befördere man mit Hilfe des Austauschsatzes 2.2.10 die Vektoren $x_1, \dots, x_n$ hinein. Wegen $n < m$ bleibt ein y_j aus $y_1, \dots, y_m$ übrig. QED

2.3 Untervektorräume, direkte Summen und Quotienten

Eine Teilmenge E des $\mathbb{R}^3$ nennt man eine *Ebene* durch 0, wenn es linear unabhängige Vektoren $x, y \in \mathbb{R}^3$ gibt mit

$$E = \mathbb{R}x + \mathbb{R}y := \{\, \alpha x + \beta y \mid \alpha, \beta \in \mathbb{R} \,\}.$$

Eine Teilmenge G der $\mathbb{R}^3$ heißt *Gerade* durch 0, wenn es einen von 0 verschiedenen Vektor $x \in \mathbb{R}^3$ gibt mit

$$G = \mathbb{R}x := \{\, \alpha x \mid \alpha \in \mathbb{R} \,\}.$$

Wir wollen uns überlegen, daß im $\mathbb{R}^3$ jede Ebene E durch 0 und auch jede Gerade G durch 0 bezüglich den auf $\mathbb{R}^3$ definierten Operationen $+$ und $\cdot$ ein $\mathbb{R}$-Vektorraum ist. Sei etwa $E = \mathbb{R}x + \mathbb{R}y$. Offenbar genügt es zu zeigen, daß E gegen die auf $\mathbb{R}^3$ definierten Operationen $+$ und $\cdot$ abgeschlossen ist und das Nullelement 0 enthält. (Die Abgeschlossenheit gegen die Bildung des Negativen muß dann nicht extra geprüft werden, da wir ja schon $(-1)x = -x$ gezeigt haben). Dies ist aber klar, denn

$$\begin{aligned}(\alpha_1 x + \beta_1 y) + (\alpha_2 x + \beta_2 y) &= (\alpha_1 + \alpha_2)x + (\beta_1 + \beta_2)y,\\ \gamma(\alpha x + \beta y) &= \gamma\alpha x + \gamma\beta y,\\ 0 &= 0x + 0y \in E.\end{aligned}$$

Für eine Gerade G ist der Beweis ähnlich. Entsprechend definieren wir jetzt den allgemeinen Begriff eines Untervektorraums.

2.3.1. DEFINITION. Sei V ein K-Vektorraum und $U \subseteq V$. U heißt *Untervektorraum*, wenn für alle $x, y \in V$ und $\alpha \in K$ gilt

1. $0 \in U$.
2. Wenn $x, y \in U$, so ist $x + y \in U$.
3. Wenn $x \in U$, so ist $\alpha x \in U$.

Statt Untervektorraum sagt man auch kurz *Unterraum*.

BEISPIELE. 1. Ebenen durch 0 und Geraden durch 0 im $\mathbb{R}^3$ sind Unterräume.

2. Sei $1 \leq m \leq n$. Dann ist

$$\{ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \mid \xi_1 = \cdots = \xi_m = 0 \} \subseteq K^n$$

ein Unterraum.

3. Die Menge aller Polynome aus $\mathbb{R}[t]$ mit nur geraden Potenzen von t ist Unterraum von $\mathbb{R}[t]$.

4. Seien V, W K-Vektorräume. Dann ist

$$\{ (x, 0) \mid x \in V \} \subseteq V \times W$$

ein Unterraum.

5.

$$C(\mathbb{R}) := \{ f \mid f \colon \mathbb{R} \to \mathbb{R} \text{ stetig} \} \subseteq \mathbb{R}^{\mathbb{R}}$$
$$D(\mathbb{R}) := \{ f \mid f \colon \mathbb{R} \to \mathbb{R} \text{ differenzierbar} \} \subseteq \mathbb{R}^{\mathbb{R}}$$

sind Unterräume.

6. Sei V K-Vektorraum. 0 und V sind dann (triviale) Unterräume von V.

BEMERKUNG. Sei V ein K-Vektorraum und $(U_i)_{i \in I}$ eine nicht leere Familie von Unterräumen von V. Dann ist $\bigcap_{i \in I} U_i$ Unterraum von V.

BEWEIS. 1. $0 \in \bigcap_{i \in I} U_i$, da $0 \in U_i$ für alle $i \in I$.

2. Seien $x, y \in \bigcap_{i \in I} U_i$. Dann gilt

$$\begin{aligned} x, y &\in U_i && \text{für alle } i \in I \\ x + y &\in U_i && \text{für alle } i \in I \\ x + y &\in \bigcap_{i \in I} U_i. \end{aligned}$$

3. Sei $x \in \bigcap_{i \in I} U_i$ und $\alpha \in K$. Dann zeigt man wie eben $\alpha x \in \bigcap_{i \in I} U_i$. QED

2.3.2. DEFINITION. Sei V ein K-Vektorraum, $M \subseteq V$ eine Teilmenge und $(U_i)_{i \in I}$ eine Familie von Unterräumen von V.

1. $\langle M \rangle := \{ x \in V \mid \exists n \in \mathbb{N}, x_1, \ldots, x_n \in M, \alpha_1, \ldots, \alpha_n \in K \; x = \alpha_1 x_1 + \cdots + \alpha_n x_n \}$ heißt der *von M erzeugte* (oder *von M aufgespannte*) Unterraum von V. Insbesondere ist $\langle \emptyset \rangle = 0$.
2. $\sum_{i \in I} U_i := \langle \bigcup_{i \in I} U_i \rangle$ heißt *Summe der Unterräume* $(U_i)_{i \in I}$. Im Fall $I = \{1, \ldots, n\}$ schreibt man

$$\sum_{i \in I} U_i =: \sum_{i=1}^{n} U_i =: U_1 + \cdots + U_n.$$

Statt $\langle \{x_1, \ldots, x_n\} \rangle$ schreiben wir kurz $\langle x_1, \ldots, x_n \rangle$.

Bemerkung. Sei V ein K-Vektorraum, $M \subseteq V$ eine Teilmenge und $(U_i)_{i \in I}$ eine Familie von Unterräumen von V. Dann gilt

1. $\langle M \rangle$ ist Unterraum von V.
2. $\langle M \rangle$ ist der kleinste M umfassende Unterraum von V, d.h. ist $U \subseteq V$ Unterraum von V mit $M \subseteq U$, so ist $\langle M \rangle \subseteq U$.
3. $\sum_{i \in I} U_i = \{ x \in V \mid \exists i_1, \dots, i_n \in I \text{ verschieden}, y_{i_1} \in U_{i_1}, \dots, y_{i_n} \in U_{i_n}\ x = y_{i_1} + \dots + y_{i_n} \}$, insbesondere also für $I = \{1, \dots, n\}$
$$U_1 + \dots + U_n = \{ y_1 + \dots + y_n \mid y_1 \in U_1, \dots, y_n \in U_n \}.$$

Beweis. 1. $0 \in \langle M \rangle$ ist klar. Seien $x, y \in \langle M \rangle$. Dann ist $x = \alpha_1 x_1 + \dots + \alpha_n x_n$ und $y = \beta_1 y_1 + \dots + \beta_m y_m$, also $x + y = \alpha_1 x_1 + \dots + \alpha_n x_n + \beta_1 y_1 + \dots + \beta_m y_m \in \langle M \rangle$. Aus $x \in \langle M \rangle$ folgt offenbar $\alpha x \in \langle M \rangle$.

2. Sei $U \subseteq V$ Unterraum und $M \subseteq U$. Zu zeigen ist $\langle M \rangle \subseteq U$. Sei also $x \in \langle M \rangle$, etwa $x = \alpha_1 x_1 + \dots + \alpha_n x_n$ mit $x_1, \dots, x_n \in M$. Offenbar ist dann $x \in U$.

3. $\subseteq$. Sei $x \in \sum_{i \in I} U_i = \langle \bigcup_{i \in I} U_i \rangle$. Dann ist $x = \alpha_1 x_1 + \dots + \alpha_n x_n$ mit $x_j \in \bigcup_{i \in I} U_i$ für $1 \le j \le n$. Jedes x_j ist aus einem U_{i_j} und damit auch $\alpha_j x_j \in U_{i_j}$. Ferner kann man offenbar annehmen, daß $i_1, \dots, i_n$ verschieden sind.

$\supseteq$. Sei $x = y_{i_1} + \dots + y_{i_n}$ mit $y_{i_j} \in U_{i_j}$. Dann ist auch $x = 1 y_{i_1} + \dots + 1 y_{i_n}$ mit $y_{i_j} \in \bigcup_{i \in I} U_i$, also $x \in \langle \bigcup_{i \in I} U_i \rangle$. Qed

2.3.3. *Dimension und Basis von Unterräumen.* Wir wollen uns jetzt überlegen, daß jeder Unterraum eines endlichdimensionalen Vektorraums eine endliche Basis besitzt, die zu einer Basis des Gesamtraums ergänzt werden kann. Der Beweis ist unerwartet schwierig: man muß entweder die allgemeine (mit Hilfe des Zornschen Lemmas bewiesene) Aussage benutzen, daß jeder Vektorraum eine Basis besitzt, oder aber eine Basis schrittweise konstruieren.

Satz. *Sei V ein n-dimensionaler K-Vektorraum und $U \subseteq V$ ein Unterraum. Dann gilt*

1. $\dim(U) \le n$.
2. *Es gibt eine Basis $x_1, \dots, x_m$ von U, die sich zu einer Basis $x_1, \dots, x_m, x_{m+1}, \dots, x_n$ von V ergänzen läßt.*

Beweis. Offenbar genügt es, (2) zu zeigen. Wir überlegen uns zunächst, daß es eine Basis $x_1, \dots, x_m$ von U gibt. Dies folgt offenbar aus 2.2.8, wo wir mit Hilfe des Zornschen Lemmas gezeigt haben, daß jeder Vektorraum eine Basis besitzt. In dem hier vorliegenden Spezialfall eines n-dimensionalen Vektorraums können wir einen Beweis führen, der das Zornsche Lemma nicht verwendet.

Es ist $0 \subseteq U$. Gilt $0 = U$, so sind wir fertig. Gilt $0 \subsetneq U$, so wähle man ein $x_1 \in U$ mit $x_1 \neq 0$.

Es ist $\langle x_1 \rangle \subseteq U$. Gilt $\langle x_1 \rangle = U$, so sind wir fertig. Gilt $\langle x_1 \rangle \subsetneq U$, so wähle man ein $x_2 \in U \setminus \langle x_1 \rangle$. Dann sind x_1, x_2 linear unabhängig.

Es ist $\langle x_1, x_2 \rangle \subseteq U$. Gilt $\langle x_1, x_2 \rangle = U$, so sind wir fertig. Gilt $\langle x_1, x_2 \rangle \subsetneq U$, so wähle man ein $x_3 \in U \setminus \langle x_1, x_2 \rangle$. Dann sind x_1, x_2, x_3 linear unabhängig.

Dieses Verfahren setze man fort. Es muß abbrechen, da nach Satz 2.2.9(1) je $n+1$ Vektoren in V linear abhängig sind. Für ein $m \le n$ erhält man also eine Basis $x_1, \dots, x_m$ von U. Nach dem Basisergänzungssatz 2.2.7 läßt sie sich zu einer Basis $x_1, \dots, x_m, x_{m+1}, \dots, x_n$ von V ergänzen. Qed

2.3.4. *Zerlegung von Vektorräumen.* Unser nächstes Ziel ist es, die Möglichkeit der Zerlegung von Vektorräumen in kleinere zu studieren.

Definition. Sei V ein K-Vektorraum und $U_1, \ldots, U_n$ Unterräume. V heißt *direkte Summe* von $U_1, \ldots, U_n$ (in Zeichen: $V = U_1 \oplus \cdots \oplus U_n$ oder $V = \bigoplus_{i=1}^n U_i$), wenn

1. $V = U_1 + \cdots + U_n$ ("Summe") und
2. Hat man $x_1 \in U_1, \ldots, x_n \in U_n$ mit $x_1 + \cdots + x_n = 0$, so sind $x_1, \ldots, x_n$ alle $= 0$. ("Direktheit").

Bemerkungen. 1. Unter den Voraussetzungen der Definition gilt $V = U_1 \oplus \cdots \oplus U_n$ genau dann, wenn sich jedes $x \in V$ eindeutig schreiben läßt in der Form $x = x_1 + \cdots + x_n$ mit $x_i \in U_i$ für $1 \leq i \leq n$.

Beweis. $\Longrightarrow$. Die Existenz der Darstellung folgt aus Teil (1) der Definition. Die Eindeutigkeit ergibt sich wie folgt. Nehmen wir an, mit $x_i, x_i' \in U$ hätten wir

$$x = x_1 + \cdots + x_n = x_1' + \cdots + x_n'.$$

Dann ist $0 = (x_1 - x_1') + \cdots + (x_n - x_n')$ mit $x_i - x_i' \in U$. Aus Teil (2) der Definition folgt jetzt $x_i = x_i'$.

$\Longleftarrow$. 1. $V = U_1 + \cdots + U_n$ ist klar. 2. Seien $x_1 \in U_1, \ldots, x_n \in U_n$ mit $x_1 + \cdots + x_n = 0$ gegeben. Da auch $0 + \cdots + 0 = 0$ gilt, folgt aus der Eindeutigkeit $x_1 = \cdots = x_n = 0$. Qed

2. Sei $x_1, \ldots, x_n$ eine Basis von V und $m \leq n$. Dann gilt

$$\begin{aligned} V &= Kx_1 \oplus \cdots \oplus Kx_n \\ &= (Kx_1 \oplus \cdots \oplus Kx_m) \oplus (Kx_{m+1} \oplus \cdots \oplus Kx_n). \end{aligned}$$

Beweis. Wir beginnen mit dem Beweis der ersten Gleichung. 1. Es ist $V = \sum_{i=1}^n Kx_i$, da $x_1, \ldots, x_n$ ein Erzeugendensystem ist. 2. Seien $y_i = \alpha_i x_i \in Kx_i$ für $1 \leq i \leq n$ und es gelte $\sum_{i=1}^n \alpha_i x_i = 0$. Dann ist $\alpha_1 = \cdots = \alpha_n = 0$, da $x_1, \ldots, x_n$ linear unabhängig ist. Also ist $y_1 = \cdots = y_n = 0$.

Wir beweisen jetzt die zweite Gleichung. Sei $U := \langle x_1, \ldots, x_m \rangle$ und $U' := \langle x_{m+1}, \ldots, x_n \rangle$. Nach der ersten Gleichung ist

$$\begin{aligned} U &= Kx_1 \oplus \cdots \oplus Kx_m, \\ U' &= Kx_{m+1} \oplus \cdots \oplus Kx_n. \end{aligned}$$

Zu zeigen ist $V = U \oplus U'$. 1. $V = U + U'$ ist klar. 2. Seien $y \in U$ und $y' \in U'$ mit $y + y' = 0$ gegeben. Dann ist $y = \sum_{i=1}^m \alpha_i x_i$ und $y' = \sum_{i=m+1}^n \alpha_i x_i$ mit $\alpha_1, \ldots, \alpha_n \in K$. Wegen $0 = y + y' = \sum_{i=1}^n \alpha_i x_i$ folgt $\alpha_1 = \cdots = \alpha_n = 0$, da $x_1, \ldots, x_n$ linear unabhängig ist. Also ist $y = y' = 0$. Qed

3. Seien U und U' Unterräume von V mit $V = U \oplus U'$. Falls dann $x_1, \ldots, x_m$ eine Basis von U und $x_{m+1}, \ldots, x_n$ eine Basis von U' ist, so ist $x_1, \ldots, x_m, x_{m+1}, \ldots, x_n$ eine Basis von V. Ist also $V = U \oplus U'$ endlichdimensional, so gilt

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

Beweis. $x_1, \ldots, x_n$ erzeugt $U+U'$, also V. Seien jetzt $\alpha_1, \ldots, \alpha_n \in K$ mit $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$ gegeben. Dann ist

$$0 = \underbrace{\alpha_1 x_1 + \cdots + \alpha_m x_m}_{x \in U} + \underbrace{\alpha_{m+1} x_{m+1} + \cdots + \alpha_n x_n}_{x' \in U'}.$$

Wegen $V = U \oplus U'$ folgt $x = x' = 0$. Also ist $\alpha_1 = \cdots = \alpha_m = 0$, da $x_1, \ldots, x_m$ Basis von U ist. Genauso ergibt sich $\alpha_{m+1} = \cdots = \alpha_n = 0$. Qed

4. Seien U, U' Unterräume von V. Dann gilt

$$V = U \oplus U' \quad \equiv \quad \begin{cases} 1. & V = U + U' \\ 2. & U \cap U' = 0 \end{cases}$$

Beweis. $\Longrightarrow$. (1) gilt nach Definition. 2. Sei $x \in U \cap U'$. Dann ist $x \in U$ und $-x \in U'$, also

$$0 = \underbrace{x}_{\in U} + \underbrace{(-x)}_{\in U'}.$$

Aus der Definition der direkten Summe folgt $x = 0$.

$\Longleftarrow$. 1. $V = U + U'$ ist vorausgesetzt. 2. Seien $x \in U$ und $x' \in U'$ mit $x + x' = 0$ gegeben. Dann ist $x = -x' \in U'$. Es folgt $x \in U \cap U'$, also $x = 0$ und damit $x' = 0$. Qed

5. V, W seien K-Vektorräume. $V \times W$ ist mit komponentenweiser Addition und Skalarmultiplikation ein K-Vektorraum (nach einem Beispiel in 2.3.1), und es gilt

$$V \times W = (V \times 0) \oplus (0 \times W).$$

Beweis. Klar. Qed

2.3.5. *Quotienten von Vektorräumen.* Sei V ein K-Vektorraum und $U \subseteq V$ ein Unterraum. Im allgemeinen gibt es dann viele Unterräume $U' \subseteq V$ mit $V = U \oplus U'$. Betrachten wir zum Beispiel im $\mathbb{R}^2$ den Unterraum

$$U = \{ \begin{pmatrix} \alpha \\ 0 \end{pmatrix} \mid \alpha \in \mathbb{R} \},$$

also die "x-Achse". Dann hat jede andere Gerade U' durch den Nullpunkt die Eigenschaft $V = U \oplus U'$; eine natürliche (d.h. basisunabhängige) Auswahl ist nicht möglich. Man kann jedoch zeigen, daß alle U' mit $V = U \oplus U'$ "im wesentlichen gleiche" Unterräume sind. Im nächsten Kapitel werden wir diesen Begriff präzisieren und dann (in Beispiel (4) in 3.1.1) diese Aussage beweisen. Zum Beweis werden wir zeigen, daß jeder solche Unterraum U' "im wesentlichen gleich" ist einem aus V und seinem Unterraum U auf natürliche Weise konstruierten *Quotientenraum* V/U. – Hier wollen wir die Konstruktion von V/U durchführen und zeigen, daß jeder Unterraum $U' \subseteq V$ mit $V = U \oplus U'$ in einer natürlichen Weise bijektiv auf V/U abgebildet werden kann. (Im nächsten Kapitel werden wir sehen, daß diese Abbildung ein sog. "Isomorphismus" ist). In unserem Beispiel ist V/U die Menge der horizontalen Linien.

Für Teilmengen $M, N \subseteq V$ definieren wir die *Komplexaddition* von M und N durch

$$M + N := \{ x + y \mid x \in M, y \in N \}.$$

Dann gilt

$$M + N = N + M,$$
$$(M + N) + P = M + (N + P).$$

Statt $\{x\} + M$ schreiben wir $x + M$.

Satz. *Sei V ein K-Vektorraum und $U \subseteq V$ ein Unterraum. Dann kann man auf*

$$V/U := \{\, x + U \mid x \in V \,\}$$

eine Addition und eine Skalarmultiplikation so erklären, daß gilt

$$\begin{aligned}(x+U)+(y+U) &= (x+y)+U,\\ \alpha(x+U) &= \alpha x + U,\end{aligned}$$

und V/U mit dieser Addition und Skalarmultiplikation ein K-Vektorraum wird.

Beweis. Es gilt für die Komplexaddition

$$(x+U)+(y+U) = x+y+U+U = (x+y)+U, \quad \text{da } U+U=U.$$

Die obige Addition ist also wohldefiniert. Wir zeigen jetzt, daß auch die Skalarmultiplikation wohldefiniert ist. Sei also $x+U = x'+U$. Zu zeigen ist $\alpha x + U = \alpha x' + U$. Im Fall $\alpha = 0$ ist dies klar. Im Fall $\alpha \neq 0$ argumentieren wir wie folgt. Es genügt, $\alpha x + U \subseteq \alpha x' + U$ zu zeigen. Sei also $y \in U$. Gesucht ist ein $z \in U$ mit $\alpha x + y = \alpha x' + z$. Nun gilt

$$\begin{aligned}& x + \frac{1}{\alpha} y \in x + U = x' + U, && \\ & x + \frac{1}{\alpha} y = x' + u && \text{für ein } u \in U,\\ & \alpha x + y = \alpha x' + \alpha u = \alpha x' + z && \text{mit } z := \alpha u \in U.\end{aligned}$$

Offenbar ist $(V/U, +)$ eine abelsche Gruppe; neutrales Element ist $0+U$, also U, und das inverse Element zu $x+U$ ist $(-x)+U$. Ferner gilt

$$\alpha\big((x+U)+(y+U)\big) = \alpha(x+U) + \alpha(y+U),$$

denn es ist

$$\begin{aligned}\alpha\big((x+U)+(y+U)\big) &= \alpha((x+y)+U)\\ &= \alpha(x+y)+U\\ &= (\alpha x + \alpha y) + U\\ &= (\alpha x + U) + (\alpha y + U)\\ &= \alpha(x+U) + \alpha(y+U).\end{aligned}$$

Die restlichen Vektorraumeigenschaften beweist man ähnlich. Qed

Satz. *Sei V ein K-Vektorraum und $U, U' \subseteq V$ Unterräume mit $V = U \oplus U'$. Sei g die* natürliche Abbildung

$$\begin{aligned} g\colon V &\to V/U\\ x &\mapsto x + U.\end{aligned}$$

Dann ist die Einschränkung $f := g{\restriction}U'$ von g auf U', also

$$\begin{aligned} f\colon U' &\to V/U\\ x &\mapsto x + U.\end{aligned}$$

eine bijektive Abbildung.

BEWEIS. f ist injektiv: Seien $x, y \in U'$ mit $x + U = y + U$ gegeben. Dann gilt

$$\begin{aligned} x + 0 &= y + u && \text{für ein } u \in U \\ 0 &= \underbrace{(y - x)}_{\in U'} + \underbrace{u}_{\in U} \\ y - x &= 0 && \text{wegen } V = U \oplus U' \\ x &= y. \end{aligned}$$

f ist surjektiv: Sei $x + U \in V/U$ mit $x \in V$. Wegen $V = U \oplus U'$ ist dann $x = u + y$ mit $u \in U$ und $y \in U'$. Also ist

$$x + U = (u + y) + U = y + U, \quad \text{da } u + U = U.$$

Mit unserem $y \in U'$ haben wir also $x + U = f(y)$. QED

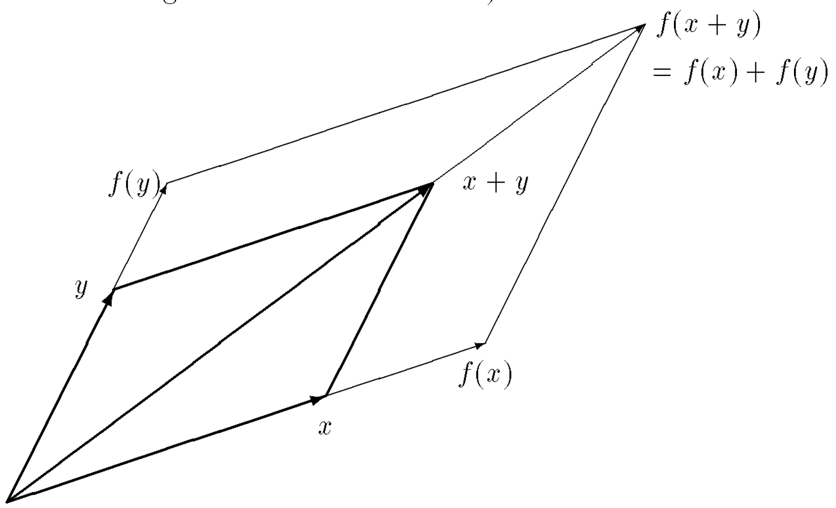
Kapitel 3

Lineare Abbildungen

3.1 Lineare Abbildungen

Wir betrachten die folgenden (zunächst anschaulich gegebenen) Abbildungen $f\colon \mathbb{R}^2 \to \mathbb{R}^2$:

1. Streckung (in allen Richtungen um denselben Faktor).



2. Streckung in einer Richtung.
3. Spiegelung am Nullpunkt.
4. Spiegelung an einer Geraden durch den Nullpunkt.
5. Drehung um den Winkel α um den Nullpunkt.

Wie man sich leicht geometrisch klarmacht, gilt in allen diesen Fällen stets

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

Wir abstrahieren hieraus den folgenden allgemeinen Begriff einer linearen Abbildung.

3.1.1. Definition. Seien V und W K-Vektorräume und $f\colon V \to W$ eine Abbildung. f heißt *linear*, wenn für alle $x, y \in V$ und alle $\alpha \in K$ gilt

1. $f(x+y) = f(x) + f(y)$ und
2. $f(\alpha x) = \alpha f(x)$.

Bemerkungen. Seien V und W K-Vektorräume und $f\colon V \to W$ eine lineare Abbildung.

1. Es gilt

$$f(0) = 0,$$
$$f(-x) = -f(x) \quad \text{für alle } x \in V.$$

Beweis. Es ist $f(0) = f(0x) = 0f(x) = 0$, und $f(x) + f(-x) = f(x - x) = f(0) = 0$. Qed

2. Sind $x_1, \dots, x_n$ linear abhängig, so sind auch $f(x_1), \dots, f(x_n)$ linear abhängig.

Beweis. Es gelte $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$, wobei nicht alle α_i verschwinden. Dann ist auch $f(\alpha_1 x_1 + \cdots + \alpha_n x_n) = \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n) = 0$, d.h. auch $f(x_1), \dots, f(x_n)$ sind linear abhängig. Qed

3. Ist $V' \subseteq V$ ein Unterraum, so ist auch $f(V') \subseteq W$ ein Unterraum. Insbesondere ist also $f(V)$ ein Unterraum von W. Man nennt den Unterraum $f(V)$ das *Bild* der linearen Abbildung f und bezeichnet ihn mit $\mathrm{Bild}(f)$.

Beweis. Sei $V' \subseteq V$ ein Unterraum, also $f(V') \subseteq W$. Wir zeigen die drei Unterraumeigenschaften aus der Definition 2.3.1 von Unterräumen.

$0 \in f(V')$: Es ist $0 \in V'$, also $0 = f(0) \in f(V')$.

Aus $u, v \in f(V')$ folgt $u + v \in f(V')$: Seien $u, v \in f(V')$, also $u = f(x)$ und $v = f(y)$ für $x, y \in V'$. Dann ist $u + v = f(x) + f(y) = f(x + y) \in f(V')$, denn aus $x, y \in V'$ folgt $x + y \in V'$.

Aus $u \in f(V')$ folgt $\alpha u \in f(V')$: Sei $u \in f(V')$, also $u = f(x)$ mit $x \in V'$. Dann ist $\alpha u = \alpha f(x) = f(\alpha x) \in f(V')$, denn aus $x \in V'$ folgt $\alpha x \in V'$. Qed

4. Bilden $x_1, \dots, x_n$ ein Erzeugendensystem von V, so bilden $f(x_1), \dots, f(x_n)$ ein Erzeugendensystem von $f(V)$.

Beweis. Sei $y \in f(V)$, also $y = f(x)$ für ein $x \in V$. Dann kann man x schreiben in der Form $x = \alpha_1 x_1 + \cdots + \alpha_n x_n$, also auch $f(x)$ in der Form $f(x) = \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n)$. Qed

5. Ist $W' \subseteq W$ ein Unterraum, so ist auch $f^{-1}(W') \subseteq V$ ein Unterraum. Insbesondere ist also $f^{-1}(0)$ ein Unterraum von V. Man nennt den Unterraum $f^{-1}(0)$ den *Kern* der linearen Abbildung f und bezeichnet ihn mit $\mathrm{Kern}(f)$.

Beweis. Sei $W' \subseteq W$ ein Unterraum, also $f^{-1}(W') \subseteq V$. Wir zeigen die drei Unterraumeigenschaften aus der Definition 2.3.1 von Unterräumen.

$0 \in f^{-1}(W')$: Es ist $f(0) = 0 \in W'$, also $0 \in f^{-1}(W')$.

Aus $x, y \in f^{-1}(W')$ folgt $x + y \in f^{-1}(W')$: Seien $x, y \in f^{-1}(W')$, also $f(x), f(y) \in W'$. Dann ist $f(x) + f(y) = f(x + y) \in W'$, also $x + y \in f^{-1}(W')$.

Aus $x \in f^{-1}(W')$ folgt $\alpha x \in f^{-1}(W')$: Sei $x \in f^{-1}(W')$, also $f(x) \in W'$. Dann ist $\alpha f(x) = f(\alpha x) \in W'$, also $\alpha x \in f^{-1}(W')$. Qed

Beispiele. 1. Die Abbildungen

$$0\colon V \to W \qquad\qquad \mathrm{id}_V\colon V \to V$$
$$x \mapsto 0 \qquad\qquad\qquad x \mapsto x$$

sind linear, nicht jedoch die Abbildung

$$f\colon V \to W$$
$$x \mapsto x_0 \neq 0.$$

2. Die Abbildung

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \alpha x \end{aligned}$$

ist linear, denn es gilt

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

Wir zeigen jetzt, daß *jede* lineare Abbildung $f\colon \mathbb{R} \to \mathbb{R}$ diese Form hat.

