Lineare Algebra und Analytische Geometrie

Helmut Schwichtenberg

Mathematisches Institut der Universität München
Wintersemester 2004/2005 und Sommersemester 2005

Inhaltsverzeichnis

Vorwort iii

Kapitel 1. Lineare Gleichungssysteme 1
1.1. Mengen 1
1.2. Gruppen und Körper 6
1.3. Matrizen 15
1.4. Lineare Gleichungssysteme 18

Kapitel 2. Vektorräume 29
2.1. Definition von Vektorräumen 29
2.2. Lineare Unabhängigkeit, Basis, Dimension 31
2.3. Untervektorräume, direkte Summen und Quotienten 39

Kapitel 3. Lineare Abbildungen 47
3.1. Lineare Abbildungen 47
3.2. Kern und Bild 55
3.3. Rang von Matrizen 57
3.4. Lineare Abbildungen und Matrizen 60

Kapitel 4. Affine Geometrie 65
4.1. Affine Unterräume und lineare Varietäten 65
4.2. Quadratische Hyperflächen 70

Kapitel 5. Euklidische Geometrie 79
5.1. Euklidische Vektorräume 79
5.2. Orthonormalisierung 84
5.3. Orthogonale Abbildungen 90
5.4. Hauptachsentransformation nach Jacobi 95

Kapitel 6. Determinanten 103
6.1. Existenz und Eindeutigkeit der Determinante 103
6.2. Eigenschaften der Determinante 107
6.3. Anwendungen 113

Kapitel 7. Eigenwerte und Eigenvektoren, Diagonalisierung 119
7.1. Polynome 119
7.2. Eigenwerte und Eigenvektoren 123
7.3. Jordansche Normalform 128
7.4. Diagonalisierung 137

Kapitel 8. Komplexifizierung und unitäre Vektorräume 141
8.1. Komplexifizierung 141
8.2. Anwendung auf die reelle Jordansche Normalform 149

8.3. Unitäre Vektorräume 152

Kapitel 9. Endomorphismen 165
9.1. Normale Endomorphismen 165
9.2. Zerlegung von Endomorphismen; geometrische Deutung 168
9.3. Orthogonale Endomorphismen 179

Kapitel 10. Gruppen, Moduln, Tensorprodukt 195
10.1. Gruppen 195
10.2. Faktorgruppen 208
10.3. Moduln, Satz von Cayley-Hamilton 214
10.4. Tensorprodukt 222

Literaturverzeichnis 231

Index 233

Vorwort

Das vorliegende Skriptum gibt den Inhalt einer vierstündigen Vorlesung wieder, die ich im akademischen Jahr 2004/2005 am Mathematischen Institut der Universität München gehalten habe. Hauptsächliche Quellen waren Vorlesungen meines Kollegen Hans-Jürgen Schneider sowie das Buch von Gerd Fischer [**2**] über lineare Algebra.

Bedanken möchte ich mich bei Oliver Deiser für seine Unterstützung bei der Vorbereitung der Vorlesung. Für die gute Zusammenarbeit und viele hilfreiche Anmerkungen bedanke ich mich bei den Tutoren und Korrektoren.

München, im Juli 2005

Helmut Schwichtenberg

KAPITEL 1

Lineare Gleichungssysteme

Bei unserem Studium der linearen Algebra wollen wir von einem konkreten Problem ausgehen: 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

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

gilt. Das System (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. Wir wollen uns hier auf eine Umgebung konzentrieren, bei der die Zahlen des Gleichungssystems Elemente eines sogenannten *Körpers* sind. Der Begriff des Körpers baut auf dem einer *Gruppe* auf, den 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 entwickeln, und uns dabei auf eine inhaltliche Vorstellung des Mengenbegriffs stützen. Für eine genauere Diskussion der historischen Entwicklung der Mengenlehre und insbesondere eine präzise axiomatische Entwicklung der Mengenlehre müssen wir auf die Literatur (etwa das Buch von Deiser [**1**]) oder Vorlesungen über mathematische Logik verweisen.

1.1.1. Cantors Definition. 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.

Insbesondere ist eine Menge durch ihre Elemente vollständig bestimmt (Extensionalitätsprinzip).

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

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

bilden, die Menge aller Objekte x aus V mit der Eigenschaft $A(x)$. Man beachte, daß $\{\, x \mid A(x) \,\}$ wieder ein Objekt aus V ist. Beispiele für Eigenschaften sind etwa

(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: $\mathrm{Mg}(y) \wedge x \in y$.

Da man hier alle Objekte mit einer gewissen Eigenschaft zusammenfaßt, spricht man von einem (naiven) *Komprehensionsprinzip*.

Cantors Definition – oder genauer unsere naive Auffassung davon – ist jedoch so nicht haltbar, da sie zu Widersprüchen führt. Am bekanntesten ist die *Russellsche Antinomie*: Sei

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

Dann gilt[1] $x \in x_0 \equiv \mathrm{Mg}(x) \wedge x \notin x$ 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. Einen Grund für diesen Widerspruch kann man darin sehen, daß wir hier – unter Verwendung des naiven Komprehensionsprinzips – von der Vorstellung einer fertigen Gesamtheit aller Mengen ausgegangen sind. 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 die 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:

Stufe 0:	–
Stufe 1:	$\emptyset$
Stufe 2:	$\emptyset$, $\{\emptyset\}$
Stufe 3:	$\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}\}$
	„und so weiter“.

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.

[1] Wir verwenden $\equiv$ für „genau dann, wenn“.

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 eine beliebige Gesamtheit von Mengen. Jede Menge ist dann insbesondere eine Klasse. Ferner gibt es Klassen, die keine Mengen sind, z.B. die Klasse V aller Mengen; solche Klassen nennt man *echte Klassen*.

Bemerkung (Einfache Typen). Für die meisten Zwecke in der Mathematik und insbesondere auch für diese Vorlesung würde es ausreichen, mit einem einfacheren und konkreteren Mengenbegriff zu arbeiten, nämlich mit Mengen von Objekten eines festen „einfachen“ Typs τ. *Einfache Typen* sind aus einem 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 in diesem Kontext 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 dann immer eine Gesamtheit von Objekten eines festen Typs τ. Meist ist der Typ aus dem Zusammenhang 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.

Im folgenden verwenden wir den naiven Mengenbegriff.

1.1.3. Teilmengen, Durchschnitt, Vereinigung. Sind M und N Mengen, 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, so definieren wir

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

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

$$M \cap N = N \cap M \qquad M \cup N = N \cup M,$$
$$(M \cap N) \cap K = M \cap (N \cap K) \qquad (M \cup N) \cup K = M \cup (N \cup K),$$
$$M \cap (N \cup K) = (M \cap N) \cup (M \cap K) \qquad M \cup (N \cap K) = (M \cup N) \cap (M \cup K).$$

sowie

$$M \setminus (N \cap K) = (M \setminus N) \cup (M \setminus K),$$
$$M \setminus (N \cup K) = (M \setminus N) \cap (M \setminus K).$$

Zwei Mengen M und N heißen *disjunkt*, wenn sie keine gemeinsamen Elemente haben, also wenn $M \cap N = \emptyset$.

Sind M und N Mengen, so nennt man die Menge

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

das *kartesische Produkt* der Mengen M und N. Hierbei ist (a, b) das *geordnete Paar*[2] von a und b.

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. Ist speziell $M = N$, so spricht man von einer *Relation auf* M.

Beispiel. Die Teilbarkeitsrelation $m|n$ auf $\mathbb{N}$ ist die Menge $\{\, (m, n) \mid \exists k{\in}\mathbb{N}\colon m \cdot k = n \,\} \subseteq \mathbb{N} \times \mathbb{N}$. Hierbei ist „$\exists k{\in}\mathbb{N}\colon m \cdot k = n$“ zu lesen als „es gibt ein $k \in \mathbb{N}$ mit $m \cdot k = n$“. „$\exists$“ heißt *Existenzquantor*.

1.1.4. Abbildungen. Im folgenden werden Relationen R zwischen M und N eine besondere Rolle spielen, für die es zu jedem $a \in M$ höchstens ein $b \in N$ gibt mit aRb. Solche Relationen nennt man *rechtseindeutig* oder auch *Funktionen*.

Unter einer *Abbildung* verstehen wir ein Tripel (M, f, N), wobei M, N Mengen sind und f eine rechtseindeutige Relation zwischen M und N ist (d.h. daß aus afb und afc folgt $b = c$), und ferner es zu jedem $a \in M$ ein $b \in N$ gibt mit afb. Das eindeutig bestimmte b mit afb wird mit $f(a)$ bezeichnet. Eine Abbildung (M, f, N) wird meist in der Form

$$f\colon M \to N$$

mitgeteilt; man nennt M den *Definitionsbereich* und N den *Wertebereich* der Abbildung.

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

$$f\colon M' \to N$$
$$a \mapsto f(a)$$

definierte Abbildung (also $(M', \{\, (a, f(a)) \mid a \in M' \,\}, N)$) 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 Funktionswerte die Gleichheit der Argumente folgt, also wenn

$$\forall a_1, a_2{\in}M\colon (f(a_1) = f(a_2) \to a_1 = a_2)$$

(zu lesen als „für alle $a_1, a_2 \in M$ gilt: wenn $f(a_1) = f(a_2)$, so $a_1 = a_2$“; „$\forall$“ heißt *Allquantor*);

(2) *surjektiv*, wenn jedes $b \in N$ als Funktionswert auftritt, also wenn

$$\forall b{\in}N \exists a{\in}M\colon f(a) = b;$$

[2]Man kann den Begriff des geordneten Paares innerhalb der Mengenlehre definieren. Wir verzichten hier jedoch darauf und verwenden ihn als Grundbegriff.

(3) *bijektiv*, wenn sie injektiv und surjektiv ist.

Zu jeder bijektiven Abbildung $f\colon M \to N$ gibt es 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 definiert 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

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

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* $\mathrm{id}_M\colon M \to M$ 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 $f^{-1} \circ f = \mathrm{id}_M$.

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 $C \subseteq N$

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

$f[A]$ heißt das *Bild* von A unter f, und $f^{-1}[C]$ das *Urbild* von C unter f. 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 und seien $A, B \subseteq M$ 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.

1.1.5. Familien von Mengen, Potenzmenge. 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 $I = \{1, \ldots, n\}$, so hat man eine endliche Familie und schreibt[3] $(f_1, \ldots, f_n)$

[3]Diese Schreibweise kollidiert für $n = 2$ mit der oben eingeführten Schreibweise (a, b) für geordnete Paare. Es sollte jedoch jeweils aus dem Kontext klar sein, was gemeint ist.

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, so definieren wir

$$\bigcap_{i\in I} M_i := \{\, a \mid \forall i{\in}I\colon a \in M_i \,\} \quad (\textit{Durchschnitt} \text{ der Familie } (M_i)_{i\in I}),$$

$$\bigcup_{i\in I} M_i := \{\, a \mid \exists i{\in}I\colon a \in M_i \,\} \quad (\textit{Vereinigung} \text{ der Familie } (M_i)_{i\in I}).$$

Sei M eine Menge. Jeder Teilmenge $X \subseteq M$ kann man eine Abbildung $c_X\colon M \to \{0,1\}$ zuordnen:

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

c_X heißt die *charakteristische Funktion* der Menge X. Offenbar enthält c_X dieselbe Information wie X; man kann deshalb die Menge X mit ihrer charakteristischen Funktion identifizieren.

Die Menge aller Teilmengen einer gegebenen Menge M bezeichnet man mit

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

Bemerkung. Für beliebige Mengen M, N gilt offenbar

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

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, Äquivalenzrelationen. *Ganze Zahlen* kann man darstellen als Paare (m, n) natürlicher Zahlen. Gemeint ist mit (m, n) die Differenz $m - n$. Dann sind zum Beispiel $(m + 1, n + 1)$ und (m, n) als gleich zu betrachten. Allgemein definieren wir eine Relation $\sim_{\mathbb{Z}}$ zwischen Paaren natürlicher Zahlen durch

$$\forall m,n,p,q{\in}\mathbb{N}\colon \big((m,n) \sim_{\mathbb{Z}} (p,q) :\equiv m + q = p + n\big).$$

Es gilt dann:

$\forall m{\in}\mathbb{N}\colon (m,n) \sim_{\mathbb{Z}} (m,n)$,

$\forall m,n,p,q{\in}\mathbb{N}\colon \big((m,n) \sim_{\mathbb{Z}} (p,q) \to (p,q) \sim_{\mathbb{Z}} (m,n)\big)$,

$\forall m,n,p,q,k,l{\in}\mathbb{N}\colon \big((m,n) \sim_{\mathbb{Z}} (p,q) \wedge (p,q) \sim_{\mathbb{Z}} (k,l) \to (m,n) \sim_{\mathbb{Z}} (k,l)\big)$.

Die ersten beiden Aussagen sind klar. Zum Beweis der dritten betrachten wir beliebige $m, n, p, q, k, l \in \mathbb{N}$. 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 $m+l=k+n$. Damit ist der Beweis abgeschlossen.

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

$$\begin{aligned}
&\forall a{\in}M\colon a\sim a && (\sim \text{ ist } \textit{reflexiv}),\\
&\forall a,b{\in}M\colon \big(a\sim b\to b\sim a\big) && (\sim \text{ ist } \textit{symmetrisch}),\\
&\forall a,b,c{\in}M\colon \big(a\sim b\wedge b\sim c\to a\sim c\big) && (\sim \text{ ist } \textit{transitiv}).
\end{aligned}$$

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 „Äquivalenzklasse“ hat sich eingebürgert).

Sei M eine Menge und $\sim$ eine Äquivalenzrelation auf M. Eine nicht leere Teilmenge N von M heißt eine *Äquivalenzklasse* der Relation $\sim$, wenn gilt:

($\sim$1) $\forall a,b{\in}M\colon \big(a\in N\wedge b\sim a\to b\in N\big)$ (N ist $\sim$-*abgeschlossen*),
($\sim$2) $\forall a,b{\in}M\colon \big(a,b\in N\to a\sim b\big)$.

BEMERKUNGEN. (1) Jedes $a\in M$ liegt in genau einer Äquivalenzklasse; sie wird mit $[a]$ bezeichnet.

BEWEIS. 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 ($\sim$1). 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 ($\sim$2). 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 N_1,N_2 Äquivalenzklassen mit $a\in N_1$ und $a\in N_2$. Zu zeigen ist $N_1=N_2$; aus Symmetriegründen genügt $N_1\subseteq N_2$. Sei also $c\in N_1$. Zu zeigen ist $c\in N_2$. Wegen $c,a\in N_1$ folgt $c\sim a$ nach ($\sim$2) für N_1, also auch $c\in N_2$ nach ($\sim$1) für N_2. □

(2) Je zwei verschiedene Äquivalenzklassen sind disjunkt.

BEWEIS. Zu zeigen ist, daß aus $a\in N_1,N_2$ folgt $N_1=N_2$. Das ist aber gerade der Eindeutigkeitsteil der Aussage (1). □

Eine Äquivalenzrelation auf M zerlegt also die „Trägermenge“ M vollständig in paarweise disjunkte Äquivalenzklassen. Die Menge aller Äquivalenzklassen wird mit $M/\!\sim$ bezeichnet, also $M/\!\sim:=\{\,[a]\mid a\in M\,\}$.

Wir lassen im folgenden die Allquantoren am Anfang einer Aussage meist weg, schreiben also z.B. $a\sim a$ statt $\forall a{\in}M\colon a\sim a$.

1.2.2. Addition ganzer Zahlen, Verträglichkeit mit $\sim_{\mathbb{Z}}$. Wir definieren die *Addition ganzer Zahlen* durch

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

(Hier haben wir die Allquantoren am Anfang weggelassen). Im folgenden schreiben wir kurz $\sim$ für die oben definierte Äquivalenzrelation $\sim_{\mathbb{Z}}$ auf $\mathbb{N}\times\mathbb{N}$. Die von $\sim$ erzeugten Äquivalenzklassen bezeichnen wir mit α, β, γ.

Die Addition ist verträglich mit $\sim$, 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

$$\begin{aligned} &\text{aus } [a]=[a'] \text{ und } [b]=[b'] \text{ folgt } [a+b]=[a'+b'], \text{ und} \\ &\text{aus } [a]=[a'] \text{ folgt } [-a]=[-a']. \end{aligned}$$

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

$$\begin{aligned} &\text{aus } a \sim a' \text{ und } b \sim b' \text{ folgt } a+b \sim a'+b', \text{ und} \\ &\text{aus } a \sim a' \text{ folgt } -a \sim -a'. \end{aligned}$$

Das ist 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. Gruppen. 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, der an vielen Stellen in der Mathematik auftaucht.

DEFINITION. Seien G eine Menge und $\circ\colon G \times G \to G$ eine Abbildung. $(G,\circ)$ (oder oft nur kurz G) 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 $e \in G$ (genannt *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 kurz $\mathbb{Z}$) der *ganzen Zahlen* konstruiert.

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

Beweis. Gezeigt wird

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

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 $(\alpha_1 \circ \alpha_2 \circ \cdots \circ \alpha_i) \circ (\alpha_{i+1} \circ \cdots \circ \alpha_n \circ \alpha_{n+1})$ mit $1 \leq i < n$ (dies kann man oBdA annehmen; andernfalls folgt die Behauptung sofort aus der Induktionsvoraussetzung). Dann gilt

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

Damit ist der Beweis abgeschlossen. □

Bemerkungen. Sei G eine Gruppe.

(1) Seien $e \in G$ ein neutrales Element und $\alpha, \alpha' \in G$ derart, daß $\alpha' \circ \alpha = e$. Dann gilt auch $\alpha \circ \alpha' = e$.

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

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

   Das war zu zeigen. □

(2) Sei $e \in G$ ein neutrales Element. Dann gilt $\alpha \circ e = \alpha$ für alle $\alpha \in G$.

   Beweis. $\alpha \circ e = \alpha \circ \alpha' \circ \alpha = e \circ \alpha = \alpha$ nach (1). □

(3) Es gibt genau ein neutrales Element $e \in G$.

   Beweis. Seien e, e^* neutrale Elemente von G. Dann gilt $e^* = e \circ e^* = e$ nach (2). □

(4) Zu jedem $\alpha \in G$ gibt es genau ein inverses Element $\alpha' \in G$.

BEWEIS. Seien α', α^* inverse Elemente zu $\alpha \in G$. Dann gilt

$$\alpha^* = \alpha^* \circ e = \alpha^* \circ \alpha \circ \alpha' = e \circ \alpha = \alpha'.$$

Hierbei haben wir (1) - (3) benutzt. □

Das zu α eindeutig bestimmte inverse Element wird mit α^{-1} bezeichnet. Wir schreiben α^n für $\alpha \circ \cdots \circ \alpha$ und α^{-n} für $\alpha^{-1} \circ \cdots \circ \alpha^{-1}$, jeweils mit n Vorkommen von α bzw. α^{-1}. Ferner sei $\alpha^0 := e$.

LEMMA. *Seien 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) $\forall \alpha, \beta \in G \exists \gamma \in G\colon \alpha \circ \gamma = \beta$.
(b) $\forall \alpha, \beta \in G \exists \delta \in G\colon \delta \circ \alpha = \beta$.

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

$\Leftarrow$. G erfülle (1) und (2). Zu zeigen ist, daß G eine Gruppe ist. Wir müssen also ein neutrales Element e finden derart daß gilt

$$e \circ \alpha = \alpha \text{ für alle } \alpha \in G, \tag{1.2}$$

$$\forall \alpha \in G \exists \alpha' \in G\colon \alpha' \circ \alpha = e \tag{1.3}$$

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. Wir zeigen jetzt (1.2). Sei $\alpha \in G$ beliebig. Wähle $\gamma \in G$ mit $\alpha_0 \circ \gamma = \alpha$. Dann gilt

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

(1.3) folgt aus (2b). □

BEMERKUNG. Seien G eine Gruppe und $\alpha, \beta \in G$. Dann gilt

$$(\alpha^{-1})^{-1} = \alpha \tag{1.4}$$

$$(\alpha \circ \beta)^{-1} = \beta^{-1} \circ \alpha^{-1} \tag{1.5}$$

BEWEIS. Zu (1.4). Es gilt $(\alpha^{-1})^{-1} \circ \alpha^{-1} = e$ und auch $\alpha \circ \alpha^{-1} = e$. Die Behauptung folgt aus der Eindeutigkeit des Inversen.

Zu (1.5). $\beta^{-1} \circ \alpha^{-1} \circ \alpha \circ \beta = \beta^{-1} \circ e \circ \beta = e$. □

BEISPIEL. Seien 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. $(S(M), \circ)$ heißt 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 abelsch ist. Sei zum Beispiel

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

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

Dann gilt $(g \circ f)(0) = 1$, aber $(f \circ g)(0) = 2$, also $g \circ f \neq f \circ g$.

Im Spezialfall $M = \{1, \dots, n\}$ schreibt man S_n statt $S(M)$. Jede Abbildung $\sigma \in S_n$ heißt eine *Permutation* der Zahlen $1, \dots, n$.

1.2.4. Rationale Zahlen. *Rationale Zahlen* stellen wir dar als Paare $((m,n),k)$ einer ganzen Zahl (m,n) und einer positiven 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

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

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

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

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

$$\begin{aligned} &((m,n),k) \cdot \big(((p,q),l) + ((p',q'),l')\big) = \\ &\big(((m,n),k) \cdot ((p,q),l)\big) + \big(((m,n),k) \cdot ((p',q'),l')\big). \end{aligned}$$

1.2.5. Körper. 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. Seien K eine Menge und $+\colon K \times K \to K$, $\cdot\colon K \times K \to K$ Abbildungen. Das Tripel $(K, +, \cdot)$ (oder kurz K) heißt *Körper*, wenn gilt:

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

$$\alpha(\beta+\gamma) = \alpha\beta + \alpha\gamma \text{ und } (\alpha+\beta)\gamma = \alpha\gamma + \beta\gamma \quad (\textit{Distributivgesetze}).$$

Wir schreiben $\alpha + \beta$ für $+(\alpha,\beta)$ und $\alpha\beta$ oder $\alpha \cdot \beta$ für $\cdot(\alpha,\beta)$. Ferner soll die Multiplikation $\cdot$ stärker binden als $+$.

Beispiel. In 1.2.4 haben wir den Körper $(\mathbb{Q},+,\cdot)$ (oder kurz $\mathbb{Q}$) der *rationalen Zahlen* konstruiert.

Bezeichnungen. (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. Für $\alpha \cdot \beta^{-1}$ schreiben wir $\frac{\alpha}{\beta}$.

(3) Wir schreiben

$$\sum_{i=1}^{n} \alpha_i \quad \text{für } \alpha_1 + \alpha_2 + \cdots + \alpha_n,$$

$$\prod_{i=1}^{n} \alpha_i \quad \text{für } \alpha_1 \cdot \alpha_2 \cdot \ldots \cdot \alpha_n,$$

$$n\alpha \quad \text{für } \alpha + \cdots + \alpha \text{ (mit } n \text{ Vorkommen von } \alpha),$$

$$\alpha^n \quad \text{für } \alpha \cdot \ldots \cdot \alpha \text{ (mit } n \text{ Vorkommen von } \alpha).$$

Im Fall $n = 0$ setzen wir $\sum_{i=1}^{0} \alpha_i := 0$, $\prod_{i=1}^{n} \alpha_i := 1$, $0\alpha := 0$ und $\alpha^0 := 1$.

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

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

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

(1) $0 \cdot \alpha = 0$ und $\alpha \cdot 0 = 0$.

Beweis. Z.B. ist $0\cdot\alpha = (0+0)\cdot\alpha = 0\cdot\alpha + 0\cdot\alpha$, also $0\cdot\alpha = 0$. □

(2) Aus $\alpha \neq 0$ und $\beta \neq 0$ folgt $\alpha\beta \neq 0$.

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

(3) $-\alpha = (-1)\alpha$.

Beweis. Nach (1) ist $1\alpha = \alpha$ für alle α. Wegen $1 + (-1) = 0$ folgt $\alpha + (-1)\alpha = 0$ und damit $-\alpha = (-1)\alpha$. □

(4) $(-\alpha)\cdot(-\beta) = \alpha\beta$.

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

BEISPIELE. (1) *Reelle Zahlen* kann man auffassen als Äquivalenzklassen von *Cauchyfolgen* (s. Analysis-Vorlesung) rationaler Zahlen, also von Folgen $(\alpha_n)_{n\in\mathbb{N}}$ mit $\alpha_n \in \mathbb{Q}$ derart daß

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

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

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

Addition und Multiplikation definiert man komponentenweise. Man kann leicht zeigen, daß Addition und Multiplikation mit der Äquivalenzrelation $\sim_{\mathbb{R}}$ 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 für beliebige Körper beweisen.

(3) Seien K ein Körper und $\xi \in K$ kein Quadrat in K (d.h. es gibt kein $\alpha \in K$ mit $\alpha^2 = \xi$). Wir wollen einen Körper $K \times K$ definieren. Stellt man sich (außerhalb der Legalität) (α, β) als $\alpha + \beta\sqrt{\xi}$ vor, so liegt die folgende Definition nahe:

$$\begin{aligned}(\alpha, \beta) + (\alpha', \beta') &:= (\alpha + \alpha', \beta + \beta'),\\ (\alpha, \beta) \cdot (\alpha', \beta') &:= (\alpha\alpha' + \beta\beta'\xi, \alpha\beta' + \beta\alpha').\end{aligned}$$

Wir wollen uns überlegen, daß $K \times K$ mit dieser Addition $+$ und Multiplikation $\cdot$ zu einem Körper wird. Man bezeichnet ihn mit $(K(\sqrt{\xi}), +, \cdot)$ oder kurz $K(\sqrt{\xi})$.

BEWEIS. 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 $(-\alpha, -\beta)$.

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. Zu (a). Es gilt

$(1,0)\cdot(\alpha,\beta)=(\alpha,\beta)$. Zu (b). Gegeben sei $(\alpha,\beta)\neq(0,0)$. Gesucht ist (α',β') mit

$$(\alpha',\beta')(\alpha,\beta)=(1,0) \tag{1.6}$$

Ansatz: Nehmen wir an, wir hätten schon (α',β') mit (1.6). 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 (1.6).

Die Distributivität ergibt sich durch Nachrechnen. □

Zwei interessante Spezialfälle wollen wir besonders betrachten. Seien $K=\mathbb{R}$ 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

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

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

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

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 $\xi=2$.

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

BEWEIS. Andernfalls wäre $2 = \frac{m^2}{n^2}$ 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 ein gemeinsamer Teiler von m und n: Widerspruch. □

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 vereinfachen.

1.3.1. Definition von Matrizen, Schreibweisen.

DEFINITION. Seien 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. Mit $K^{m \times n}$ bezeichnen wir die Menge aller $(m \times n)$-Matrizen 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:

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

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

1.3.2. Algebraische Operationen mit Matrizen.

DEFINITION. Seien 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 $A \cdot C$.

Die Addition von Matrizen und die Multiplikation einer Matrix mit einem Körperelement sind „komponentenweise“ definiert und leicht zu verstehen. Dagegen scheint die Multiplikation von Matrizen willkürlich definiert zu sein; wir werden jedoch (in 1.3.4) sehen, daß dieser Eindruck täuscht.

BEISPIELE. (1) Es ist oft hilfreich, sich die Matrizenmultiplikation graphisch vorzustellen:

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

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

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

1.3.3. Rechenregeln für Matrizen.

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

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

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

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

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

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

Das war zu zeigen. □

(3) *Neutrales Element* bei der Matrizenmultiplikation. Für $m \geq 1$ sei die *Einheitsmatrix* definiert durch

$$E := E_m := (\delta_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq m}} \quad \text{mit} \quad \delta_{ij} := \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j, \end{cases}$$

also

$$E = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}.$$

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

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

$BE = B$ zeigt man analog. □

(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 Distributivgesetz für Körperelemente.

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

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

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.

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

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

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. □

Wir haben schon gesehen, daß die Komposition von Abbildungen nicht kommutativ ist. Deshalb ist es nicht verwunderlich, daß auch die Matrizenmultiplikation ist i.a. *nicht* kommutativ ist. Beispiel:

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

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

Insbesondere gibt es also Matrizen $A \neq 0$, $B \neq 0$ mit $BA = 0$.

Halten wir also fest: Das Gleichungssystem (1.1) läßt sich in Matrizenschreibweise darstellen als $Ax = b$ mit

$$A = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} \in K^{m\times n},$$

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^{n\times 1}, \quad b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \in K^{m\times 1}.$$

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 Lösungsverfahren 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. Zeilenstufenmatrizen. Eine besondere Rolle bei der allgemeinen Formulierung des Lösungsverfahrens werden Matrizen in Zeilenstufenform spielen.

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

BEISPIEL.

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

sind Zeilenstufenmatrizen, nicht aber

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

DEFINITION. Seien K ein Körper und $A, A' \in K^{m\times n}$. Wir sagen: A' entsteht aus A durch eine (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. Hilfssätze. Wir stellen einige Hilfssätze bereit, die die Essenz des Gaußschen Lösungsverfahrens darstellen. Seien K ein Körper, $A \in K^{m\times n}$ und $b \in K^{m\times 1}$. Wir betrachten die Gleichung (1.1), also $Ax = b$.

HILFSSATZ (Zeilenumformungen verändern 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

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

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

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

Addition beider Zeilen ergibt

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

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

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

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

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

Das war zu zeigen. □

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

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

$$(A\ b) = \begin{pmatrix} 0 & \dots & 0 & \alpha_{1j_1} & \dots & \beta_1 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \alpha_{mj_1} & \dots & \beta_m \end{pmatrix} \mapsto \begin{pmatrix} 0 & \dots & 0 & \alpha'_{1j_1} & \dots & \beta'_1 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \alpha'_{mj_1} & \dots & \beta'_m \end{pmatrix},$$

wobei $\alpha'_{1j_1} \neq 0$. Addiert man zu jeder Zeile mit von Null verschiedenem j_1-ten Element ein Vielfaches der ersten Zeile mit einem jeweils passenden $\lambda \in K$, $\lambda \neq 0$, 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. □

Wir können also oBdA annehmen, daß A eine Zeilenstufenmatrix ist, d.h. $(A\ b)$ die folgende Form hat:

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

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

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

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

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

Hilfssatz (Frei wählbare Komponenten). *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 $j \notin \{j_1, \dots, j_r\}$. 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 $\xi_{j_2}, \dots, \xi_{j_r}$. □

1.4.3. Gaußsches Lösungsverfahren. Wir fassen unsere Überlegungen wie folgt zusammen.

SATZ. *Seien K ein Körper, $A \in K^{m\times n}$ und $b \in K^{m\times 1}$. Man betrachte das lineare Gleichungssystem $Ax = b$. Dann kann man durch Umformung von $(A\ b)$ in Zeilenstufenform entscheiden, ob $Ax = b$ lösbar ist.*

Wir wollen uns jetzt noch im Fall der Lösbarkeit eine Übersicht über die Gesamtheit aller Lösungen verschaffen.