Beweis. Sei $f\colon \mathbb{R} \to \mathbb{R}$ linear. Dann gilt

$$f(x) = f(x1) = xf(1) = \underbrace{f(1)}_{\alpha}x. \quad \text{QED}$$

3. Die Abbildung

$$\begin{aligned} f\colon K^n &\to K^m \\ x &\mapsto Ax \end{aligned}$$

mit $A \in K^{m\times n}$ ist linear, denn es gilt

$$\begin{aligned} &f(x+y) = A(x+y) = Ax + Ay = f(x) + f(y), \\ &f(\alpha x) = A\alpha x = \alpha Ax = \alpha f(x). \end{aligned}$$

Wir zeigen jetzt, daß *jede* lineare Abbildung $f\colon K^n \to K^m$ diese Form hat.

Beweis. Sei $f\colon K^n \to K^m$ linear. Es gibt $\alpha_{ij} \in K$ für $1 \le i \le m$ und $1 \le j \le n$ so daß

$$f(e_j) = \sum_{i=1}^{m} \alpha_{ij} e_i \qquad \text{für } 1 \le j \le n.$$

Setze $A := (\alpha_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$. Dann ist

$$\begin{aligned} f(x) &= f\Big(\sum_{j=1}^{n} \xi_j e_j\Big) \\ &= \sum_{j=1}^{n} \xi_j f(e_j) \\ &= \sum_{j=1}^{n} \xi_j \Big(\sum_{i=1}^{m} \alpha_{ij} e_i\Big) \\ &= \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} \alpha_{ij}\xi_j\Big) e_i \\ &= \begin{pmatrix} \sum_{j=1}^{n} \alpha_{1j}\xi_j \\ \vdots \\ \sum_{j=1}^{n} \alpha_{mj}\xi_j \end{pmatrix} \\ &= Ax. \quad \text{QED} \end{aligned}$$

Wir wollen noch einen Zusammenhang zur Theorie der linearen Gleichungssysteme herstellen. Sei wieder

$$\begin{aligned} f\colon K^n &\to K^m \\ x &\mapsto Ax \end{aligned}$$

mit $A \in K^{m\times n}$. Dann ist die Lösungsmenge des *homogenen* linearen Gleichungssystems $Ax = 0$ der Unterraum $f^{-1}(0)$ des K^n; nach Abschnitt 1.4 wissen wir, daß er $(n-r)$-dimensional ist mit einer damals aus A bestimmten Zahl r. Ferner ist (im Falle der Lösbarkeit) die Lösungsmenge des *inhomogenen* linearen Gleichungssystems $Ax = b$ von der Form $x_0 + f^{-1}(0)$, wobei x_0 eine spezielle Lösung ist.

4. Sei V ein K-Vektorraum und $U \subseteq V$ ein Unterraum. Dann ist die in 2.3.5 definierte *natürliche Abbildung*

$$\begin{aligned} g\colon V &\to V/U \\ x &\mapsto x + U. \end{aligned}$$

eine lineare Abbildung.

Beweis.

$$\begin{aligned} g(x+y) &= x + y + U \\ &= (x+U) + (y+U) \quad \text{nach Definition von } + \text{ auf } V/U \\ &= g(x) + g(y) \end{aligned}$$

und

$$\begin{aligned} g(\alpha x) &= \alpha x + U \\ &= \alpha(x+U) \qquad\qquad \text{nach Definition von } \cdot \text{ auf } V/U \\ &= \alpha g(x). \quad \text{QED} \end{aligned}$$

5. Sei V ein K-Vektorraum, I eine nicht leere Indexmenge und

$$V^I := \{\, f \mid f\colon I \to V \text{ Abbildung} \,\}$$

der in Beispiel 2.1.2(3) eingeführte K-Vektorraum. Sei ferner $h\colon I \to I$ eine beliebige Abbildung. Dann ist

$$\begin{aligned} f\colon V^I &\to V^I \\ g &\mapsto g \circ h \end{aligned}$$

eine lineare Abbildung.

Beweis. $f(g_1 + g_2) = f(g_1) + f(g_2)$:

$$\begin{aligned} (f(g_1+g_2))(x) &= ((g_1+g_2)\circ h)(x) \\ &= (g_1+g_2)(h(x)) \\ &= g_1(h(x)) + g_2(h(x)) \\ &= (g_1\circ h)(x) + (g_2 \circ h)(x) \\ &= (f(g_1))(x) + (f(g_2))(x) \\ &= (f(g_1) + f(g_2))(x). \end{aligned}$$

$f(\alpha g) = \alpha f(g)$ beweist man entsprechend. QED

6. Sei $D(\mathbb{R}) := \{ f \mid f\colon \mathbb{R} \to \mathbb{R} \text{ differenzierbar} \}$ der Vektorraum aus Beispiel 2.1.2(4). Dann ist

$$\begin{aligned} D(\mathbb{R}) &\to \mathbb{R}^{\mathbb{R}} \\ f &\mapsto f' \end{aligned}$$

linear. Ferner ist für festes $x_0 \in \mathbb{R}$ die folgende Abbildung linear.

$$\begin{aligned} D(\mathbb{R}) &\to \mathbb{R} \\ f &\mapsto f'(x_0) \end{aligned}$$

3.1.2. Wir wollen uns jetzt mit der Frage befassen, wie man eine lineare Abbildung f konstruieren kann. Es wird sich zeigen, daß man die Werte von f auf einer Basis beliebig vorgeben kann, und daß f damit eindeutig bestimmt ist.

BEMERKUNG. Sei $(x_i)_{i\in I}$ eine Basis von V. Jedes $x \in V$ läßt sich dann darstellen in der Form

$$x = \sum_{i\in I} \alpha_i x_i$$

mit einer Familie $(\alpha_i)_{i\in I}$ von Skalaren, in der fast alle α_i (d.h. alle bis auf endlich viele) verschwinden. Wir wollen uns überlegen, daß diese Darstellung eindeutig ist. Hat man nämlich noch

$$x = \sum_{i\in I} \beta_i x_i,$$

so folgt $0 = \sum_{i\in I}(\alpha_i - \beta_i)x_i$. Da die Familie $(x_i)_{i\in I}$ linear unabhängig ist, folgt $\alpha_i - \beta_i = 0$ für alle $i \in I$, also $\alpha_i = \beta_i$.

SATZ. *(Lineare Fortsetzung). Seien V, W K-Vektorräume, $(x_i)_{i\in I}$ eine Basis von V und $(y_i)_{i\in I}$ eine beliebige Familie von Vektoren aus W. Dann gibt es genau eine lineare Abbildung $f\colon V \to W$ mit $f(x_i) = y_i$ für alle $i \in I$. Weiter gilt*

1. *$f(V) = \langle \{ y_i \mid i \in I \} \rangle$,*
2. *f ist injektiv genau dann, wenn die Familie $(y_i)_{i\in I}$ linear unabhängig ist.*

BEWEIS. Aufgrund der Vorbemerkung können wir definieren

$$\begin{aligned} f\colon V &\to W \\ \sum_{i\in I} \alpha_i x_i &\mapsto \sum_{i\in I} \alpha_i y_i \end{aligned}$$

1. f ist linear. $f(x + y) = f(x) + f(y)$: Sei $x = \sum_{i\in I} \alpha_i x_i$ und $y = \sum_{i\in I} \beta_i x_i$. Dann erhält man

$$\begin{aligned} f(x+y) &= f\Big(\sum_{i\in I}(\alpha_i + \beta_i)x_i\Big) \\ &= \sum_{i\in I}(\alpha_i + \beta_i)y_i \\ &= \sum_{i\in I} \alpha_i y_i + \sum_{i\in I} \beta_i y_i \\ &= f(x) + f(y). \end{aligned}$$

$f(\alpha x) = \alpha f(x)$ beweist man ähnlich.

2. f ist eindeutig bestimmt. Sei noch $g\colon V \to W$ linear mit $g(x_i) = y_i$ für alle $i \in I$. Dann gilt für jedes $x = \sum_{i\in I} \alpha_i x_i \in V$

$$f(x) = f\Big(\sum_{i\in I} \alpha_i x_i\Big) = \sum_{i\in I} \alpha_i f(x_i) = \sum_{i\in I} \alpha_i y_i = \sum_{i\in I} \alpha_i g(x_i) = g\Big(\sum_{i\in I} \alpha_i x_i\Big) = g(x).$$

1. Wir zeigen jetzt $f(V) = \langle \{\, y_i \mid i \in I \,\} \rangle$.

$$\begin{aligned} y \in f(V) &\equiv \exists (\alpha_i)_{i\in I} \big(\text{fast alle } \alpha_i = 0 \text{ und } y = f(\textstyle\sum_{i\in I} \alpha_i x_i)\big) \\ &\equiv \exists (\alpha_i)_{i\in I} \big(\text{fast alle } \alpha_i = 0 \text{ und } y = \textstyle\sum_{i\in I} \alpha_i f(x_i)\big) \\ &\equiv \exists (\alpha_i)_{i\in I} \big(\text{fast alle } \alpha_i = 0 \text{ und } y = \textstyle\sum_{i\in I} \alpha_i y_i\big) \\ &\equiv y \in \langle \{\, y_i \mid i \in I \,\} \rangle. \end{aligned}$$

2. Zu zeigen bleibt, daß f injektiv ist genau dann, wenn die Familie $(y_i)_{i\in I}$ linear unabhängig ist.

$$\begin{aligned} f \text{ ist injektiv} &\equiv \forall x, y{\in}V\, (f(x) = f(y) \to x = y) \\ &\equiv \forall x, y{\in}V\, (f(x - y) = 0 \to x - y = 0) \\ &\equiv \forall z{\in}V\, (f(z) = 0 \to z = 0) \\ &\equiv \forall (\alpha_i)_{i\in I} \text{ mit fast allen } \alpha_i = 0 \Big(f(\sum_{i\in I} \alpha_i x_i) = 0 \to \sum_{i\in I} \alpha_i x_i = 0\Big) \\ &\equiv \forall (\alpha_i)_{i\in I} \text{ mit fast allen } \alpha_i = 0 \Big(\sum_{i\in I} \alpha_i y_i = 0 \to \forall i{\in}I\, \alpha_i = 0\Big) \\ &\equiv (y_i)_{i\in I} \text{ ist linear unabhängig.} \end{aligned}$$

FOLGERUNG. *Seien V, W K-Vektorräume und $f\colon V \to W$ eine bijektive lineare Abbildung (d.h. ein Isomorphismus (s.u.)). Ist dann $(x_i)_{i\in I}$ eine Basis von V, so ist $(f(x_i))_{i\in I}$ eine Basis von W.*

BEWEIS. $(f(x_i))_{i\in I}$ ist linear unabhängig, da f injektiv ist, und ein Erzeugendensystem, da f surjektiv ist. QED

3.1.3. *Spezielle lineare Abbildungen.* Wir führen zunächst einige oft verwendete Bezeichnungen und Schreibweisen ein. Sei K ein Körper und V, W K-Vektorräume. Eine lineare Abbildung $f\colon V \to W$ nennt man auch einen *Vektorraumhomomorphismus* oder kurz *Homomorphismus*. Man schreibt

$$\mathrm{Hom}_K(V, W) := \mathrm{Hom}(V, W) := \{\, f\colon V \to W \mid f \text{ linear} \,\}.$$

Sei $f\colon V \to W$ ein Homomorphismus. f heißt

Monomorphismus,	wenn f injektiv ist,
Epimorphismus,	wenn f surjektiv ist,
Isomorphismus,	wenn f bijektiv ist,
Endomorphismus,	wenn $V = W$ ist, und
Automorphismus,	wenn $V = W$ und f bijektiv ist.

Man schreibt $V \cong W$, wenn es einen Isomorphismus $f\colon V \to W$ gibt.

Bemerkungen. 1. Seien V, W, U K-Vektorräume und $f\colon V \to W$, $g\colon W \to U$ Vektorraumhomomorphismen. Dann ist auch

$$\begin{aligned} g \circ f\colon V &\to W \\ x &\mapsto g(f(x)) \end{aligned}$$

ein Vektorraumhomomorphismus.

Beweis. Es gilt

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

$(g \circ f)(\alpha x) = \alpha(g \circ f)(y)$ beweist man ähnlich. Qed

2. Seien V, W K-Vektorräume und $f\colon V \to W$ ein Isomorphismus. Dann ist auch $f^{-1}\colon W \to V$ ein Isomorphismus.

Beweis. Seien $u, v \in W$ gegeben. Setze $x := f^{-1}(u)$ und $y := f^{-1}(v)$. Dann gilt

$$\begin{aligned} f^{-1}(u+v) &= f^{-1}(f(x) + f(y)) \\ &= f^{-1}(f(x+y)) \\ &= x + y \\ &= f^{-1}(u) + f^{-1}(v). \end{aligned}$$

Ferner gilt

$$f^{-1}(\alpha u) = f^{-1}(\alpha f(x)) = f^{-1}(f(\alpha x)) = \alpha x = \alpha f^{-1}(u). \quad \text{Qed}$$

3. $\mathrm{Aut}(V) := \{\, f\colon V \to V \mid f \text{ Automorphismus} \,\}$ ist bezüglich der Komposition $\circ$ eine Gruppe. Neutrales Element ist die Identität id_V, und inverses Element zu $f \in \mathrm{Aut}(V)$ ist f^{-1}. Dies folgt unmittelbar aus (1) und (2).

4. Seien V, W K-Vektorräume.

$$\mathrm{Hom}_K(V, W) = \{\, f\colon V \to W \mid f \text{ linear} \,\}$$

ist Untervektorraum des in Beispiel 2.1.2(3) eingeführten K-Vektorraums

$$W^V = \{\, f\colon V \to W \mid f \text{ Abbildung} \,\}.$$

Beweis. Wir zeigen, daß aus $f, g \in \mathrm{Hom}_K(V, W)$ folgt $f + g \in \mathrm{Hom}_K(V, W)$.

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

Ähnlich zeigt man, daß mit f auch $\alpha f \in \mathrm{Hom}_K(V, W)$. Ferner ist offenbar $0 \in \mathrm{Hom}_K(V, W)$. Qed

3.1.4. *Isomorphie und Dimension.*

LEMMA. *Seien V, W, U K-Vektorräume. Dann gilt*

$$\begin{aligned}&V \cong V,\\ &\text{wenn } V \cong W, \text{ so ist } W \cong V,\\ &\text{wenn } V \cong W \text{ und } W \cong U, \text{ so ist } V \cong U.\end{aligned}$$

BEWEIS. Die Behauptung folgt sofort aus den Bemerkungen (1) und (2) in 3.1.3.

LEMMA. *Seien V, W K-Vektorräume. Dann gilt*

1. *Wenn $V \cong W$, so ist $\dim(V) = \dim(W)$.*
2. *Wenn $\dim(V) = \dim(W) < \infty$, so ist $V \cong W$.*

BEWEIS. 1. Sei $V \cong W$, also nach dem vorigen Lemma auch $W \cong V$. *Fall* $\dim(V) = n < \infty$. Dann ist nach der Folgerung in 3.1.2 auch $\dim(W) = n < \infty$. Den Fall $\dim(W) = n < \infty$ behandelt man entsprechend. *Fall* $\dim(V) = \dim(W) = \infty$. Dann ist nichts zu zeigen.

2. Sei $\dim(V) = \dim(W) = n$. Nach dem Satz 3.1.2 von der linearen Fortsetzung ist dann $V \cong K^n$ und ebenso $W \cong K^n$, also $V \cong W$. QED

FOLGERUNG. *Sei V ein endlichdimensionaler K-Vektorraum und $U \subseteq V$ ein Unterraum. Dann gilt*

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

BEWEIS. Nach Satz 2.3.3 wähle man eine Basis $x_1, \ldots, x_m$ von U, die sich zu einer Basis $x_1, \ldots, x_m, x_{m+1}, \ldots, x_n$ von V ergänzen läßt. Sei $U' := \langle x_{m+1}, \ldots, x_n \rangle$. Dann ist $V = U \oplus U'$ nach Bemerkung (4) in 2.3.4, also $\dim(V) = \dim(U) + \dim(U')$ nach Bemerkung (3) in 2.3.4. Wegen $U' \cong V/U$ (dies folgt aus dem zweiten Satz in 2.3.5 zusammen mit Beispiel (4) in 3.1.1) folgt aus Teil (1) des vorigen Lemmas die Behauptung. QED

3.2 Kern und Bild

Seien V, W K-Vektorräume und $f\colon V \to W$ linear. In den Beispielen (3) und (5) von 3.1.1 hatten wir den *Kern* und das *Bild* von f definiert durch

$$\begin{aligned}\mathrm{Kern}(f) &:= \{\, x \in V \mid f(x) = 0 \,\},\\ \mathrm{Bild}(f) &:= f(V) \subseteq W.\end{aligned}$$

Wir hatten bereits bewiesen, daß $\mathrm{Kern}(f)$ ein Unterraum von V und $\mathrm{Bild}(f)$ ein Unterraum von W ist.

3.2.1. LEMMA. *Seien V, W K-Vektorräume und $f\colon V \to W$ linear. Dann ist f injektiv genau dann, wenn $\mathrm{Kern}(f) = 0$ ist.*

BEWEIS.

$$\begin{aligned} f \text{ ist injektiv} &\equiv \forall x, y{\in}V\,(f(x) = f(y) \to x = y) \\ &\equiv \forall x, y{\in}V\,(f(x-y) = 0 \to x - y = 0) \\ &\equiv \forall z{\in}V\,(f(z) = 0 \to z = 0) \\ &\equiv \operatorname{Kern}(f) = 0. \quad \text{QED} \end{aligned}$$

3.2.2. SATZ. *(Homomorphiesatz). Seien V, W K-Vektorräume, $f\colon V \to W$ surjektiver Homomorphismus und $g\colon V \to V/\operatorname{Kern}(f)$ der in 2.3.5 definierte natürliche Homomorphismus. Dann gibt es genau eine Abbildung $h\colon V/\operatorname{Kern}(f) \to W$ mit $h \circ g = f$. Dieses eindeutig bestimmte h ist ein Isomorphismus.*

BEWEIS. Existenz von h. Wir wollen h definieren durch

$$\begin{aligned} h\colon V/\operatorname{Kern}(f) &\to W \\ x + \operatorname{Kern}(f) &\mapsto f(x). \end{aligned}$$

Zu zeigen ist, daß h dadurch wohldefiniert ist. Gelte also $x + \operatorname{Kern}(f) = y + \operatorname{Kern}(f)$. Zu zeigen ist $f(x) = f(y)$. Es ist

$$\begin{aligned} x &= y + z \quad \text{für ein } z \in \operatorname{Kern}(f) \\ x - y &= z \\ f(x-y) &= 0 \\ f(x) &= f(y). \end{aligned}$$

h ist injektiv: Gelte $f(x) = f(y)$. Zu zeigen ist $x + \operatorname{Kern}(f) = y + \operatorname{Kern}(f)$. Aus Symmetriegründen genügt es, $\subseteq$ zu beweisen. Sei also $z \in \operatorname{Kern}(f)$. Zu zeigen ist $x + z \in y + \operatorname{Kern}(f)$, also $x - y + z \in \operatorname{Kern}(f)$. Dies folgt aus

$$f(x - y + z) = \underbrace{f(x) - f(y)}_{0} + \underbrace{f(z)}_{0} = 0.$$

h ist surjektiv: Sei $w \in W$. Dann gilt $w = f(x)$ für ein $x \in V$, also $h(x + \operatorname{Kern}(f)) = f(x) = w$.

h ist linear:

$$\begin{aligned} &h((x + \operatorname{Kern}(f)) + (y + \operatorname{Kern}(f))) && h(\alpha(x + \operatorname{Kern}(f))) \\ &= h(x + y + \operatorname{Kern}(f)) && = h(\alpha x + \operatorname{Kern}(f)) \\ &= f(x + y) && = f(\alpha x) \\ &= f(x) + f(y) && = \alpha f(x) \\ &= h(x + \operatorname{Kern}(f)) + h(y + \operatorname{Kern}(f)) && = \alpha h(x + \operatorname{Kern}(f)). \end{aligned}$$

Eindeutigkeit von h. Wegen der Forderung $h \circ g = f$ muß gelten $h(x + \operatorname{Kern}(f)) = h(g(x)) = f(x)$. Also ist h eindeutig bestimmt. QED

3.2.3. FOLGERUNG. *(Dimensionssatz). Es seien V ein endlichdimensionaler K-Vektorraum und $f\colon V \to W$ eine lineare Abbildung. Dann ist*

$$\dim(V) = \dim(\operatorname{Kern}(f)) + \dim(\operatorname{Bild}(f)).$$

BEWEIS. Nach dem Homomorphiesatz ist $V/\operatorname{Kern}(f) \cong \operatorname{Bild}(f)$, und nach der Folgerung in 3.1.4 gilt

$$\dim(V) = \dim(V/\operatorname{Kern}(f)) + \dim(\operatorname{Kern}(f)).$$

Folgerung. *Sei V ein endlichdimensionaler K-Vektorraum und $f\colon V \to V$ linear. Dann sind folgende Aussagen äquivalent.*

1. *f ist Monomorphismus.*

2. *f ist Epimorphismus.*

3. *f ist Automorphismus.*

Beweis. Es genügt zu zeigen, daß aus (2) stets (1) folgt (denn offensichtlich folgt (3) aus (1), und auch (1) und (2) aus (3)). Sei also f Epimorphismus. Dann ist $\mathrm{Bild}(f) = V$, also $\dim(\mathrm{Bild}(f)) = \dim(V)$ und deshalb nach dem Dimensionssatz $\dim(\mathrm{Kern}(f)) = 0$, also $\mathrm{Kern}(f) = 0$. Also ist f ein Monomorphismus. Qed

Folgerung. *Sei V endlichdimensionaler K-Vektorraum und U, U' Unterräume von V. Dann gilt*

$$\dim(U + U') = \dim(U) + \dim(U') - \dim(U \cap U').$$

Beweis. Wir betrachten die folgende Abbildung.

$$\begin{aligned} f\colon U \times U' &\to U + U' \\ (x, x') &\mapsto x + x'. \end{aligned}$$

Offenbar ist f linear und surjektiv. Nach dem Dimensionssatz gilt

$$\dim(U \times U') = \dim(\mathrm{Kern}(f)) + \dim(U + U').$$

Wegen $\dim(U \times U') = \dim(U) + \dim(U')$ genügt es zu zeigen, daß $\dim(\mathrm{Kern}(f)) = \dim(U \cap U')$ ist. Dazu zeigen wir, daß die folgende Abbildung g ein Isomorphismus ist.

$$\begin{aligned} g\colon U \cap U' &\to \mathrm{Kern}(f) \\ x &\mapsto (x, -x). \end{aligned}$$

g ist wohldefiniert: Sei $x \in U \cap U'$. Dann ist $f(x, -x) = x - x = 0$, also $(x, -x) \in \mathrm{Kern}(f)$. Offenbar ist g linear und injektiv. Zu zeigen bleibt, daß g auch surjektiv ist. Sei also $(x, y) \in \mathrm{Kern}(f)$. Dann ist $x \in U$, $y \in U'$ und $f(x, y) = x + y = 0$. Also $x = -y$ und damit $x \in U \cap U'$ und $(x, y) = g(x)$. Qed

3.3 Rang von Matrizen

In diesem Abschnitt sei K stets ein Körper.

3.3.1. Definition. Sei $A \in K^{m \times n}$ und $a_1, \ldots, a_m$ die Zeilen von A sowie $b_1, \ldots, b_n$ die Spalten von A. Der Unterraum $\langle a_1, \ldots, a_m \rangle \subseteq K^{1 \times n}$ heißt *Zeilenraum* von A und $\dim(\langle a_1, \ldots, a_m \rangle)$ der *Zeilenrang* von A. Entsprechend heißt der Unterraum $\langle b_1, \ldots, b_n \rangle$ der *Spaltenraum* von A und $\dim(\langle b_1, \ldots, b_n \rangle)$ der *Spaltenrang* von A.

Bemerkungen. 1. Bei elementaren Zeilenumformungen ändert sich der Zeilenraum nicht (also auch nicht der Zeilenrang).

Beweis. Für das Vertauschen zweier Zeilen ist das klar, ebenso für das Multiplizieren einer Zeile mit $\alpha \neq 0$, $\alpha \in K$. Im Fall der Addition einer Zeile zu einer anderen beachte man, daß

$$\langle a_1, \dots, a_i, \dots, a_j, \dots, a_m \rangle = \langle a_1, \dots, a_i, \dots, a_i + a_j, \dots, a_m \rangle. \quad \text{Qed}$$

2. Ist A Stufenmatrix mit r Zeilen $\neq 0$, so ist r der Zeilenrang.

Beweis. Die r Zeilen $\neq 0$ sind linear unabhängig (da A Stufenmatrix ist) und offenbar ein Erzeugendensystem des Zeilenraums. Qed

3. Ein Berechnungsverfahren für den Zeilenrang von A ergibt sich wie folgt. Man überführe A durch elementare Zeilenumformungen in eine Stufenmatrix A'. Aus A' lese man die Anzahl r der Zeilen $\neq 0$ ab.

4. Der Zeilenrang von A ist die Maximalzahl linear unabhängiger Vektoren unter $a_1, \dots, a_m$.

Beweis. Nach einer Charakterisierung des Basisbegriffs in Satz 2.2.5 ist eine maximale linear unabhängige Teilfamilie von $(a_1, \dots, a_n)$ eine Basis. Qed

5. Für den Spaltenraum gelten (1)–(4) entsprechend.

3.3.2. Satz. *Für jedes $A \in K^{m \times n}$ ist der Zeilenrang von A gleich dem Spaltenrang von A; diese Zahl wird der* Rang *von A genannt, und mit* $\mathrm{rg}(A)$ *bezeichnet.*

Beweis. Wir betrachten die lineare Abbildung

$$\begin{aligned} f_A \colon K^n &\to K^m \\ x &\mapsto Ax. \end{aligned}$$

Schritt 1. $\mathrm{Bild}(f_A)$ ist der Spaltenraum von A, denn für die kanonische Basis $e_1, \dots, e_n$ des K^n ist $\mathrm{Bild}(f_A) = \langle f(e_1), \dots, f(e_n) \rangle$ nach dem Satz 3.1.2 über lineare Fortsetzung, und $f(e_j) = Ae_j$ ist die j-te Spalte von A.

Schritt 2. Ferner ist $\dim(\mathrm{Kern}(f_A)) = n - r$, wobei r den Zeilenrang von A bedeutet. Denn $\mathrm{Kern}(f_A) = \{\, x \in K^n \mid Ax = 0 \,\}$ hat nach Abschnitt 1.4 eine Basis $x_1, \dots, x_{n-r}$, wobei r die Anzahl der Zeilen $\neq 0$ einer Matrix A' ist, die durch elementare Zeilenumformungen aus A entsteht; die Behauptung folgt jetzt aus obiger Bemerkung.

Schritt 3. Nach dem Dimensionssatz ist

$$n = \dim(K^n) = \dim(\mathrm{Kern}(f_A)) + \dim(\mathrm{Bild}(f_A)) = n - r + \dim(\mathrm{Bild}(f_A)),$$

und $\dim(\mathrm{Bild}(f_A))$ ist der Spaltenrang von A. Qed

Bemerkung. Offenbar ist $\mathrm{rg}(A) = \mathrm{rg}(A^t)$.