SATZ. *Seien K ein Körper, $A \in K^{m\times n}$ und $b \in K^{m\times 1}$. Wir setzen voraus, daß das lineare Gleichungssystem $Ax = b$ lösbar ist. Dann findet man eine spezielle Lösung $x_0 \in K^{n\times 1}$ und weiter ein $r \in \mathbb{N}$ mit $0 \le r \le n$ und $x_1, \dots, x_{n-r} \in K^{n\times 1}$ so, daß*

(1) $X := \{\, x_0 + \lambda_1 x_1 + \cdots + \lambda_{n-r} x_{n-r} \mid \lambda_1, \dots, \lambda_{n-r} \in K \,\}$ die Menge der Lösungen von $Ax = b$ ist, und
(2) die Darstellung der Lösungen in X eindeutig ist.

BEWEIS. Der Beweis des Satzes wird in fünf Schritten geführt.

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

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

gemäß dem dritten Hilfssatz, so daß $Ax_0 = b$.

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

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

für $1 \le j \le n$, $j \notin \{j_1, \dots, j_r\}$ mit $Ax'_j = 0$ und bilden daraus $x_1, \dots, x_{n-r}$ durch Umnumerierung. Wir setzen für $1 \le l \le n$, $l \notin \{j_1, \dots, j_r\}$

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

und konstruieren daraus x'_j gemäß dem dritten Hilfssatz so daß $Ax'_j = 0$.

Schritt 3. ($X \subseteq$ Lösungsmenge). Seien $\lambda_1, \dots, \lambda_{n-r} \in K$, $x = x_0 + \sum_{j=1}^{n-r} \lambda_j x_j$. 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 $\lambda'_1, \dots, \lambda'_{n-r}$ existieren mit

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

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 des dritten Hilfssatzes zu zeigen, daß für $1 \leq l \leq n$ mit $l \notin \{j_1, \dots, j_r\}$ gilt $\lambda_l = \mu_l$. Für $1 \leq j, l \leq n$ mit $j, l \notin \{j_1, \dots, j_r\}$ gilt aber $\xi'_{jl} = \delta_{jl}$, also

$$\mu_l = \sum_{\substack{j=1,\dots,n \\ j \neq j_1,\dots,j_r}} \lambda_j \xi'_{jl} = \lambda_l.$$

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

$$0 = \sum_{\substack{j=1,\dots,n \\ j \neq j_1,\dots,j_r}} (\lambda_j - \mu_j) \xi'_{jl} = \lambda_l - \mu_l.$$

Damit ist der Beweis abgeschlossen. □

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

$$\left(\begin{array}{ccc|c} 0 & 2 & 3 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 2 \end{array}\right) \begin{array}{r} \\ \mapsto (-2)\cdot Z_2 + Z_1 \\ (-1)\cdot Z_3 + Z_1 \\ \\ \mapsto (-1)\cdot Z_3 + Z_2 \end{array} \begin{array}{c} \\ \left(\begin{array}{ccc|c} 0 & 2 & 3 & 1 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 3 & -1 \end{array}\right) \\ \\ \left(\begin{array}{ccc|c} 0 & 2 & 3 & 1 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right) \end{array}$$

($Z_i = i$-te Zeile). 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

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

1.4.4. Folgerungen. Aus dem Gaußschen Eliminationsverfahren ergeben sich leicht einige interessante Folgerungen. Zunächst sieht man leicht, daß im Fall von mehr Unbekannten als Gleichungen stets eine nichttriviale Lösung von $Ax = 0$ existiert.

BEHAUPTUNG. *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. □

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

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

Dann nennt man

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

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

$$(A+B)^t = A^t + B^t \quad \text{und} \quad (\lambda A)^t = \lambda A^t.$$

Beweis. Dies ist klar, da die Matrizenaddition und die Multiplikation einer Matrix mit einem Körperelement komponentenweise definiert sind. □

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

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

Beweis. Sei

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

Dann ist

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

Es folgt

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

Also ist $(AB)^t = B^tA^t$. □

1.4.5. Invertierbare Matrizen. Als weitere Folgerung gewinnen wir aus dem Gaußschen Eliminationsverfahren ein Kriterium, wann und wie sich Matrizen „invertieren“ lassen.

Definition. $A \in K^{n\times n}$ heißt *invertierbar*, wenn es ein $B \in K^{n\times n}$ gibt mit $BA = E$.

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

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:

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

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

Beweis. (1) $\Rightarrow$ (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) $\Rightarrow$ (3). Nach Annahme hat man

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

mit $A' = E$. Seien $e_1, \ldots, e_n$ die Spalten von E und $e'_1, \ldots, e'_n$ die Spalten von E'. Dann gilt nach dem ersten Hilfssatz für $1 \leq i \leq n$

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

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

(3) $\Rightarrow$ (4). Gelte $yA = 0$. Dann ist $y = yE = yAX = 0$.

(4) $\Rightarrow$ (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) $\Rightarrow$ (2) schon bewiesen ist, folgt (5).

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

(6) $\Rightarrow$ (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) $\Rightarrow$ (3). □

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

Bemerkungen. (1) Gilt $A^{-1}A = E$, so ist nach obigem Korollar auch $AA^{-1} = E$.

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

BEWEIS. Man definiere $(AB)^{-1} := B^{-1}A^{-1}$. Dann ergibt sich $AB(AB)^{-1} = ABB^{-1}A^{-1} = E$. □

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

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

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. Definition von Vektorräumen

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 gegebenen 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 $(\mathbb{R}^{m\times n}, +)$ ist eine). Das neutrale Element ist der Nullvektor

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

und

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

(2) Es gilt

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

2.1.2. Vektorräume. Für viele geometrische Untersuchungen kommt es nur auf die obigen Eigenschaften (1) und (2) des $\mathbb{R}^n$ an. Sie geben Anlaß zur folgenden allgemeinen Definition eines Vektorraums.

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 gilt:

(1) $(V,+)$ ist eine abelsche Gruppe ist.
(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 von V nennt man *Vektoren*, und die Elemente von K heißen *Skalare*.

Wir schreiben wie üblich $x-y$ für $x+(-y)$ und nx für $x+x+\cdots+x$ (mit n Vorkommen von x)

BEISPIELE. Sei K ein Körper.

(1) $V=K^n$ $(:=K^{n\times 1})$ bildet mit folgender Addition und Skalarmultiplikation einen K-Vektorraum:

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

(2) $V = K[t]$, also die Menge aller Polynome in einer Variablen t mit Koeffizienten aus K, bildet mit folgender Addition und Skalarmultiplikation einen K-Vektorraum:

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

Polynome werden hier als formale Ausdrücke (nicht als Funktionen) aufgefaßt, d.h. als Folgen $(\alpha_i)_{i\in\mathbb{N}}$ aus Köperelementen, von denen nur endlich viele ungleich Null sind. Das neutrale Element 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) Seien 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

$$\begin{aligned}&f + g\colon I \to V && \alpha f\colon I \to V\\ &(f+g)(i) := f(i) + g(i)' && (\alpha f)(i) := \alpha \cdot f(i).\end{aligned}$$

(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 punktweise 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 $x + (-1)x = 1x + (-1)x = (1 + (-1))x = 0x = 0$. □

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 $\alpha, \beta \in \mathbb{R}$ gilt:

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

$x, y \in \mathbb{R}^3$ heißen linear abhängig, wenn das Gegenteil 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) $\exists \beta \in \mathbb{R}\colon (x = \beta y)$ oder $\exists \alpha \in \mathbb{R}\colon (y = \alpha x)$.

BEWEIS. (1) $\Rightarrow$ (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) $\Rightarrow$ (1). Sei etwa $x = \beta y$. Dann ist $1x - \beta y = 0$. □

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 $x, y, z \in \mathbb{R}^3$ folgende Aussagen äquivalent sind.

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

2.2.2. Lineare Unabhängigkeit. Wir definieren jetzt allgemein die Begriffe der linearen Unabhängigkeit und der linearen Abhängigkeit für Familien von Vektoren. Wir tun dies in zwei Schritten: zunächst für endliche Familien, und anschließend allgemein.

DEFINITION. Seien 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 $\alpha_1, \ldots, \alpha_n \in K$ gilt:

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

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

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

Die lineare Unabhängigkeit von $(x_1, \dots, x_n)$ ist offenbar unabhängig von der Aufzählung der Vektoren $x_1, \dots, x_n$. Man sagt 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. Seien V ein K-Vektorraum, I eine Indexmenge und $(x_i)_{i\in I}$ eine Familie von Vektoren aus V. Wir nennen $(x_i)_{i\in I}$ *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*.

Offenbar ist jede Teilfamilie einer linear unabhängigen Familie 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 die Vektoren $e_1, \dots, e_n$ linear unabhängig, denn aus

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

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

(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 $\alpha_0 + \alpha_1 t + \cdots + \alpha_n t^n = 0$. Dann ist offenbar $\alpha_0 = \alpha_1 = \cdots = \alpha_n = 0$.

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

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

BEWEIS.

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

Daraus folgt die Behauptung aufgrund der Definitionen. ∎

Als wichtige Folgerung ergibt sich hieraus, daß mehr als n Vektoren im K^n stets linear abhängig sind; dies folgt aus unseren Kenntnissen über das Gauß-Verfahren.

DEFINITION. Seien 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 $x = \alpha_1 x_1 + \cdots + \alpha_n x_n$.

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

LEMMA (Charakterisierung der linearen Abhängigkeit). *Seien 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.

$\Rightarrow$. Gelte $\beta_1 y_1 + \cdots + \beta_m y_m + \alpha_1 x_1 + \cdots + \alpha_n x_n = 0$, 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

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

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

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

Das war zu zeigen. □

2.2.3. Basis, Erzeugendensystem.

DEFINITION. Seien V ein K-Vektorraum, I eine Indexmenge und $(x_i)_{i \in I}$ eine Familie von Vektoren aus V.

(1) $(x_i)_{i \in I}$ heißt *Erzeugendensystem* von V wenn sich jedes $x \in V$ aus einer endlichen Teilfamilie von $(x_i)_{i \in I}$ linear kombinieren läßt.
(2) $(x_i)_{i \in I}$ heißt *Basis* von V, wenn $(x_i)_{i \in I}$ ein linear unabhängiges Erzeugendensystem von V ist.

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

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

□

(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 $\alpha_0+\alpha_1 t+\cdots+\alpha_n t^n \in \mathbb{R}[t]$ Linearkombination von endlich vielen t^m. □

3. Die leere Familie $(x_i)_{i\in\emptyset}$ ist Basis von 0, dem nur aus dem Nullelement bestehenden Vektorraum.

Wir geben jetzt einige nützliche Charakterisierungen des Basisbegriffs.

SATZ. *Seien 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 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. Wir verwenden mehrfach die Charakterisierung der linearen Abhängigkeit in 2.2.2.

(1) ⇒ (2). Seien $x_1, \ldots, x_n$ Basis und $y \in V$. Dann ist y Linearkombination von $x_1, \ldots, x_n$, also $x_1, \ldots, x_n, y$ linear abhängig.

(2) ⇒ (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 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) ⇒ (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) ⇒ (1). Die Behauptung folgt unmittelbar aus der Charakterisierung der linearen Abhängigkeit in 2.2.2. □

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

SATZ (Längenvergleichsatz). *Seien $x_1, \dots, x_n$ linear unabhängig und $y_1, \dots, y_m$ ein Erzeugendensystem von V. Dann ist $n \le m$.*

BEWEIS. Wir nehmen $m < n$ an und leiten daraus einen Widerspruch her. Für $1 \le j \le n$ 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 1.4.4 existieren deshalb $\beta_1, \dots, \beta_n \in K$, die nicht alle verschwinden, mit

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

Also ist

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

Dies widerspricht der linearen Unabhängigkeit von $x_1, \dots, x_n$. □

KOROLLAR. *In einem endlichdimensionalen Vektorraum haben je zwei Basen dieselbe endliche Länge.*

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

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

2.2.5. Basisergänzungssatz. Wir zeigen, daß sich in einem endlichdimensionalen Vektorraum jede linear unabhängige Familie aus einem gegebenen Erzeugendensystem heraus zu einer Basis ergänzen läßt.

SATZ (Basisergänzungssatz). *Seien V ein endlichdimensionaler K-Vektorraum, $x_1, \dots, x_n$ linear unabhängig und $y_1, \dots, y_m$ ein Erzeugendensystem. Dann 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.*

BEWEIS. Offenbar genügt es, folgendes zu zeigen: 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 der Charakterisierung der linearen Abhängigkeit in 2.2.2 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. □

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

2.2.6. Jeder Vektorraum hat eine Basis; Zornsches Lemma. 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. Wie eben folgt daraus, 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*).

Oft nennt man auch das Paar $(M, \leq)$ oder kurz M eine partielle Ordnung.

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 ein *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 eine *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 gilt:

(1) b ist obere Schranke von X.
(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 eine obere Schranke in M. Dann gibt es ein maximales Element in M.*

Beweis. Man benötigt das *Auswahlsaxiom* (es läßt 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. □

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. Die Definitionen übertragen sich 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. *Seien V ein K-Vektorraum, $X \subseteq V$ 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 $\mathcal{T} \subseteq \mathcal{S}$ eine obere Schranke in $\mathcal{S}$ besitzt. Sei also $\mathcal{T} \subseteq \mathcal{S}$ linear geordnet. Wir zeigen, daß dann $\bar{Z} := \bigcup_{Z \in \mathcal{T}} Z$ eine obere Schranke 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 $z_1, \ldots, z_n \in \bar{Z}$ verschieden 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 $Z_0 \subsetneq Z_0 \cup \{y_j\} \subseteq X \cup Y$. Dies widerspricht aber der Maximalität von Z_0. □

2.2.7. Lineare Abhängigkeit in n-dimensionalen Vektorräumen.

SATZ. *Sei V ein n-dimensionaler K-Vektorraum. Dann gilt:*

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

(2) Sei $(x_1, \ldots, x_n)$ eine Familie von Vektoren 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) ⇒ (b) ist klar. (b) ⇒ (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) ⇒ (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 Satz in 2.2.4.

□

SATZ (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.*

BEWEIS. $m \leq n$ gilt nach dem Längenvergleichsatz in 2.2.4. Die restliche Aussage folgt aus dem Basisergänzungssatz in 2.2.5. 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 vermeiden und den Beweis (der restlichen Aussage) durch Induktion nach m führen. *Basis* $m = 0$. Klar. *Schritt* $m-1 \Rightarrow m$. 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

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

Aus der linearen Unabhängigkeit von $y_1, \ldots, y_{m-1}, x_m, \ldots, x_n$ 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 $\mu_1 = \cdots = \mu_{m-1} = \lambda_{m+1} = \cdots = \lambda_n = 0$.

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

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

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

2.3. Untervektorräume, direkte Summen und Quotienten

2.3.1. Motivation. 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 des $\mathbb{R}^3$ 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 für alle $\alpha_1, \beta_1, \alpha_2, \beta_2, \alpha, \beta, \gamma \in \mathbb{R}$ und $x, y \in \mathbb{R}^3$ gilt:

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

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

2.3.2. Untervektorräume.

Definition. Seien 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 ein K-Vektorraum. 0 ($:= \{0\}$) und V sind dann (triviale) Unterräume von V.

BEMERKUNG. Seien 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) Seien $x \in \bigcap_{i\in I} U_i$ und $\alpha \in K$. Dann zeigt man wie eben $\alpha x \in \bigcap_{i\in I} U_i$.

□

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

(1) $\langle M\rangle := \{\, x \in V \mid \exists n\in\mathbb{N}, x_1, \dots, x_n \in M, \alpha_1, \dots, \alpha_n \in K \colon x = \alpha_1 x_1 + \cdots + \alpha_n x_n \,\}$ 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 $I = \{1, \dots, n\}$ schreibt man

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

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

BEMERKUNG. Seien 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) $\sum_{i\in I} U_i$ ist die Menge der $x \in V$ für die es $y_{i_1} \in U_{i_1}, \dots, y_{i_n} \in U_{i_n}$ gibt ($i_1, \dots, i_n \in I$ verschieden) so daß $x = y_{i_1} + \cdots + y_{i_n}$, insbesondere also für $I = \{1, \dots, n\}$

$$U_1 + \cdots + U_n = \{\, y_1 + \cdots + y_n \mid y_1 \in U_1, \dots, y_n \in U_n \,\}.$$

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

(2) Seien $U \subseteq V$ Unterraum und $M \subseteq U$. Zu zeigen ist $\langle M\rangle \subseteq U$. Sei also $x \in \langle M\rangle$, etwa $x = \alpha_1 x_1 + \cdots + \alpha_n x_n$ 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 $x = \alpha_1 x_1 + \cdots + \alpha_n x_n$ 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 $x = y_{i_1} + \cdots + y_{i_n}$ mit $y_{i_j} \in U_{i_j}$. Dann ist auch $x = 1y_{i_1} + \cdots + 1y_{i_n}$ mit $y_{i_j} \in \bigcup_{i\in I} U_i$, also $x \in \langle \bigcup_{i\in I} U_i\rangle$.

□

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

(1) $\dim U \leq n$.

(2) Es gibt 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.

BEWEIS. Es genügt, (2) zu zeigen. Wir überlegen uns zunächst, daß es eine Basis $x_1, \ldots, x_m$ von U gibt. Dies folgt offenbar aus 2.2.6, 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 2.2.7 je $n+1$ Vektoren in V linear abhängig sind. Für ein $m \leq n$ erhält man also eine Basis $x_1, \ldots, x_m$ von U. Nach dem Basisergänzungssatz 2.2.5 läßt sie sich zu einer Basis $x_1, \ldots, x_m, x_{m+1}, \ldots, x_n$ von V ergänzen. □

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

DEFINITION. Seien 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 gilt:

(1) $V = U_1 + \cdots + U_n$ („Summe“).

(2) Hat man $x_1 \in U_1, \ldots, x_n \in U_n$ mit $x_1 + \cdots + x_n = 0$, so gilt $x_1 = \cdots = x_n = 0$. („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. $\Rightarrow$. Die Existenz der Darstellung folgt aus dem ersten Teil der Definition. Die Eindeutigkeit ergibt sich wie folgt. Nehmen wir an, mit $x_i, x_i' \in U_i$ 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 $x_i - x_i' \in U_i$. Aus dem zweiten Teil der Definition folgt jetzt $x_i = x_i'$.

$\Leftarrow$. $V = U_1 + \cdots + U_n$ ist klar. Seien nun $x_1 \in U_1, \ldots, x_n \in U_n$ mit $x_1 + \cdots + x_n = 0$ gegeben. Da auch $0 + \cdots + 0 = 0$ gilt, folgt aus der Eindeutigkeit $x_1 = \cdots = x_n = 0$. □

(2) Seien $x_1, \ldots, x_n$ eine Basis von V und $m \leq n$. Mit $Kx := \langle x \rangle$ 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. Es ist $V = \sum_{i=1}^n Kx_i$, da $x_1, \ldots, x_n$ ein Erzeugendensystem ist. Seien nun $y_i = \alpha_i x_i \in Kx_i$ für $1 \leq i \leq n$ und es gelte $\sum_{i=1}^n \alpha_i x_i = 0$. Dann ist $\alpha_1 = \cdots = \alpha_n = 0$ wegen der linearen Unabhängigkeit von $x_1, \ldots, x_n$, also $y_1 = \cdots = y_n = 0$.

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 $V = U \oplus U'$. $V = U + U'$ ist klar. Seien nun $y \in U$ 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 $\alpha_1 = \cdots = \alpha_n = 0$ wegen der linearen Unabhängigkeit von $x_1, \ldots, x_n$, also $y = y' = 0$. □

(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 auch 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 $\alpha_{m+1} = \cdots = \alpha_n = 0$. □

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

$$(V = U \oplus U') \equiv \begin{cases} V = U + U' \\ U \cap U' = 0 \end{cases}$$

BEWEIS. $\Rightarrow$. $V = U + U'$ gilt nach Definition. 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 $x = 0$.

$\Leftarrow$. $V = U + U'$ ist vorausgesetzt. Seien nun $x \in U$ 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 $x' = 0$. □

(5) V, W seien K-Vektorräume. $V \times W$ ist mit komponentenweiser Addition und Skalarmultiplikation ein K-Vektorraum (nach einem Beispiel in 2.3.2), und es gilt

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

Beweis. Klar. □

2.3.5. Quotienten von Vektorräumen. Seien 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 von der x-Achse verschiedene 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 3.1.3) 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. *Seien 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

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

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. □

SATZ. *Seien 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 $f := g|U'$ 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 $x = y+u$ mit $y \in U'$ und $u \in U$. Also ist

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

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

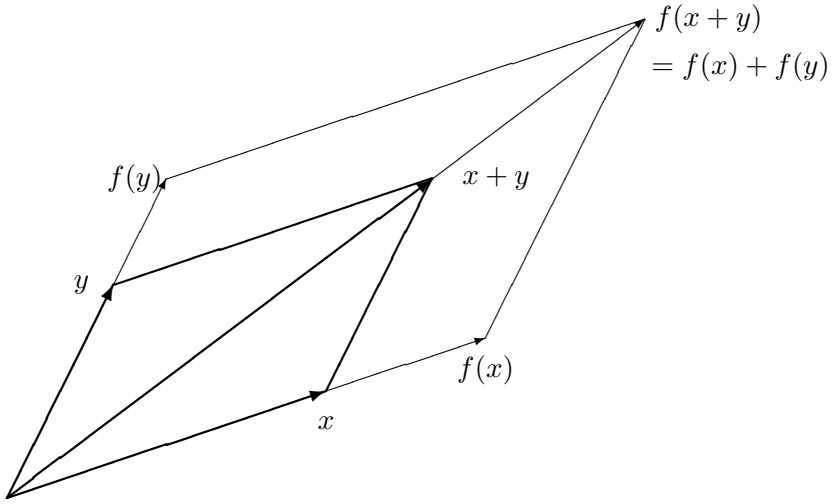
KAPITEL 3

Lineare Abbildungen

3.1. Lineare Abbildungen

3.1.1. Motivation. 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

$$f(x+y) = f(x) + f(y),$$
$$f(\alpha x) = \alpha f(x).$$

Wir abstrahieren hieraus den folgenden allgemeinen Begriff einer linearen Abbildung.

3.1.2. Definition linearer Abbildungen; Kern und Bild.

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 $f(x)+f(-x) = f(x-x) = f(0) = 0$. □

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

(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 $\operatorname{im} f$.

BEWEIS. Sei $V' \subseteq V$ ein Unterraum, also $f[V'] \subseteq W$. Wir zeigen die drei Unterraumeigenschaften aus der Definition von Unterräumen in 2.3.2.

$0 \in f[V']$: 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 $\alpha x \in V'$. □

(4) Bilden $x_1, \ldots, x_n$ ein Erzeugendensystem von V, so bilden auch $f(x_1), \ldots, f(x_n)$ 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 $f(x) = \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n)$. □

(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 $\ker f$.

BEWEIS. Sei $W' \subseteq W$ ein Unterraum, also $f^{-1}[W'] \subseteq V$. Wir zeigen wieder die drei Unterraumeigenschaften.

$0 \in f^{-1}[W']$: 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 jetzt $x, y \in f^{-1}[W']$, 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 $\alpha x \in f^{-1}[W']$. □

3.1.3. Beispiele linearer Abbildungen.

BEISPIELE. (1) Die Abbildungen

$$\begin{aligned} 0\colon V &\to W \\ x &\mapsto 0 \end{aligned} \qquad\qquad \begin{aligned} \mathrm{id}_V\colon V &\to V \\ x &\mapsto x \end{aligned}$$

sind linear, nicht jedoch die Abbildung

$$\begin{aligned} f\colon V &\to W \\ x &\mapsto x_0 \neq 0. \end{aligned}$$

(2) Die Abbildung

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

ist linear, denn es gilt

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

Jede lineare Abbildung $f\colon \mathbb{R} \to \mathbb{R}$ hat diese Form, denn ist $f\colon \mathbb{R} \to \mathbb{R}$ linear, so gilt $f(x) = f(x1) = xf(1)$.

(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 wieder, 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 $1 \leq i \leq m$ und $1 \leq j \leq n$ so daß

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

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

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

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

Das war zu zeigen. □

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 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) Seien V ein K-Vektorraum und $U \subseteq V$ ein Unterraum. Dann ist die im ersten Satz von 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

$$g(\alpha x) = \alpha x + U$$

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

Das war zu zeigen. □

(5) Seien V ein K-Vektorraum, I eine nicht leere Indexmenge und

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

der in 2.1.2, Beispiel (3) eingeführte K-Vektorraum. Sei $h\colon I \to I$ 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. □

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

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

3.1.4. Lineare Fortsetzung. 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

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

f ist linear: $f(x+y) = f(x) + f(y)$: Seien $x = \sum_{i\in I}\alpha_i x_i$ 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.

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

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

(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.

f ist injektiv

$\equiv \forall x, y\in V\colon (f(x) = f(y) \to x = y)$

$\equiv \forall x, y\in V\colon (f(x - y) = 0 \to x - y = 0)$

$\equiv \forall z\in V\colon (f(z) = 0 \to z = 0)$

$\equiv \forall (\alpha_i)_{i\in I}$ mit fast allen $\alpha_i = 0\colon \Big(f(\sum_{i\in I} \alpha_i x_i) = 0 \to \sum_{i\in I} \alpha_i x_i = 0\Big)$

$\equiv \forall (\alpha_i)_{i\in I}$ mit fast allen $\alpha_i = 0\colon \Big(\sum_{i\in I} \alpha_i y_i = 0 \to \forall i\in I\colon \alpha_i = 0\Big)$

$\equiv (y_i)_{i\in I}$ ist linear unabhängig.

Das war zu zeigen. □

KOROLLAR. *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. □

3.1.5. Spezielle lineare Abbildungen. Wir führen zunächst einige oft verwendete Bezeichnungen und Schreibweisen ein. Seien 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) Es seien V, W, U K-Vektorräume und $f\colon V \to W$, $g\colon W \to U$ Vektorraumhomomorphismen. Dann ist auch

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

ein Vektorraumhomomorphismus.

BEWEIS. Es gilt

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

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

(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

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

Das war zu zeigen. ∎

(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 $f, g \in \mathrm{Hom}_K(V,W)$ stets 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 $0 \in \mathrm{Hom}_K(V,W)$. ∎

3.1.6. Isomorphie und Dimension.

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

$V \cong V$.
Wenn $V \cong W$, so ist $W \cong V$.
Wenn $V \cong W$ und $W \cong U$, so ist $V \cong U$.

BEWEIS. Die Behauptung folgt sofort aus den Bemerkungen (1) und (2) in 3.1.5. □

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 dem Korollar in 3.1.4 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 von der linearen Fortsetzung in 3.1.4 ist dann $V \cong K^n$ und ebenso $W \cong K^n$, also $V \cong W$. □

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

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

BEWEIS. Nach 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 3.1.3) folgt aus Teil (1) des vorigen Lemmas die Behauptung. □

3.2. Kern und Bild

Seien V, W K-Vektorräume und $f\colon V \to W$ linear. In den Beispielen (3) und (5) von 3.1.2 hatten wir den *Kern* und das *Bild* von f definiert durch

$$\ker f := \{\, x \in V \mid f(x) = 0 \,\},$$
$$\operatorname{im} f := f[V] \subseteq W.$$

Wir hatten bereits gesehen, daß $\ker f$ ein Unterraum von V und $\operatorname{im} f$ ein Unterraum von W ist.

3.2.1. Charakterisierung von Monomorphismen.

LEMMA. *Seien V, W K-Vektorräume und $f\colon V \to W$ linear. Dann ist f injektiv genau dann, wenn $\ker f = 0$ ist.*

BEWEIS.

$$\begin{aligned} f \text{ injektiv} &\equiv \forall x, y \in V\colon (f(x) = f(y) \to x = y) \\ &\equiv \forall x, y \in V\colon (f(x - y) = 0 \to x - y = 0) \\ &\equiv \forall z \in V\colon (f(z) = 0 \to z = 0) \\ &\equiv (\ker f = 0). \end{aligned}$$

Das war zu zeigen. □

3.2.2. Homomorphiesatz.

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

BEWEIS. Existenz von h. Wir wollen h definieren durch

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

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

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

h ist injektiv: Gelte $f(x) = f(y)$. Zu zeigen ist $x+\ker f = y+\ker f$. Aus Symmetriegründen genügt es, $\subseteq$ zu beweisen. Sei also $z \in \ker f$. Zu zeigen ist $x + z \in y + \ker f$, also $x - y + z \in \ker f$. 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 $h(x + \ker f) = f(x) = w$.

h ist linear:

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

Eindeutigkeit von h. Wegen der Forderung $h \circ g = f$ gilt $h(x+\ker f) = h(g(x)) = f(x)$. Also ist h eindeutig bestimmt. ∎

3.2.3. Dimensionssatz.

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

$$\dim V = \dim \ker f + \dim \operatorname{im} f.$$

BEWEIS. Nach dem Homomorphiesatz ist $V/\ker f \cong \operatorname{im} f$, und nach dem Korollar in 3.1.6 gilt $\dim V = \dim V/\ker f + \dim \ker f$. ∎

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

(1) f ist Monomorphismus.
(2) f ist Epimorphismus.
(3) f ist Automorphismus.

BEWEIS. Nach dem Dimensionssatz ist jede der beiden Aussagen (1) und (2) zu

$$\dim V = \dim \operatorname{im} f$$

äquivalent. Deshalb ist jede von ihnen auch zu (3) äquivalent. □

KOROLLAR. *Seien 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.

$$f\colon U \times U' \to U + U'$$
$$(x, x') \mapsto x + x'.$$

Offenbar ist f linear und surjektiv. Nach dem Dimensionssatz gilt

$$\dim (U \times U') = \dim \ker f + \dim (U + U').$$

Wegen $\dim U \times U' = \dim U + \dim U'$ genügt es zu zeigen, daß $\dim \ker f = \dim U \cap U'$ ist. Dazu zeigen wir, daß die folgende Abbildung g ein Isomorphismus ist.

$$g\colon U \cap U' \to \ker f$$
$$x \mapsto (x, -x).$$

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

3.3. Rang von Matrizen

In Kapitel 1 haben wir unter den $n \times n$-Matrizen diejenigen besonders betrachtet, die invertierbar waren. Wir wollen jetzt diese Unterscheidung verfeinern und den Begriff des Rangs einer $n \times m$-Matrix einführen. – In diesem Abschnitt sei K stets ein Körper.

3.3.1. Zeilenraum und Spaltenraum.

DEFINITION. Seien $A \in K^{m \times n}$ 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

$$\langle a_1, \ldots, a_i, \ldots, a_j, \ldots, a_m \rangle = \langle a_1, \ldots, a_i, \ldots, a_i + a_j, \ldots, a_m \rangle.$$

Daraus folgt die Behauptung. □

(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. □

(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 $a_1, \ldots, a_m$.

BEWEIS. Nach unserer Charakterisierung des Begriffs einer Basis in 2.2.3 ist eine maximale linear unabhängige Teilfamilie von $(a_1, \ldots, a_n)$ eine Basis. □

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

3.3.2. Definition des Rangs. Wir zeigen jetzt, daß Zeilen- und Spaltenrang einer Matrix stets übereinstimmen.

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* $\operatorname{rg}(A)$ *bezeichnet.*

BEWEIS. Wir betrachten die lineare Abbildung

$$f_A \colon K^n \to K^m$$
$$x \mapsto Ax.$$

Schritt 1. $\operatorname{im} f_A$ ist der Spaltenraum von A, denn für die kanonische Basis $e_1, \ldots, e_n$ des K^n ist $\operatorname{im} f_A = \langle f_A(e_1), \ldots, f_A(e_n)\rangle$ nach dem Satz über die lineare Fortsetzung in 3.1.4, und $f_A(ne_j) = Ae_j$ ist die j-te Spalte von A.