3.3.3. Lemma. *Eine $n \times n$-Matrix A ist genau dann invertierbar, wenn $\mathrm{rg}(A) = n$ ist.*

Beweis. Es gilt

$$\begin{aligned} \mathrm{rg}(A) = n &\equiv \dim(\mathrm{Bild}(f_A)) = n && \text{nach dem Beweis des vorangehenden Satzes} \\ &\equiv f_A \text{ ist surjektiv} \\ &\equiv f_A \text{ ist injektiv} && \text{nach der Folgerung aus 3.2.3} \\ &\equiv \forall x \in K^{n \times 1} (Ax = 0 \to x = 0) && \text{nach 3.2.1} \\ &\equiv A \text{ ist invertierbar} && \text{nach 1.4.10. Qed} \end{aligned}$$

3.3.4. LEMMA.

1. *Für Matrizen $A \in K^{m\times n}$ und $B \in K^{n\times p}$ gilt*

$$\mathrm{rg}(A) + \mathrm{rg}(B) - n \leq \mathrm{rg}(AB) \leq \min(\mathrm{rg}(A), \mathrm{rg}(B)).$$

2. *Für Matrizen $A \in K^{n\times n}$, $B \in K^{n\times p}$ und $C \in K^{p\times p}$ mit A, C invertierbar gilt*

$$\mathrm{rg}(ABC) = \mathrm{rg}(B).$$

3. *Für Matrizen $A, B \in K^{m\times n}$ gilt*

$$\mathrm{rg}(A+B) \leq \mathrm{rg}(A) + \mathrm{rg}(B).$$

BEWEIS. 1. Zum Beweis der zweiten Ungleichung betrachten wir

$$K^p \xrightarrow{f_B} K^n \xrightarrow{f_A} K^m.$$

Nach Abschnitt 1.3 gilt $f_A \circ f_B = f_{AB}$. Also ist $\mathrm{Bild}(f_{AB}) \subseteq \mathrm{Bild}(f_A)$, also $\mathrm{rg}(AB) \leq \mathrm{rg}(A)$. Ferner gilt allgemein für eine lineare Abbildung $g\colon V \to W$, daß $\dim(g(V)) \leq \dim(V)$; dies folgt aus dem Dimensionssatz 3.2.3. Speziell für $g :=$ Einschränkung von f_A auf $\mathrm{Bild}(f_B)$ folgt

$$\dim(\underbrace{f_A(\mathrm{Bild}(f_B))}_{\mathrm{Bild}(f_{AB})}) \leq \dim(\mathrm{Bild}(f_B)),$$

also $\mathrm{rg}(AB) \leq \mathrm{rg}(B)$. Zum Beweis der ersten Ungleichung betrachten wir dasselbe g wie eben. Nach dem Dimensionssatz ist

$$\dim(\mathrm{Bild}(f_B)) = \dim(\mathrm{Kern}(g)) + \dim(\mathrm{Bild}(g)).$$

Wir schätzen beide Ausdrücke auf der rechten Seite ab. Wegen $\mathrm{Kern}(g) \subseteq \mathrm{Kern}(f_A)$ ist

$$\dim(\mathrm{Kern}(g)) \leq \dim(\mathrm{Kern}(f_A)) = n - \dim(\mathrm{Bild}(f_A)) = n - \mathrm{rg}(A).$$

Weiter ist $\mathrm{Bild}(g) = f_A(\mathrm{Bild}(f_B)) = \mathrm{Bild}(f_{AB})$, also $\dim(\mathrm{Bild}(g)) = \mathrm{rg}(AB)$. Damit ergibt sich

$$\mathrm{rg}(B) \leq n - \mathrm{rg}(A) + \mathrm{rg}(AB).$$

2. Nach Teil (1) ist

$$\underbrace{\mathrm{rg}(A)}_{n} + \mathrm{rg}(BC) - n \leq \mathrm{rg}(ABC) \leq \mathrm{rg}(BC), \quad \text{also } \mathrm{rg}(BC) = \mathrm{rg}(ABC).$$

Wieder nach Teil (1) ist

$$\mathrm{rg}(B) + \underbrace{\mathrm{rg}(C)}_{p} - p \leq \mathrm{rg}(BC) \leq \mathrm{rg}(B), \qquad \text{also } \mathrm{rg}(B) = \mathrm{rg}(BC).$$

3. Wegen $f_{A+B} = f_A + f_B$ ist $\mathrm{Bild}(f_{A+B}) \subseteq \mathrm{Bild}(f_A) + \mathrm{Bild}(f_B)$. Daraus ergibt sich

$$\dim(\mathrm{Bild}(f_{A+B})) \leq \dim(\mathrm{Bild}(f_A) + \mathrm{Bild}(f_B)) \leq \dim(\mathrm{Bild}(f_A)) + \dim(\mathrm{Bild}(f_B))$$

nach einer Folgerung aus dem Dimensionssatz 3.2.3, also $\mathrm{rg}(A+B) \leq \mathrm{rg}(A) + \mathrm{rg}(B)$. QED

Bemerkung. Mit Hilfe des Rangbegriffs kann man verschiedene Kriterien für die Lösbarkeit eines linearen Gleichungssystems $Ax = b$ formulieren. Es sei K ein Körper, $b \in K^m$ und $A \in K^{m\times n}$. Man betrachte die folgenden drei Aussagen.

L. (Lösbar). Es gibt ein $x \in K^n$ mit $Ax = b$.

EL. (Eindeutig lösbar). Es gibt genau ein $x \in K^n$ mit $Ax = b$.

UL. (Universell lösbar). Für alle $c \in K^m$ gibt es ein $x \in K^n$ mit $Ax = c$.

Dann hat man

1. (L) gilt genau dann, wenn $\mathrm{rg}(\begin{pmatrix}A & b\end{pmatrix}) = \mathrm{rg}(A)$.
2. (EL) gilt genau dann, wenn $\mathrm{rg}(\begin{pmatrix}A & b\end{pmatrix}) = \mathrm{rg}(A) = n$.
3. (UL) gilt genau dann, wenn $\mathrm{rg}(A) = m$.

Beweis. Dies ist nur eine Umformulierung der Ergebnisse von Abschnitt 1.4.

3.4 Lineare Abbildungen und Matrizen

In unseren einführenden Betrachtungen zu linearen Abbildungen in 3.1.1 (Beispiel 3) haben wir gesehen, daß sich jede lineare Abbildung $f\colon K^n \to K^m$ eindeutig darstellen läßt in der Form $f(x) = Ax$ mit $A \in K^{m\times n}$, wobei $A = (\alpha_{ij})$ bestimmt ist durch

$$f(e_j) = \sum_{i=1}^{m} \alpha_{ij} e_i \qquad \text{für } 1 \le j \le n.$$

Wir haben also auf diese Weise abstrakte Objekte (wie lineare Abbildungen) durch konkrete (wie Matrizen) beschrieben. Jetzt wollen wir diese Art von "Konkretisierung" auf lineare Abbildungen $f\colon V \to W$ mit beliebigen endlichdimensionalen Vektorräumen V, W übertragen. Um ein solches f durch eine Matrix darstellen zu können, müssen wir vorher Basen in V und in W festlegen.

3.4.1. Satz. *Sei V ein n-dimensionaler und W ein m-dimensionaler K-Vektorraum. Ferner seien Basen $\mathcal{A} = (x_1, \ldots, x_n)$ von V und $\mathcal{B} = (y_1, \ldots, y_m)$ von W gegeben. Dann ist*

$$\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}\colon \mathrm{Hom}_K(V, W) \to K^{m\times n}$$

$$\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}(f) = \text{die Matrix } A = (\alpha_{ij}) \text{ mit } f(x_j) = \sum_{i=1}^{m} \alpha_{ij} y_i \text{ für } 1 \le j \le n$$

ein Isomorphismus. $\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}(f)$ heißt die darstellende Matrix *der linearen Abbildung $f\colon V \to W$ bzgl. der Basen $\mathcal{A}$ von V und $\mathcal{B}$ von W.*

Beweis. 1. $\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}$ ist wohldefiniert, da $\mathcal{B} = (y_1, \ldots, y_n)$ eine Basis von W ist.

2. $\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}(f + g) = \mathcal{M}^{\mathcal{B}}_{\mathcal{A}}(f) + \mathcal{M}^{\mathcal{B}}_{\mathcal{A}}(g)$: Sei $\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}(f) = (\alpha_{ij})$ und $\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}(g) = (\beta_{ij})$. Dann gilt

$$\begin{aligned}(f + g)(x_j) &= f(x_j) + g(x_j)\\ &= \sum_{i=1}^{m} \alpha_{ij} y_i + \sum_{i=1}^{m} \beta_{ij} y_i\\ &= \sum_{i=1}^{m} (\alpha_{ij} + \beta_{ij}) y_i,\end{aligned}$$

also $\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}(f+g) = (\alpha_{ij}+\beta_{ij})$. $\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}(\alpha f) = \alpha \mathcal{M}^{\mathcal{B}}_{\mathcal{A}}(f)$ beweist man entsprechend.

3. $\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}$ ist surjektiv. Sei $A = (\alpha_{ij}) \in K^{m\times n}$. Nach dem Satz 3.1.2 von der linearen Fortsetzung existiert ein $f \in \mathrm{Hom}(V, W)$ mit $f(x_j) = \sum_{i=1}^m \alpha_{ij} y_i$ für $1 \leq j \leq n$, also $\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}(f) = A$.

4. $\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}$ ist injektiv. Es genügt zu zeigen, daß $\mathrm{Kern}(\mathcal{M}^{\mathcal{B}}_{\mathcal{A}}) = 0$. Gelte also $f(x_j) = 0$ für $1 \leq j \leq n$. Dann ist $f = 0$, da die lineare Fortsetzung eindeutig bestimmt ist. □

Bemerkungen. 1. Die definierenden Gleichungen für $A = (\alpha_{ij})$, also $f(x_j) = \sum_{i=1}^m \alpha_{ij} y_i$ für $1 \leq j \leq n$, kann man als Matrizengleichung formulieren:

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^t \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

2. Seien V ein n-dimensionaler und W ein m-dimensionaler K-Vektorraum, $\mathcal{A} = (x_1, \ldots, x_n)$ eine Basis von V und $\mathcal{B} = (y_1, \ldots, y_m)$ eine Basis von W. Ferner sei $f \in \mathrm{Hom}(V, W)$ und $A = \mathcal{M}^{\mathcal{B}}_{\mathcal{A}}(f)$. $x \in V$ habe bzgl. der Basis $(x_1, \ldots, x_n)$ die Koordinaten $\xi_1, \ldots, \xi_n$. Dann hat $f(x) \in W$ bzgl. der Basis $(y_1, \ldots, y_m)$ die Koordinaten

$$A\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix},$$

denn es ist $f(x) = f(\sum_{j=1}^n \xi_j x_j) = \sum_{j=1}^n \xi_j f(x_j) = \sum_{j=1}^n \xi_j \sum_{i=1}^m \alpha_{ij} y_i = \sum_{i=1}^m (\sum_{j=1}^n \alpha_{ij}\xi_j) y_i$.

3.4.2. *Darstellende Matrizen einer linearen Abbildung.* Wir wollen uns jetzt eine Übersicht über alle darstellenden Matrizen einer gegebenen linearen Abbildung $f\colon V \to W$ verschaffen, also bzgl. beliebiger Basen von V und W.

Satz. *Sei V ein n-dimensionaler und W ein m-dimensionaler K-Vektorraum und $f\colon V \to W$ linear. Dann gilt für alle $A \in K^{m\times n}$: A stellt f bzgl. geeigneter Basen von V und W dar genau dann, wenn* $\mathrm{rg}(A) = \mathrm{rg}(f) := \dim(\mathrm{Bild}(f))$. $\mathrm{rg}(f)$ *heißt der* Rang *von f.*

Beweis. *Vorbemerkung* 1. Sei $(x_1, \ldots, x_n)$ Basis von V. Dann sind auch Basen von V:

a. $(x_1, \ldots, \alpha x_i, \ldots, x_n)$ mit $\alpha \in K$, $\alpha \neq 0$,

b. $(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n)$, wobei x_j an der i-ten Stelle und x_i an der j-ten Stelle steht,

c. $(x_1, \ldots, x_i, \ldots, x_j + x_i, \ldots, x_n)$.

(a) und (b) sieht man sofort, und für (c) zeigt man leicht, daß $(x_1, \ldots, x_i, \ldots, x_j + x_i, \ldots, x_n)$ linear unabhängig ist und ein Erzeugendensystem bildet.

Vorbemerkung 2. A' entstehe aus A durch eine elementare Zeilen- und Spaltenumformung. Dann gilt:

$$\begin{aligned} &A \text{ stellt } f \text{ bzgl. geeigneter Basen von } V \text{ und } W \text{ dar} \\ \equiv\ &A' \text{ stellt } f \text{ bzgl. geeigneter Basen von } V \text{ und } W \text{ dar.} \end{aligned}$$

Dies sieht man leicht mit Hilfe von Vorbemerkung 1.

Jetzt können wir den Satz beweisen. Jedes $A \in K^{m \times n}$ läßt sich durch elementare Zeilen- und Spaltenumformungen bringen in die Form

$$\begin{pmatrix} 1 & & & 0 & 0 & \dots & 0 \\ & \ddots & & & & & \\ 0 & & & 1 & 0 & \dots & 0 \\ 0 & & & 0 & 0 & \dots & 0 \\ \vdots & & & & & & \\ 0 & & & 0 & 0 & \dots & 0 \end{pmatrix}.$$

$\mathrm{rg}(A)$ bleibt dabei unverändert. OBdA sei also A von dieser Form. Demnach ist zu zeigen: Es gibt Basen $(x_1, \dots, x_r, x_{r+1}, \dots, x_n)$ von V und $(y_1, \dots, y_r, y_{r+1}, \dots, y_m)$ von W so daß $f(x_1) = y_1, \dots, f(x_r) = y_r, f(x_{r+1}) = \dots = f(x_n) = 0$ genau dann, wenn $\mathrm{rg}(f) = r$.

$\Longrightarrow$. $\mathrm{rg}(f) = \dim(\mathrm{Bild}(f)) = \dim(\langle y_1, \dots, y_r \rangle) = r$.

$\Longleftarrow$. Gelte $\mathrm{rg}(f) = r$. Man wähle eine Basis $(x_{s+1}, \dots, x_n)$ von $\mathrm{Kern}(f)$, und ergänze sie zu einer Basis $(x_1, \dots, x_s, x_{s+1}, \dots, x_n)$ von V. Dann gilt

$$\langle x_1, \dots, x_s \rangle \cong V/\mathrm{Kern}(f) \cong \mathrm{Bild}(f)$$

und damit $r = s$. Ferner ist $(f(x_1), \dots, f(x_r))$ Basis von $\mathrm{Bild}(f)$, denn wegen $f(x_{r+1}) = \dots = f(x_n) = 0$ ist es ein Erzeugendensystem, und es ist $\dim(\mathrm{Bild}(f)) = r$. Man setze jetzt $y_1 = f(x_1), \dots, y_r = f(x_r)$ und ergänze diese Vektoren zu einer Basis $y_1, \dots, y_r, y_{r+1}, \dots, y_m$ von W. QED

3.4.3. *Darstellende Matrizen eines Endomorphismus.* Wir betrachten jetzt den Spezialfall eines Endomorphismus. Sei also V ein n-dimensionaler K-Vektorraum und $\mathcal{A} = (x_1, \dots, x_n)$ eine Basis von V. Dann ist

$$\mathcal{M}_{\mathcal{A}} \colon \mathrm{Hom}_K(V, V) \to K^{n \times n}$$
$$\mathcal{M}_{\mathcal{A}}(f) = \text{die Matrix } A = (\alpha_{ij}) \text{ mit } f(x_j) = \sum_{i=1}^{n} \alpha_{ij} x_i \text{ für } 1 \le j \le n$$

ein Isomorphismus. $\mathcal{M}_{\mathcal{A}}(f)$ heißt die *darstellende Matrix* des Endomorphismus $f \colon V \to V$ bzgl. der Basis $\mathcal{A} = (x_1, \dots, x_n)$ von V.

Wir wollen uns eine Übersicht über alle darstellenden Matrizen eines gegebenen Endomorphismus $f \colon V \to V$ verschaffen. Dieses Problem ist wesentlich schwieriger als das eben behandelte, da wir nur *eine* Basis frei wählen können. Eine Lösung werden wir später mit Hilfe des Begriffs der *Jordan-Normalform* erhalten. Hier wollen wir uns eine gewisse erste Übersicht verschaffen.

Definition. Zwei Matrizen $A, B \in K^{n \times n}$ heißen *ähnlich* (geschrieben $A \sim B$), wenn es eine invertierbare Matrix $C \in K^{n \times n}$ gibt mit $B = CAC^{-1}$.

Bemerkung. $\sim$ ist eine Äquivalenzrelation auf $K^{n \times n}$.

Beweis. Reflexivität: Es ist $A = EAE^{-1}$. Symmetrie: Wenn gilt $B = CAC^{-1}$, so folgt $A = C^{-1}BC$. Transitivität: Aus $A_2 = CA_1C^{-1}$ und $A_3 = DA_2D^{-1}$ folgt $A_3 = DCA_1C^{-1}D^{-1} = DCA_1(DC)^{-1}$. QED

SATZ. *Sei V ein n-dimensionaler K-Vektorraum und $A, B \in K^{n\times n}$. Dann gilt $A \sim B$ genau dann, wenn A und B bezüglich geeigneter Basen dasselbe $f \in \mathrm{Hom}_K(V,V)$ darstellen.*

BEWEIS. $\Longleftarrow$. Wir nehmen zunächst an, daß $A = (\alpha_{ij})$ und $B = (\beta_{ij})$ dasselbe $f \in \mathrm{Hom}_K(V,V)$ darstellen. Seien also $\mathcal{A} = (x_1, \ldots, x_n)$ und $\mathcal{B} = (y_1, \ldots, y_n)$ Basen von V und

$$f(x_j) = \sum_{i=1}^{m} \alpha_{ij} x_i \qquad\qquad f(y_j) = \sum_{i=1}^{m} \beta_{ij} y_i$$

also

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad\qquad f\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = B^t \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Wir wählen uns $C, D \in K^{n\times n}$ mit

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = C \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = D \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Dann gilt

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = CD \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

also $CD = E$ und damit $D = C^{-1}$. Man erhält

$$f\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = Cf \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = CA^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = CA^tC^{-1} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

also $B^t = CA^tC^{-1}$, d.h. $A^t \sim B^t$ und damit auch $A \sim B$.

$\Longrightarrow$. Gelte $A \sim B$. Dann ist auch $A^t \sim B^t$ und damit $B^t = CA^tC^{-1}$ mit einer invertierbaren Matrix C. Sei $\mathcal{A} = (x_1, \ldots, x_n)$ eine Basis von V. Wir definieren ein $f \in \mathrm{Hom}_K(V,V)$ durch

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

und $g, h \in \mathrm{Hom}_K(V,V)$ durch

$$g\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = C \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad h\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = C^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Dann ist

$$(h \circ g)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = Ch \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = CC^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

also $h \circ g = \mathrm{id}_V$ und ebenso $g \circ h = \mathrm{id}_V$, d.h. g und h sind Isomorphismen. Also ist

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := g \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

eine Basis von V. Man erhält

$$f \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = Cf \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = CA^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = CA^tC^{-1}C \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = CA^tC^{-1} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Also stellt B auch f dar. QED

FOLGERUNG. *Sei V ein n-dimensionaler K-Vektorraum und $f \in \mathrm{Hom}_K(V, V)$. Dann bilden alle f darstellenden Matrizen $A \in K^{n \times n}$ eine Äquivalenzklasse bzgl. $\sim$.*

Kapitel 4

Affine Geometrie

In der affinen Geometrie werden in beliebigen K-Vektorräumen V (nicht nur im $\mathbb{R}^2$, $\mathbb{R}^3$) Eigenschaften geometrischer Figuren (z.B. Parallelität, Teilverhältnisse) studiert, die bei allen "affinen" Abbildungen (z.B. Parallelprojektionen, Streckungen, Drehungen, Spiegelungen) erhalten bleiben.

Die geometrischen Figuren (d.h. die betrachteten Teilmengen von V) werden Lösungsmengen einer oder mehrerer linearer Gleichungen oder Lösungsmengen einer quadratischen Gleichung sein. Es ergeben sich z.B. Punkte, Geraden und Ebenen bzw. z.B. Ellipsen und Kegel.

4.1 Affine Unterräume und lineare Varietäten

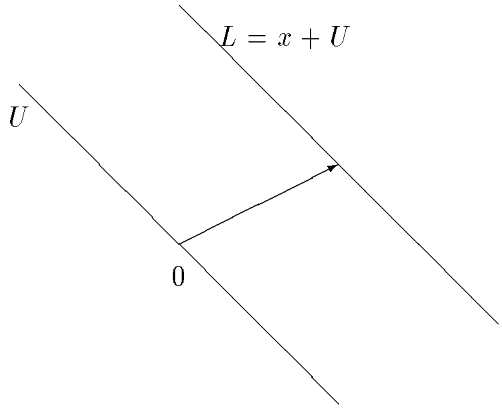
4.1.1. *Affine Unterräume.* Unter einem affinen Unterraum von V verstehen wir einen um einen Vektor x vom Nullpunkt verschobenen Unterraum.

DEFINITION. Sei V ein K-Vektorraum, $U \subseteq V$ ein Unterraum und $x \in V$.

$$L := x + U = \{\, x + y \mid y \in U \,\}$$

heißt *affiner Unterraum* von V. Außerdem heißt auch $\emptyset$ affiner Unterraum von V.

BEISPIEL. Sei $K = \mathbb{R}$ und $V = \mathbb{R}^2$.



LEMMA. *Sei V ein K-Vektorraum, $U, U' \subseteq V$ Unterräume und $x, x' \in V$. Dann gilt $x + U \subseteq x' + U'$ genau dann, wenn $U \subseteq U'$ und $x - x' \in U'$.*

BEWEIS. $\Longrightarrow$. Zunächst zeigen wir $U \subseteq U'$. Sei also $y \in U$. Dann gilt $x + y = x' + y'$ für ein $y' \in U'$ und auch $x = x' + z'$ für ein $z' \in U'$. Also ist

$$\begin{aligned} x' + z' + y &= x' + y' \\ y &= y' - z' \in U'. \end{aligned}$$

Zu zeigen bleibt $x - x' \in U'$. Es ist $x = x' + y'$ für ein $y' \in U'$, also $x - x' = y' \in U'$.

$\Longleftarrow$. Sei $U \subseteq U'$ und $x - x' \in U'$. Zu zeigen ist $x + U \subseteq x' + U'$. Sei also $z \in U$. Dann gilt

$$x + z = x' + (x - x') + z \in x' + U'.$$ QED

FOLGERUNG. *Sei V ein K-Vektorraum, $U, U' \subseteq V$ Unterräume und $x, x' \in V$. Dann gilt $x + U = x' + U'$ genau dann, wenn $U = U'$ und $x - x' \in U$.*

DEFINITION. Sei V ein K-Vektorraum, U, U' Unterräume von V und $x, x' \in V$, also $L = x + U$, $L' = x' + U'$ affine Unterräume von V.

1. $\dim(L) := \dim(U)$ heißt die *Dimension* von L. Ferner setzt man $\dim(\emptyset) := -1$.
2. L heißt *Punkt*, *Gerade* bzw. *Ebene* von V, wenn $\dim(L) = 0$, $= 1$ bzw. $= 2$ ist. L heißt *Hyperebene* von V, wenn $\dim(V) = n < \infty$ und $\dim(L) = n - 1$ ist.
3. U heißt die *Richtung* von L.
4. L heißt *parallel* zu L' (geschrieben $L \parallel L'$), wenn $U \subseteq U'$ oder $U' \subseteq U$ ist.

4.1.2. *Affine Hülle.* Wir wollen die affine Hülle einer Teilmenge M von V definieren als den kleinsten M umfassenden affinen Unterraum von V.

DEFINITION. Sei V ein K-Vektorraum, $M \subseteq V$ eine nicht leere Teilmenge und $x \in M$.

$$A(M) := x + \langle \{ x - y \mid y \in M \} \rangle$$

heißt die *affine Hülle* von M (in V).

Im folgenden Lemma zeigen wir, daß $A(M)$ unabhängig von x ist.

LEMMA. *Sei V ein K-Vektorraum und $M \subseteq V$ eine nicht leere Teilmenge. Dann gilt*

$$A(M) = \{ y \in V \mid \exists n \geq 1, y_1, \ldots, y_n \in M, \alpha_1, \ldots, \alpha_n \in K \big(y = \sum_{i=1}^{n} \alpha_i y_i \text{ und } \sum_{i=1}^{n} \alpha_i = 1 \big) \},$$

und $A(M)$ ist der kleinste affine Unterraum von V, der M enthält.

BEWEIS. 1. Setze $A_x(M) := x + \langle \{ x - y \mid y \in M \} \rangle$ für alle $x \in M$. Wir zeigen, daß $A_x(M)$ gleich der im Lemma angegebenen (und von x unabhängigen) Menge ist.

$\subseteq$. Jedes Element aus $A_x(M)$ hat die Form $x + \sum_{i=1}^{n} \alpha_i (x - y_i)$ mit $\alpha_i \in K$ und $y_i \in M$. Es ist also $= \big(1 + \sum_{i=1}^{n} \alpha_i\big) x + \sum_{i=1}^{n} (-\alpha_i) y_i$ und damit aus der obigen Menge $A(M)$, denn es gilt $1 + \sum_{i=1}^{n} \alpha_i + \sum_{i=1}^{n} (-\alpha_i) = 1$.

$\supseteq$. Seien $y_i \in M$, $\alpha_i \in K$ und $\sum_{i=1}^n \alpha_i = 1$. Dann gilt

$$\sum_{i=1}^{n} \alpha_i y_i = x + \sum_{i=1}^{n} \alpha_i y_i - \big(\sum_{i=1}^{n} \alpha_i\big)x = x + \sum_{i=1}^{n}(-\alpha_i)(x - y_i) \in A_x(M).$$

2. Nach Definition ist $A(M)$ ein affiner Unterraum von M; ferner gilt offenbar $M \subseteq A(M)$. Sei nun L ein beliebiger affiner Unterraum von V mit $M \subseteq L$. Zu zeigen ist $A(M) \subseteq L$. Sei etwa $L = x' + U$ mit $x' \in V$ und einem Unterraum U von V. Sei noch $x \in M$. Dann ist $A(M) = x + \langle \{ x - y \mid y \in M \} \rangle$. Wegen $x \in M \subseteq L = x' + U$ ist $x - x' \in U$, also nach der Folgerung in 4.1.1 $L = x + U$. Also ist

$$\begin{aligned} \{ y - x \mid y \in M \} &\subseteq \{ y - x \mid y \in L \} \subseteq U \\ \langle \{ x - y \mid y \in M \} \rangle &\subseteq U \\ A(M) &\subseteq L. \quad \text{QED} \end{aligned}$$

Bemerkung. Sei V ein K-Vektorraum und $x_0, \ldots, x_n \in V$. Dann gilt

$$\begin{aligned} A(x_0, \ldots, x_n) &:= A(\{x_0, \ldots, x_n\}) \\ &= x_0 + \langle x_0 - x_1, \ldots, x_0 - x_n \rangle \\ &= \{ \sum_{i=0}^{n} \alpha_i x_i \mid \alpha_i \in K, \sum_{i=0}^{n} \alpha_i = 1 \} \\ &= \text{kleinster affiner Unterraum von } V\text{, der } x_0, \ldots, x_n \text{ enthält.} \end{aligned}$$

Im Spezialfall zweier Vektoren $x, y \in V$ ergibt sich

$$\begin{aligned} A(x, y) &= x + K(y - x) \\ &= \{ \alpha x + (1 - \alpha)y \mid \alpha \in K \} \\ &= \text{die Gerade durch } x, y, \text{ falls } x \neq y. \end{aligned}$$

4.1.3. *Lineare Varietäten.* Sei V ein n-dimensionaler K-Vektorraum. $L \subseteq V$ heißt *lineare Varietät*, wenn es einen Isomorphismus $f\colon V \to K^n$ gibt so daß gilt

$$f(L) = \{ x \in K^n \mid Ax = b \}$$

für eine Matrix $A \in K^{m \times n}$ und ein $b \in K^m$, d.h. so daß $f(L)$ das Nullstellengebilde von endlich vielen linearen Gleichungen ist.

Bemerkungen. 1. Sei V ein n-dimensionaler K-Vektorraum und $L \subseteq V$ lineare Varietät. Ist $g\colon V \to K^n$ ein beliebiger Isomorphismus, so hat auch $g(L)$ die Form $\{ x \in K^n \mid Bx = c \}$ mit geeigneten $B \in K^{m \times n}$ und $c \in K^m$.

2. Sei $h\colon K^n \to K^n$ ein Isomorphismus, $A \in K^{m \times n}$ und $b \in K^m$. Dann ist $h(\{ x \in K^n \mid Ax = b \}) = \{ x \in K^n \mid Bx = c \}$ mit geeigneten $B \in K^{m \times n}$ und $c \in K^m$.

Beweis. 2. Jeder Isomorphismus $h\colon K^n \to K^n$ hat die Form $h(x) = Cx$ mit einer invertierbaren Matrix $C \in K^{n \times n}$. Also gilt

$$h(\{ x \in K^n \mid Ax = b \}) = \{ Cx \mid Ax = b \} = \{ y \mid AC^{-1}y = b \}.$$

1. Sei $L \subseteq V$ lineare Varietät. Dann gibt es einen Isomorphismus $f\colon V \to K^n$ so daß gilt $f(L) = \{ x \in K^n \mid Ax = b \}$. Sei nun $g\colon V \to K^n$ ein beliebiger Isomorphismus. Dann ist $h := g \circ f^{-1}\colon K^n \to K^n$ ein Isomorphismus, und es gilt nach (2)

$$g(L) = h(f(L)) = h(\{ x \in K^n \mid Ax = b \}) = \{ y \mid AC^{-1}y = b \}.$$

4.1.4. SATZ. *Sei V ein n-dimensionaler K-Vektorraum. $L \subseteq V$ ist genau dann eine lineare Varietät, wenn L affiner Unterraum von V ist.*

BEWEIS. OBdA sei $V = K^n$ und $L \neq \emptyset$.

$\Longrightarrow$. Sei $L = \{ x \in K^n \mid Ax = b \}$ mit $A \in K^{m\times n}$ und $b \in K^m$. Wegen $L \neq \emptyset$ hat L nach Abschnitt 1.4 die Form

$$L = x_0 + (Kx_1 \oplus \cdots \oplus Kx_{n-r}).$$

Mit $U := Kx_1 \oplus \cdots \oplus Kx_{n-r}$ (bzw. $U = 0$ im Fall $n = r$) ist $L = x_0 + U$, also L affiner Unterraum.

$\Longleftarrow$. Sei $L = x + U$ mit $x \in K^n$ und $U \subseteq K^n$ Unterraum. Im Fall $L = K^n$ sind wir sofort fertig ($A = 0$ und $b = 0$). Sei also $\dim(L) = \dim(U) = d < n$. Man wähle eine Basis $x_1, \ldots, x_n$ von U und betrachte das lineare Gleichungssystem

$$\begin{pmatrix} x_1^t \\ \vdots \\ x_d^t \end{pmatrix} y = 0.$$

Wegen $\operatorname{rg}\begin{pmatrix} x_1^t \\ \vdots \\ x_d^t \end{pmatrix} = d$ hat nach 1.4 der Lösungsraum die Gestalt $Ka_1 \oplus \cdots \oplus Ka_{n-d}$. Setze

$$A = \begin{pmatrix} a_1^t \\ \vdots \\ a_{n-d}^t \end{pmatrix}.$$

Wir zeigen jetzt $U = \{ y \in K^n \mid Ay = 0 \}$. Sei x_i einer der obigen Basisvektoren von U. Dann gilt

$$\begin{aligned} x_i^t a_j &= 0 \quad \text{für alle } j \\ a_j^t x_i &= 0 \quad \text{für alle } j \\ Ax_i &= 0 \end{aligned}$$

Also ist $U \subseteq \{ y \in K^n \mid Ay = 0 \}$. Die Gleichheit folgt jetzt aus $\dim(\{ y \in K^n \mid Ay = 0 \}) = n - \operatorname{rg}(A) = n - (n-d) = d = \dim(U)$. Man erhält

$$\begin{aligned} L &= x + U \\ &= x + \{ y \in K^n \mid Ay = 0 \} \\ &= \{ x + y \mid Ay = 0 \} \\ &= \{ z \mid A(z - x) = 0 \} \\ &= \{ z \mid Az = Ax \} \end{aligned}$$

Also ist L lineare Varietät. QED

4.1.5. *Affine Abbildungen.* Seien W ein K-Vektorraum und $y \in W$. Dann heißt

$$\begin{aligned} T_y \colon W &\to W \\ x &\mapsto x + y \end{aligned}$$

Translation von W um y. Sei noch $f\colon V \to W$ linear.

$$\begin{aligned} g := T_y \circ f \colon V &\to W \\ x &\mapsto f(x) + y \end{aligned}$$

heißt *affine Abbildung* von V nach W. Ist speziell $V = W$, so nennt man jede bijektive affine Abbildung $g\colon V \to V$ eine *Affinität*.

Bemerkungen. 1. Seien U, V, W K-Vektorräume und $g\colon U \to V$, $h\colon V \to W$ affine Abbildungen. Dann ist auch $h \circ g\colon U \to W$ eine affine Abbildung.

Beweis. Sei $g = T_y \circ f$ und $h = T_z \circ f$. Dann gilt für jedes $x \in U$

$$\begin{aligned} (h \circ g)(x) &= h(f(x) + y) \\ &= f'(f(x) + y) + z \\ &= (f' \circ f)(x) + f'(y) + z, \end{aligned}$$

also $h \circ g = T_{f'(y)+z} \circ (f' \circ f)$. Qed

2. Sei V ein K-Vektorraum. Ist $g\colon V \to V$ eine Affinität, so auch g^{-1}.

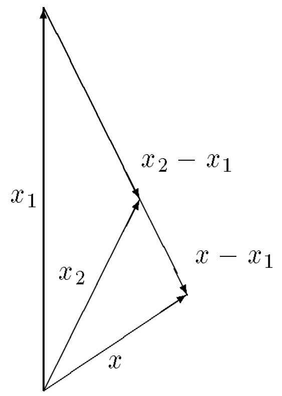
Beweis. Sei $g = T_y \circ f$ mit einem Isomorphismus $f\colon V \to V$. Setze $h := T_{f^{-1}(-y)} \circ f^{-1}$. Dann gilt

$$\begin{aligned} (g \circ h)(x) &= g(f^{-1}(x) - f^{-1}(y)) = x - y + y = x \\ (h \circ g)(x) &= h(f(x) + y) = x + f^{-1}(y) - f^{-1}(y) = x, \end{aligned}$$

also $g \circ h = h \circ g = \mathrm{id}_V$, d.h. $h = g^{-1}$. Qed

3. Sei V ein K-Vektorraum. Dann ist $\mathrm{Aff}(V) := \{\, g\colon V \to V \mid g \text{ Affinität} \,\}$ eine Gruppe, die *affine Gruppe* von V. Dies ergibt sich unmittelbar aus (1) und (2).

4.1.6. *Teilverhältnis.* Sei V ein K-Vektorraum und $x_1, x_2, x \in V$ mit $x_1 \neq x_2$ drei Vektoren aus V, die auf einer Geraden liegen, d.h. so daß $x = x_1 + \alpha(x_2 - x_1)$ $(= \alpha x_2 + (1 - \alpha)x_1)$.



Dann ist α eindeutig bestimmt (denn aus $x = x_1 + \alpha(x_2 - x_1) = x_1 + \beta(x_2 - x_1)$ folgt $(\alpha - \beta)(x_2 - x_1) = 0$, also $\alpha = \beta$). Man nennt $\mathrm{TV}(x_1, x_2, x) := \alpha$ das *Teilverhältnis* von x_1, x_2, x.

LEMMA. *(Invarianz des Teilverhältnisses unter affinen Abbildungen). Seien V, W K-Vektorräume, $g\colon V \to W$ eine affine Abbildung und $x_1, x_2, x \in V$ auf einer Geraden mit $x_1 \neq x_2$ und $g(x_1) \neq g(x_2)$. Dann liegen auch $g(x_1), g(x_2), g(x)$ auf einer Geraden und es gilt*

$$\mathrm{TV}(x_1, x_2, x) = \mathrm{TV}(g(x_1), g(x_2), g(x)).$$

BEWEIS. Sei $g = T_y \circ f$ mit $y \in V$ und $f\colon V \to W$ linear. Aus $x = x_1 + \alpha(x_2 - x_1)$ folgt

$$\begin{aligned} g(x_1) + \alpha(g(x_2) - g(x_1)) &= f(x_1) + y + \alpha(f(x_2) + y - f(x_1) - y) \\ &= f(x_1) + \alpha(f(x_2) - f(x_1)) + y \\ &= f(x_1 + \alpha(x_2 - x_1)) + y \\ &= f(x) + y \\ &= g(x). \end{aligned}$$

Also ist $\alpha = \mathrm{TV}(x_1, x_2, x) = \mathrm{TV}(g(x_1), g(x_2), g(x))$. QED

4.2 Quadratische Hyperflächen

Wir definieren quadratische Hyperflächen als Nullstellengebilde quadratischer Funktionen und klassifizieren sie bezüglich affiner Äquivalenz. In diesem Abschnitt sei stets K ein Körper mit $2 = 1 + 1 \neq 0$, also zum Beispiel $K = \mathbb{R}$.

4.2.1. *Quadratische Funktionen.* Seien $A \in K^{n\times n}$, $a \in K^n$, $\alpha \in K$ und $F\colon K^n \to K$ definiert durch $F(x) := x^t A x + a^t x + \alpha$. Das heißt mit

$$A = (\alpha_{ij}), \quad a = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

gilt

$$F(x) = F\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \sum_{i,j=1}^{n} \alpha_{ij}\xi_i\xi_j + \sum_{i=1}^{n} \alpha_i\xi_i + \alpha.$$

Ist $A \neq 0$, so heißt F *quadratische Funktion* und $V(F) := \{\, x \in K^n \mid F(x) = 0 \,\}$ *quadratische Hyperfläche* oder auch *Quadrik*.

BEMERKUNG. Man könnte wie oben wieder diese Definition auf beliebige n-dimensionale Vektorräume übertragen; gebraucht wird dazu die folgende Bemerkung 3. Das bringt aber höchstens eine Komplikation.

BEISPIELE. Sei $K = \mathbb{R}$ und $n = 2$.

$F(x) = \xi_1^2 + \xi_2^2 - 1$	$V(F)$ *Kreis*
$F(x) = \dfrac{\xi_1^2}{\alpha_1^2} + \dfrac{\xi_2^2}{\alpha_2^2} - 1$ mit $\alpha_1, \alpha_2 \neq 0$	$V(F)$ *Ellipse*
$F(x) = \xi_1\xi_2 - 1$	$V(F)$ *Hyperbel*
$F(x) = \xi_1^2 + \xi_2$	$V(F)$ *Parabel*

Bemerkungen. 1. In der obigen Definition kann A symmetrisch (d.h. $A = A^t$) angenommen werden.

Beweis. $x^tAx = (x^tAx)^t = x^tA^tx$, also

$$x^tAx = \frac{1}{2}(x^tAx + x^tA^tx) = x^t\underbrace{\left(\frac{1}{2}(A + A^t)\right)}_{B}x \quad \text{mit } B \text{ symmetrisch. Qed}$$

2. (Eindeutigkeit der Darstellung). Seien $A, A' \in K^{n\times n}$ symmetrisch, $a, a' \in K^n$, $\alpha, \alpha' \in K$ und für alle $x \in K^n$

$$\begin{aligned} F(x) &= x^tAx + a^tx + \alpha, \\ F'(x) &= x^tA'x + (a')^tx + \alpha'. \end{aligned}$$

Gilt dann $F(x) = F'(x)$ für alle $x \in K^n$, so ist $A = A'$, $a = a'$ und $\alpha = \alpha'$.

Beweis. Seien $A = (\alpha_{ij})$, $A' = (\alpha'_{ij})$, $a = (\alpha_j)$, $a' = (\alpha'_j)$.

a. Aus $F(0) = F'(0)$ folgt $\alpha = \alpha'$.

b. Seien $1 \leq i \leq n$ beliebig. Dann ist e_i^tA die i-te Zeile von A und $e_i^tAe_j = \alpha_{ij}$. Es gilt

$$\begin{aligned} F(e_i) &= e_i^tAe_i + a^te_i + \alpha = \alpha_{ii} + \alpha_i + \alpha \\ F(2e_i) &= 4\alpha_{ii} + 2\alpha_i + \alpha \end{aligned}$$

also nach Voraussetzung

$$\begin{aligned} \alpha_{ii} + \alpha_i + \alpha &= \alpha'_{ii} + \alpha'_i + \alpha', \\ 4\alpha_{ii} + 2\alpha_i + \alpha &= 4\alpha'_{ii} + 2\alpha'_i + \alpha'. \end{aligned}$$

Wegen $\alpha = \alpha'$ folgt daraus $2\alpha_{ii} = 2\alpha'_{ii}$, also (da $2 \neq 0$ in K) auch $\alpha_{ii} = \alpha'_{ii}$. Damit gilt aber auch $\alpha_i = \alpha'_i$.

c. Seien i, j beliebig mit $i \neq j$ und $1 \leq i, j \leq n$. Es gilt

$$\begin{aligned} F(e_i + e_j) &= (e_i + e_j)^tA(e_i + e_j) + a^t(e_i + e_j) + \alpha \\ &= e_i^tAe_i + e_j^tAe_i + e_i^tAe_j + e_j^tAe_j + a^te_i + a^te_j + \alpha \\ &= \alpha_{ii} + 2\alpha_{ij} + \alpha_{jj} + \alpha_i + \alpha_j + \alpha \qquad \text{da } A \text{ symmetrisch ist.} \end{aligned}$$

Aus $F(e_i + e_j) = F'(e_i + e_j)$ folgt mit (a) und (b) $2\alpha_{ij} = 2\alpha'_{ij}$, also auch $\alpha_{ij} = \alpha'_{ij}$. Qed

3. (Verhalten bei Affinitäten). Sei $F\colon K^n \to K$ eine quadratische Funktion und $g\colon K^n \to K^n$ eine Affinität. Dann ist auch $F \circ g$ eine quadratische Funktion und es gilt

$$V(F \circ g) = g^{-1}(V(F)).$$

Beweis. a. Sei $F(x) = x^tAx + a^tx + \alpha$ mit $A \in K^{n\times n}$ symmetrisch und $\neq 0$, $a \in K^n$ und $\alpha \in K$. Sei ferner

$$\begin{aligned} g\colon K^n &\to K^n \\ x &\to Cx + b \text{ mit } C \in \mathrm{GL}(n, K) \end{aligned}$$

Dann gilt

$$\begin{aligned}
F(g(x)) &= (Cx+b)^t A(Cx+b) + a^t(Cx+b) + \alpha \\
&= x^t C^t A C x + \underbrace{x^t C^t A b}_{b^t A C x} + b^t A C x + b^t A b + a^t C x + a^t b + \alpha \\
&= x^t \underbrace{C^t A C}_{=:\,A'} x + \underbrace{(2b^t A C + a^t C)}_{=:\,(a')^t} x + \underbrace{b^t A b + a^t b + \alpha}_{=:\,\alpha'}.
\end{aligned}$$

A' ist wieder symmetrisch und $\neq 0$. b.

$$\begin{aligned}
x \in V(F \circ g) &\equiv F(g(x)) = 0 \\
&\equiv g(x) \in V(F) \\
&\equiv x \in g^{-1}(V(F)).
\end{aligned}$$

4.2.2. *Affine Äquivalenz.* $F, F' \colon K^n \to K$ seien quadratische Funktionen, $Q = V(F)$ und $Q' = V(F')$. F heißt *affin äquivalent* zu F' genau dann, wenn es eine Affinität $g\colon K^n \to K^n$ gibt mit $F' = F \circ g$. Eine Teilmenge $Q \subseteq K^n$ heißt *affin äquivalent* zu $Q' \subseteq K^n$ genau dann, wenn es eine Affinität $g\colon K^n \to K^n$ gibt mit $g(Q') = Q$.

Wir betrachten jetzt folgendes Problem. Gegeben sei eine quadratische Funktion $F\colon K^n \to K$. Gesucht ist eine Affinität $g\colon K^n \to K^n$ so daß $F \circ g$ eine "möglichst einfache Form" hat. Das wesentliche Hilfsmittel hierfür ist der folgende Satz.

Satz. *Sei $A \in K^{n\times n}$ symmetrisch. Dann existieren $C \in \mathrm{GL}(n,K)$ und $\gamma_1, \ldots, \gamma_n \in K$ so daß*

$$C^t A C = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

(Hier und im folgenden sollen nicht ausgefüllte Stellen von Matrizen 0 sein).

Beweis. OBdA sei A nicht von Diagonalgestalt (sonst kann man $C = E$ wählen). Sei i minimal mit $\alpha_{ij} \neq 0$ für ein $j > i$.

$$A = \begin{pmatrix} \alpha_{11} & & & & \\ & \ddots & & & \\ & & \alpha_{ii} & \ldots & \alpha_{in} \\ & & \vdots & & \vdots \\ & & \alpha_{ni} & \ldots & \alpha_{nn} \end{pmatrix}$$

Fall 1. $\alpha_{ii} \neq 0$. Setze

$$C = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & \frac{-\alpha_{i,i+1}}{\alpha_{ii}} & \ldots & \frac{-\alpha_{in}}{\alpha_{ii}} \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \in \mathrm{GL}(n,K).$$

Dann gilt

$$C^t = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \frac{-\alpha_{i,i+1}}{\alpha_{ii}} & 1 & & \\ & & \vdots & & \ddots & \\ & & \frac{-\alpha_{in}}{\alpha_{ii}} & & & 1 \end{pmatrix},$$

also

$$C^t A = \begin{pmatrix} \alpha_{11} & & & & & \\ & \ddots & & & & \\ & & \alpha_{ii} & & \dots & \alpha_{in} \\ & & 0 & * & \dots & * \\ & & \vdots & \vdots & & \vdots \\ & & 0 & * & \dots & * \end{pmatrix}$$

und damit

$$C^t AC = \begin{pmatrix} \alpha_{11} & & & & & \\ & \ddots & & & & \\ & & \alpha_{ii} & 0 & \dots & 0 \\ & & 0 & * & \dots & * \\ & & \vdots & \vdots & & \vdots \\ & & 0 & * & \dots & * \end{pmatrix}.$$

Fall 2. $\alpha_{ii} = 0$, aber $a_{jj} \neq 0$ für ein $j > i$. Setze

$$C = \begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & 0 & & & & 1 & & \\ & & & & 1 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & 1 & & & \\ & & & 1 & & & & 0 & & \\ & & & & & & & & 1 & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{pmatrix} \in \mathrm{GL}(n, K).$$

$C^t AC = CAC$ entsteht aus A durch Vertauschen der i-ten mit der j-ten Zeile und Spalte. Jetzt kann man wie im Fall (1) fortfahren.

Fall 3. $\alpha_{ii} = \cdots = \alpha_{nn} = 0$ und $\alpha_{ij} \neq 0$ für ein $j > i$. Setze

$$C = \begin{pmatrix}
1 &&&&&&&&&& \\
& \ddots &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& 1 &&&& 1 &&& \\
&&&& 1 &&&&&& \\
&&&&& \ddots &&&&& \\
&&&&&& 1 &&&& \\
&&& 1 &&&& -1 &&& \\
&&&&&&&& 1 && \\
&&&&&&&&& \ddots & \\
&&&&&&&&&& 1
\end{pmatrix} \in \mathrm{GL}(n, K).$$

Dann gilt

$$C^t A = CA = \begin{pmatrix}
\alpha_{11} &&&&&&& \\
& \ddots &&&&&& \\
&& \alpha_{i-1,i-1} &&&&& \\
&&& \underbrace{\alpha_{ii}}_{0} + \alpha_{ji} & \dots & \alpha_{ij} + \underbrace{\alpha_{jj}}_{0} & \dots & \alpha_{in} + \alpha_{jn} \\
&&& * & \dots & * & \dots & * \\
&&& \vdots &&&& \vdots \\
&&& * & \dots & * & \dots & *
\end{pmatrix},$$

also

$$CAC = \begin{pmatrix}
\alpha_{11} &&&&&& \\
& \ddots &&&&& \\
&& \alpha_{i-1,i-1} &&&& \\
&&& 2\alpha_{ij} & * & \dots & * \\
&&& * & * && \\
&&& \vdots &&& \vdots \\
&&& * && \dots & *
\end{pmatrix}.$$

Wegen $\alpha_{ij} \neq 0$ ist auch $2\alpha_{ij} \neq 0$ und wir können wie im Fall (1) fortfahren.

Durch Iteration dieses Verfahrens erhält man eine Matrix von Diagonalgestalt. Die im Satz geforderte Matrix C ist dann das Produkt aller verwendeten Matrizen. QED

BEMERKUNG. Wie im Fall 2 des Beweises ergibt sich, daß man zu gegebenem symmetrischen $A \in K^{n \times n}$ ein $C \in \mathrm{GL}(n, K)$ und $\gamma_1, \dots, \gamma_r \in K \setminus \{0\}$ finden kann so daß

$$C^t AC = \begin{pmatrix}
\gamma_1 &&&&& \\
& \ddots &&&& \\
&& \gamma_r &&& \\
&&& 0 && \\
&&&& \ddots & \\
&&&&& 0
\end{pmatrix}.$$

4.2.3. SATZ. *Jede quadratische Hyperfläche $Q \subseteq K^n$ ist affin äquivalent zu einer Fläche des Typs $V(F)$, wobei für F folgende Formen auftreten.*

$$F(x) = \sum_{i=1}^{r} \gamma_i \xi_i^2 + 1 \qquad \text{Mittelpunktsfläche} \tag{4.1}$$

$$F(x) = \sum_{i=1}^{r} \gamma_i \xi_i^2 \qquad \text{Kegelfläche} \tag{4.2}$$

$$F(x) = \sum_{i=1}^{r} \gamma_i \xi_i^2 + \xi_{r+1} \qquad \text{parabolische Fläche.} \tag{4.3}$$

Dabei ist $r \le n$ in (4.1), (4.2), $r < n$ in (4.3), $\gamma_1, \ldots, \gamma_r \in K$ und $x = \big(\xi_1, \ldots, \xi_n\big)^t$. Im Fall $K = \mathbb{R}$ kann $\gamma_i \in \{\pm 1\}$ gewählt werden, und im Fall $K = \mathbb{C}$ kann $\gamma_i = 1$ gewählt werden.

BEWEIS. Sei $Q = V(G)$ und $G = x^t A x + a^t x + \alpha$ mit $A \in K^{n \times n}$ symmetrisch, $a \in K^n$ und $\alpha \in K$. Nach dem Satz in 4.2.2 gibt es $C \in \mathrm{GL}(n, K)$ und $\gamma_1, \ldots, \gamma_n \in K$ mit

$$C^t A C = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

Dann gilt

$$\begin{aligned} G(Cx) &= x^t C^t A C x + a^t C x + \alpha \\ &= \sum_{i=1}^{n} \gamma_i \xi_i^2 + \sum_{i=1}^{n} \beta_i \xi_i + \alpha \\ &=: G_1(x). \end{aligned}$$

G_1 ist affin äquivalent zu G, da $x \mapsto Cx$ eine Affinität ist. Nach obiger Bemerkung kann man annehmen, daß $\gamma_1, \ldots, \gamma_r \neq 0$ und $\gamma_{r+1} = \cdots = \gamma_n = 0$ sind für ein r mit $1 \le r \le n$.

$$\begin{aligned} G_1(x) &= \sum_{i=1}^{r} \gamma_i \xi_i^2 + \sum_{i=1}^{n} \beta_i \xi_i + \alpha \\ &= \sum_{i=1}^{r} \gamma_i \big(\xi_i + \frac{\beta_i}{2\gamma_i}\big)^2 + \sum_{i=r+1}^{n} \beta_i \xi_i + \alpha - \sum_{i=1}^{r} \frac{\beta_i^2}{4\gamma_i} \\ &=: G_2(x') \quad \text{mit } \xi_i' := \begin{cases} \xi_i + \dfrac{\beta_i}{2\gamma_i} & \text{für } 1 \le i \le r \\ \xi_i & \text{für } r+1 \le i \le n. \end{cases} \end{aligned}$$

G_2 ist affin äquivalent zu G_1, da $x \mapsto x'$ eine Translation ist, und es gilt

$$G_2(x') = \sum_{i=1}^{r} \gamma_i (\xi_i')^2 + \sum_{i=r+1}^{n} \beta_i \xi_i' + \alpha' = G_1(x) \quad \text{mit } \alpha' = \alpha - \sum_{i=1}^{r} \frac{\beta_i^2}{4\gamma_i}.$$

Fall 1. $\beta_{r+1} = \cdots = \beta_n = 0$, $\alpha' \neq 0$. OBdA sei $\alpha' = 1$ (bei Division durch α' bleibt $V(G_2)$ unverändert). Dann sind G und G_2 affin äquivalent, also auch $Q = V(G)$ und $V(G_2)$ affin äquivalent, und $V(G_2)$ ist vom Typ (4.1).

Fall 2. $\beta_{r+1} = \cdots = \beta_n = 0$, $\alpha' = 0$. Q ist bis auf affine Äquivalenz vom Typ (4.2).

Fall 3. $\beta_i \neq 0$ für ein i mit $r+1 \leq i \leq n$. Zur Vereinfachung der Schreibweise nehmen wir $\beta_{r+1} \neq 0$ an.

$$\begin{aligned} G_2(x') &= \sum_{i=1}^{r} \gamma_i(\xi_i'')^2 + \xi_{r+1}'' \\ &=: G_3(x'') \quad \text{mit } \xi_i'' := \xi_i' \text{ für } i \neq r+1 \text{ und } \xi_{r+1}'' := \sum_{i=r+1}^{n} \beta_i \xi_i' + \alpha'. \end{aligned}$$

G_3 ist affin äquivalent zu G_2, da

$$x' \mapsto \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & \beta_{r+1} & & \cdots & \beta_n \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} x' + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \alpha' \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

eine Affinität ist. Also ist Q bis auf affine Äquivalenz vom Typ (4.3).

Sei jetzt $K = \mathbb{R}$. Man wähle $\lambda_i \in \mathbb{R}$ mit

$$\begin{aligned} \gamma_i &= \lambda_i^2 \quad && \text{für } \gamma_i > 0, \\ \gamma_i &= -\lambda_i^2 && \text{für } \gamma_i < 0. \end{aligned}$$

Mit $\xi_i' = \lambda_i \xi_i$ folgt $\gamma_i \xi_i^2 = \pm(\lambda_i \xi_i)^2 = \pm(\xi_i')^2$.

Im Fall $K = \mathbb{C}$ wähle man $\lambda_i \in \mathbb{C}$ mit $\gamma_i = \lambda_i^2$. Mit $\xi_i' = \lambda_i \xi_i$ folgt $\gamma_i \xi_i^2 = (\lambda_i \xi_i)^2 = (\xi')^2$. QED

Bemerkung. Aus dem Beweis ergibt sich noch, daß jede quadratische Funktion F affin äquivalent ist zu einer quadratischen Funktion G ist, die bis auf einen konstanten Faktor von einer der Formen (4.1) – (4.3) ist.

4.2.4. *Klassifikation der quadratischen Hyperflächen $\neq \emptyset$ im $\mathbb{R}^2$ und $\mathbb{R}^3$.* Im $\mathbb{R}^2$ ist jede quadratische Hyperfläche $\neq \emptyset$ affin äquivalent zu einem $V(F_i)$, $i = 1, \ldots, 7$:

Mittelpunktsflächen
$r = 2$

	$F_1(x) = -\xi_1^2 - \xi_2^2 + 1$	Ellipse,
	$F_2(x) = -\xi_1^2 + \xi_2^2 + 1$	Hyperbel,

$r = 1$

	$F_3(x) = -\xi_1^2 + 1$	zwei parallele Geraden

Kegelflächen
$r = 2$

	$F_4(x) = \xi_1^2 + \xi_2^2$	Punkt
	$F_5(x) = \xi_1^2 - \xi_2^2$	zwei sich schneidende Geraden

$r = 1$

	$F_6(x) = \xi_1^2$	Gerade

Parabolische Flächen
$r = 1$

	$F_7(x) = \xi_1^2 + \xi_2$	Parabel

Im $\mathbb{R}^3$ ist jede quadratische Hyperfläche $\neq \emptyset$ affin äquivalent zu einem $V(F_i)$, $i = 1, \ldots, 14$:

Mittelpunktsflächen

$r = 3$

$F_1(x) = -\xi_1^2 - \xi_2^2 - \xi_3^2 + 1$	Ellipsoid
$F_2(x) = -\xi_1^2 - \xi_2^2 + \xi_3^2 + 1$	einschaliges Hyperboloid
$F_3(x) = -\xi_1^2 + \xi_2^2 + \xi_3^2 + 1$	zweischaliges Hyperboloid

$r = 2$

$F_4(x) = -\xi_1^2 - \xi_2^2 + 1$	elliptischer Zylinder
$F_5(x) = -\xi_1^2 + \xi_2^2 + 1$	hyperbolischer Zylinder

$r = 1$

$F_6(x) = -\xi_1^2 + 1$	zwei parallele Ebenen

Kegelflächen

$r = 3$

$F_7(x) = \xi_1^2 + \xi_2^2 + \xi_3^2$	Punkt
$F_8(x) = \xi_1^2 + \xi_2^2 - \xi_3^2$	Kegel

$r = 2$

$F_9(x) = \xi_1^2 + \xi_2^2$	Gerade
$F_{10}(x) = \xi_1^2 - \xi_2^2$	zwei sich schneidende Ebenen

$r = 1$

$F_{11}(x = \xi_1^2$	Ebene

Parabolische Flächen

$r = 2$

$F_{12}(x) = \xi_1^2 + \xi_2^2 + \xi_3$	elliptisches Paraboloid
$F_{13}(x) = \xi_1^2 - \xi_2^2 + \xi_3$	hyperbolisches Paraboloid

$r = 1$

$F_{14}(x) = \xi_1^2 + \xi_2$	parabolischer Zylinder

4.2.5. Bemerkungen. Man kann zeigen:

1. In $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ sind je zwei der angegebenen $V(F_i)$ nicht affin äquivalent.
2. Allgemein sind der Typ (d.h. (4.1), (4.2) oder (4.3)) und die Zahl r affine Invarianten.
3. Im Fall $K = \mathbb{R}$ ist die Anzahl der i mit $\gamma_i = +1$ eine affine Invariante (Sylvestersches Trägheitsgesetz).

4.2.6. Bemerkung. In Maple kann man mit `implicitplot3d` diese Flächen graphisch darstellen. Zum Beispiel liefert der Aufruf

```
> with(plots): implicitplot3d(x^2+2*y^2+3*z^2-3,x=-2..2,y=-2..2,z=-2..2);
```

ein Ellipsoid.

Kapitel 5

Euklidische Geometrie

In der euklidischen Geometrie unterscheidet man feiner als in der affinen Geometrie. Etwa ein Kreis wird als verschieden von einer Ellipse aufgefaßt. Ferner wird die Messung von Längen und Winkeln möglich.

5.1 Euklidische Vektorräume

Wir wollen Vektorräume mit einer zusätzlichen Struktur versehen, die die Messung von Längen und Winkeln möglich macht. Zur Motivation betrachten wir den $\mathbb{R}^2$. Die Länge eines Vektors $x = \left(\begin{smallmatrix}\xi_1\\ \xi_2\end{smallmatrix}\right) \in \mathbb{R}^2$ sei erklärt durch

$$\|x\| := \sqrt{\xi_1^2 + \xi_2^2}.$$

Der aus der Schule bekannte sogenannte Cosinussatz sagt folgendes aus. Seien $x, y \neq 0$ und φ der Winkel zwischen x und y. Dann gilt

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| \cos\varphi.$$

Man erhält für $y = \left(\begin{smallmatrix}\eta_1\\ \eta_2\end{smallmatrix}\right)$

$$\begin{aligned}
\|x\| \cdot \|y\| \cos\varphi &= \frac{1}{2}\big(\|x+y\|^2 - \|x\|^2 - \|y\|^2\big)\\
&= \frac{1}{2}\big((\xi_1+\eta_1)^2 + (\xi_2+\eta_2)^2 - \xi_1^2 - \xi_2^2 - \eta_1^2 - \eta_2^2\big)\\
&= \xi_1\eta_1 + \xi_2\eta_2\\
&:= \sigma(x, y).
\end{aligned}$$

Mit Hilfe des so definierten $\sigma\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ kann man also Längen und Winkel messen:

$$\begin{aligned}
\|x\| &= \sqrt{\sigma(x,x)}\\
\cos\varphi &= \frac{\sigma(x,y)}{\|x\| \cdot \|y\|}.
\end{aligned}$$

In unserem allgemeinen Kontext gehen wir von einer axiomatischen Beschreibung des Skalarprodukts σ aus und erklären die Länge eines Vektors und den Winkel zwischen zwei Vektoren durch die obigen Beziehungen.