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

Schritt 3. Nach dem Dimensionssatz ist

$$n = \dim K^n = \dim \ker f_A + \dim \operatorname{im} f_A = n - r + \dim \operatorname{im} f_A,$$

und $\dim \operatorname{im} f_A$ ist der Spaltenrang von A. □

BEMERKUNG. Offenbar ist $\operatorname{rg}(A) = \operatorname{rg}(A^t)$.

3.3.3. Eigenschaften des Rangs.

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

BEWEIS. Es gilt

$\operatorname{rg}(A) = n$
$\equiv \dim \operatorname{im} f_A = n$ nach dem Beweis des vorigen Satzes
$\equiv f_A$ ist surjektiv

$\equiv f_A$ ist injektiv — nach dem Korollar aus 3.2.3

$\equiv \forall x{\in}K^{n\times 1}(Ax = 0 \to x = 0)$ — nach 3.2.1

$\equiv A$ ist invertierbar — nach 1.4.5.

Das war zu zeigen. □

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

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

(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

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

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

$$\operatorname{rg}(A + B) \le \operatorname{rg}(A) + \operatorname{rg}(B).$$

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 $\operatorname{im} f_{AB} \subseteq \operatorname{im} f_A$, also $\operatorname{rg}(AB) \le \operatorname{rg}(A)$. Ferner gilt allgemein für eine lineare Abbildung $g\colon V \to W$, daß $\dim g[V] \le \dim V$; dies folgt aus dem Dimensionssatz 3.2.3. Speziell für $g :=$ Einschränkung von f_A auf $\operatorname{im} f_B$ folgt

$$\dim \underbrace{f_A(\operatorname{im} f_B)}_{\operatorname{im} f_{AB}} \le \dim \operatorname{im} f_B,$$

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

$$\dim \operatorname{im} f_B = \dim \ker g + \dim \operatorname{im} g.$$

Wir schätzen beide Ausdrücke auf der rechten Seite ab. Wegen $\ker g \subseteq \ker f_A$ ist

$$\dim \ker g \le \dim \ker f_A = n - \dim \operatorname{im} f_A = n - \operatorname{rg}(A).$$

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

$$\operatorname{rg}(B) \le n - \operatorname{rg}(A) + \operatorname{rg}(AB).$$

(2) Nach Teil (1) ist

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

Wieder nach Teil (1) ist

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

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

$$\dim \operatorname{im} f_{A+B} \leq \dim (\operatorname{im} f_A + \operatorname{im} f_B) \leq \dim \operatorname{im} f_A + \dim \operatorname{im} f_B$$

nach einem Korollar aus dem Dimensionssatz in 3.2.3, also $\operatorname{rg}(A+B) \leq \operatorname{rg}(A) + \operatorname{rg}(B)$.

□

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

BEWEIS. Dies ist nur eine Umformulierung der Ergebnisse von 1.4. □

3.4. Lineare Abbildungen und Matrizen

In unseren einführenden Betrachtungen zu linearen Abbildungen in 3.1.3 (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

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

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. Darstellende Matrix einer linearen Abbildung.

SATZ. *Seien V, W K-Vektorräume der Dimensionen n, m. 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*

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

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

ein Isomorphismus. $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)$ 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. *Schritt* 1. $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}$ ist wohldefiniert, da $\mathcal{B} = (y_1, \ldots, y_m)$ eine Basis von W ist.

Schritt 2. $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f+g) = \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) + \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(g)$: Seien $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) = (\alpha_{ij})$ und $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(g) = (\beta_{ij})$. 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 $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f+g) = (\alpha_{ij} + \beta_{ij})$. $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(\alpha f) = \alpha \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)$ beweist man entsprechend.

Schritt 3. $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}$ ist surjektiv. Sei $A = (\alpha_{ij}) \in K^{m \times n}$. Nach dem Satz 3.1.4 von der linearen Fortsetzung existiert ein $f \in \operatorname{Hom}(V, W)$ mit $f(x_j) = \sum_{i=1}^{m} \alpha_{ij} y_i$ für $1 \le j \le n$, also $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f) = A$.

Schritt 4. $\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}$ ist injektiv. Es genügt zu zeigen, daß $\ker \mathcal{M}_{\mathcal{A}}^{\mathcal{B}} = 0$. Gelte also $f(x_j) = 0$ für $1 \le j \le n$. 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 $1 \le j \le n$, 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, W K-Vektorräume der Dimensionen n, m, und es seien $\mathcal{A} = (x_1, \ldots, x_n)$ Basis von V und $\mathcal{B} = (y_1, \ldots, y_m)$ Basis von W. Ferner sei $f \in \operatorname{Hom}(V, W)$ und $A = \mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(f)$. $x \in V$ 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

$$\begin{aligned}f(x) = f(\sum_{j=1}^{n} \xi_j x_j) &= \sum_{j=1}^{n} \xi_j f(x_j)\\ = \sum_{j=1}^{n} \xi_j \sum_{i=1}^{m} \alpha_{ij} y_i &= \sum_{i=1}^{m} (\sum_{j=1}^{n} \alpha_{ij} \xi_j) y_i.\end{aligned}$$

3.4.2. Rang 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. *Seien V, W K-Vektorräume der Dimensionen n, m, 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* $\operatorname{rg}(A) = \operatorname{rg}(f) := \dim \operatorname{im} f$. *Man nennt* $\operatorname{rg}(f)$ *den* Rang *von* f.

BEWEIS. *Vorbemerkung* 1. Sei $(x_1, \dots, x_n)$ Basis von V. Dann sind auch Basen von V:

(1) $(x_1, \dots, \alpha x_i, \dots, x_n)$ mit $\alpha \in K$, $\alpha \neq 0$,
(2) $(x_1, \dots, x_j, \dots, x_i, \dots, x_n)$, wobei x_j an der i-ten Stelle und x_i an der j-ten Stelle steht,
(3) $(x_1, \dots, x_i, \dots, x_j + x_i, \dots, x_n)$.

(1) und (2) sieht man sofort, und für (3) zeigt man leicht, daß

$$(x_1, \dots, x_i, \dots, x_j + x_i, \dots, x_n)$$

linear unabhängig ist und ein Erzeugendensystem bildet.

Vorbemerkung 2. A' entstehe aus A durch eine elementare Zeilen- oder 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

$$\begin{pmatrix} 1 & & 0 & 0 & \dots & 0 \\ & \ddots & & & & \\ 0 & & 1 & 0 & \dots & 0 \\ 0 & & 0 & 0 & \dots & 0 \\ \vdots & & & & & \\ 0 & & 0 & 0 & \dots & 0 \end{pmatrix}.$$

$\operatorname{rg}(A)$ bleibt dabei unverändert. OBdA sei also A von dieser Form. Demnach ist zu zeigen: Es gibt Basen

$$\begin{aligned} &(x_1, \dots, x_r, x_{r+1}, \dots, x_n) \quad \text{von } V \text{ und} \\ &(y_1, \dots, y_r, y_{r+1}, \dots, y_m) \quad \text{von } W \end{aligned}$$

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 $\operatorname{rg}(f) = r$.

$\Rightarrow$. $\operatorname{rg}(f) = \dim \operatorname{im} f = \dim \langle y_1, \dots, y_r \rangle = r$.

$\Leftarrow$. Gelte $\operatorname{rg}(f) = r$. Man wähle eine Basis $(x_{s+1}, \dots, x_n)$ von $\ker f$, und ergänze sie zu einer Basis $(x_1, \dots, x_s, x_{s+1}, \dots, x_n)$ von V. Dann gilt

$$\langle x_1, \dots, x_s \rangle \cong V/\ker f \cong \operatorname{im} f$$

und damit $r = s$. Ferner ist $(f(x_1), \dots, f(x_r))$ Basis von $\operatorname{im} f$, denn wegen $f(x_{r+1}) = \dots = f(x_n) = 0$ ist es ein Erzeugendensystem, und es ist $\dim \operatorname{im} f = r$. 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. ∎

3.4.3. Ähnliche Matrizen. Wir betrachten jetzt den Spezialfall eines Endomorphismus. Seien also V ein n-dimensionaler K-Vektorraum und $\mathcal{A} = (x_1, \ldots, x_n)$ 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 $\mathcal{A} = (x_1, \ldots, x_n)$ 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 erhalten wir nur ein sehr vorläufiges Resultat.

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 $A_3 = DCAC^{-1}D^{-1} = DCA(DC)^{-1}$. □

Satz. *Seien 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. $\Leftarrow$. 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

$$f(x_j) = \sum_{i=1}^{n} \alpha_{ij} x_i \qquad\qquad f(y_j) = \sum_{i=1}^{n} \beta_{ij} y_i$$

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

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = C \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = D \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix};$$

Dies ist möglich, da $x_1, \ldots, x_n$ und $y_1, \ldots, y_n$ Erzeugendensysteme sind. Es ergibt sich

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = CD \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

also $CD = E$ und damit $D = C^{-1}$. Man erhält (wegen der Linearität von f)

$$f\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = Cf\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = CA^t\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = CA^tC^{-1}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

also $B^t = CA^tC^{-1}$, d.h. $A^t \sim B^t$ und damit auch $A \sim B$.

$\Rightarrow$. 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. □

KOROLLAR. *Seien V ein n-dimensionaler K-Vektorraum und f ein Endomorphismus von V. Dann bilden die 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

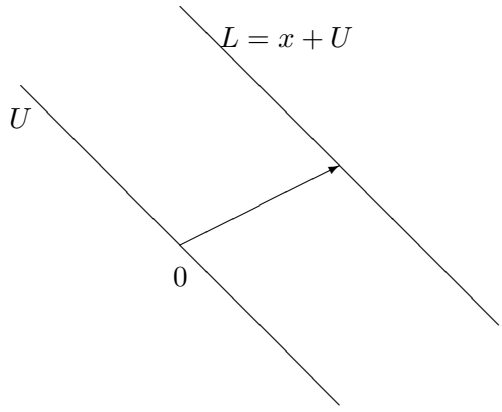
4.1.1. Affine Unterräume. Unter einem affinen Unterraum von V verstehen wir einen um einen Vektor x vom Nullpunkt verschobenen Unterraum.

Definition. Seien 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. Seien $K = \mathbb{R}$ und $V = \mathbb{R}^2$.



Lemma. *Seien 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. $\Rightarrow$. 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 $x - x' = y' \in U'$.

$\Leftarrow$. Seien $U \subseteq U'$ und $x - x' \in U'$. Zu zeigen ist $x + U \subseteq x' + U'$. Sei also $z \in U$. Dann gilt $x + z = x' + (x - x') + z \in x' + U'$. □

KOROLLAR. *Seien V ein K-Vektorraum, ferner $U, U' \subseteq V$ Unterräume und $x, x' \in V$. Dann gilt*

$$x + U = x' + U' \text{ genau dann, wenn } U = U' \text{ und } x - x' \in U.$$

DEFINITION. Seien 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 definieren die affine Hülle einer Teilmenge M von V als den kleinsten M umfassenden affinen Unterraum von V.

DEFINITION. Seien 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. *Seien V ein K-Vektorraum und $M \subseteq V$ eine nicht leere Teilmenge. Dann gilt*

(1) $A(M)$ besteht aus allen $y \in V$, die sich schreiben lassen in der Form

$$y = \sum_{i=1}^{n} \alpha_i y_i \quad \text{mit} \quad \sum_{i=1}^{n} \alpha_i = 1 \quad (y_1, \ldots, y_n \in M,\ \alpha_1, \ldots, \alpha_n \in K.)$$

(2) $A(M)$ 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

$$\begin{aligned} \{ y - x \mid y \in M \} &\subseteq \{ y - x \mid y \in L \} \subseteq U \\ \langle \{ x - y \mid y \in M \} \rangle &\subseteq U \\ A(M) &\subseteq L. \end{aligned}$$

Das war zu zeigen. □

BEMERKUNG. Seien 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) Seien $h\colon K^n \to K^n$ 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. 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

$$h[\{ x \in K^n \mid Ax = b \}] = \{ Cx \mid Ax = b \} = \{ y \mid AC^{-1}y = b \}.$$

Das war zu zeigen. □

(2) Seien 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 $c \in K^m$.

BEWEIS. 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 (1)

$$g[L] = h[f[L]] = h[\{ x \in K^n \mid Ax = b \}] = \{ y \mid AC^{-1}y = b \}.$$

Das war zu zeigen. □

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 $L \neq \emptyset$.

$\Rightarrow$. 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 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.

$\Leftarrow$. Seien $L = x + U$ 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 $x_1, \ldots, x_d$ 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. ∎

4.1.4. 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. Seien $g = T_y \circ f$ 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 $h \circ g = T_{f'(y)+z} \circ (f' \circ f)$. □

(2) Sei V ein K-Vektorraum. Ist $g \colon V \to V$ eine Affinität, so auch g^{-1}.

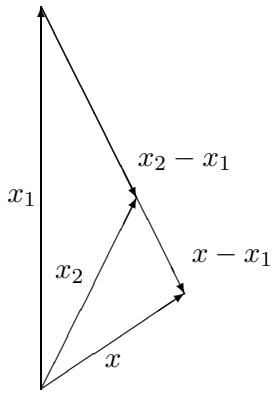
BEWEIS. Sei $g = T_y \circ f$ mit einem Isomorphismus $f \colon V \to V$. Setze $h := T_{f^{-1}(-y)} \circ f^{-1}$. Dann gilt

$$\begin{aligned} (g \circ h)(x) &= g(f^{-1}(x) - f^{-1}(y)) = x - y + y = x \\ (h \circ g)(x) &= h(f(x) + y) = x + f^{-1}(y) - f^{-1}(y) = x, \end{aligned}$$

also $g \circ h = h \circ g = \mathrm{id}_V$, d.h. $h = g^{-1}$. □

(3) Sei V ein K-Vektorraum. Dann ist die Menge $\mathrm{Aff}(V) := \{\, g \colon V \to V \mid g \text{ Affinität} \,\}$ eine Gruppe, die *affine Gruppe* von V. Dies ergibt sich unmittelbar aus (1) und (2).

4.1.5. Teilverhältnis. Seien 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. Seien $g = T_y \circ f$ mit $y \in W$ 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 $\alpha = \mathrm{TV}(x_1, x_2, x) = \mathrm{TV}(g(x_1), g(x_2), g(x))$. □

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.

BEISPIEL. Seien $K = \mathbb{R}$ 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*

$F(x) = \xi_1^2 + \xi_2$ $\qquad\qquad$ $V(F)$ *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

$$x^tAx = \frac{1}{2}(x^tAx + x^tA^tx) = x^t \underbrace{(\frac{1}{2}(A + A^t))}_{B} x.$$

Offenbar ist B symmetrisch. □

(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 $x \in K^n$

$$F(x) = x^tAx + a^tx + \alpha,$$
$$F'(x) = x^tA'x + (a')^tx + \alpha'.$$

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 $1 \le i \le n$ beliebig. Dann ist e_i^tA die i-te Zeile von A und $e_i^tAe_j = \alpha_{ij}$. Es gilt

$$F(e_i) = e_i^tAe_i + a^te_i + \alpha = \alpha_{ii} + \alpha_i + \alpha$$
$$F(2e_i) = 4\alpha_{ii} + 2\alpha_i + \alpha$$

also nach Voraussetzung

$$\alpha_{ii} + \alpha_i + \alpha = \alpha'_{ii} + \alpha'_i + \alpha',$$
$$4\alpha_{ii} + 2\alpha_i + \alpha = 4\alpha'_{ii} + 2\alpha'_i + \alpha'.$$

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 $1 \le i, j \le n$. Es gilt

$$\begin{aligned} &F(e_i + e_j) \\ &= (e_i + e_j)^tA(e_i + e_j) + a^t(e_i + e_j) + \alpha \\ &= e_i^tAe_i + e_j^tAe_i + e_i^tAe_j + e_j^tAe_j + a^te_i + a^te_j + \alpha \\ &= \alpha_{ii} + 2\alpha_{ij} + \alpha_{jj} + \alpha_i + \alpha_j + \alpha \qquad \text{da } A \text{ symm.} \end{aligned}$$

Aus $F(e_i + e_j) = F'(e_i + e_j)$ folgt mit (a) und (b) $2\alpha_{ij} = 2\alpha'_{ij}$, also auch $\alpha_{ij} = \alpha'_{ij}$.