5.1.1. Definition. Sei V ein $\mathbb{R}$-Vektorraum und $\sigma\colon V \times V \to \mathbb{R}$ eine Abbildung. σ heißt *euklidisches Skalarprodukt*, wenn σ *bilinear* ist, d.h. für alle $x, y, z \in V$ und alle $\alpha, \beta \in \mathbb{R}$ gilt

$$\sigma(\alpha x + \beta y, z) = \alpha\sigma(x, z) + \beta\sigma(y, z),$$
$$\sigma(x, \alpha y + \beta z) = \alpha\sigma(x, y) + \beta\sigma(y, z).$$

ferner σ *symmetrisch* ist, d.h. für alle $x, y \in V$ gilt

$$\sigma(x, y) = \sigma(y, x),$$

und σ *positiv definit* ist, d.h. für alle $x \in V$ gilt

$$\sigma(x, x) > 0 \quad \text{für } x \neq 0.$$

(V, σ) heißt *euklidischer Vektorraum*, wenn σ ein euklidisches Skalarprodukt auf V ist.

Beispiele. 1. $V = \mathbb{R}^n$, und für

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}$$

setzt man

$$\sigma(x, y) := x^t y = \xi_1\eta_1 + \cdots + \xi_n\eta_n.$$

Man zeigt dann leicht, daß damit ein euklidisches Skalarprodukt definiert ist. Man nennt σ das *kanonische Skalarprodukt* im $\mathbb{R}^n$.

2. Sei (V, σ) euklidischer Vektorraum, W ein $\mathbb{R}$-Vektorraum und $f\colon W \to V$ ein Monomorphismus. Setze

$$\sigma_f(y_1, y_2) := \sigma(f(y_1), f(y_2))$$

für alle $y_1, y_2 \in W$. Dann ist (W, σ_f) ein euklidischer Vektorraum.

3. Sei

$$V := \{\, (\alpha_0, \alpha_1, \alpha_2, \dots) \in \mathbb{R}^{\mathbb{N}} \mid \sum_{i=0}^{\infty} \alpha_i^2 < \infty \,\}$$

und

$$\sigma(a, b) = \sum_{i=0}^{\infty} \alpha_i\beta_i$$

für $a = (\alpha_i)$ und $b = (\beta_i) \in \mathbb{R}^{\mathbb{N}}$. Dann ist (V, σ) ein euklidischer Vektorraum, der sogenannte Hilbertsche Folgenraum (Beweis, insbesondere daß σ wohldefiniert: s. Analysis).

4. $V = \{\, f\colon [a, b] \to \mathbb{R} \mid f \text{ stetig} \,\}$

$$\sigma(f, g) := \int_a^b f(t)g(t)\,dt$$

für $f, g \in V$. Dann ist (V, σ) euklidischer Vektorraum. (Beweis: s. Analysis).

Bemerkung. Sei (V,σ) ein euklidischer Vektorraum.

1. Für alle $x \in V$ gilt

$$\sigma(x,0) = 0,$$

denn aufgrund der Bilinearität ist $\sigma(x,0) = \sigma(x,0) + \sigma(x,0)$.

2. Sei $x_1,\ldots,x_n$ eine Basis von V, $x = \sum_{i=1}^n \xi_i x_i$ und $y = \sum_{i=1}^n \eta_i x_i$. Dann gilt

$$\sigma(x,y) = (\xi_1,\ldots,\xi_n)A\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix} \quad \text{mit } A = (\sigma(x_i,x_j)) \in \mathbb{R}^n.$$

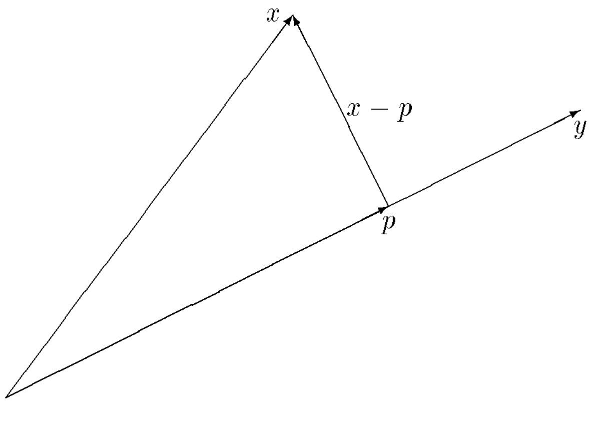
Beweis

$$\begin{aligned}\sigma(x,y) &= \sigma(\sum_{i=1}^n \xi_i x_i, \sum_{i=1}^n \eta_i x_i)\\ &= \sum_{i=1}^n\sum_{j=1}^n \xi_i\eta_j\sigma(x_i,x_j)\\ &= (\xi_1,\ldots,\xi_n)A\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix}. \quad \text{Qed}\end{aligned}$$

5.1.2. *Orthogonale Projektion auf einen Vektor.* Sei (V,σ) ein euklidischer Vektorraum und $x,y \in V$. Dann nennt man $\|x\| := \sqrt{\sigma(x,x)} \geq 0$ die *Länge* von x. Ferner nennt man x *orthogonal* zu y und schreibt $x \perp y$, wenn gilt $\sigma(x,y) = 0$. Ist noch $y \neq 0$, so nennt man ein $p \in V$ *orthogonale Projektion* von x auf y, wenn gilt

1. $p \in \mathbb{R}y$ und
2. $x - p \perp y$.

$h := x - p$ heißt *Lot* von x auf y.



Satz. *Sei (V,σ) ein euklidischer Vektorraum, $x,y \in V$ und $y \neq 0$. Es gibt genau eine orthogonale Projektion von x auf y, und zwar*

$$p = \frac{\sigma(x,y)}{\sigma(y,y)}y.$$

Beweis. 1. Eindeutigkeit. Sei p orthogonale Projektion von x auf y, also $p = \alpha y$ und $x - p \perp y$. Dann gilt $\sigma(x-p, y) = 0$, also $\sigma(x, y) = \alpha\sigma(y, y)$ und daher $\alpha = \frac{\sigma(x,y)}{\sigma(y,y)}$, da $y \neq 0$.

2. Existenz. Setze $p = \frac{\sigma(x,y)}{\sigma(y,y)}y$. Dann ist $p \in \mathbb{R}y$ und

$$\sigma(x-p, y) = \sigma(x, y) - \frac{\sigma(x, y)}{\sigma(y, y)}\sigma(y, y) = 0. \qquad \text{QED}$$

5.1.3. Satz. *Sei (V, σ) ein euklidischer Vektorraum und $x, y \in V$. Dann gilt*

1. $\|x\| = 0$ *genau dann, wenn* $x = 0$.
2. $\|\alpha x\| = |\alpha| \cdot \|x\|$ *für alle* $\alpha \in \mathbb{R}$.
3. *(Schwarzsche Ungleichung).* $|\sigma(x, y)| \leq \|x\| \cdot \|y\|$. *Die Gleichheit gilt genau dann, wenn x und y linear abhängig sind.*
4. *(Dreiecksungleichung).* $\|x + y\| \leq \|x\| + \|y\|$.

Beweis. 1. Die Behauptung folgt aus $\|x\|^2 = \sigma(x, x)$.

2. $\|\alpha x\|^2 = \sigma(\alpha x, \alpha x) = \alpha^2\sigma(x, x) = |\alpha|^2\|x\|^2$.

3. OBdA sei $y \neq 0$. Sei $p = \frac{\sigma(x,y)}{\sigma(y,y)}y$ die orthogonale Projektion von x auf y und $h := x - p$ das Lot von x auf y. Dann gilt

$$\begin{aligned}
0 &\leq \sigma(h, h) \\
&= \sigma(x-p, x-p) \\
&= \sigma(x, x) - 2\sigma(x, p) + \sigma(p, p) \\
&= \sigma(x, x) - 2\frac{\sigma(x, y)}{\sigma(y, y)}\sigma(x, y) + \frac{\sigma(x, y)^2}{\sigma(y, y)^2}\sigma(y, y) \\
&= \frac{\sigma(x, x)\sigma(y, y) - \sigma(x, y)^2}{\sigma(y, y)},
\end{aligned}$$

also $\sigma(x, y)^2 \leq \sigma(x, x)\sigma(y, y)$. Die Gleichheit gilt genau dann, wenn $h = 0$, also wenn $x = p = \frac{\sigma(x,y)}{\sigma(y,y)}y$. Im Fall $h = 0$ sind also x, y linear abhängig. Umgekehrt hat man

$$\begin{aligned}
x, y \text{ sind linear abhängig} &\implies \alpha x + \beta y = 0 \quad \text{mit } \alpha \neq 0 \text{ (da } y \neq 0) \\
&\implies x = \gamma y \\
&\implies \sigma(x, y) = \gamma\sigma(y, y).
\end{aligned}$$

Also ist

$$x = \frac{\sigma(x, y)}{\sigma(y, y)}y.$$

4. Man erhält aufgrund der Schwarzschen Ungleichung

$$\begin{aligned}
\sigma(x+y, x+y) &= \sigma(x, x) + 2\sigma(x, y) + \sigma(y, y) \\
&\leq \sigma(x, x) + 2\|x\| \cdot \|y\| + \sigma(y, y) \\
&= \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2. \quad \text{QED}
\end{aligned}$$

5.1.4. *Winkel.* Wir können jetzt Winkel wie folgt erklären. Sei (V, σ) ein euklidischer Vektorraum, $x, y \in V$ und $x, y \neq 0$. Sei weiter $\varphi \in \mathbb{R}$ mit $0 \leq \varphi \leq \pi$. φ heißt *Winkel* zwischen x und y, wenn gilt

$$\cos \varphi = \frac{\sigma(x, y)}{\|x\| \cdot \|y\|}.$$

Bemerkungen. Sei (V, σ) ein euklidischer Vektorraum und $x, y \in V$.

1. Für $x, y \neq 0$ ist der Winkel zwischen x und y wohldefiniert, denn nach dem Satz gilt

$$-1 \leq \frac{\sigma(x, y)}{\|x\| \cdot \|y\|} \leq +1.$$

Also gibt es (s. Analysis) ein eindeutig bestimmtes $\varphi \in \mathbb{R}, 0 \leq \varphi \leq \pi$ mit $\cos \varphi = \frac{\sigma(x,y)}{\|x\| \cdot \|y\|}$.

2. Sei $x, y \neq 0$ und p die orthogonale Projektion von x auf y. Dann gilt $\frac{\|p\|}{\|x\|} = \cos \varphi$, wobei φ der Winkel zwischen x und y ist.

Beweis.

$$\frac{\|p\|}{\|x\|} = \frac{|\sigma(x, y)|}{\|y\| \cdot \|x\|} = \cos \varphi. \quad \text{QED}$$

3. (Cosinussatz) Seien $x, y \neq 0$ und φ der Winkel zwischen x und y. Dann gilt

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| \cos \varphi.$$

Insbesondere ist $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ genau dann, wenn $x \perp y$ (Satz des Pythagoras).

Beweis. $\sigma(x + y, x + y) = \sigma(x, x) + \sigma(y, y) + 2\sigma(x, y)$, und $\sigma(x, y) = \|x\| \cdot \|y\| \cos \varphi$. QED

4. Als Beispiel wollen wir den Winkel zwischen einer Kante x und der Diagonale y des Einheitswürfels im $\mathbb{R}^4$ berechnen, wobei wir das kanonische Skalarprodukt σ im $\mathbb{R}^4$ zugrunde legen. Es sei also

$$x = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Dann hat man $\sigma(x, y) = 1$, sowie $\|x\| = 1$ und $\|y\| = \sqrt{4} = 2$. Für den gesuchten Winkel φ gilt also

$$\cos \varphi = \frac{\sigma(x, y)}{\|x\| \cdot \|y\|} = \frac{1}{2}, \quad \text{d.h.} \quad \varphi = \frac{\pi}{3}.$$

Es handelt sich demnach um einen Winkel von 60 Grad.

5. σ läßt sich durch $\| \cdot \|$ ausdrücken:

$$\sigma(x + y, x + y) = \sigma(x, x) + \sigma(y, y) + 2\sigma(x, y),$$

also

$$\sigma(x, y) = \frac{1}{2}\big(\|x + y\|^2 - \|x\|^2 - \|y\|^2\big).$$

5.2 Orthonormalisierung

5.2.1. *Orthonormalsystem und Orthonormalbasis.* Sei (V, σ) ein euklidischer Vektorraum und $(e_i)_{i \in I}$ eine Familie von Vektoren aus V. $(e_i)_{i \in I}$ heißt *Orthonormalsystem* (ONS) in V, wenn gilt

$$\begin{aligned} \sigma(e_i, e_j) &= 0 \quad \text{für } i, j \in I,\ i \neq j, \\ \sigma(e_i, e_i) &= 1 \quad \text{für } i \in I. \end{aligned}$$

$(e_i)_{i \in I}$ heißt *Orthonormalbasis* (ONB) in V, wenn zusätzlich $(e_i)_{i \in I}$ Basis von V ist.

Bemerkungen. 1. Im Standardbeispiel $(\mathbb{R}^n, \sigma)$ mit $\sigma(x, y) = x^t y$ ist $(e_1, \ldots, e_n)$ Orthonormalbasis.

2. Ist $(e_i)_{i \in I}$ Orthonormalsystem, so ist $(e_i)_{i \in I}$ linear unabhängig. Beweis: Sei $J \subseteq I$, J endlich. Gelte

$$\sum_{i \in J} \alpha_i e_i = 0.$$

Dann gilt für jedes $j \in J$

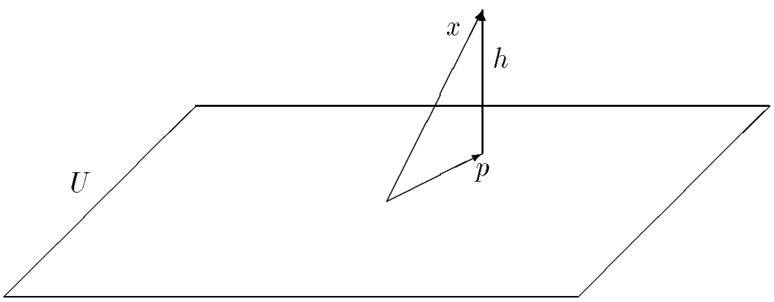
$$0 = \sigma(\sum_{i \in J} \alpha_i e_i, e_j) = \sum_{i \in J} \alpha_i \sigma(e_i, e_j) = \alpha_j.$$

Unser Ziel ist die Konstruktion einer Orthonormalbasis aus einer gegebenen Basis. Dazu behandeln wir zunächst die orthogonale Projektion auf einen Unterraum.

5.2.2. *Orthogonale Projektion auf einen Unterraum.* Sei (V, σ) ein euklidischer Vektorraum, $x \in V$ und $U \subseteq V$ Unterraum. $p \in V$ heißt *orthogonale Projektion* von x auf U, wenn gilt

1. $p \in U$ und
2. $x - p \perp U$ (d.h. $x - p \perp u$ für alle $u \in U$).

$h := x - p$ heißt *Lot* von x auf U. (Für $U = \mathbb{R}y$ ist das die alte Definition).



5.2.3. Satz. *Sei (V, σ) ein euklidischer Vektorraum, $x \in V$, $U \subseteq V$ Unterraum und $(e_1, \ldots, e_n)$ eine Orthonormalbasis von U. Dann gibt es genau eine orthogonale Projektion von x auf U, und zwar*

$$p = \sum_{i=1}^{n} \sigma(x, e_i) e_i.$$

BEWEIS. 1. Eindeutigkeit. Annahme: p ist orthogonale Projektion von x auf U. Dann gilt

$$\begin{aligned}
&p = \sum_{i=1}^{n} \alpha_i e_i \\
&x - p \perp e_j \quad \text{für } j = 1, \ldots, n \\
&\sigma(x - p, e_j) = 0 \\
&\sigma(x, e_j) = \sigma(p, e_j) = \sigma(\sum_{i=1}^{n} \alpha_i e_i, e_j) = \sum_{i=1}^{n} \alpha_i \sigma(e_i, e_j) = \alpha_j.
\end{aligned}$$

2. Existenz. Setze $p = \sum_{i=1}^{n} \sigma(x, e_i) e_i$. Dann ist $p \in U$, und es gilt

$$\begin{aligned}
&\sigma(x - p, e_j) = \sigma(x, e_j) - \sum_{i=1}^{n} \sigma(x, e_i)\sigma(e_i, e_j) = \sigma(x, e_j) - \sigma(x, e_j) = 0 \\
&\sigma(x - p, u) = 0 \quad \text{für alle } u \in U, \text{ also} \\
&x - p \perp U. \quad \text{QED}
\end{aligned}$$

BEMERKUNGEN. 1. $\sigma(x, e_i)$ heißt auch *i-ter* FOURIER*koeffizient* von x bzgl. $(e_1, \ldots, e_n)$.

2. Die Länge $\|h\|$ des Lotes von x auf U ist der *Abstand* des Vektors x von dem Unterraum U, definiert durch

$$d(x, U) := \min\{\, \|x - u\| \mid u \in U \,\}.$$

BEWEIS. Es gilt $h = x - p$, wobei p die orthogonale Projektion von x auf U ist. Zu zeigen ist $\|x - p\| = \min\{\, \|x - u\| \mid u \in U \,\}$. Wegen $p \in U$ gilt $\geq$. Umgekehrt gilt für beliebiges $u \in U$

$$\begin{aligned}
x - u &= x - p + p - u \\
\|x - u\|^2 &= \|x - p\|^2 + \|p - u\|^2 \quad \text{da } (x - p) \perp (p - u) \\
\|x - u\| &\geq \|x - p\|. \quad \text{QED}
\end{aligned}$$

5.2.4. SCHMIDT*sches Orthonormalisierungsverfahren.* Wir geben jetzt ein Verfahren an, mit dessen Hilfe man aus einer gegebenen (endlichen oder abzählbar unendlichen) Basis eine Orthonormalbasis herstellen kann.

SATZ. *Sei (V, σ) ein euklidischer Vektorraum, $x_1, x_2, x_3, \ldots$ linear unabhängig (endlich oder abzählbar unendlich viele Vektoren). Dann kann man ein Orthonormalsystem $e_1, e_2, e_3, \ldots$ konstruieren (soviele wie die x_n's) so daß für jedes n gilt: $(e_1, \ldots, e_n)$ ist Orthonormalbasis von $U_n := \langle x_1, \ldots, x_n \rangle$.*

BEWEIS. Wir konstruieren $e_1, e_2, e_3, \ldots$ induktiv.

1. *Basis* $n = 1$. Setze $e_1 := \frac{x_1}{\|x_1\|}$; man beachte hierbei, daß $x_1 \neq 0$. Dann ist $\sigma(e_1, e_1) = 1$.

2. *Schritt* $n \to n + 1$. Seien $e_1, \ldots, e_n$ schon konstruiert. $(e_1, \ldots, e_n)$ ist Orthonormalbasis von $U_n = \langle x_1, \ldots, x_n \rangle$. Sei p_{n+1} die orthogonale Projektion von x_{n+1} auf U_n, also

$$p_{n+1} = \sum_{i=1}^{n} \sigma(x_{n+1}, e_i) e_i.$$

Dann ist $h_{n+1} := x_{n+1} - p_{n+1}$ das Lot von x_{n+1} auf U_n. Es ist $h_{n+1} \neq 0$, da $x_{n+1} \notin U_n = \langle x_1, \dots, x_n \rangle$ und $p_{n+1} \in U_n$. Setze $e_{n+1} := \frac{h_{n+1}}{\|h_{n+1}\|}$. Dann gilt

$$\begin{aligned} \sigma(e_i, e_{n+1}) &= \frac{1}{\|h_{n+1}\|}\sigma(e_i, h_{n+1}) = 0 \quad \text{für } 1 \leq i \leq n \\ \sigma(e_{n+1}, e_{n+1}) &= 1. \end{aligned}$$

Nach Induktionsvoraussetzung ist also $e_1, \dots, e_n, e_{n+1}$ Orthonormalsystem. Zu zeigen bleibt $\langle e_1, \dots, e_{n+1} \rangle = \langle x_1, \dots, x_{n+1} \rangle$ $(= U_{n+1})$.

$\subseteq$. Nach Induktionsvoraussetzung ist $\langle e_1, \dots, e_n \rangle = \langle x_1, \dots, x_n \rangle$ $(= U_n)$. Ferner gilt

$$e_{n+1} = \frac{1}{\|h_{n+1}\|}(x_{n+1} - p_{n+1}) \quad \text{und} \quad p_{n+1} \in U_n.$$

$\supseteq$. $x_{n+1} = h_{n+1} + p_{n+1} = \|h_{n+1}\| e_{n+1} + p_{n+1}$ und $p_{n+1} \in U_n = \langle e_1, \dots, e_n \rangle$. QED

FOLGERUNG. *Jeder endlichdimensionale euklidische Vektorraum besitzt eine Orthonormalbasis.* QED

BEMERKUNG. Sei (V, σ) ein euklidischer Vektorraum, $U \subseteq V$ endlichdimensionaler Unterraum und $x \in V$. Die orthogonale Projektion p von x auf U kann man wie folgt berechnen.

1. Falls eine Orthonormalbasis $e_1, \dots, e_n$ von U bekannt ist, hat man

$$p = \sum_{i=1}^{n} \sigma(x, e_i) e_i.$$

2. Falls nur ein Erzeugendensystem $x_1, \dots, x_m$ von U bekannt ist, setze man

$$p = \sum_{i=1}^{m} \alpha_i x_i$$

und bestimme $\alpha_1, \dots, \alpha_m$ aus

$$A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = b \quad \text{mit } A = (\sigma(x_i, x_j)) \text{ und } b = (\sigma(x, x_i)).$$

Dann ist p die orthogonale Projektion von x auf U.

BEWEIS. $p \in U$ ist klar. Ferner sind folgende Aussagen äquivalent.

$$\begin{aligned} & x - p \perp U \\ & \sigma(x - p, x_i) = 0 \quad \text{für } i = 1, \dots, m \\ & \sigma(x, x_i) = \sigma(p, x_i) = \sum_{j=1}^{n} \alpha_j \sigma(x_j, x_i) \quad \text{für } i = 1, \dots, m \\ & b = A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}. \end{aligned}$$

Damit ist das Problem auf die Lösung eines linearen Gleichungssystems reduziert.

5.2.5. *Orthogonales Komplement.* Sei V ein euklidischer Vektorraum und $M \subseteq V$ eine Teilmenge. Dann ist das *orthogonale Komplement* $M^\perp$ von M (gelesen: "M senkrecht") definiert durch

$$M^\perp := \{ x \in V \mid x \perp M \},$$

wobei $x \perp M$ bedeutet $x \perp y$ für alle $y \in M$.

Bemerkungen. 1. $M^\perp$ ist Unterraum, da σ bilinear ist.

2. Sei $U \subseteq V$ Unterraum und $x_1, \ldots, x_n$ Erzeugendensystem von U. Dann ist

$$U^\perp = \{x_1, \ldots, x_n\}^\perp,$$

da σ bilinear ist.

3. Es gilt $M \subseteq M^{\perp\perp}$, da $x \perp M^\perp$ für alle $x \in M$.

Satz. *Sei (V, σ) ein euklidischer Vektorraum und $U \subseteq V$ Unterraum.*

1. *Ist U endlichdimensional, so gilt $V = U \oplus U^\perp$.*
2. *Ist V endlichdimensional, so gilt $U = U^{\perp\perp}$.*

Beweis. 1. $V = U + U^\perp$: Sei $x \in V$ und p die orthogonale Projektion von x auf U. Dann ist

$$x = p + (x - p) \in U + U^\perp.$$

$U \cap U^\perp = 0$: Sei $x \in U \cap U^\perp$. Dann ist $x \perp x$, also $\sigma(x, x) = 0$ und damit $x = 0$.

2. Wegen

$$\begin{aligned} \dim(V) &= \dim(U) + \dim(U^\perp) \\ \dim(V) &= \dim(U^\perp) + \dim(U^{\perp\perp}) \end{aligned}$$

folgt $\dim(U) = \dim(U^{\perp\perp})$. Mit $U \subseteq U^{\perp\perp}$ ergibt sich die Behauptung. Qed

Bemerkung. Teil (2) des Satzes gilt auch im Fall $\dim(U) < \infty$ und beliebigem V; dies wird hier aber nicht gezeigt.

5.2.6. *Abstand affiner Unterräume.* Als erste einfache Anwendung der bisher entwickelten Theorie wollen wir zeigen, daß der Abstand zweier affiner Unteräume gleich der Länge ihres gemeinsamen Lotes ist. Diese Begriffe sind zunächst zu erklären.

Definition. Sei (V, σ) ein euklidischer Vektorraum und L, L' affine Unterräume, etwa $L = x + U$, $L' = x' + U'$ mit $x, x' \in V$ und $U, U' \subseteq V$ Unterräume. $h \in V$ heißt *gemeinsames Lot* von L und L', wenn es $p \in L$ und $p' \in L'$ gibt mit

$$\begin{aligned} &h = p - p', \\ &h \perp U \quad \text{und} \quad h \perp U'. \end{aligned}$$

Ist h gemeinsames Lot von L und L', so heißt $\|h\| =: d(L, L')$ der *Abstand* von L und L'. (Diese Definition ist sinnvoll aufgrund des folgenden Satzes).

Beispiel. Sei $V = \mathbb{R}^3$ und L, L' *windschiefe Geraden* im $\mathbb{R}^3$, d.h. $L \not\parallel L'$. Dann ist h die kürzeste Verbindung zwischen L und L'. Sind L, L' parallele Geraden im $\mathbb{R}^3$, so ist h die lotrechte Verbindung zwischen L und L'.

Satz. *Sei (V, σ) ein euklidischer Vektorraum und L, L' affine Unterräume, etwa $L = x + U$, $L' = x' + U'$ mit $x, x' \in V$ und $U, U' \subseteq V$ Unterräume. Dann gilt:*

1. *Es gibt höchstens ein gemeinsames Lot von L und L'.*
2. *Sind L und L' endlichdimensional, so existiert das gemeinsame Lot.*
3. *Ist h gemeinsames Lot von L und L', so gilt*

$$\|h\| = \min\{\, \|q - q'\| \mid q \in L \text{ und } q' \in L' \,\}.$$

Beweis. 1. Seien h und k gemeinsame Lote von L und L'. Dann gilt

$$\begin{array}{ll} h = p - p' \quad \text{mit } p \in L,\ p' \in L' \qquad & k = q - q' \quad \text{mit } q \in L,\ q' \in L' \\ h \perp U, U' & k \perp U, U' \end{array}$$

Dann ist $h - k \in U + U'$, denn es gilt

$$h - k = (p - p') - (q - q') = \underbrace{(p - q)}_{\in U} - \underbrace{(p' - q')}_{\in U'}.$$

Weiter ist $h - k \in (U + U')^\perp$, denn es gilt $h \perp U, U'$ und $k \perp U, U'$, und σ ist bilinear. Also ist $\sigma(h - k, h - k) = 0$ und damit $h - k = 0$.

2. Seien $L = x + U$ und $L' = x' + U'$ endlichdimensional. Mit U, U' ist dann auch $U + U'$ endlichdimensional. Also existiert die orthogonale Projektion $\tilde{p}$ von $x - x'$ auf $U + U'$, d.h. $\tilde{p} \in U + U'$ und $h := (x - x') - \tilde{p}$ ist das Lot von $x - x'$ auf $U + U'$. Sei $\tilde{p} = q + q'$ für $q \in U$ und $q' \in U'$. Man erhält

$$\begin{array}{ll} h = \underbrace{(x - q)}_{\in L} - \underbrace{(x' + q')}_{\in L'} & \\ h \perp U, U' \qquad\qquad & \text{da } h \perp U + U'. \end{array}$$

Das war zu zeigen.

3. Sei h gemeinsames Lot von L und L', also

$$\begin{array}{l} h = p - p' \quad \text{mit } p \in L,\ p' \in L' \\ h \perp U, U'. \end{array}$$

Sei noch $q \in L$, $q' \in L'$. Dann ist $p - q \in U$, also $h \perp p - q$, und ebenso $p' - q' \in U'$, also $h \perp p' - q'$. Weiter ist

$$q - q' = p - \underbrace{(p - q)}_{\in U} - p' + \underbrace{(p' - q')}_{\in U'} = \underbrace{p - p'}_{h} + y \quad \text{mit } y \in U + U'.$$

Wegen $h \perp U, U'$ ist $h \perp y$. Nach dem Satz des Pythagoras folgt

$$\|q - q'\|^2 = \|h\|^2 + \|y\|^2 \geq \|h\|^2. \quad \text{Qed}$$

5.2.7. HESSE*sche Normalform für Hyperebenen.* Als weitere Anwendung leiten wir eine nützliche Darstellung von Hyperebenen her.

SATZ. *Sei (V, σ) ein endlichdimensionaler euklidischer Vektorraum und $H \subseteq V$.*

1. *H ist genau dann Hyperebene, wenn sich H schreiben läßt in der Form*
$$H = \{ x \in V \mid \sigma(n, x) = \beta \}$$
mit $n \in V$, $\|n\| = 1$ und $\beta \in \mathbb{R}$. n heißt Normalenvektor *zu H.*

2. *Sei $n \in V$, $\|n\| = 1$ und $H = \{ x \in V \mid \sigma(n, x) = \beta \}$. Dann gilt für alle $y \in V$*
$$d(y, H) = |\sigma(n, y) - \beta|.$$

BEWEIS. 1. Sei H Hyperebene, also $H = h_0 + U$ mit $U \subseteq V$ Unterraum und $\dim(U) = \dim(V) - 1$. Dann ist $\dim(U^\perp) = 1$ (da $V = U \oplus U^\perp$), also $U^\perp = \mathbb{R}v$ für ein $v \neq 0$. Setze
$$n := \frac{v}{\|v\|}.$$
Dann ist auch $U^\perp = \mathbb{R}n$ und $\|n\| = 1$. Folgende Aussagen sind dann äquivalent.
$$\begin{aligned} &h \in H \\ &h - h_0 \in U \\ &\sigma(n, h - h_0) = 0 \qquad \text{da } n \in U^\perp \\ &\sigma(n, h) = \sigma(n, h_0) =: \beta. \end{aligned}$$
Also ist $H = \{ x \in V \mid \sigma(n, x) = \beta \}$ mit $\beta := \sigma(n, h_0)$.

Sei nun umgekehrt $H = \{ x \in V \mid \sigma(n, x) = \beta \}$. Setze $h_0 := \beta n \in H$. Dann ist
$$\begin{aligned} H - h_0 &= \{ x - h_0 \mid \sigma(n, x) = \beta \} \\ &= \{ x - h_0 \mid \sigma(n, x) = \sigma(n, h_0) \} \\ &= \{ y \mid \sigma(n, y) = 0 \} \\ &= (\mathbb{R}n)^\perp. \end{aligned}$$
$H - h_0$ ist also ein $(\dim(V) - 1)$-dimensionaler Unterraum von V. Damit ist gezeigt, daß H eine Hyperebene ist.

2. Sei $H = h_0 + U$ und $U = (\mathbb{R}n)^\perp$. Dann ist
$$\begin{aligned} d(y, H) &= \min\{ \|y - h\| \mid h \in H \} \\ &= \min\{ \|y - h_0 - u\| \mid n \in U \} \\ &= d(y - h_0, U) \\ &= \text{Länge des Lots von } y - h_0 \text{ auf } U. \end{aligned}$$
Sei p die orthogonale Projektion von $y - h_0$ auf U, also
$$\begin{aligned} y - h_0 - p &\perp U, \\ y - h_0 - p &= \lambda n. \end{aligned}$$

Dann ist

$$\begin{aligned}
|\sigma(n,y)-\beta| &= |\sigma(n,y)-\sigma(n,h_0)| \\
&= |\sigma(n,y-h_0)| \\
&= |\sigma(n,p+(y-h_0-p))| \\
&= |\sigma(n,p+\lambda n)| \\
&= |\sigma(n,p)+\lambda| \\
&= \lambda && \text{da } p\in U \\
&= \text{Länge des Lots von } y-h_0 \text{ auf } U. && \text{QED}
\end{aligned}$$

5.3 Orthogonale Abbildungen

Euklidische Vektorräume enthalten neben der Vektorraumstruktur noch zusätzlich das Skalarprodukt σ. Es liegt jetzt nahe, lineare Abbildungen zu betrachten, die das Skalarprodukt erhalten. Solche Abbildungen nennen wir Bewegungen oder auch orthogonale Abbildungen.

In diesem Abschnitt sei stets (V,σ) ein euklidischer Vektorraum endlicher Dimension.

5.3.1. SATZ. *Sei $(e_1,\ldots,e_n)$ Orthonormalbasis von V, $f\colon V\to V$ ein Endomorphismus und $A\in\mathbb{R}^{n\times n}$ die darstellende Matrix von f bzgl. der Basis $(e_1,\ldots,e_n)$ von V, also*

$$f(e_j)=\sum_{i=1}^{n}\alpha_{ij}e_i.$$

Dann sind folgende Aussagen äquivalent.

1. $\|f(x)\|=\|x\|$ *für alle* $x\in V$.
2. $\sigma(f(x),f(y))=\sigma(x,y)$ *für alle* $x,y\in V$.
3. $f(e_1),\ldots,f(e_n)$ *ist Orthonormalbasis von* V.
4. $A^tA=E$.

BEWEIS. Die Äquivalenz von (1) und (2) folgt aus

$$\begin{aligned}
\|x\| &= \sqrt{\sigma(x,x)}, \\
\sigma(x,y) &= \frac{1}{2}(\|x+y\|^2-\|x\|^2-\|y\|^2).
\end{aligned}$$

(2) $\Longrightarrow$ (3). Es gilt $\sigma(f(e_i),f(e_j))=\sigma(e_i,e_j)=\delta_{ij}$. Also ist auch $f(e_1),\ldots,f(e_n)$ ein Orthonormalsystem und deshalb linear unabhängig. Mit $\dim(V)=n$ folgt, daß $f(e_1),\ldots,f(e_n)$ auch eine Orthonormalbasis von V ist.

(3) $\Longrightarrow$ (2). Sei $x = \sum_{i=1}^n \alpha_i e_i$ und $y = \sum_{j=1}^n \beta_j e_j$. dann gilt

$$\begin{aligned}
\sigma(x,y) &= \sigma\Big(\sum_{i=1}^n \alpha_i e_i, \sum_{j=1}^n \beta_j e_j\Big)\\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_i\beta_j \sigma(e_i, e_j)\\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_i\beta_j \delta_{ij}\\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_i\beta_j \sigma(f(e_i), f(e_j))\\
&= \sigma\Big(\sum_{i=1}^n \alpha_i f(e_i), \sum_{j=1}^n \beta_j f(e_j)\Big)\\
&= \sigma(f(x), f(y)).
\end{aligned}$$

Die Äquivalenz von (3) und (4) ergibt sich wie folgt. Zunächst ist

$$\begin{aligned}
\sigma(f(e_i), f(e_j)) &= \sigma\Big(\sum_{k=1}^n \alpha_{ki} e_k, \sum_{\ell=1}^n \alpha_{\ell j} e_\ell\Big)\\
&= \sum_{k=1}^n \sum_{\ell=1}^n \alpha_{ki}\alpha_{\ell j} \sigma(e_k, e_\ell)\\
&= \sum_{k=1}^n \alpha_{ki}\alpha_{kj} \qquad \text{da } \sigma(e_k, e_\ell) = \delta_{k\ell}.
\end{aligned}$$

Also sind folgende Aussagen äquivalent.

$$\begin{aligned}
&f(e_1), \dots, f(e_n) \text{ ist Orthonormalbasis von } V\\
&\sigma(f(e_i), f(e_j)) = \delta_{ij} \quad \text{für alle } i, j \text{ mit } 1 \le i, j \le n\\
&\sum_{k=1}^n \alpha_{ki}\alpha_{kj} = \delta_{ij} \quad \text{für alle } i, j \text{ mit } 1 \le i, j \le n\\
&A^t A = E. \quad \text{QED}
\end{aligned}$$

DEFINITION. Sei f wie im Satz und $(e_1, \dots, e_n)$ eine Orthonormalbasis von V. Der Endomorphismus f heißt *orthogonal*, wenn f eine der Bedingungen (1)-(3) (und damit alle) des Satzes erfüllt. Die Matrix $A \in K^{n\times n}$ heißt *orthogonal*, wenn $A^t A = E$. Man nennt

$$O(n) := \{\, A \in \mathrm{GL}(n, K) \mid A^t A = E \,\}$$

die *orthogonale Gruppe*.

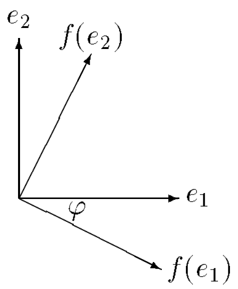
BEMERKUNGEN. 1. Sei $(e_1, \dots, e_n)$ eine Orthonormalbasis von V. Ein Endomorphismus f von V ist nach dem Satz genau dann orthogonal, wenn die f darstellende Matrix (bzgl. $(e_1, \dots, e_n)$) orthogonal ist.

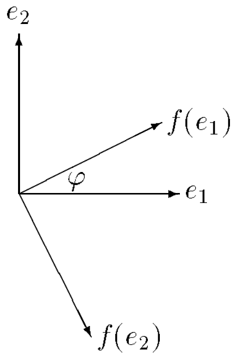
2. Eine Teilmenge U einer Gruppe $(G, \circ)$ heißt *Untergruppe*, wenn die Gruppenstruktur auf G eine Gruppenstruktur auf U induziert, d.h. wenn gilt (1) $\alpha \circ \beta \in U$ für alle $\alpha, \beta \in U$, (2) $e \in U$ und (3) $\alpha^{-1} \in U$ für alle $\alpha \in U$. – $O(n)$ ist Untergruppe von $\mathrm{GL}(n, K)$.

3. Wichtige Beispiele orthogonaler Matrizen im Fall $n = 2$ sind:

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \quad \text{Drehung}$$
$$B = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \quad \text{Drehspiegelung.}$$

Man zeigt leicht, daß $A, B \in O(2)$ sind. Die zugehörigen Endomorphismen kann man sich wie folgt veranschaulichen.





5.3.2. *Bewegungen, euklidische Äquivalenz.* Das Gegenstück zu affinen Abbildungen in der affinen Geometrie sind in der euklidischen Geometrie die Bewegungen.

Sei (wie stets in diesem Abschnitt) (V, σ) ein euklidischer Vektorraum endlicher Dimension. Eine Abbildung $g \colon V \to V$ heißt *Bewegung*, wenn es eine orthogonale Abbildung $f \colon V \to V$ und ein $y \in V$ gibt so daß für alle $x \in V$ gilt

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

Bemerkung. $\{g \colon V \to V \mid g \text{ Bewegung}\}$ ist eine Untergruppe der affinen Gruppe $\mathrm{Aff}(V)$.

Wir wollen die Überlegungen aus Abschnitt 4.2.2 – insbesondere die Klassifikation der Quadriken – auf den Fall der euklidischen Geometrie übertragen. Im folgenden betrachten wir also den speziellen euklidischen Raum $V = \mathbb{R}^n$ mit dem kanonischen Skalarprodukt $\sigma(x, y) = x^t y$.

Quadratische Funktionen werden wie in 4.2.1 definiert. Wie dort ergibt sich für sie das folgende Verhalten bei Bewegungen. Sei $F \colon \mathbb{R}^n \to \mathbb{R}$ eine quadratische Funktion und $g \colon \mathbb{R}^n \to \mathbb{R}^n$ eine Bewegung. Dann ist auch $F \circ g$ eine quadratische Funktion und es gilt

$$V(F \circ g) = g^{-1}(V(F)).$$

$F, F' \colon \mathbb{R}^n \to \mathbb{R}$ seien quadratische Funktionen, $Q = V(F)$ und $Q' = V(F')$. F heißt *euklidisch äquivalent* zu F', wenn es eine Bewegung $g \colon \mathbb{R}^n \to \mathbb{R}^n$ gibt mit $F' = F \circ g$. Eine Teilmenge $Q \subseteq \mathbb{R}^n$ heißt *euklidisch äquivalent* zu $Q' \subseteq \mathbb{R}^n$, wenn es eine Bewegung $g \colon \mathbb{R}^n \to \mathbb{R}^n$ gibt mit $g(Q') = Q$.

Wir betrachten wieder folgendes Problem. Gegeben sei eine quadratische Funktion $F \colon \mathbb{R}^n \to \mathbb{R}$. Gesucht ist eine Bewegung $g \colon \mathbb{R}^n \to \mathbb{R}^n$ so daß $F \circ g$ eine "möglichst einfache Form" hat. Das wesentliche Hilfsmittel hierfür ist die folgende Verschärfung des Satzes in 4.2.2, die im Fall des Grundkörpers $\mathbb{R}$ möglich ist.

SATZ. *(Hauptachsentransformation). Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Dann existieren $C \in O(n)$ und $\gamma_1, \dots, \gamma_n \in \mathbb{R}$ so daß*

$$C^t AC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

(Nicht ausgefüllte Stellen von Matrizen sollen wieder 0 sein).

Den Beweis werden wir im nächsten Abschnitt 5.4 führen.

5.3.3. SATZ. *Jede quadratische Hyperfläche $Q \subseteq \mathbb{R}^n$ ist euklidisch äquivalent zu einer Fläche des Typs $V(F)$, wobei für F folgende Formen auftreten.*

$$F(x) = \sum_{i=1}^{r} \frac{\xi_i^2}{\gamma_i^2} - \sum_{i=r+1}^{m} \frac{\xi_i^2}{\gamma_i^2} + 1 \qquad \text{Mittelpunktsfläche} \tag{5.1}$$

$$F(x) = \sum_{i=1}^{r} \frac{\xi_i^2}{\gamma_i^2} - \sum_{i=r+1}^{m} \frac{\xi_i^2}{\gamma_i^2} \qquad \text{Kegelfläche} \tag{5.2}$$

$$F(x) = \sum_{i=1}^{r} \frac{\xi_i^2}{\gamma_i^2} - \sum_{i=r+1}^{m} \frac{\xi_i^2}{\gamma_i^2} + \xi_{m+1} \qquad \text{parabolische Fläche.} \tag{5.3}$$

Dabei ist $r \leq m \leq n$ in (5.1), (5.2), $r < n$ in (5.3), $\gamma_1, \dots, \gamma_r \in \mathbb{R}$ und $x^t = \begin{pmatrix} \xi_1 & \dots & \xi_n \end{pmatrix}$.

BEWEIS. Wir wiederholen im wesentlichen den Beweis aus 4.2.3, wobei wir an einigen Stellen die entsprechenden Modifikationen anbringen.

Sei $Q = V(G)$ und $G = x^t Ax + a^t x + \alpha$ mit $A \in \mathbb{R}^{n\times n}$ symmetrisch, $a \in \mathbb{R}^n$ und $\alpha \in \mathbb{R}$. Nach dem Satz über die Hauptachsentransformation in Abschnitt 5.4 gibt es $C \in O(n)$ und $\gamma_1, \dots, \gamma_n \in \mathbb{R}$ mit

$$C^t AC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

Dann gilt

$$\begin{aligned} G(Cx) &= x^t C^t ACx + a^t Cx + \alpha \\ &= \sum_{i=1}^{n} \gamma_i \xi_i^2 + \sum_{i=1}^{n} \beta_i \xi_i + \alpha \\ &=: G_1(x). \end{aligned}$$

G_1 ist euklidisch äquivalent zu G, da $x \mapsto Cx$ eine Bewegung ist. Wir können wieder annehmen, daß $\gamma_1, \dots, \gamma_r \neq 0$ und $\gamma_{r+1} = \dots = \gamma_n = 0$ sind für ein r mit $1 \leq r \leq n$.

$$\begin{aligned} G_1(x) &= \sum_{i=1}^{r} \gamma_i \xi_i^2 + \sum_{i=1}^{n} \beta_i \xi_i + \alpha \\ &= \sum_{i=1}^{r} \gamma_i \big(\xi_i + \frac{\beta_i}{2\gamma_i}\big)^2 + \sum_{i=r+1}^{n} \beta_i \xi_i + \alpha - \sum_{i=1}^{r} \frac{\beta_i^2}{4\gamma_i} \\ &=: G_2(x') \quad \text{mit } \xi_i' := \begin{cases} \xi_i + \dfrac{\beta_i}{2\gamma_i} & \text{für } 1 \leq i \leq r \\ \xi_i & \text{für } r+1 \leq i \leq n. \end{cases} \end{aligned}$$

G_2 ist euklidisch äquivalent zu G_1, da $x \mapsto x'$ eine Translation ist, und es gilt

$$G_2(x') = \sum_{i=1}^{r} \gamma_i(\xi_i')^2 + \sum_{i=r+1}^{n} \beta_i \xi_i' + \alpha' = G_1(x) \quad \text{mit } \alpha' = \alpha - \sum_{i=1}^{r} \frac{\beta_i^2}{4\gamma_i}.$$

Fall 1. $\beta_{r+1} = \cdots = \beta_n = 0$, $\alpha' \neq 0$. OBdA sei $\alpha' = 1$ (bei Division durch α' bleibt $V(G_2)$ unverändert). Dann sind G und G_2 euklidisch äquivalent, also auch $Q = V(G)$ und $V(G_2)$ euklidisch äquivalent, und $V(G_2)$ ist vom Typ (5.1).

Fall 2. $\beta_{r+1} = \cdots = \beta_n = 0$, $\alpha' = 0$. Q ist bis auf euklidische Äquivalenz vom Typ (5.2).

Fall 3. $\beta_i \neq 0$ für ein i mit $r+1 \leq i \leq n$. Setze $b^t := \begin{pmatrix}\beta_{r+1} & \dots & \beta_n\end{pmatrix}$. OBdA sei $\|b\| = 1$ (sonst dividiere man durch $\|b\|$).

$$\begin{aligned} G_2(x') &= \sum_{i=1}^{r} \gamma_i(\xi_i'')^2 + \xi_{r+1}'' \\ &=: G_3(x'') \quad \text{mit } \xi_i'' := \begin{cases} \xi_i' & \text{für } 1 \leq i \leq r \\ \sum_{i=r+1}^{n} \beta_i \xi_i' + \alpha' & \text{für } i = r+1 \\ \text{beliebig} & \text{für } r+1 < i \leq n. \end{cases} \end{aligned}$$

G_3 ist euklidisch äquivalent zu G_2, da

$$x' \mapsto \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \beta_{r+1} & \dots & \beta_n \\ & & * & \dots & * \\ & & \vdots & & \vdots \\ & & * & \dots & * \end{pmatrix} x' + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \alpha' \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

als Bewegung gewählt werden kann. Um dies zu sehen, beachte man, daß oBdA

$$\begin{pmatrix} \beta_{r+1} & \dots & \beta_n \\ * & \dots & * \\ \vdots & & \vdots \\ * & \dots & * \end{pmatrix}$$

orthogonal ist (also $\in O(n-r)$), da sich b zu einer Orthonormalbasis von $\mathbb{R}^{n-r}$ ergänzen läßt. Also ist Q bis auf euklidische Äquivalenz vom Typ (5.3). QED

BEMERKUNG. Aus dem Beweis ergibt sich noch, daß jede quadratische Funktion F euklidisch äquivalent ist zu einer quadratischen Funktion G ist, die bis auf einen konstanten Faktor von einer der Formen (5.1) – (5.3) ist.

5.4 Hauptachsentransformation nach Jacobi

In diesem Abschnitt sei wieder (V, σ) ein euklidischer Vektorraum endlicher Dimension. Behandelt wird die sogenannte reelle Hauptachsentransformation nach einem von JACOBI 1842 angegebenen Verfahren. Anschaulich gesprochen handelt es sich um die Konstruktion einer Basis $\mathcal{B}$ des $\mathbb{R}^n$ so daß in dem neuen, durch $\mathcal{B}$ definierten Koordinatensystem eine gegebene lineare Abbildung f_A mit symmetrischer Matrix A auf den Koordinatenachsen als Streckung um einen Faktor wirkt; $\mathcal{B}$ kann als Orthonormalbasis gewählt werden.

Wir zeigen, daß es für jede symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ eine orthogonale Matrix $C \in O(n)$ gibt so daß

$$C^t AC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

(Nicht ausgefüllte Stellen von Matrizen sollen wieder 0 sein). Für $n = 1$ ist das offensichtlich; es sei also im folgenden $n \geq 2$.

Man beachte, daß die Voraussetzung der Symmetrie von A notwendig ist. Gilt nämlich $C^tAC = D$ für ein $C \in O(n)$ und eine Diagonalmatrix D, so folgt $A = CDC^t$ und damit die Symmetrie von A.

Die Idee des Jacobischen Verfahrens kann man wie folgt beschreiben. Für $1 \leq i < j \leq n$ und $\varphi \in [0, 2\pi]$ setzt man $c := \cos\varphi$, $s := \sin\varphi$ und

$$R_{ij}(\varphi) := \begin{pmatrix} 1 &&&&&&&&& \\ & \ddots &&&&&&&& \\ && 1 &&&&&&& \\ &&& c &&& s &&& \\ &&&& 1 &&&&& \\ &&&&& \ddots &&&& \\ &&&&&& 1 &&& \\ &&& -s &&& c &&& \\ &&&&&&& 1 && \\ &&&&&&&& \ddots & \\ &&&&&&&&& 1 \end{pmatrix} \in O(n),$$

wobei die i-te und j-te Zeile und Spalte zwei Einträge haben. $R_{ij}(\varphi)$ heißt Jacobi*sche Rotationsmatrix*. Ihre Orthogonalität ergibt sich genau so wie die am Ende von 5.3.1 gezeigte Orthogonalität von $\left(\begin{smallmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{smallmatrix}\right)$.

Man sucht eine Rotationsmatrix $R_{ij}(\varphi)$ so, daß beim Übergang von A zu $R_{ij}(\varphi)^t AR_{ij}(\varphi)$ ein Koeffizient größten Betrages außerhalb der Diagonalen $= 0$ wird und versucht, diesen Schritt zu iterieren. (Dabei können Koeffizienten, die schon 0 waren, wieder $\neq 0$ werden, aber die Beträge werden immer kleiner).

5.4.1. Lemma. *Sei $A = A^t \in \mathbb{R}^{n\times n}$ und $1 \leq i < j \leq n$. Dann existiert ein $\varphi \in [0, 2\pi]$ so daß der (i, j)-te Koeffizient von $R_{ij}(\varphi)^t AR_{ij}(\varphi)$ Null ist.*

Beweis. Für eine beliebige Matrix $A = (\alpha_{kl}) \in \mathbb{R}^{n\times n}$ ist

$$A \cdot R_{ij}(\varphi) = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1i}c - \alpha_{1j}s & \dots & \alpha_{1i}s + \alpha_{1j}c & \dots & \alpha_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \alpha_{n1} & \dots & \alpha_{ni}c - \alpha_{nj}s & \dots & \alpha_{ni}s + \alpha_{nj}c & \dots & \alpha_{nn} \end{pmatrix},$$

wobei die i-te und die j-te Spalte modifiziert sind. Analog gilt für ein beliebiges $C = (\gamma_{kl})$

$$R_{ij}(\varphi)^t \cdot C = \begin{pmatrix} \gamma_{11} & \dots & \gamma_{1n} \\ \vdots & & \vdots \\ c\gamma_{i1} - s\gamma_{j1} & \dots & c\gamma_{in} - s\gamma_{jn} \\ \vdots & & \vdots \\ s\gamma_{i1} + c\gamma_{j1} & \dots & s\gamma_{in} + c\gamma_{jn} \\ \vdots & & \vdots \\ \gamma_{n1} & \dots & \gamma_{nn} \end{pmatrix},$$

wobei die i-te und die j-te Zeile modifiziert sind. Die Koeffizienten von $R_{ij}(\varphi)^t A R_{ij}(\varphi) =: (\beta_{kl})$ ergeben sich durch Einsetzen der Koeffizienten von der ersten Matrix in die zweite. Wegen $(R_{ij}(\varphi)^t A R_{ij}(\varphi))^t = R_{ij}(\varphi)^t A^t R_{ij}(\varphi)$ und $A = A^t$ ist (β_{kl}) symmetrisch. Man erhält

1. für $k \neq i, j$ und $l \neq i, j$: $\beta_{kl} = \alpha_{kl}$.
2. für $l \neq i, j$: $\beta_{il} = c\alpha_{il} - s\alpha_{jl}$.
3. für $l \neq i, j$: $\beta_{jl} = s\alpha_{il} + c\alpha_{jl}$.
4.

$$\begin{aligned} \beta_{ij} &= c\gamma_{ij} - s\gamma_{jj} \\ &= c(\alpha_{ii}s + \alpha_{ij}c) - s(\alpha_{ji}s + \alpha_{jj}c) \\ &= cs(\alpha_{ii} - \alpha_{jj}) + \alpha_{ij}(c^2 - s^2). \end{aligned}$$

Zu zeigen bleibt, daß für ein geeignetes $\varphi \in [0, 2\pi]$ gilt $\beta_{ij} = 0$. Dazu beachte man zunächst, daß

$$cs(\alpha_{ii} - \alpha_{jj}) + \alpha_{ij}(c^2 - s^2) = \frac{\alpha_{ii} - \alpha_{jj}}{2} \sin 2\varphi + \alpha_{ij} \cos 2\varphi,$$

denn bekanntlich ist

$$\begin{aligned} \sin(\varphi_1 + \varphi_2) &= \sin\varphi_1 \cos\varphi_2 + \cos\varphi_1 \sin\varphi_2, \\ \cos(\varphi_1 + \varphi_2) &= \cos\varphi_1 \cos\varphi_2 - \sin\varphi_1 \sin\varphi_2. \end{aligned}$$

Fall $\alpha_{ii} = \alpha_{jj}$. Setze $\varphi = \frac{\pi}{4}$; dann ist $\cos 2\varphi = \cos\frac{\pi}{2} = 0$.
Fall $\alpha_{ii} \neq \alpha_{jj}$. Folgende Aussagen sind äquivalent.

$$\begin{aligned} \frac{\alpha_{ii} - \alpha_{jj}}{2} \sin 2\varphi + \alpha_{ij} \cos 2\varphi = 0 \\ \tan 2\varphi = \frac{-2\alpha_{ij}}{\alpha_{ii} - \alpha_{jj}} \\ \varphi = \frac{1}{2} \arctan \frac{-2\alpha_{ij}}{\alpha_{ii} - \alpha_{jj}}. \end{aligned} \quad \text{QED}$$

5.4.2. Unser Verfahren kann durchaus Koeffizienten $\alpha_{kl} = 0$ von A in Koeffizienten $\beta_{kl} \neq 0$ überführen. Wir können aber zeigen, daß dabei die Koeffizienten betragsmäßig kleiner werden.

Lemma. *Sei $A = A^t \in \mathbb{R}^{n\times n}$, $A = (\alpha_{kl})$. Für jedes $B = (\beta_{kl}) \in \mathbb{R}^{n\times n}$ sei $w(B) = \sum_{i\neq j} \beta_{ij}^2$ die Quadratsumme der Koeffizienten von B außerhalb der Diagonalen. Wenn man $i < j$ so wählt, daß $|\alpha_{ij}|$ maximal unter allen $|\alpha_{kl}|$ mit $k \neq l$ ist und $R_{ij}(\varphi)$ wie im vorangehenden Lemma bestimmt, so gilt*

$$w\big(R_{ij}(\varphi)^t A R_{ij}(\varphi)\big) \leq \Big(1 - \frac{2}{n^2 - n}\Big) \cdot w(A).$$

Beweis. $A = (\alpha_{kl})$ und $B = (\beta_{kl}) := R_{ij}(\varphi)^t A R_{ij}(\varphi)$ sind symmetrisch. Zunächst ist

$$\beta_{il}^2 + \beta_{jl}^2 = \alpha_{il}^2 + \alpha_{jl}^2 \quad \text{für } l \neq i, j, \tag{5.4}$$

denn es gilt

$$\begin{aligned}
\beta_{il}^2 + \beta_{jl}^2 &= (c\alpha_{il} - s\alpha_{jl})^2 + (s\alpha_{il} + c\alpha_{jl})^2 \\
&= c^2\alpha_{il}^2 - 2cs\alpha_{il}\alpha_{jl} + s^2\alpha_{jl}^2 + s^2\alpha_{il}^2 + 2sc\alpha_{il}\alpha_{jl} + c^2\alpha_{jl}^2 \\
&= (c^2 + s^2)(\alpha_{il}^2 + \alpha_{jl}^2) \\
&= \alpha_{il}^2 + \alpha_{jl}^2 && \text{da } c^2 + s^2 = 1.
\end{aligned}$$

Ferner folgt nach Wahl von (i, j), also aus der Voraussetzung $|\alpha_{kl}| \leq |\alpha_{ij}|$ für $k \neq l$, daß

$$w(A) = \sum_{\substack{k,l=1 \\ k\neq l}}^{n} \alpha_{kl}^2 \leq (n^2 - n)\alpha_{ij}^2. \tag{5.5}$$

Daraus folgt

$$\begin{aligned}
w(B) &= 2 \cdot \sum_{k<l} \beta_{kl}^2 \\
&= \sum_{\substack{k\neq i,j \\ l\neq i,j \\ k\neq l}} \beta_{kl}^2 + 2 \cdot \sum_{l\neq i,j} \beta_{il}^2 + 2 \cdot \sum_{l\neq i,j} \beta_{jl}^2 + \underbrace{\beta_{ij}^2 + \beta_{ji}^2}_{=0} \\
&= \sum_{\substack{k\neq i,j \\ l\neq i,j \\ k\neq l}} \alpha_{kl}^2 + 2 \cdot \sum_{l\neq i,j} \alpha_{il}^2 + 2 \cdot \sum_{l\neq i,j} \alpha_{jl}^2 && \text{nach (5.4)} \\
&= w(A) - 2\alpha_{ij}^2 \\
&\leq w(A) \cdot \Big(1 - \frac{2}{n^2 - n}\Big),
\end{aligned}$$

denn nach (5.5) ist

$$-2\alpha_{ij}^2 \leq -\frac{2}{n^2 - n} w(A), \quad \text{also} \quad w(A) - 2\alpha_{ij}^2 \leq w(A) \cdot \Big(1 - \frac{2}{n^2 - n}\Big). \quad \text{QED}$$

5.4.3. *Einige Begriffe aus der Analysis.* Um beweisen zu können, daß unser Verfahren tatsächlich zu einer Diagonalmatrix führt, werden einige Begriffe aus der Analysis gebraucht.

Definition.