□

(3) (Verhalten bei Affinitäten). Seien $F\colon K^n \to K$ 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

$$\begin{aligned} F(g(x)) &= (Cx+b)^tA(Cx+b) + a^t(Cx+b) + \alpha \\ &= x^tC^tACx + \underbrace{x^tC^tAb}_{b^tACx} + b^tACx + b^tAb + a^tCx + a^tb + \alpha \\ &= x^t \underbrace{C^tAC}_{=:\,A'}x + \underbrace{(2b^tAC + a^tC)}_{=:\,(a')^t}x + \underbrace{b^tAb + a^tb + \alpha}_{=:\,\alpha'}. \end{aligned}$$

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. 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^tAC in* Diagonalgestalt *ist, d.h.*

$$C^tAC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

(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

$$C = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \frac{-\alpha_{i,i+1}}{\alpha_{ii}} & \cdots & \frac{-\alpha_{in}}{\alpha_{ii}} & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 & \end{pmatrix} \in \mathrm{GL}(n, K).$$

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

$$C^t A = \begin{pmatrix} \alpha_{11} & & & & & \\ & \ddots & & & & \\ & & \alpha_{ii} & & \cdots & \alpha_{in} \\ & & 0 & * & \cdots & * \\ & & \vdots & \vdots & & \vdots \\ & & 0 & * & \cdots & * \end{pmatrix}$$

und damit

$$C^t AC = \begin{pmatrix} \alpha_{11} & & & & & \\ & \ddots & & & & \\ & & \alpha_{ii} & 0 & \cdots & 0 \\ & & 0 & * & \cdots & * \\ & & \vdots & \vdots & & \vdots \\ & & 0 & * & \cdots & * \end{pmatrix}.$$

Fall 2. $\alpha_{ii} = 0$, aber $a_{jj} \neq 0$ für ein $j > i$. Setze

$$C = \begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & 0 & & & 1 & & & \\ & & & & 1 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & 1 & & & \\ & & & 1 & & & & 0 & & \\ & & & & & & & & 1 & \\ & & & & & & & & & \ddots \\ & & & & & & & & & & 1 \end{pmatrix} \in \mathrm{GL}(n, K).$$

$C^tAC = CAC$ 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

$$C = \begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & 1 & & & 1 & & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & 1 & & & & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{pmatrix} \in \mathrm{GL}(n, K).$$

Dann gilt $C^tA = CA =$

$$\begin{pmatrix} \alpha_{11} & & & & & & & & \\ & \ddots & & & & & & & \\ & & \alpha_{i-1,i-1} & & & & & & \\ & & & \underbrace{\alpha_{ii}}_{0} + \alpha_{ji} & \dots & \alpha_{ij} + \underbrace{\alpha_{jj}}_{0} & \dots & \alpha_{in} + \alpha_{jn} \\ & & & * & \dots & * & \dots & * \\ & & & \vdots & & & & \vdots \\ & & & * & \dots & * & \dots & * \end{pmatrix},$$

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. ∎

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

$$C^tAC = \begin{pmatrix} \gamma_1 & & & & & \\ & \ddots & & & & \\ & & \gamma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

4.2.3. Klassifikation von Quadratiken.

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 $r \leq n$ in (4.1), (4.2), $r < n$ in (4.3), $\gamma_1, \ldots, \gamma_r \in K$ und $x = (\xi_1, \ldots, \xi_n)^t$. 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. Seien $Q = V(G)$ und $G = x^tAx + a^tx + \alpha$ 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

$$C^tAC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

Dann gilt

$$\begin{aligned} G(Cx) &= x^tC^tACx + a^tCx + \alpha \\ &= \sum_{i=1}^{n} \gamma_i \xi_i^2 + \sum_{i=1}^{n} \beta_i \xi_i + \alpha \\ &=: G_1(x). \end{aligned}$$

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 $r+1 \le i \le n$. 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 $\gamma_i \xi_i^2 = (\lambda_i \xi_i)^2 = (\xi_i')^2$. ∎

BEMERKUNG. Aus dem Beweis ergibt sich noch, daß jede quadratische Funktion F affin äquivalent ist zu einer quadratischen Funktion G, die bis auf einen konstanten Faktor von einer der Formen (4.1) – (4.3) ist.

4.2.4. Klassifikation der quadratischen Hyperflächen; affin. Im $\mathbb{R}^2$ ist jede quadratische Hyperfläche $\neq \emptyset$ affin äquivalent zu einem $V(F_i)$, $i = 1, \ldots, 7$:

Mittelpunktsfl.		
$r = 2$		
	$F_1(x) = -\xi_1^2 - \xi_2^2 + 1$	Ellipse,
	$F_2(x) = -\xi_1^2 + \xi_2^2 + 1$	Hyperbel,
$r = 1$		
	$F_3(x) = -\xi_1^2 + 1$	zwei parallele Geraden
Kegelfl.		
$r = 2$		
	$F_4(x) = \xi_1^2 + \xi_2^2$	Punkt
	$F_5(x) = \xi_1^2 - \xi_2^2$	zwei sich schneidende Geraden
$r = 1$		
	$F_6(x) = \xi_1^2$	Gerade
Parabolische Fl.		
$r = 1$		
	$F_7(x) = \xi_1^2 + \xi_2$	Parabel

Im $\mathbb{R}^3$ ist jede quadratische Hyperfläche $\neq \emptyset$ affin äquivalent zu einem $V(F_i)$, $i = 1, \ldots, 14$:

(1) 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

(2) 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
	$F_{10}(x) = \xi_1^2 - \xi_2^2$	2 sich schneidende Ebenen
$r = 1$		
	$F_{11}(x) = \xi_1^2$	Ebene

(3) 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

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).

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

5.1.1. Motivation. 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 $x = \binom{\xi_1}{\xi_2} \in \mathbb{R}^2$ 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 $y = \binom{\eta_1}{\eta_2}$

$$\begin{aligned}\|x\| \cdot \|y\| \cos\varphi &= \frac{1}{2}\big(\|x + y\|^2 - \|x\|^2 - \|y\|^2\big) \\ &= \frac{1}{2}\big((\xi_1 + \eta_1)^2 + (\xi_2 + \eta_2)^2 - \xi_1^2 - \xi_2^2 - \eta_1^2 - \eta_2^2\big) \\ &= \xi_1\eta_1 + \xi_2\eta_2 \\ &=: \sigma(x, y).\end{aligned}$$

Mit Hilfe des so definierten „Skalarprodukts“ $\sigma\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ kann man also Längen und Winkel messen:

$$\|x\| = \sqrt{\sigma(x, x)},$$
$$\cos\varphi = \frac{\sigma(x, y)}{\|x\| \cdot \|y\|}.$$

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.

5.1.2. Euklidisches Skalarprodukt.

Definition. Seien 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

$$\sigma(\alpha x + \beta y, z) = \alpha\sigma(x, z) + \beta\sigma(y, z),$$
$$\sigma(x, \alpha y + \beta z) = \alpha\sigma(x, y) + \beta\sigma(x, z).$$

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

$$\sigma(x, x) > 0 \quad \text{falls } x \neq 0.$$

(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 leicht, daß damit ein euklidisches Skalarprodukt definiert ist, und nennt es das *kanonische Skalarprodukt* im $\mathbb{R}^n$.

(2) Seien (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).

BEMERKUNGEN. 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) Seien $x_1, \dots, x_n$ 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

$$\sigma(x,y) = (\xi_1, \dots, \xi_n) A \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} \quad \text{mit } A = (\sigma(x_i, x_j)) \in \mathbb{R}^{n \times n}.$$

BEWEIS.

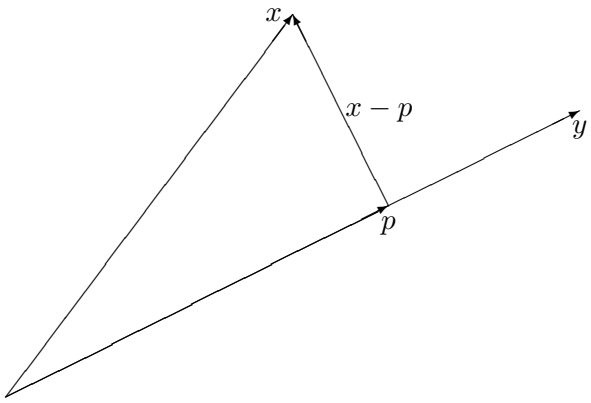
$$\begin{aligned} \sigma(x,y) &= \sigma(\sum_{i=1}^n \xi_i x_i, \sum_{j=1}^n \eta_j x_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n \xi_i \eta_j \sigma(x_i, x_j) \\ &= (\xi_1, \dots, \xi_n) A \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}. \end{aligned}$$

Das war zu zeigen. □

5.1.3. Orthogonale Projektion auf einen Vektor. Seien (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. *Seien (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. 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 $y \neq 0$.

Existenz. Setze $p = \frac{\sigma(x,y)}{\sigma(y,y)}y$. Dann ist $p \in \mathbb{R}y$ und

$$\sigma(x-p,y) = \sigma(x,y) - \frac{\sigma(x,y)}{\sigma(y,y)}\sigma(y,y) = 0.$$

Das war zu zeigen. □

5.1.4. Schwarzsche Ungleichung.

SATZ. *Seien (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 $y \neq 0$. Seien $p = \frac{\sigma(x,y)}{\sigma(y,y)}y$ 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

$$\begin{aligned}
x, y \text{ sind linear abhängig} &\Rightarrow \alpha x + \beta y = 0 \quad \text{mit } \alpha \neq 0 \text{ (da } y \neq 0) \\
&\Rightarrow x = \gamma y \\
&\Rightarrow \sigma(x,y) = \gamma\sigma(y,y).
\end{aligned}$$

Also ist

$$x = \frac{\sigma(x,y)}{\sigma(y,y)}y.$$

(4) Man erhält aufgrund der Schwarzschen Ungleichung

$$\begin{aligned}
\sigma(x+y, x+y) &= \sigma(x,x) + 2\sigma(x,y) + \sigma(y,y) \\
&\leq \sigma(x,x) + 2\|x\| \cdot \|y\| + \sigma(y,y) \\
&= \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2.
\end{aligned}$$

Das war zu zeigen.

□

5.1.5. Winkel. Wir können jetzt Winkel wie folgt erklären. Seien (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. Seien (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) Seien $x, y \neq 0$ 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. $\frac{\|p\|}{\|x\|} = \frac{|\sigma(x,y)|}{\|y\| \cdot \|x\|} = |\cos\varphi|$. □

(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 $\sigma(x,y) = \|x\| \cdot \|y\| \cos\varphi$. □

(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. Seien (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

$$\sigma(e_i, e_j) = 0 \quad \text{für } i, j \in I,\ i \neq j,$$
$$\sigma(e_i, e_i) = 1 \quad \text{für } i \in I.$$

$(e_i)_{i\in I}$ 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 $j \in J$

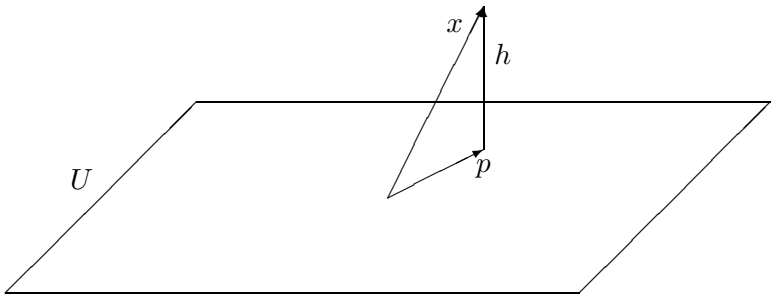
$$0 = \sigma(\sum_{i\in J} \alpha_i e_i, e_j) = \sum_{i\in J} \alpha_i \sigma(e_i, e_j) = \alpha_j.$$

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. Es seien (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).



Satz. *Seien (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. Eindeutigkeit. Annahme: p ist orthogonale Projektion von x auf U. Dann gilt

$$p = \sum_{i=1}^{n} \alpha_i e_i$$
$$x - p \perp e_j \quad \text{für } j = 1, \ldots, n$$
$$\sigma(x - p, e_j) = 0$$
$$\sigma(x, e_j) = \sigma(p, e_j) = \sigma(\sum_{i=1}^{n} \alpha_i e_i, e_j) = \sum_{i=1}^{n} \alpha_i \sigma(e_i, e_j) = \alpha_j.$$

Existenz. Setze $p = \sum_{i=1}^{n} \sigma(x, e_i) e_i$. Dann ist $p \in U$, und es gilt

$$\sigma(x - p, e_j) = \sigma(x, e_j) - \sum_{i=1}^{n} \sigma(x, e_i)\sigma(e_i, e_j) = \sigma(x, e_j) - \sigma(x, e_j) = 0$$
$$\sigma(x - p, u) = 0 \quad \text{für alle } u \in U, \text{ also}$$
$$x - p \perp U.$$

Das war zu zeigen. □

BEMERKUNGEN. (1) $\sigma(x, e_i)$ heißt auch *i-ter Fourierkoeffizient* 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 $u \in U$

$$x - u = x - p + p - u$$
$$\|x - u\|^2 = \|x - p\|^2 + \|p - u\|^2 \quad \text{da } (x - p) \perp (p - u)$$
$$\|x - u\| \geq \|x - p\|.$$

Das war zu zeigen. □

5.2.3. Schmidtsches Orthonormalisierungsverfahren. Wir geben jetzt ein Verfahren an, mit dessen Hilfe man aus einer gegebenen (endlichen oder abzählbar unendlichen) Basis eine Orthonormalbasis herstellen kann.

SATZ. *Seien (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.

Basis $n = 1$. Setze $e_1 := \frac{x_1}{\|x_1\|}$; man beachte hierbei, daß $x_1 \neq 0$. Dann ist $\sigma(e_1, e_1) = 1$.

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 $x_{n+1} \notin U_n = \langle x_1, \ldots, x_n \rangle$ 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 $e_1, \ldots, e_n, e_{n+1}$ Orthonormalsystem. Zu zeigen bleibt $\langle e_1, \ldots, e_{n+1} \rangle = \langle x_1, \ldots, x_{n+1} \rangle$ $(= U_{n+1})$.