1. Eine Folge $(\alpha_n)_{n\in\mathbb{N}}$ reeller Zahlen heißt *beschränkt*, wenn es $a, b \in \mathbb{Q}$ gibt mit
$$a \leq \alpha_n \leq b \quad \text{für alle } n \in \mathbb{N}.$$

2. Eine Folge (α_n) reeller Zahlen heißt *konvergent* gegen $\alpha \in \mathbb{R}$ (in Zeichen $\lim_{n\to\infty} \alpha_n = \alpha$), falls gilt
$$\forall \varepsilon{>}0 \exists N{\in}\mathbb{N} \forall n{\geq}N \; |\alpha_n - \alpha| \leq \varepsilon.$$

3. Eine Folge $(A_m)_{m\in\mathbb{N}}$ von reellen $n \times n$-Matrizen heißt *beschränkt* bzw. *konvergent*, wenn jede der Koeffizientenfolgen $(\alpha_{ij}^{(m)})_{m\in\mathbb{N}}$ beschränkt bzw. konvergent ist. Im zweiten Fall setzt man
$$\lim_{m\to\infty} A_m := \begin{pmatrix} \lim_{m\to\infty} \alpha_{11}^{(m)} & \dots & \lim_{m\to\infty} \alpha_{1n}^{(m)} \\ \vdots & & \vdots \\ \lim_{m\to\infty} \alpha_{n1}^{(m)} & \dots & \lim_{m\to\infty} \alpha_{nn}^{(m)} \end{pmatrix} \in \mathbb{R}^{n\times n}$$

Satz.

1. *Falls die Folgen $(\alpha_n)_{n\in\mathbb{N}}$ und $(\beta_n)_{n\in\mathbb{N}}$ reeller Zahlen konvergieren, so auch die Folgen $(\alpha_n + \beta_n)_{n\in\mathbb{N}}$ und $(\alpha_n \cdot \beta_n)_{n\in\mathbb{N}}$, und es gilt*
$$\lim_{n\to\infty}(\alpha_n + \beta_n) = \lim_{n\to\infty} \alpha_n + \lim_{n\to\infty} \beta_n \quad \text{und} \quad \lim_{n\to\infty}(\alpha_n \cdot \beta_n) = \lim_{n\to\infty} \alpha_n \cdot \lim_{n\to\infty} \beta_n.$$

2. *Für konvergente Folgen $(A_m)_{m\in\mathbb{N}}$ und $(B_m)_{m\in\mathbb{N}}$ reeller $n \times n$-Matrizen gilt entsprechend*
$$\lim_{m\to\infty}(A_m + B_m) = \lim_{m\to\infty} A_m + \lim_{m\to\infty} B_m \quad \text{und} \quad \lim_{m\to\infty}(A_m \cdot B_m) = \lim_{m\to\infty} A_m \cdot \lim_{m\to\infty} B_m.$$

3. *Die Folge $(A_m)_{m\in\mathbb{N}}$ reeller $n \times n$-Matrizen konvergiert genau dann, wenn $(A_m^t)_{m\in\mathbb{N}}$ konvergiert, und in diesem Fall gilt $\lim_{m\to\infty}(A_m^t) = (\lim_{m\to\infty} A_m)^t$.*

4. *Eine konvergente Folge orthogonaler Matrizen konvergiert gegen eine orthogonale Matrix.*

5. *(Satz von* Bolzano-Weierstrass*). Jede beschränkte Folge reeller Zahlen besitzt eine konvergente Teilfolge.*

6. *Jede beschränkte Folge von Matrizen besitzt eine konvergente Teilfolge.*

7. *Jede Folge orthogonaler Matrizen besitzt eine konvergente Teilfolge.*

Beweis. 1. Bekannt aus der Analysis.

2. Addition und Multiplikation von Matrizen sind über die Koeffizienten definiert, und für diese gilt die Behauptung nach (1).

3. Der Beweis ist offensichtlich.

4. Sei $(A_m)_{m\in\mathbb{N}}$ eine konvergente Folge orthogonaler $n \times n$-Matrizen und $A = \lim_{m\to\infty} A_m$. Dann ist
$$A^t A = (\lim_{m\to\infty} A_m)^t \lim_{m\to\infty} A_m = \lim_{m\to\infty} A_m^t \cdot \lim_{m\to\infty} A_m = \lim_{m\to\infty} A_m^t A_m = \lim_{m\to\infty} E = E,$$

d.h. A ist auch orthogonal.

5. Bekannt aus der Analysis. Wegen seiner Wichtigkeit sei der Beweis hier noch einmal wiederholt. Sei also (α_n) eine beschränkte Folge reeller Zahlen. Es gibt also $a, b \in \mathbb{Q}$ mit

$$a \leq \alpha_n \leq b \quad \text{für alle } n \in \mathbb{N}.$$

Wir definieren jetzt rekursiv $a_k, b_k \in \mathbb{Q}$ und $n_k \in \mathbb{N}$ mit

1. $a = a_0 \leq a_1 \leq \cdots \leq a_k < b_k \leq \cdots \leq b_1 \leq b_0 = b$ und $b_k = a_k + \frac{1}{2^k}(b-a)$,
2. $a_k \leq \alpha_{n_k} \leq b_k$ und $n_0 < n_1 < \cdots < n_k$,
3. es gibt unendlich viele $n \geq n_k$ mit $a_k \leq \alpha_n \leq b_k$.

Eine so definierte Teilfolge (α_{n_k}) wird eine Cauchyfolge sein. Denn sei etwa $l \geq k$. Im Fall $\alpha_{n_l} \leq \alpha_{n_k}$ hat man dann

$$0 \leq \alpha_{n_k} - \alpha_{n_l} \leq b_k - a_l \leq b_k - a_k = \frac{1}{2^k}(b-a),$$

und im Fall $\alpha_{n_k} \leq \alpha_{n_l}$

$$0 \leq \alpha_{n_l} - \alpha_{n_k} \leq b_l - a_k \leq b_k - a_k = \frac{1}{2^k}(b-a).$$

Für $k = 0$ setzen wir $a_0 = a$, $b_0 = b$ und $n_0 = 0$. Nehmen wir jetzt an, die rekursive Konstruktion wäre bis k ausgeführt. Wir unterteilen das Intervall $[a_k, b_k]$ von a_k bis b_k in zwei Hälften: $[a_k, c]$ und $[c, b_k]$ mit $c := \frac{1}{2}(a_k + b_k)$. Da nach (3) unendlich viele α_n mit $n \geq n_k$ im Intervall $[a_k, b_k]$ liegen, müssen auch in einer der beiden Hälften unendlich viele solche α_n liegen.

Fall (i). Für unendlich viele $n \geq n_k$ gilt $a_k \leq \alpha_n \leq c$. Dann setzen wir $a_{k+1} := a_k$, $b_{k+1} := c$ und wählen n_{k+1} als das erste $n > n_k$ mit $a_k \leq \alpha_n \leq c$.

Fall (ii). Für unendlich viele $n \geq n_k$ gilt $c \leq \alpha_n \leq b_k$. Dann setzen wir $a_{k+1} := c$, $b_{k+1} := b_k$ und wählen n_{k+1} als das erste $n > n_k$ mit $c \leq \alpha_n \leq b_k$. QED

6. Dies folgt sofort aus (5).

7. Nach (6) genügt es zu zeigen, daß jede Folge orthogonaler Matrizen beschränkt ist. Aus $A^t A = E$ folgt aber $\sum_{j=1}^{n} \alpha_{ij}^2 = 1$, also $|\alpha_{ij}| \leq 1$ für alle i, j und damit die Behauptung. QED

5.4.4. SATZ. *(Hauptachsentransformation nach* JACOBI*). Gegeben sei $A = A^t \in \mathbb{R}^{n\times n}$. Wir definieren rekursiv eine Folge $(R_i)_{i\in\mathbb{N}}$ von* JACOBI*schen Rotationsmatrizen wie folgt. R_m sei zu $R_{m-1}^t \cdots R_0^t A R_0 \cdots R_{m-1}$ gewählt wie in 5.4.2 (insbesondere also R_0 zu A gewählt wie in 5.4.2). Für alle $m \in \mathbb{N}$ sei $C_m := R_0 \cdots R_m$. Dann hat die Folge $(C_m)_{m\in\mathbb{N}}$ eine konvergente Teilfolge $(C_{\mu(m)})_{m\in\mathbb{N}}$, und $C := \lim_{m\to\infty} C_{\mu(m)}$ ist orthogonal und es gibt $\gamma_1, \ldots, \gamma_n \in \mathbb{R}$ mit*

$$C^t A C = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

BEWEIS. Die Behauptungen über $(C_m)_{m\in\mathbb{N}}$ und C folgen aus 5.4.3 (7) und (4). Zu zeigen bleibt, daß $C^t A C$ Diagonalmatrix ist. Sei also $A_0 := A$ und $A_{m+1} := C_m^t A C_m = R_m^t \cdots R_0^t A R_0 \cdots R_m$. Dann gilt mit 5.4.3(2) und (3) $\lim_{m\to\infty} A_{\mu(m)} = C^t A C$. Nach dem Hauptlemma 5.4.2 hat man

$$w(A_{m+1}) \leq \left(1 - \frac{2}{n^2 - n}\right) \cdot w(A_m),$$

also

$$w(A_m) \leq \left(1 - \frac{2}{n^2 - n}\right)^m \cdot w(A).$$

Setzt man $A_m =: (\alpha_{ij}^{(m)})$, so ist also für alle $i \neq j$ die Folge $(\alpha_{ij}^{(m)})$ eine Nullfolge. Nach der Limesdefinition für Matrizen ist deshalb in $C^tAC = \lim_{m\to\infty} A_{\mu(m)}$ der (i,j)-te Koeffizient für $i \neq j$ gleich Null, d.h. C^tAC ist eine Diagonalmatrix. QED

BEMERKUNG. Dieser Satz ist zu einer konstruktiven numerischen Berechnung der Diagonalmatrix anwendbar: Man führt das Verfahren bis zu einem hinreichend großen m_0 durch, so daß der Fehler klein genug wird.

Kapitel 6

Determinanten

Determinanten sind in den letzten Jahrzehnten in der linearen Algebra weniger wichtig geworden; man arbeitet stattdessen besser mit dem Begriff der linearen Abhängigkeit. Eine wesentliche Rolle spielen Determinanten immer noch bei der Berechnung von Eigenwerten.

6.1 Existenz und Eindeutigkeit der Determinante

6.1.1. *Ringe.* Wir führen zunächst als Verallgemeinerung des Körperbegriffs den Begriff eines *Ringes* ein. Der Grund ist, daß viele Eigenschaften der Determinante die Möglichkeit der Division nicht benötigen. Diese Allgemeinheit wird uns später von Nutzen sein (etwa beim Beweis des Satzes von CAYLEY/HAMILTON).

DEFINITION. Sei R eine Menge und $+\colon R \times R \to R$, $\cdot\colon R \times R \to R$ Abbildungen sowie $1 \in R$ ($+$ heißt Addition, $\cdot$ Multiplikation; Schreibweise: $\alpha + \beta := +(\alpha, \beta), \alpha\beta := \alpha \cdot \beta := \cdot(\alpha, \beta)$). R (oder genauer $(R, +, \cdot)$) heißt *Ring*, wenn gilt

1. $(R, +)$ ist eine abelsche Gruppe (ihr neutrales Element wird mit 0, das zu $\alpha \in R$ inverse Element mit $-\alpha$ bezeichnet).
2. Für alle $\alpha, \beta, \gamma \in R$ gilt
$$\begin{aligned} &\alpha(\beta\gamma) = (\alpha\beta)\gamma && \text{Assoziativgesetz} \\ &\alpha 1 = \alpha = 1\alpha && \text{1 ist beidseitig neutrales Element.} \end{aligned}$$
3. Für alle $\alpha, \beta, \gamma \in R$ gelten die *Distributivgesetze*
$$\begin{aligned} (\alpha + \beta)\gamma &= \alpha\gamma + \beta\gamma, \\ \alpha(\beta + \gamma) &= \alpha\beta + \alpha\gamma. \end{aligned}$$

Ein Ring heißt *kommutativ*, wenn für alle $\alpha, \beta \in R$ gilt $\alpha\beta = \beta\alpha$.

BEMERKUNGEN. Sei R ein Ring.

1. 0 und 1 sind eindeutig bestimmt. Für 0 gilt dies, da $(R, +)$ eine Gruppe ist; für 1 hat man $1 = 1 \cdot 1' = 1'$.

2. Wie bei Körpern zeigt man, daß für alle $\alpha, \beta \in R$ gilt
$$\begin{aligned} 0 \cdot \alpha &= 0, \\ -\alpha &= (-1)\alpha, \\ (-\alpha)(-\beta) &= \alpha\beta. \end{aligned}$$

3. In einem (kommutativen) Ring hat nicht notwendig jedes $\alpha \neq 0$ hat ein (multiplikatives) Inverses. Falls es existiert, so ist es in einem kommutativen Ring jedoch eindeutig bestimmt. Ferner muß $0 \neq 1$ *nicht* gelten (zum Beispiel ist $\{0\}$ ein kommutativer Ring).

BEISPIELE. 1. Jeder Körper ist ein (kommutativer) Ring.

2. $\mathbb{Z}$ ist (kommutativer) Ring.

3. Sei K ein Körper. $K^{n\times n}$ ist mit der Matrizenaddition und -multiplikation ein Ring. Für $n > 1$ ist $K^{n\times n}$ *nicht kommutativ*, und es gibt *Nullteiler*, d.h. $A, B \neq 0$ mit $AB = 0$ (siehe 1.3.3).

Im Rest dieses Abschnitts sei K ein kommutativer Ring.

6.1.2. *Existenz der Determinante.* Sei $n \in \mathbb{N}$, $n \geq 1$. Eine Abbildung $\det\colon K^{n\times n} \to K$ (geschrieben als $\det(A) = \det(a_1, \ldots, a_n)$, wobei $a_1, \ldots, a_n \in K^n$ die Spalten von $A \in K^{n\times n}$ sind) heißt *Determinante*, wenn gilt

1. det ist *multilinear*, d.h. für alle i mit $1 \leq i \leq n$ und alle $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n \in K^n$ ist

$$f\colon K^n \to K$$
$$f(x) = \det(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n)$$

linear.

2. det ist *alternierend*, d.h. für alle i mit $1 \leq i < n$ und alle $a_1, \ldots, a_n \in K^n$ mit $a_i = a_{i+1}$ ist

$$\det(a_1, \ldots, a_n) = 0.$$

3. det ist *normiert*, d.h. $\det(E) = 1$. ($E \in K^{n\times n}$ ist die Einheitsmatrix).

SATZ. *Für jedes $n \in \mathbb{N}$, $n \geq 1$ existiert eine Determinante* $\det\colon K^{n\times n} \to K$.

BEWEIS. Induktion über n. *Basis* $n = 1$. Die Identität id_K hat offenbar die verlangten Eigenschaften.

Schritt $n - 1 \Rightarrow n$. Sei $d\colon K^{(n-1)\times(n-1)} \to K$ eine Determinante (Induktionsvoraussetzung). Zu definieren ist eine multilineare, alternierende und normierte Abbildung

$$\det\colon K^{n\times n} \to K.$$

Wähle i mit $1 \leq i \leq n$ fest. Sei $A = (\alpha_{ij}) \in K^{n\times n}$. Setze A_{ij} := die Matrix $\in K^{(n-1)\times(n-1)}$, die aus A durch Streichen der i-ten Zeile und der j-ten Spalte entsteht. Wir definieren jetzt

$$\det(A) := \sum_{j=1}^{n} \underbrace{(-1)^{i+j}\alpha_{ij} d(A_{ij})}_{=:f_j(a_1,\ldots,a_n)},$$

wobei $a_n, \ldots, a_n$ Spalten von A sind.

1. det ist multilinear. Offenbar genügt: f_j ist multilinear, d.h. $f_j(a_1, \ldots, a_n)$ ist linear in a_l, für $1 \leq l \leq n$. *Fall* $l \neq j$. Klar, da $d(A_{ij})$ linear in a_l ist und α_{ij} unabhängig von a_l ist. *Fall* $l = j$. Klar, da $d(A_{il})$ unabhängig von a_l ist und α_{ij} linear ist in a_l.

2. det ist alternierend. Sei $1 \leq l < n$ und $a_1, \ldots, a_n \in K^n$ mit $a_l = a_{l+1}$. Sei $A = (a_1, \ldots, a_n)$. Zu zeigen ist $\det(A) = 0$. Für $j \neq l, l+1$ ist $f_j(a_1, \ldots, a_n) = 0$, da $d(A_{ij}) = 0$ (denn d ist alternierend, und A_{ij} enthält zwei gleiche aufeinander folgende Spalten). Also:

$$\begin{aligned}
\det(A) &= f_l(a_1, \ldots, a_n) + f_{l+1}(a_1, \ldots, a_n) \\
&= (-1)^{i+l}\alpha_{il}d(A_{il}) + (-1)^{i+l+1}\underbrace{\alpha_{i\,l+1}}_{=\alpha_{il}}\underbrace{d(A_{i\,l+1})}_{=d(A_{il})} \\
&= 0
\end{aligned}$$

3. det ist normiert.

$$\begin{aligned}
\det(E) &= \sum_{j=1}^{n}(-1)^{i+j}\delta_{ij}d(E_{ij}) \\
&= (-1)^{2i}d(E') && \text{mit } E' \in K^{(n-1)\times(n-1)} \text{ Einheitsmatrix} \\
&= 1. && \text{QED}
\end{aligned}$$

6.1.3. *Eindeutigkeit der Determinante.* Die Eindeutigkeit der Determinante wird sich aus den folgenden drei Lemmata ergeben. Wir beweisen eine etwas allgemeinere Aussage für multilineare und alternierende (also nicht notwendig normierte) Abbildungen; dies wird später zu einem einfachen Beweis des Determinantenproduktsatzes führen.

Lemma. *Sei $d\colon K^{n\times n} \to K$ multilinear und alternierend. Dann gilt für alle $a_1, \ldots, a_n \in K^n$ und alle i, j mit $1 \leq i < j \leq n$:*

1. *$d(a_1, \ldots, a_n) = 0$, falls $a_i = a_j$.*
2. *$d(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n) = -d(a_1, \ldots, a_j, \ldots, a_i, \ldots, a_n)$.*

Beweis zu (2) für $j = i+1$. Sei $f\colon K^n \times K^n \to K$ definiert durch

$$f(x, y) := d(a_1, \ldots, a_{i-1}, x, y, a_{i+2}, \ldots, a_n).$$

Da d multilinear und alternierend ist, gilt

$$0 = f(x+y, x+y) = \underbrace{f(x,x)}_{=0} + \underbrace{f(y,y)}_{=0} + f(x,y) + f(y,x),$$

also $f(x,y) = -f(y,x)$.

Zu (1). Gegegen sei $(a_1, \ldots, a_n)$ mit $a_i = a_j$. Man vertausche a_i mit a_{i+1}, dann mit a_{i+2} usw.; schließlich mit a_{j-1}. Es entsteht $(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{j-1}, a_i, a_j, \ldots, a_n)$. Also

$$\begin{aligned}
&d(a_1, \ldots, a_{i-1}, a_i, a_{i-1}, \ldots, a_{j-1}, a_j, \ldots, a_n) \\
&= \pm d(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{j-1}, a_i, a_j, \ldots, a_n) \quad \text{nach dem obigen Spezialfall von (2)} \\
&= 0 \quad \text{da } d \text{ alternierend ist.}
\end{aligned}$$

Die allgemeine Form von (2) folgt jetzt aus (1) genauso wie oben der Spezialfall $j = i+1$ aus der Voraussetzung folgte, daß d alternierend ist. QED

LEMMA. *Sei $d\colon K^{n\times n} \to K$ multilinear und alternierend. Dann gilt für alle $A = (\alpha_{ij}) \in K^{n\times n}$*

$$d(A) = \sum_{\sigma\in S_n} \alpha_{\sigma(1)1} \dots \alpha_{\sigma(n)n} d(e_{\sigma(1)}, \dots, e_{\sigma(n)}).$$

(Erinnerung: $S_n := S(\{1,\dots,n\}) := \{\,\sigma\colon \{1,\dots,n\} \to \{1,\dots,n\} \mid \sigma$ bijektiv$\,\}$ ist die Gruppe der Permutationen *von $\{1,\dots,n\}$).*

BEWEIS. Seien $a_1,\dots,a_n$ die Spalten von A. Offenbar gilt $a_j = \sum_{i=1}^n \alpha_{ij} e_i$, also

$$\begin{aligned} d(A) &= d(a_1,\dots,a_n) \\ &= d(\sum_{i_1=1}^n \alpha_{i_1 1} e_{i_1}, a_2,\dots,a_n) \\ &= \sum_{i_1=1}^n \alpha_{i_1 1} d(e_{i_1}, a_2,\dots,a_n) \\ &= \sum_{i_1=1}^n \cdots \sum_{i_n=1}^n \alpha_{i_1 1}\dots\alpha_{i_n n} d(e_{i_1},\dots,e_{i_n}) \\ &= \sum_{\sigma\in S_n} \alpha_{\sigma(1)1}\dots\alpha_{\sigma(n)n} d(e_{\sigma(1)},\dots,e_{\sigma(n)}), \end{aligned}$$

denn $d(e_{i_1},\dots,e_{i_n}) = 0$ falls $i_j = i_l$ für ein j,l mit $j \neq l$. QED

LEMMA. *Sei $d\colon K^{n\times n} \to K$ multilinear und alternierend, und $\sigma \in S_n$. Dann gilt*

$$d(e_{\sigma(1)},\dots,e_{\sigma(n)}) = (-1)^{m_\sigma} d(e_1,\dots,e_n)$$

mit einer von d unabhängigen Zahl m_σ.

BEWEIS. Wir zeigen zunächst, daß sich $(\sigma(1),\dots,\sigma(n))$ durch eine endliche, von σ abhängige Anzahl m_σ von Vertauschungen zweier Elemente in $(1,\dots,n)$ überführen läßt. Sei etwa $\sigma(i) = n$ für ein $i < n$. Vertausche $\sigma(i) = n$ mit $\sigma(n)$. Ergebnis: $(\sigma(1),\dots,\sigma(i-1),\sigma(n),\sigma(i+1),\dots,\sigma(i))$. Wende jetzt dasselbe Verfahren auf das restliche $(n-1)$-Tupel (ohne das letzte Element) an.

Mit dem ersten Lemma ergibt sich hieraus $d(e_{\sigma(1)},\dots,e_{\sigma(n)}) = (-1)^{m_\sigma} d(e_1,\dots,e_n)$. Insbesondere ist also $d(e_{\sigma(1)},\dots,e_{\sigma(n)}) = (-1)^{m_\sigma}$ im Fall einer Determinante d; dann hat man nämlich $d(e_1,\dots,e_n) = 1$). QED

Aus diesen drei Lemmata ergibt sich unmittelbar das gesuchte Eindeutigkeitsresultat:

SATZ. *Sei $n \in \mathbb{N}$, $n \geq 1$ und $d, d'\colon K^{n\times n} \to K$ Determinanten. Dann ist $d = d'$.* QED

SCHREIBWEISE. Ist $A = (\alpha_{ij})$, so schreibt man für $\det(A)$ auch

$$\begin{vmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \dots & \alpha_{nn} \end{vmatrix}$$

6.1.4. *Permutationen und Transpositionen.* Als Folgerung aus dem Beweis notieren wir einfache Eigenschaften der Permutationsgruppe S_n. Insbesondere führen wir den Begriff des *Vorzeichens* ε_σ einer Permutation $\sigma \in S_n$ ein; er wird später gebraucht.

$\tau \in S_n$ heißt *Transposition*, wenn τ zwei Elemente vertauscht, also wenn es l, j mit $1 \leq l < j \leq n$ gibt mit

$$\tau(i) = \begin{cases} i & \text{für } i \neq j, l \\ l & \text{für } i = j \\ j & \text{für } i = l. \end{cases}$$

Lemma. *Sei n eine natürliche Zahl ≥ 1.*

1. *S_n hat $n!$ ($:= 1 \cdot 2 \cdot \ldots \cdot n$) Elemente.*
2. *Zu jedem $\sigma \in S_n$ gibt es Transpositionen $\tau_1, \ldots, \tau_m$ so daß*

$$\sigma = \tau_1 \circ \cdots \circ \tau_m.$$

m ist dabei entweder stets gerade oder stets ungerade, unabhängig von der Wahl der Transpositionen.

Beweis zu (1) durch Induktion über n. *Basis* $n = 1$. Klar. *Schritt* $n \Rightarrow n + 1$. Nach Induktionsvoraussetzung gibt es genau $n!$ Permutationen $\sigma \in S_{n+1}$ mit $\sigma(n+1) = i$ (für festes i mit $1 \leq i \leq n+1$). Also insgesamt gibt es $(n+1)n! = (n+1)!$ viele.

Zu (2). Nach dem Beweis des dritten Lemmas in 6.1.3 gibt es zu $\sigma \in S_n$ Transpositionen $\tau_1, \ldots, \tau_m$ mit

$$\begin{aligned} \tau_m \circ \cdots \circ \tau_1 \circ \sigma &= \mathrm{id} \\ \sigma &= \tau_1^{-1} \circ \cdots \circ \tau_m^{-1} \\ &= \tau_1 \circ \cdots \circ \tau_m \quad (\text{da } \tau_i^{-1} = \tau_i). \end{aligned}$$

Auch umgekehrt folgt aus $\sigma = \tau_1 \circ \cdots \circ \tau_m$, daß $\tau_m \circ \cdots \circ \tau_1 \circ \sigma = \mathrm{id}$ ist, und daraus wie im Beweis des dritten Lemmas

$$\det(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = (-1)^m.$$

(Beachte: det ist eindeutig bestimmt, etwa bzgl. $K = \mathbb{Z}$ oder $= \mathbb{Q}$). Also ist m entweder stets gerade oder stets ungerade. Qed

Definition. Für $\sigma \in S_n$ und Transpositionen $\tau_1, \ldots, \tau_m \in S_n$ mit $\sigma = \tau_1 \circ \cdots \circ \tau_m$ heißt $\varepsilon_\sigma := (-1)^m$ das *Vorzeichen* von σ; σ heißt *gerade*, falls $\varepsilon_\sigma = 1$, und *ungerade*, falls $\varepsilon_\sigma = -1$.

Bemerkung. Für $\sigma, \sigma' \in S_n$ gilt $\varepsilon_{\sigma \circ \sigma'} = \varepsilon_\sigma \varepsilon_{\sigma'}$. Ferner gilt $\varepsilon_{\mathrm{id}} = 1$. Die Abbildung $\sigma \mapsto \varepsilon_\sigma$ ist also ein Gruppenhomomorphismus von S_n in die multiplikative Gruppe $\{-1, 1\}$.

Beweis. Sei $\sigma = \tau_1 \circ \cdots \circ \tau_m$, also $\varepsilon_\sigma = (-1)^m$, und $\sigma' = \tau_1' \circ \cdots \circ \tau_{m'}'$, also $\varepsilon_{\sigma'} = (-1)^{m'}$. Dann ist $\sigma \circ \sigma' = \tau_1 \circ \cdots \circ \tau_m \circ \tau_1' \circ \cdots \circ \tau_{m'}'$, also $\varepsilon_{\sigma \circ \sigma'} = (-1)^{m+m'} = \varepsilon_\sigma \varepsilon_{\sigma'}$. Ferner ist $\varepsilon_{\mathrm{id}} = (-1)^\circ = 1$. Qed

6.2 Eigenschaften der Determinante

In diesem Abschnitt sei K stets ein kommutativer Ring.

6.2.1. Satz. *Sei $n \in \mathbb{N}$, $n \geq 1$ und $A = (\alpha_{ij}) \in K^{n\times n}$ mit Spalten $a_1, \ldots, a_n$. Dann gilt*

1. $\det(A) = \det(A^t) = \sum_{\sigma\in S_n} \varepsilon_\sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)}$
2. *(Entwicklung nach der i-ten Zeile;* Laplacescher Entwicklungssatz*). Für festes i mit $1 \leq i \leq n$ ist*
$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \alpha_{ij} \det(A_{ij})$$
(A_{ij} entsteht aus A durch Streichen der i-ten Zeile und j-ten Spalte).
3. *(Entwicklung nach der j-ten Spalte). Für festes j mit $1 \leq j \leq n$ ist*
$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \alpha_{ij} \det(A_{ij}).$$
4. *Seien $b_i \in K^n$, $\sigma \in S_n$ und $\alpha, \beta \in K$.*
   a. $\det(a_1, \ldots, a_{i-1}, \alpha a_i + \beta b_i, \alpha_{i+1}, \ldots, a_n) =$
   $\alpha \det(a_1, \ldots, a_n) + \beta \det(a_1, \ldots, a_{i-1}, b_i, a_{i+1}, \ldots, a_n)$.
   b. $\det(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) = \varepsilon_\sigma \det(a_1, \ldots, a_n)$.
   c. $\det(a_1, \ldots, a_n) = 0$ *falls $a_i = a_j$ für ein Paar i, j mit $1 \leq i < j \leq n$.*
5. *Ist $d\colon K^{n\times n} \to K$ multilinear und alternierend, so gilt $d(A) = \det(A) \cdot d(E)$.*

Beweis. 1. Nach dem zweiten und dritten Lemma aus 6.1.3 gilt

$$\begin{aligned}
\det(A) &= \sum_{\sigma\in S_n} \varepsilon_\sigma \alpha_{\sigma(1)1} \cdots \alpha_{\sigma(n)n} \\
&= \sum_{\sigma\in S_n} \varepsilon_\sigma \alpha_{1\sigma^{-1}(1)} \cdots \alpha_{n\sigma^{-1}(n)} && \text{da } K \text{ kommutativ ist} \\
&= \sum_{\sigma\in S_n} \varepsilon_\sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)} && \text{da } {}^{-1}\colon S_n \to S_n \text{ bijektiv ist und } \varepsilon_{\sigma^{-1}} = \varepsilon_\sigma, \\
& && \text{denn } \varepsilon_{\sigma^{-1}} \cdot \varepsilon_\sigma = \varepsilon_{\mathrm{id}} = 1 \\
\det(A^t) &= \sum_{\sigma\in S_n} \varepsilon_\sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)}.
\end{aligned}$$

2. Dies wurde im Beweis des Existenzsatzes 6.1.2 gezeigt.

3.
$$\begin{aligned}
\det(A) &= \det(A^t) && \text{nach Teil 1} \\
&= \sum_{j=1}^{n} (-1)^{i+j} \alpha_{ji} \det(A_{ji}) && \text{nach Teil 2, da } (A^t)_{ij} = (A_{ji})^t.
\end{aligned}$$

4. a. Nach Definition der Determinante. b. Man wiederhole den Beweis des dritten Lemmas aus 6.1.3 mit a_i statt e_i. c. Dies war bereits als Teil 1 des ersten Lemmas in 6.1.3 bewiesen worden.

5. Dies folgt aus dem zweiten und dritten Lemma in 6.1.3 zusammen mit Teil 1. Qed

Folgerung. *(Multiplikationssatz). Für $A, B \in K^{n\times n}$ gilt* $\det(AB) = \det(A)\det(B)$.

Beweis. Sei $A \in K^{n\times n}$ fest. Definiere

$$\begin{aligned}
&d\colon K^{n\times n} \to K,\\
&d(B) = \det(AB) = \det(Ab_1, \ldots, Ab_n) \quad (b_1, \ldots, b_n \text{ Spalten von } B).
\end{aligned}$$

d ist aufgrund der Definition multilinear und alternierend. Also gilt nach Teil 5 des Satzes

$$\det(AB) = d(B) = \det(B)d(E) = \det(B)\det(A). \quad \text{Qed}$$

Bemerkung. Wir wollen uns noch überlegen, welche Auswirkungen die in die in 1.4.1 eingeführten elementaren Spaltenumformungen einer Matrix auf ihre Determinante haben.

1. Vertauschen zweier Spalten verändert das Vorzeichen (nach Teil 4b des Satzes).
2. Multiplikation einer Spalte mit einem $\lambda \in K$ führt zum λ-fachen der Determinante (nach Teil 4a des Satzes).
3. Addition einer Spalte zu einer anderen verändert den Wert der Determinante nicht, denn es gilt nach den Teilen 4a,c des Satzes

$$\begin{aligned}
&\det(a_1, \ldots, a_i, \ldots, aa_i{+}_j, \ldots, a_n)\\
&= \det(a_1, \ldots, a_i, \ldots, a_i, \ldots, a_n) + \det(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n)\\
&= 0 + \det(a_1, \ldots, a_n).
\end{aligned}$$

Wegen $\det(A) = \det(A^t)$ gelten dieselben Aussagen auch für elementare Zeilenumformungen.

6.2.2. *Die komplementäre Matrix.* Wir zeigen jetzt eine Aussage, die gelegentlich beim Berechnen des Inversen einer Matrix hilft und die uns später beim Beweis des Satzes von Cayley/Hamilton gute Dienste leisten wird. Dafür benötigen wir einen etwas technischen Begriff. Die *komplementäre Matrix* $A^{\#}$ einer Matrix $A \in K^{n\times n}$ ist definiert durch

$$A^{\#} = (\alpha^{\#}_{ij}), \quad \text{wobei } \alpha^{\#}_{ij} := (-1)^{i+j}\det(A_{ji}).$$

Satz. *Für alle Matrizen $A \in K^{n\times n}$ gilt*

$$A^{\#} \cdot A = \det(A) \cdot E.$$

Beweis. Zu zeigen ist

$$\sum_{j=1}^{n} \alpha^{\#}_{ij}\alpha_{jl} = \begin{cases} \det(A) & \text{für } i = l,\\ 0 & \text{für } i \neq l.\end{cases}$$

Fall $i = l$. $\sum_{j=1}^{n} \alpha^{\#}_{ij}\alpha_{jl} = \sum_{j=1}^{n}(-1)^{i+j}\alpha_{ji}\det(A_{ji}) = \det(A)$. *Fall $i \neq l$:* A' entstehe aus A durch Ersetzen der i-ten Spalte durch die l-te. Dann ist

$$A_{ji} = A'_{ji} \quad \text{und} \quad \alpha_{jl} = \alpha'_{ji}.$$

Man erhält:

$$\begin{aligned}\sum_{j=1}^{n}\alpha_{ij}^{\#}\alpha_{jl} &= \sum_{j=1}^{n}(-1)^{i+j}\alpha_{jl}\det(A_{ji})\\ &= \sum_{j=1}^{n}(-1)^{i+j}\alpha'_{ji}\det(A'_{ji})\\ &= \det(A')\\ &= 0, \quad \text{da die } i\text{-te und die } l\text{-te Spalte von } A' \text{ übereinstimmen. QED}\end{aligned}$$

FOLGERUNG. *Sei K ein Körper, $A \in K^{n\times n}$. Folgende Aussagen sind äquivalent:*

1. *A ist invertierbar.*
2. $\det(A) \neq 0$.
3. $\mathrm{rg}(A) = n$.

In diesem Falle gilt außerdem

$$A^{-1} = \frac{1}{\det(A)}\cdot A^{\#}, \tag{6.1}$$

$$\det(A^{-1}) = \frac{1}{\det(A)}. \tag{6.2}$$

BEWEIS. Die Äquivalenz von (1) und (3) war schon in 3.3.3 gezeigt worden. $1 \Longrightarrow 2$. Die Invertierbarkeit von A besagt $A^{-1}\cdot A = E$, also $\det(A^{-1})\det(A) = \det(A^{-1}A) = \det(E) = 1$; dies beweist auch (6.2). $2 \Longrightarrow 1$. Ist $\det(A) \neq 0$, so ergibt sich aus dem Satz

$$\left(\frac{1}{\det(A)}\cdot A^{\#}\right)\cdot A = E;$$

daraus folgt auch (6.1). QED

6.2.3. *Berechnung von Determinanten.* Häufig benutzte Möglichkeiten dazu sind die Entwicklung nach Zeilen oder Spalten, sowie die Verwendung elementarer Zeilen- oder Spaltenumformungen. Durch Entwicklung nach der ersten Zeile sieht man zum Beispiel, daß

$$\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix} = \alpha\delta - \beta\gamma.$$

Ebenso beweist man die *Regel von* SARRUS

$$\begin{vmatrix}\alpha_{11} & \alpha_{12} & \alpha_{13}\\ \alpha_{21} & \alpha_{22} & \alpha_{23}\\ \alpha_{31} & \alpha_{32} & \alpha_{33}\end{vmatrix} = \alpha_{11}\alpha_{22}\alpha_{33} + \alpha_{12}\alpha_{23}\alpha_{31} + \alpha_{13}\alpha_{21}\alpha_{32} - \alpha_{11}\alpha_{23}\alpha_{32} - \alpha_{12}\alpha_{21}\alpha_{33} - \alpha_{13}\alpha_{22}\alpha_{31}.$$

Ferner ist die folgende Beobachtung oft nützlich. Sei $A \in K^{n\times n}$ von der Gestalt

$$A = \begin{pmatrix}A_1 & & *\\ & \ddots & \\ 0 & & A_r\end{pmatrix}$$

mit $A_i \in K^{n_i \times n_i}$ und $\sum_{i=1}^{r} n_i = n$. Dann gilt

$$\det(A) = \det(A_1) \cdot \ldots \cdot \det(A_r).$$

Insbesondere ist also

$$\begin{vmatrix} \gamma_1 & & * \\ & \ddots & \\ 0 & & \gamma_n \end{vmatrix} = \gamma_1 \cdot \ldots \cdot \gamma_n.$$

Den Beweis führen wir durch Induktion nach r. *Basis* $r = 1$. In diesem Fall ist nichts zu zeigen. *Schritt* $r \Longrightarrow r + 1$. Sei

$$A = \begin{pmatrix} A_1 & & & * \\ & \ddots & & \\ & & A_r & \\ 0 & & & A_{r+1} \end{pmatrix}$$

mit $A_i \in K^{n_i \times n_i}$ und $n_1 + \cdots + n_{r+1} = n$. Definiere

$$d\colon K^{n_{r+1} \times n_{r+1}} \to K,$$

$$d(B) = \begin{vmatrix} A_1 & & & * \\ & \ddots & & \\ & & A_r & \\ 0 & & & B \end{vmatrix}$$

d ist multilinear und alternierend, da det es ist. Also gilt nach Satz 6.2.1(5)

$$\det(A) = d(A_{r+1}) = \det(A_{r+1}) d(E) = \det(A_{r+1}) \cdot \det(A_1) \cdot \ldots \cdot \det(A_r),$$

denn nach Induktionsvoraussetzung ist

$$d(E) = \begin{vmatrix} A_1 & & & * \\ & \ddots & & \\ & & A_r & \\ 0 & & & E \end{vmatrix} = \begin{vmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_r \end{vmatrix} = \det(A_1) \cdot \ldots \cdot \det(A_r),$$

wie man durch n_{r+1}-malige Entwicklung nach der letzten Zeile sieht. QED

6.2.4. CRAMER*sche Regel.* Sei $A \in K^{n \times n}$, $b \in K^n$ und $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^n$ mit $Ax = b$. Dann gilt

$$\xi_i \det(A) = \det(A_i),$$

wobei A_i aus A entsteht durch Ersetzen der i-ten Spalte durch b.

BEWEIS. Sei $A = (a_1, \ldots, a_n)$. Die Gleichung $Ax = b$ bedeutet also $\sum_{j=1}^{n} \xi_j a_j = b$. Daraus folgt

$$\det(A_i) = \det(a_1, \ldots, a_{i-1}, \sum_{j=1}^{n} \xi_j a_j, a_{i+1}, \ldots, a_n) = \xi_i \det(a_1, \ldots, a_n). \quad \text{QED}$$

6.2.5. VANDERMONDE*sche Determinante.* Für $\alpha_0, \ldots, \alpha_n \in K$ ist die VANDERMONDE*sche Determinante* definiert durch

$$D_n = \begin{vmatrix} 1 & \alpha_0 & \alpha_0^2 & \ldots & \alpha_0^n \\ 1 & \alpha_1 & \alpha_1^2 & \ldots & \alpha_1^n \\ \vdots & & & & \\ 1 & \alpha_n & \alpha_n^2 & \ldots & \alpha_n^n \end{vmatrix}$$

LEMMA. *Für $\alpha_0, \ldots, \alpha_n \in K$ gilt*

$$D_n = \prod_{i<j\leq n} (\alpha_j - \alpha_i).$$

BEWEIS. Induktion über n. *Basis.* Für $n = 1$ (und auch für $n = 0$, wenn man wie üblich das leere Produkt als 1 auffaßt) ist die Behauptung klar. *Schritt* $n-1 \Longrightarrow n$. Wir wollen zunächst das zweite bis $(n+1)$-te Glied der ersten Zeile zum Verschwinden bringen. Dazu ziehen wir von der zweiten Spalte das α_0-fache der ersten Spalte ab, ebenso von der dritten Spalte das α_0-fache der zweiten Spalte und so weiter. Es ergibt sich

$$D_n = \begin{vmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & \alpha_1 - \alpha_0 & \alpha_1(\alpha_1 - \alpha_0) & \ldots & \alpha_1^{n-1}(\alpha_1 - \alpha_0) \\ \vdots & & & & \\ 1 & \alpha_n - \alpha_0 & \alpha_n(\alpha_n - \alpha_0) & \ldots & \alpha_n^{n-1}(\alpha_n - \alpha_0) \end{vmatrix}$$

Durch Entwickeln nach der ersten Zeile und anschließendes Herausziehen der Faktoren $\alpha_i - \alpha_0$ erhält man unter Verwendung den Induktionsvoraussetzung

$$D_n = (\alpha_1 - \alpha_0) \ldots (\alpha_n - \alpha_0) \begin{vmatrix} 1 & \alpha_1 & \alpha_1^2 & \ldots & \alpha_1^{n-1} \\ \vdots & & & & \\ 1 & \alpha_n & \alpha_n^2 & \ldots & \alpha_n^{n-1} \end{vmatrix} = \prod_{i<j\leq n} (\alpha_j - \alpha_i). \quad \text{QED}$$

Wir wollen uns jetzt noch mit einem sogenannten *Interpolationsproblem* befassen, genauer mit der Frage, wie man durch $n+1$ Punkte im $\mathbb{R}^2$ mit verschiedenen Projektionen auf die x-Achse ein Polynom möglichst kleinen Grades finden kann, das durch alle gegebenen Punkte geht. Hierbei erweist sich die VANDERMONDEsche Determinante als nützlich.

LEMMA. *Zu $n+1$ gegebenen* Stützstellen $(\alpha_0, \beta_0), \ldots, (\alpha_n, \beta_n) \in \mathbb{R} \times \mathbb{R}$ *mit $\alpha_0, \ldots, \alpha_n$ paarweise verschieden gibt es genau ein Polynom n-ten Grades, das an jeder Stelle α_i den Wert β_i hat.*

BEWEIS. Gesucht sind $\gamma_0, \ldots, \gamma_n$ so daß

$$\sum_{k=0}^{n} \gamma_k \alpha_i^k = \beta_i \quad \text{für } i = 0, \ldots, n.$$

In Matrizenschreibweise bedeutet dies

$$\begin{pmatrix} 1 & \alpha_0 & \alpha_0^2 & \ldots & \alpha_0^n \\ \vdots & & & & \\ 1 & \alpha_n & \alpha_n^2 & \ldots & \alpha_n^n \end{pmatrix} \begin{pmatrix} \gamma_0 \\ \vdots \\ \gamma_n \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix}.$$

Hier entsteht also in natürlicher Weise die zur VANDERMONDEschen Determinante gehörige Matrix. Wegen

$$\begin{vmatrix} 1 & \alpha_0 & \alpha_0^2 & \dots & \alpha_0^n \\ \vdots & & & & \\ 1 & \alpha_n & \alpha_n^2 & \dots & \alpha_n^n \end{vmatrix} = \prod_{i<j\leq n} (\alpha_j - \alpha_i) \neq 0 \quad \text{nach Voraussetzung}$$

ist dieses Gleichungssystem eindeutig lösbar. Die Lösungen – also die Koeffizienten des Polynoms – kann man etwa nach der CRAMERschen Regel ausrechnen:

$$\gamma_k = \frac{1}{\prod_{i<j\leq n}(\alpha_j - \alpha_i)} \begin{vmatrix} 1 & \alpha_0 & \dots & \alpha_0^{k-1} & \beta_0 & \alpha_0^{k+1} & \dots & \alpha_0^n \\ \vdots & & & & & & & \\ 1 & \alpha_n & \dots & \alpha_n^{k-1} & \beta_n & \alpha_n^{k+1} & \dots & \alpha_n^n \end{vmatrix}. \quad \text{QED}$$

6.3 Anwendungen

Behandelt werden Vektorprodukte, das Volumen eines Parallelotops und der Begriff der Orientierung. In diesem Abschnitt sei K stets ein Körper.

6.3.1. *Vektorprodukt.* Seien $a, a_1, \dots, a_{n-1} \in K^n$. a heißt *Vektorprodukt* von $a_1, \dots, a_{n-1}$ (geschrieben $a_1 \times \dots \times a_{n-1}$), wenn für alle $x \in K^n$ gilt

$$\det(a_1, \dots, a_{n-1}, x) = a^t x.$$

BEMERKUNGEN. 1. Das Vektorprodukt $a_1 \times \dots \times a_{n-1}$ existiert und ist eindeutig bestimmt, denn $\det(a_1, \dots, a_{n-1}, x)$ ist (in x) eine lineare Abbildung des K^n in K und deshalb (nach 3.1.1) mit einem geeigneten a in der Form $a^t x$ darstellbar.

2. $\det(a_1, \dots, a_{n-1}, e_i)$ ist die i-te Komponente von $a_1 \times \dots \times a_{n-1}$.

3. Die Familie $a_1, \dots, a_{n-1}$ ist linear unabhängig genau dann, wenn $a_1 \times \dots \times a_{n-1} \neq 0$.

BEWEIS. $\Longrightarrow$. Seien $a_1, \dots, a_{n-1}$ linear unabhängig. Man ergänze sie zu Basis $a_1, \dots, a_{n-1}, x$ des K^n. Dann gilt

$$0 \neq \det(a_1, \dots, a_{n-1}, x) = (a_1 \times \dots \times a_{n-1})^t x.$$

Also ist $a_1 \times \dots \times a_{n-1} \neq 0$.

$\Longleftarrow$. Gelte $a_1 \times \dots \times a_{n-1} \neq 0$. Dann ist $0 \neq (a_1 \times \dots \times a_{n-1})^t (a_1 \times \dots \times a_{n-1}) = \det(a_1, \dots, a_{n-1}, a_1 \times \dots \times a_{n-1})$. Also sind $a_1, \dots, a_{n-1}$ sind linear unabhängig. QED

4. Für $\sigma \in S_{n-1}$ gilt

$$a_{\sigma(1)} \times \dots \times a_{\sigma(n-1)} = \varepsilon_\sigma a_1 \times \dots \times a_{n-1}.$$

5. Es gilt $(a_1 \times \dots \times a_{n-1})^t a_i = 0$ für $1 \leq i \leq n$, denn det ist alternierend.

6. Im Spezialfall $n = 3$ kann man das Vektorprodukt explizit angeben. Seien etwa $a = \left(\begin{smallmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{smallmatrix}\right)$

und $b = \left(\begin{smallmatrix}\beta_1\\ \beta_2\\ \beta_3\end{smallmatrix}\right) \in K^3$. Dann ist nach Bemerkung 2

$$a \times b = \begin{pmatrix} \begin{vmatrix} \alpha_1 & \beta_1 & 1 \\ \alpha_2 & \beta_2 & 0 \\ \alpha_3 & \beta_3 & 0 \end{vmatrix} \\ \begin{vmatrix} \alpha_1 & \beta_1 & 0 \\ \alpha_2 & \beta_2 & 1 \\ \alpha_3 & \beta_3 & 0 \end{vmatrix} \\ \begin{vmatrix} \alpha_1 & \beta_1 & 0 \\ \alpha_2 & \beta_2 & 0 \\ \alpha_3 & \beta_3 & 1 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} \alpha_2\beta_3 & -\alpha_3\beta_2 \\ -\alpha_1\beta_3 & +\alpha_3\beta_1 \\ \alpha_1\beta_2 & -\alpha_2\beta_1 \end{pmatrix}.$$

Lemma. *(Gleichung einer Hyperebene durch n Punkte des n-dimensionalen Raumes). Seien $a_1, \ldots, a_n \in K^n$, so daß $H := A(a_1, \ldots, a_n)$ Hyperebene im K^n ist. Dann gilt*

$$H = \{\, x \in K^n \mid a^t x = a^t a_1 \,\} \quad \textit{mit } a = (a_2 - a_1) \times \cdots \times (a_n - a_1).$$

Beweis. Folgende Aussagen sind äquivalent.

$$\begin{aligned} x &\in H \\ x &\in a_1 + \langle a_2 - a_1, \ldots, a_n - a_1 \rangle && \text{nach 4.1.2} \\ x - a_1 &\in \langle a_2 - a_1, \ldots, a_n - a_1 \rangle \\ 0 &= \det(a_2 - a_1, \ldots, a_n - a_1, x - a_1) \\ \det(a_2 - a_1, \ldots, a_n - a_1, x) &= \det(a_2 - a_1, \ldots, a_n - a_1, a_1) \\ a^t x &= a^t a_1. && \text{Qed} \end{aligned}$$

6.3.2. *Volumen eines Parallelotops.* (V, σ) sei euklidischer Vektorraum (also $K = \mathbb{R}$) und $a_1, \ldots, a_n \in V$ linear unabhängig.

1. $P(a_1, \ldots, a_n) := \{\, \sum_{i=1}^n \alpha_i a_i \mid \alpha_i \in \mathbb{R}, 0 \le \alpha_i \le 1$ für $1 \le i \le n \,\}$ heißt das von $a_1, \ldots, a_n$ erzeugte *Parallelotop*.
2. $\mathrm{Vol}(a_1, \ldots, a_n) := \sqrt{\det(\sigma(a_i, a_j))} \ge 0$ nennt man das (n-dimensionale) *Volumen* von $P(a_1, \ldots, a_n)$; dies ist eine korrekte und vernünftige Definition nach folgendem Satz.

Satz. *(V, σ) sei euklidischer Vektorraum und $a_1, \ldots, a_n \in V$ linear unabhängig. Ferner sei $e_1, \ldots, e_n$ Orthonormalbasis von $\langle a_1, \ldots, a_n \rangle$ und $A \in \mathbb{R}^{n \times n}$ eine Matrix mit*

$$a_i = \sum_{s=1}^n \alpha_{si} e_s$$

für $1 \le i \le n$. Dann gilt

1. $\mathrm{Vol}(a_1, \ldots, a_n)$ *ist wohldefiniert und gleich* $|\det(A)|$.
2. $\mathrm{Vol}(a_1, \ldots, a_n) = \mathrm{Vol}(a_1, \ldots, a_{n-1}) \cdot \|h\|$, *wobei h das Lot von a_n auf $\langle a_1, \ldots, a_{n-1} \rangle$ ist (Grundfläche mal Höhe).*

Beispiele. $n = 1$: $\mathrm{Vol}(a) = \sqrt{\sigma(a,a)} = \|a\|$.

$n = 2$: Für linear unabhängige $a_1, a_2 \in V$ ist

$$\mathrm{Vol}(a_1, a_2) = \|a_1\| \cdot \|h\| \quad \text{(Grundlinie $\cdot$ Höhe)}$$

$n = 3$: Für linear unabhängige $a_1, a_2, a_3 \in V$ ist

$$\mathrm{Vol}(a_1, a_2, a_3) = \mathrm{Vol}(a_1, a_2) \cdot \|h\| \quad \text{(Grundfläche $\cdot$ Höhe)}$$

Beweis. 1. Sei $A = (\alpha_{st})$. Nach Voraussetzung gilt

$$a_i = \sum_{s=1}^{n} \alpha_{si} e_s \quad \text{und} \quad a_j = \sum_{t=1}^{n} \alpha_{tj} e_t.$$

Also:

$$\begin{aligned} \sigma(a_i, a_j) &= \sum_{s,t=1}^{n} \alpha_{si}\alpha_{tj}\sigma(e_s, e_t) \\ &= \sum_{s=1}^{n} \alpha_{si}\alpha_{sj} \qquad\qquad \text{da $e_1, \dots, e_n$ Orthonormalbasis ist.} \end{aligned}$$

Man erhält $(\sigma(a_i, a_j))_{ij} = A^t A$, also

$$\begin{aligned} \det((\sigma(a_i, a_j))_{ij}) &= \det(A^t A) \\ &= \det(A^t)\det(A) \\ &= \det(A)^2. \end{aligned}$$

Also ist $\mathrm{Vol}(a_1, \dots, a_n)$ wohldefiniert und gleich $|\det(A)|$.

2. Schreibe $a_n = u + h$ mit $u \in U := \langle a_1, \dots, a_{n-1}\rangle$, $h \in U^\perp$ (dies ist möglich, da $\langle a_1, \dots, a_n\rangle \subseteq U \oplus U^\perp$). Dann gilt

$$\begin{aligned} \mathrm{Vol}(a_1, \dots, \underbrace{a_n}_{=u+h}) &= |\det(A)| \\ &= \mathrm{Vol}(a_1, \dots, a_{n-1}, h), \quad \text{da $a_1, \dots, a_{n-1}, u$ linear abhängig sind.} \end{aligned}$$

Also:

$$\begin{aligned} \mathrm{Vol}(a_1, \dots, a_n)^2 &= \mathrm{Vol}(a_1, \dots, a_{n-1}, h)^2 \\ &= \begin{vmatrix} \sigma(a_1, a_1) & \dots & \sigma(a_1, a_{n-1}) & 0 \\ \vdots & & & \vdots \\ \sigma(a_{n-1}, a_1) & \dots & \sigma(a_{n-1}, a_{n-1}) & 0 \\ 0 & \dots & 0 & \sigma(h,h) \end{vmatrix} \\ &= \det((\sigma(a_i, a_j))_{1 \le i,j \le n-1}) \cdot \sigma(h,h) \\ &= \mathrm{Vol}(a_1, \dots, a_{n-1})^2 \|h\|^2. \quad \text{QED} \end{aligned}$$

FOLGERUNG. *Seien $a_1, \ldots, a_{n-1} \in \mathbb{R}^2$ linear unabhängig, $\mathbb{R}^n$ mit dem üblichen Skalarprodukt versehen. Dann gilt*

$$\operatorname{Vol}(a_1, \ldots, a_{n-1}) = \|a_1 \times \cdots \times a_{n-1}\|.$$

BEWEIS. Sei $h := a_1 \times \cdots \times a_{n-1}$. Dann gilt $h \perp a_i$ für $1 \leq i \leq n-1$ (nach Bemerkung 5 in 6.3.1), und $h \neq 0$ (Bemerkung 3 in 6.3.1). h ist das Lot von h auf $\langle a_1, \ldots, a_{n-1} \rangle$. Man erhält

$$\begin{aligned}
\operatorname{Vol}(a_1, \ldots, a_{n-1}, h) &= \operatorname{Vol}(a_1, \ldots, a_{n-1})\|h\| && \text{nach Teil 2 des Satzes}\\
\operatorname{Vol}(a_1, \ldots, a_{n-1}, h) &= |\det(a_1, \ldots, a_{n-1}, h)| && \text{nach Teil 1 des Satzes}\\
&= |(a_1 \times \cdots \times a_n)^t h|\\
&= \|h\|^2.
\end{aligned}$$

Wegen $h \neq 0$ folgt $\operatorname{Vol}(a_1, \ldots, a_{n-1}) = \|h\|$. QED

6.3.3. *Determinante eines Endomorphismus, Orientierung.* Sei V ein n-dimensionaler K-Vektorraum und f ein Endomorphismus, also $f\colon V \to V$ linear. Seien ferner $A, B \in K^{n \times n}$ beide darstellende Matrizen von f. Nach dem Satz in 3.4.3 gilt dann

$$\begin{aligned}
A &\sim B\\
B &= CAC^{-1} && \text{für ein } C \in GL(n, K)\\
\det(B) &= \det(CAC^{-1})\\
&= \det(C)\det(A)\det(C^{-1})\\
&= \det(C)\det(A)\det(C)^{-1}\\
&= \det(A).
\end{aligned}$$

Wir können also die Determinante $\det(f)$ des Endomorphismus f definieren als $\det(A)$, wobei A eine beliebige darstellende Matrix von f ist.

DEFINITION. Sei V endlichdimensionaler $\mathbb{R}$-Vektorraum.

1. Sei $f\colon V \to V$ ein Endomorphismus. Gilt $\det(f) > 0$, so sagen wir, daß f die *Orientierung erhält*.

2. $\mathcal{A} = (a_1, \ldots, a_n)$ und $\mathcal{B} = (b_1, \ldots, b_n)$ seien zwei Basen von V und $C \in \mathbb{R}^{n \times n}$ mit

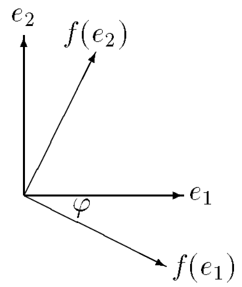
   $$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = C \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

   $\mathcal{A}$ heißt *gleichorientiert* mit $\mathcal{B}$ genau dann, wenn $\det(C) > 0$.

3. Seien $\mathcal{A}$, $\mathcal{B}$ wie eben. $\mathcal{A}$ heiße (willkürlich!) positiv orientiert. Die Basis $\mathcal{B}$ nennt man dann *positiv orientiert* genau dann, wenn $\mathcal{B}$ gleichorientiert mit $\mathcal{A}$ ist. $\mathcal{B}$ heißt *negativ orientiert* genau dann, wenn $\mathcal{B}$ nicht gleichorientiert mit $\mathcal{A}$ ist.

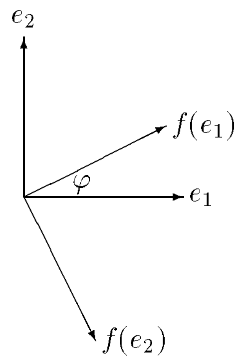
BEMERKUNG. In $\mathbb{R}^n$ nennt man *willkürlich* die Standardbasis $e_1, \ldots, e_n$ positiv orientiert.

Beispiele. 1. Sei $V = \mathbb{R}^2$ und f Drehung um den Winkel φ mit Matrix $A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$, also



Wegen $\det(A) = 1$ erhält f die Orientierung.

2. g sei die Drehspiegelung mit Matrix $B = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$, also



Wegen $\det(B) = -1$ erhält g die Orientierung nicht.

3. Sei $V = \mathbb{R}^n$ mit üblichem Skalarprodukt versehen und $a_1, \ldots, a_{n-1}$ linear unabhängig. Dann ist $a_1, \ldots, a_{n-1}, \underbrace{a_1 \times \cdots \times a_{n-1}}_{=:h}$ positiv orientiert.

Beweis.

$$\det(a_1, \ldots, a_{n-1}, h) = (a_1 \times \cdots \times a_{n-1})^t h = \|h\|^2 > 0. \quad \text{Qed}$$

Folgerung. *In $\mathbb{R}^n$ (mit dem üblichen Skalarprodukt) ist das Vektorprodukt aus $n-1$ linear unabhängigen Vektoren $a_1, \ldots, a_{n-1}$ der auf $\langle a_1, \ldots, a_{n-1} \rangle$ senkrecht stehende Vektor mit Länge $\mathrm{Vol}(a_1, \ldots, a_{n-1})$, so daß $a_1, \ldots, a_{n-1}, h$ positiv orientiert ist (d.h. gleichorientiert mit der Standard-Orthonormalbasis).*

Index

Literaturverzeichnis

[1] Gerd Fischer. *Lineare Algebra, 11-te Auflage*. Vieweg, Braunschweig, 1997.

[2] Hannes Stoppel and Birgit Griese. *Übungsbuch zur Linearen Algebra*. Vieweg, Braunschweig, 1997.