$\subseteq$. Nach Induktionsvoraussetzung ist $\langle e_1, \ldots, e_n \rangle = \langle x_1, \ldots, x_n \rangle$ $(= U_n)$. Ferner gilt

$$e_{n+1} = \frac{1}{\|h_{n+1}\|}(x_{n+1} - p_{n+1}) \quad \text{und} \quad p_{n+1} \in U_n.$$

$\supseteq$. Es gilt $x_{n+1} = h_{n+1} + p_{n+1} = \|h_{n+1}\| e_{n+1} + p_{n+1}$ und $p_{n+1} \in U_n = \langle e_1, \ldots, e_n \rangle$. ∎

KOROLLAR. *Jeder endlichdimensionale euklidische Vektorraum besitzt eine Orthonormalbasis.*

BEMERKUNGEN. Seien (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 $e_1, \ldots, e_n$ von U bekannt ist, hat man

$$p = \sum_{i=1}^{n} \sigma(x, e_i) e_i.$$

(2) Falls nur ein Erzeugendensystem $x_1, \ldots, x_m$ von U bekannt ist, setze man

$$p = \sum_{i=1}^{m} \alpha_i x_i$$

und bestimme $\alpha_1, \ldots, \alpha_m$ 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.

$$x - p \perp U$$
$$\sigma(x-p, x_i) = 0 \quad \text{für } i = 1, \ldots, m$$
$$\sigma(x, x_i) = \sigma(p, x_i) = \sum_{j=1}^{m} \alpha_j \sigma(x_j, x_i) \quad \text{für } i = 1, \ldots, m$$
$$b = A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}.$$

Damit ist das Problem auf die Lösung eines linearen Gleichungssystems reduziert. ∎

5.2.4. Orthogonales Komplement. Seien 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) Seien $U \subseteq V$ 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. *Seien (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) $V = U + U^\perp$: Seien $x \in V$ 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
$$\dim V = \dim U + \dim U^\perp$$
$$\dim V = \dim U^\perp + \dim U^{\perp\perp}$$
folgt $\dim U = \dim U^{\perp\perp}$. Mit $U \subseteq U^{\perp\perp}$ ergibt sich die Behauptung. ∎

BEMERKUNG. Teil (2) des Satzes gilt auch im Fall $\dim U < \infty$ und beliebigem V; dies wird hier aber nicht gezeigt.

5.2.5. 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. Seien (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. Seien $V = \mathbb{R}^3$ 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. *Seien (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

$$\begin{aligned} &h = p - p' \quad \text{mit } p \in L,\, p' \in L' && k = q - q' \quad \text{mit } q \in L,\, q' \in L' \\ &h \perp U, U' && k \perp U, U' \end{aligned}$$

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

$$\begin{aligned} &h = \underbrace{(x - q)}_{\in L} - \underbrace{(x' + q')}_{\in L'} \\ &h \perp U, U' \qquad\qquad \text{da } h \perp U + U'. \end{aligned}$$

Das war zu zeigen.

(3) Sei h gemeinsames Lot von L und L', also

$$h = p - p' \quad \text{mit } p \in L,\, p' \in L'$$
$$h \perp U, U'.$$

Seien 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

$$\|q - q'\|^2 = \|h\|^2 + \|y\|^2 \geq \|h\|^2.$$

Das war zu zeigen.

□

5.2.6. Hessesche Normalform für Hyperebenen. Als weitere Anwendung leiten wir eine nützliche Darstellung von Hyperebenen her.

SATZ. *Seien (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) Seien $n \in V$, $\|n\| = 1$ 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.

$$h \in H$$
$$h - h_0 \in U$$
$$\sigma(n, h - h_0) = 0 \qquad \text{da } n \in U^\perp$$
$$\sigma(n, h) = \sigma(n, h_0) =: \beta.$$

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

$$\begin{aligned} d(y, H) &= \min\{\, \|y - h\| \mid h \in H \,\} \\ &= \min\{\, \|y - h_0 - u\| \mid u \in U \,\} \\ &= d(y - h_0, U) \\ &= \text{Länge des Lots von } y - h_0 \text{ auf } U. \end{aligned}$$

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

$$\begin{aligned} |\sigma(n, y) - \beta| &= |\sigma(n, y) - \sigma(n, h_0)| \\ &= |\sigma(n, y - h_0)| \\ &= |\sigma(n, p + (y - h_0 - p))| \\ &= |\sigma(n, p + \lambda n)| \\ &= |\sigma(n, p) + \lambda| \\ &= |\lambda| && \text{da } p \in U \\ &= \text{Länge des Lots von } y - h_0 \text{ auf } U. \end{aligned}$$

Das war zu zeigen. □

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. Definition orthogonaler Abbildungen.

SATZ. *Seien $(e_1, \ldots, e_n)$ 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) $A^t A = E$.

BEWEIS. Die Äquivalenz von (1) und (2) folgt aus

$$\|x\| = \sqrt{\sigma(x,x)},$$
$$\sigma(x,y) = \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2).$$

(2)⇒(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)⇒(2). Seien $x = \sum_{i=1}^n \alpha_i e_i$ 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

$$\begin{aligned}
\sigma(f(e_i), f(e_j)) &= \sigma\Big(\sum_{k=1}^n \alpha_{ki} e_k, \sum_{\ell=1}^n \alpha_{\ell j} e_\ell\Big) \\
&= \sum_{k=1}^n \sum_{\ell=1}^n \alpha_{ki} \alpha_{\ell j} \sigma(e_k, e_\ell) \\
&= \sum_{k=1}^n \alpha_{ki} \alpha_{kj} \qquad\qquad \text{da } \sigma(e_k, e_\ell) = \delta_{k\ell}.
\end{aligned}$$

Also sind folgende Aussagen äquivalent.

$$\begin{aligned}
&f(e_1), \ldots, f(e_n) \text{ ist Orthonormalbasis von } V \\
&\sigma(f(e_i), f(e_j)) = \delta_{ij} \quad \text{für alle } i,j \text{ mit } 1 \le i,j \le n \\
&\sum_{k=1}^n \alpha_{ki}\alpha_{kj} = \delta_{ij} \quad \text{für alle } i,j \text{ mit } 1 \le i,j \le n \\
&A^t A = E.
\end{aligned}$$

Das war zu zeigen. □

DEFINITION. Seien f wie im Satz und $(e_1, \ldots, e_n)$ 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 $A^tA = E$. Man nennt

$$O(n) := \{\, A \in \mathrm{GL}(n,K) \mid A^tA = E \,\}$$

die *orthogonale Gruppe*.

BEMERKUNGEN. (1) Sei $(e_1, \ldots, e_n)$ eine Orthonormalbasis von V. Ein Endomorphismus f von V ist nach dem Satz genau dann orthogonal, wenn die f darstellende Matrix (bzgl. $(e_1, \ldots, e_n)$) orthogonal ist.

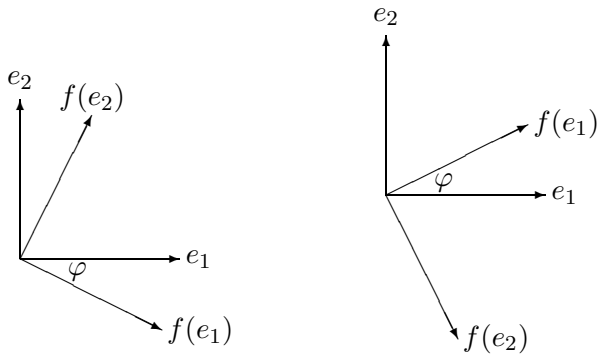
(2) Eine Teilmenge U einer Gruppe $(G, \circ)$ heißt *Untergruppe*, wenn die Gruppenstruktur auf G eine Gruppenstruktur auf U induziert, d.h. wenn gilt (1) $\alpha \circ \beta \in U$ für alle $\alpha, \beta \in U$, (2) $e \in U$ und (3) $\alpha^{-1} \in U$ für alle $\alpha \in U$. – $O(n)$ ist Untergruppe 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 den Affinitäten 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 $\sigma(x,y) = x^ty$.

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*

$$C^t A C = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

(Nicht ausgefüllte Stellen von Matrizen sollen wieder 0 sein).

Den Beweis werden wir im nächsten Abschnitt 5.4 führen.

5.3.3. Klassifikation quadratischer Hyperflächen; euklidisch.

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.*

$$F(x) = \sum_{i=1}^{m} \frac{\xi_i^2}{\gamma_i^2} - \sum_{i=m+1}^{r} \frac{\xi_i^2}{\gamma_i^2} + 1 \qquad \text{Mittelpunktsfläche} \tag{5.1}$$

$$F(x) = \sum_{i=1}^{m} \frac{\xi_i^2}{\gamma_i^2} - \sum_{i=m+1}^{r} \frac{\xi_i^2}{\gamma_i^2} \qquad \text{Kegelfläche} \tag{5.2}$$

$$F(x) = \sum_{i=1}^{m} \frac{\xi_i^2}{\gamma_i^2} - \sum_{i=m+1}^{r} \frac{\xi_i^2}{\gamma_i^2} + \xi_{r+1} \qquad \text{parabolische Fläche.} \tag{5.3}$$

Dabei ist $m \le r \le n$ in (5.1), (5.2), $m \le r < n$ 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.

Seien $Q = V(G)$ und $G = x^t A x + a^t x + \alpha$ 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 5.4 gibt es $C \in O(n)$ und $\gamma_1, \dots, \gamma_n \in \mathbb{R}$ 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 euklidisch äquivalent zu G, da $x \mapsto Cx$ eine Bewegung ist. Wir können wieder annehmen, daß $\gamma_1, \ldots, \gamma_r \neq 0$ und $\gamma_{r+1} = \cdots = \gamma_n = 0$ 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 $b^t := \begin{pmatrix} \beta_{r+1} & \ldots & \beta_n \end{pmatrix}$. OBdA sei $\|b\| = 1$ (sonst dividiere man durch $\|b\|$).

$$\begin{aligned} G_2(x') &= \sum_{i=1}^{r} \gamma_i (\xi_i'')^2 + \xi_{r+1}'' \\ &=: G_3(x'') \quad \text{mit } \xi_i'' := \begin{cases} \xi_i' & \text{für } 1 \leq i \leq r \\ \displaystyle\sum_{i=r+1}^{n} \beta_i \xi_i' + \alpha' & \text{für } i = r+1 \\ \text{beliebig} & \text{für } r+1 < i \leq n. \end{cases} \end{aligned}$$

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). □

Bemerkung. Aus dem Beweis ergibt sich noch, daß jede quadratische Funktion F euklidisch äquivalent ist zu einer quadratischen Funktion G, 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

$$C^t A C = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

(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 $C^t A C = D$ für ein $C \in O(n)$ und eine Diagonalmatrix D, so folgt $A = CDC^t$ und damit die Symmetrie von A.

5.4.1. Jacobische Rotationsmatrix. Die Idee des Jacobischen Verfahrens kann man wie folgt beschreiben. Für $1 \le i < j \le n$ 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 *Jacobische 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 $R_{ij}(\varphi)^t A R_{ij}(\varphi)$ 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).

LEMMA. *Seien $A = A^t \in \mathbb{R}^{n\times n}$ und $1 \le i < j \le n$. Dann existiert ein $\varphi \in [0, 2\pi]$ so daß der (i, j)-te Koeffizient von $R_{ij}(\varphi)^t A R_{ij}(\varphi)$ 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 eine beliebige Matrix $C = (\gamma_{kl})$

$$R_{ij}(\varphi)^t \cdot C = \begin{pmatrix} \gamma_{11} & \dots & \gamma_{1n} \\ \vdots & & \vdots \\ c\gamma_{i1} - s\gamma_{j1} & \dots & c\gamma_{in} - s\gamma_{jn} \\ \vdots & & \vdots \\ s\gamma_{i1} + c\gamma_{j1} & \dots & s\gamma_{in} + c\gamma_{jn} \\ \vdots & & \vdots \\ \gamma_{n1} & \dots & \gamma_{nn} \end{pmatrix},$$

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.

$$\begin{aligned}\frac{\alpha_{ii} - \alpha_{jj}}{2}\sin 2\varphi + \alpha_{ij}\cos 2\varphi &= 0\\ \tan 2\varphi &= \frac{-2\alpha_{ij}}{\alpha_{ii} - \alpha_{jj}}\\ \varphi &= \frac{1}{2}\arctan\frac{-2\alpha_{ij}}{\alpha_{ii} - \alpha_{jj}}.\end{aligned}$$

□

5.4.2. Hauptlemma. Unser Verfahren kann 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

$$\begin{aligned}&\beta_{il}^2 + \beta_{jl}^2\\ &= (c\alpha_{il} - s\alpha_{jl})^2 + (s\alpha_{il} + c\alpha_{jl})^2\\ &= c^2\alpha_{il}^2 - 2cs\alpha_{il}\alpha_{jl} + s^2\alpha_{jl}^2 + s^2\alpha_{il}^2 + 2sc\alpha_{il}\alpha_{jl} + c^2\alpha_{jl}^2\\ &= (c^2 + s^2)(\alpha_{il}^2 + \alpha_{jl}^2)\\ &= \alpha_{il}^2 + \alpha_{jl}^2 \qquad \text{da } c^2 + s^2 = 1.\end{aligned}$$

Ferner folgt nach Wahl von (i,j), also aus der Voraussetzung $|\alpha_{kl}| \le |\alpha_{ij}|$ für $k \neq l$, daß

$$w(A) = \sum_{\substack{k,l=1\\k\neq l}}^{n} \alpha_{kl}^2 \le (n^2-n)\alpha_{ij}^2. \tag{5.5}$$

Daraus folgt

$$\begin{aligned} w(B) &= 2 \cdot \sum_{k<l} \beta_{kl}^2 \\ &= \sum_{\substack{k\neq i,j\\ l\neq i,j\\ k\neq l}} \beta_{kl}^2 + 2 \cdot \sum_{l\neq i,j} \beta_{il}^2 + 2 \cdot \sum_{l\neq i,j} \beta_{jl}^2 + \underbrace{\beta_{ij}^2 + \beta_{ji}^2}_{=0} \\ &= \sum_{\substack{k\neq i,j\\ l\neq i,j\\ k\neq l}} \alpha_{kl}^2 + 2 \cdot \sum_{l\neq i,j} \alpha_{il}^2 + 2 \cdot \sum_{l\neq i,j} \alpha_{jl}^2 \qquad \text{nach (5.4)} \\ &= w(A) - 2\alpha_{ij}^2 \\ &\le w(A) \cdot \Big(1 - \frac{2}{n^2-n}\Big), \end{aligned}$$

denn nach (5.5) ist

$$-2\alpha_{ij}^2 \le -\frac{2}{n^2-n} w(A), \quad \text{also} \quad w(A) - 2\alpha_{ij}^2 \le w(A) \cdot \Big(1 - \frac{2}{n^2-n}\Big).$$

□

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. Eine Folge $(\alpha_n)_{n\in\mathbb{N}}$ reeller Zahlen heißt *beschränkt*, wenn es $a, b \in \mathbb{Q}$ gibt mit

$$a \le \alpha_n \le b \quad \text{für alle } n \in \mathbb{N}.$$

Eine Folge (α_n) reeller Zahlen heißt *konvergent* gegen $\alpha \in \mathbb{R}$ (in Zeichen $\lim_{n\to\infty} \alpha_n = \alpha$), falls gilt

$$\forall \varepsilon{>}0 \exists N{\in}\mathbb{N} \forall n{\ge}N \colon |\alpha_n - \alpha| \le \varepsilon.$$

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*

$$\lim_{n\to\infty}(\alpha_n+\beta_n)=\lim_{n\to\infty}\alpha_n+\lim_{n\to\infty}\beta_n$$
$$\lim_{n\to\infty}(\alpha_n\cdot\beta_n)=\lim_{n\to\infty}\alpha_n\cdot\lim_{n\to\infty}\beta_n.$$

(2) Für konvergente Folgen $(A_m)_{m\in\mathbb{N}}$ und $(B_m)_{m\in\mathbb{N}}$ reeller $n\times n$-Matrizen gilt entsprechend

$$\lim_{m\to\infty}(A_m+B_m)=\lim_{m\to\infty}A_m+\lim_{m\to\infty}B_m$$
$$\lim_{m\to\infty}(A_m\cdot B_m)=\lim_{m\to\infty}A_m\cdot\lim_{m\to\infty}B_m.$$

(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) Seien $(A_m)_{m\in\mathbb{N}}$ eine konvergente Folge orthogonaler $n\times n$-Matrizen und $A=\lim_{m\to\infty}A_m$. Dann ist

$$\begin{aligned}A^tA&=(\lim_{m\to\infty}A_m)^t\lim_{m\to\infty}A_m\\&=\lim_{m\to\infty}A_m^t\cdot\lim_{m\to\infty}A_m\\&=\lim_{m\to\infty}A_m^tA_m=\lim_{m\to\infty}E=E,\end{aligned}$$

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

$$a\le\alpha_n\le b\quad\text{für alle } n\in\mathbb{N}.$$

Wir definieren jetzt rekursiv $a_k,b_k\in\mathbb{Q}$ und $n_k\in\mathbb{N}$ mit

(a) $a=a_0\le a_1\le\cdots\le a_k<b_k\le\cdots\le b_1\le b_0=b$ und $b_k=a_k+\frac{1}{2^k}(b-a)$,

(b) $a_k\le\alpha_{n_k}\le b_k$ und $n_0<n_1<\cdots<n_k$,

(c) es gibt unendlich viele $n\ge n_k$ mit $a_k\le\alpha_n\le b_k$.

Eine so definierte Teilfolge (α_{n_k}) wird eine Cauchyfolge sein. Denn sei etwa $l\ge k$. Im Fall $\alpha_{n_l}\le\alpha_{n_k}$ hat man dann

$$0\le\alpha_{n_k}-\alpha_{n_l}\le b_k-a_l\le b_k-a_k=\frac{1}{2^k}(b-a),$$

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 $c \leq \alpha_n \leq b_k$.

(6) Dies folgt sofort aus (5).

(7) Nach (6) genügt es zu zeigen, daß jede Folge orthogonaler Matrizen beschränkt ist. Aus $A^tA = E$ folgt aber $\sum_{j=1}^n \alpha_{ij}^2 = 1$, also $|\alpha_{ij}| \leq 1$ für alle i, j und damit die Behauptung.

□

5.4.4. Hauptachsentransformation nach Jacobi.

SATZ. *Sei $A = A^t \in \mathbb{R}^{n\times n}$. Wir definieren rekursiv eine Folge $(R_i)_{i\in\mathbb{N}}$ von Jacobischen Rotationsmatrizen: 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*

$$C^tAC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

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^tAC 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) $\lim_{m\to\infty} A_{\mu(m)} = C^tAC$. 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. □

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. Seien 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 $(R, +, \cdot, 1)$) 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

$$\alpha(\beta\gamma) = (\alpha\beta)\gamma \quad \text{Assoziativgesetz}$$
$$\alpha 1 = \alpha = 1\alpha \qquad 1 \text{ ist beidseitig neutrales Element.}$$

(3) Für alle $\alpha, \beta, \gamma \in R$ gelten die *Distributivgesetze*

$$(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma,$$
$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma.$$

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

$$0 \cdot \alpha = 0,$$
$$-\alpha = (-1)\alpha,$$
$$(-\alpha)(-\beta) = \alpha\beta.$$

(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 1.3.4).

Im Rest dieses Abschnitts sei K ein kommutativer Ring.

6.1.2. Definition der Determinante.

DEFINITION. Seien $n \in \mathbb{N}$, $n \geq 1$. 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. aus $a_i = a_{i+1}$ folgt
$$\det(a_1, \ldots, a_n) = 0,$$
für alle i mit $1 \leq i < n$ und alle $a_1, \ldots, a_n \in K^n$ mit $a_i = a_{i+1}$.
(3) det ist *normiert*, d.h. $\det(E) = 1$. ($E \in K^{n\times n}$ ist die Einheitsmatrix).

6.1.3. Existenz der Determinante.

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 $a_1 \ldots, a_n$ Spalten von A sind.

(1) det ist multilinear. Es genügt zu zeigen, daß f_j multilinear ist, d.h. daß $f_j(a_1, \ldots, a_n)$ linear in a_l ist, 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. Seien $1 \leq l < n$ 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.

$$\begin{aligned}\det(E) &= \sum_{j=1}^{n}(-1)^{i+j}\delta_{ij}d(E_{ij}) \\ &= (-1)^{2i}d(E') \qquad\qquad \text{mit } E' \in K^{(n-1)\times(n-1)} \text{ Einheitsmatrix} \\ &= 1.\end{aligned}$$

□

6.1.4. 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). Gegeben sei $(a_1, \ldots, a_n)$ mit $a_i = a_j$. Vertausche a_i mit a_{i+1}, dann mit a_{i+2} und so weiter, 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. □

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:

$$S_n := S(\{1,\dots,n\}) := \{\,\sigma\colon \{1,\dots,n\} \to \{1,\dots,n\} \mid \sigma \textit{ bijektiv}\}$$

ist die Gruppe der Permutationen *von* $\{1,\dots,n\}$*).*

BEWEIS. Sind $a_1,\dots,a_n$ die Spalten von A, so ist $a_j = \sum_{i=1}^n \alpha_{ij} e_i$, 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 $j \neq l$. □

LEMMA. *Seien* $d\colon K^{n\times n} \to K$ *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 die Behauptung. □

Insbesondere ist also $d(e_{\sigma(1)},\dots,e_{\sigma(n)}) = (-1)^{m_\sigma}$ im Fall einer Determinante d; dann hat man nämlich $d(e_1,\dots,e_n) = 1$.

Aus diesen drei Lemmata ergibt sich unmittelbar das gesuchte Eindeutigkeitsresultat:

SATZ. *Seien* $n \in \mathbb{N}$*,* $n \geq 1$ *und* $d, d'\colon K^{n\times n} \to K$ *Determinanten. Dann ist* $d = d'$*.*

Ist $A = (\alpha_{ij})$, so schreibt man für $\det(A)$ auch

$$\begin{vmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \dots & \alpha_{nn} \end{vmatrix}$$

6.1.5. Permutationen und Transpositionen. Als einfache Folgerung aus dem Beweis notieren wir einige 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 so daß

$$\tau(i) = \begin{cases} i & \text{für } i \neq j, l \\ l & \text{für } i = j \\ j & \text{für } i = l. \end{cases}$$

LEMMA. *Sei n eine natürliche Zahl ≥ 1.*

(1) S_n hat $n!$ $(:= 1 \cdot 2 \cdot \cdots \cdot n)$ 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 6.1.4 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. □

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. Seien $\sigma = \tau_1 \circ \cdots \circ \tau_m$, 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 $\varepsilon_{\mathrm{id}} = (-1)^0 = 1$. □

6.2. Eigenschaften der Determinante

In diesem Abschnitt sei K stets ein kommutativer Ring.

6.2.1. Entwicklung von Determinanten.

SATZ. *Seien $n \in \mathbb{N}$, $n \geq 1$ 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 $\alpha, \beta \in K$.

(i) $\det(a_1, \ldots, a_{i-1}, \alpha a_i + \beta b_i, \alpha_{i+1}, \ldots, a_n) =$
$\alpha \det(a_1, \ldots, a_n) + \beta \det(a_1, \ldots, a_{i-1}, b_i, a_{i+1}, \ldots, a_n)$.

(ii) $\det(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) = \varepsilon_\sigma \det(a_1, \ldots, a_n)$.

(iii) $\det(a_1, \ldots, a_n) = 0$ *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 6.1.4 gilt

$$\begin{aligned}
&\det(A)\\
&= \sum_{\sigma\in S_n} \varepsilon_\sigma \alpha_{\sigma(1)1} \cdots \alpha_{\sigma(n)n}\\
&= \sum_{\sigma\in S_n} \varepsilon_\sigma \alpha_{1\sigma^{-1}(1)} \cdots \alpha_{n\sigma^{-1}(n)} && \text{da } K \text{ kommutativ ist}\\
&= \sum_{\sigma\in S_n} \varepsilon_\sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)} && \text{da } {}^{-1}\colon S_n \to S_n \text{ bijektiv, } \varepsilon_{\sigma^{-1}} = \varepsilon_\sigma,\\
& && \text{denn } \varepsilon_{\sigma^{-1}} \cdot \varepsilon_\sigma = \varepsilon_{\mathrm{id}} = 1
\end{aligned}$$

$$\det(A^t) = \sum_{\sigma\in S_n} \varepsilon_\sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)}.$$

(2) Dies wurde im Beweis des Existenzsatzes 6.1.3 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) (i) Nach Definition der Determinante. (ii) Man wiederhole den Beweis des dritten Lemmas aus 6.1.4 mit a_i statt e_i. (iii) Dies war bereits als Teil (1) des ersten Lemmas in 6.1.4 bewiesen worden.

(5) Dies folgt aus dem zweiten und dritten Lemma in 6.1.4 zusammen mit Teil (1).

□

6.2.2. Multiplikationssatz.

KOROLLAR (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

$d\colon K^{n\times n} \to K$,

$d(B) = \det(AB) = \det(Ab_1, \ldots, Ab_n)$ $(b_1, \ldots, b_n$ Spalten von $B)$.

d ist aufgrund der Definition multilinear und alternierend. Also gilt nach Teil (5) des Satzes

$$\det(AB) = d(B) = \det(B)d(E) = \det(B)\det(A).$$

Das war zu zeigen. □

6.2.3. Zeilen- und Spaltenumformungen und Determinanten. Wir wollen uns noch überlegen, welche Auswirkungen die in 1.4.1 eingeführten elementaren Spaltenumformungen einer Matrix auf ihre Determinante haben.

(1) Vertauschen zweier Spalten verändert das Vorzeichen (nach Teil (4)(ii) des Satzes).
(2) Multiplikation einer Spalte mit einem $\lambda \in K$ führt zum λ-fachen der Determinante (nach Teil (4)(i) des Satzes).
(3) Addition einer Spalte zu einer anderen verändert den Wert der Determinante nicht, denn es gilt nach den Teilen (4)(i,iii) des Satzes

$$\begin{aligned}&\det(a_1, \ldots, a_i, \ldots, a_i + a_j, \ldots, a_n)\\ &= \det(a_1, \ldots, a_i, \ldots, a_i, \ldots, a_n) + \det(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n)\\ &= 0 + \det(a_1, \ldots, a_n).\end{aligned}$$

Wegen $\det(A) = \det(A^t)$ gelten dieselben Aussagen auch für elementare Zeilenumformungen.

6.2.4. 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:

$$\begin{aligned}\sum_{j=1}^{n} \alpha_{ij}^{\#} \alpha_{jl} &= \sum_{j=1}^{n} (-1)^{i+j} \alpha_{jl} \det(A_{ji}) \\ &= \sum_{j=1}^{n} (-1)^{i+j} \alpha'_{ji} \det(A'_{ji}) \\ &= \det(A') \\ &= 0, \quad \text{da die } i\text{-te und die } l\text{-te Spalte von } A' \text{ übereinstimmen.}\end{aligned}$$

Das war zu zeigen. □

KOROLLAR. *Seien* K *ein Körper,* $A \in K^{n \times n}$. *Folgende Aussagen sind äquivalent:*

(1) A *ist invertierbar.*
(2) $\det(A) \neq 0$.
(3) $\operatorname{rg}(A) = n$.

In diesem Falle gilt außerdem

$$A^{-1} = \frac{1}{\det(A)} \cdot A^{\#}, \tag{6.1}$$

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

$$\det(A^{-1}) \det(A) = \det(A^{-1}A) = \det(E) = 1.$$

(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). □

6.2.5. 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

$$\det(A) = \det(A_1) \cdot \dots \cdot \det(A_r).$$

Insbesondere ist also

$$\begin{vmatrix} \gamma_1 & & * \\ & \ddots & \\ 0 & & \gamma_n \end{vmatrix} = \gamma_1 \cdot \dots \cdot \gamma_n.$$

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 $n_1 + \dots + n_{r+1} = n$. 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)

$$\det(A) = d(A_{r+1}) = \det(A_{r+1})d(E) = \det(A_{r+1}) \cdot \det(A_1) \cdot \dots \cdot \det(A_r),$$

denn nach Induktionsvoraussetzung ist

$$d(E) = \begin{vmatrix} A_1 & & & * \\ & \ddots & & \\ & & A_r & \\ 0 & & & E \end{vmatrix} = \begin{vmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_r \end{vmatrix} = \det(A_1) \cdot \dots \cdot \det(A_r),$$

wie man durch n_{r+1}-malige Entwicklung nach der letzten Zeile sieht.

6.2.6. Cramersche Regel. Seien $A \in K^{n\times n}$, $b \in K^n$ 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

$$\det(A_i) = \det(a_1, \ldots, a_{i-1}, \sum_{j=1}^{n} \xi_j a_j, a_{i+1}, \ldots, a_n) = \xi_i \det(a_1, \ldots, a_n).$$

Das war zu zeigen. □

6.2.7. Vandermondesche Determinante. Für $\alpha_0, \ldots, \alpha_n \in K$ ist die *Vandermondesche 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

$$D_n = (\alpha_1 - \alpha_0) \ldots (\alpha_n - \alpha_0) \begin{vmatrix} 1 & \alpha_1 & \alpha_1^2 & \ldots & \alpha_1^{n-1} \\ \vdots & & & & \\ 1 & \alpha_n & \alpha_n^2 & \ldots & \alpha_n^{n-1} \end{vmatrix} = \prod_{i<j\leq n} (\alpha_j - \alpha_i).$$

nach Induktionsvoraussetzung. □

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

$$\begin{vmatrix} 1 & \alpha_0 & \alpha_0^2 & \ldots & \alpha_0^n \\ \vdots & & & & \\ 1 & \alpha_n & \alpha_n^2 & \ldots & \alpha_n^n \end{vmatrix} = \prod_{i<j\leq n} (\alpha_j - \alpha_i) \neq 0 \quad \text{nach Voraussetzung}$$

ist dieses Gleichungssystem eindeutig lösbar. Die Lösungen – also die Koeffizienten des Polynoms – kann man etwa nach der Cramerschen Regel ausrechnen:

$$\gamma_k = \frac{1}{\prod_{i<j\leq n}(\alpha_j - \alpha_i)} \begin{vmatrix} 1 & \alpha_0 & \ldots & \alpha_0^{k-1} & \beta_0 & \alpha_0^{k+1} & \ldots & \alpha_0^n \\ \vdots & & & & & & & \\ 1 & \alpha_n & \ldots & \alpha_n^{k-1} & \beta_n & \alpha_n^{k+1} & \ldots & \alpha_n^n \end{vmatrix}.$$

□

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.

DEFINITION. Seien $a, a_1, \ldots, a_{n-1} \in K^n$. a heißt *Vektorprodukt* von $a_1, \ldots, a_{n-1}$ (geschrieben $a_1 \times \cdots \times a_{n-1}$), wenn für alle $x \in K^n$ gilt

$$\det(a_1, \ldots, a_{n-1}, x) = a^t x.$$

BEMERKUNGEN.

(1) Das Vektorprodukt $a_1 \times \cdots \times a_{n-1}$ existiert und ist eindeutig bestimmt, denn $\det(a_1, \ldots, a_{n-1}, x)$ ist (in x) eine lineare Abbildung des K^n in K und deshalb (nach 3.1.3) mit einem geeigneten a in der Form $a^t x$ darstellbar.

(2) $\det(a_1, \ldots, a_{n-1}, e_i)$ ist die i-te Komponente von $a_1 \times \cdots \times a_{n-1}$.

(3) Die Familie $a_1, \ldots, a_{n-1}$ ist linear unabhängig genau dann, wenn $a_1 \times \cdots \times a_{n-1} \neq 0$.

BEWEIS. $\Longrightarrow$. Seien $a_1, \ldots, a_{n-1}$ linear unabhängig. Man ergänze sie zu Basis $a_1, \ldots, a_{n-1}, x$ des K^n. Dann gilt

$$0 \neq \det(a_1, \ldots, a_{n-1}, x) = (a_1 \times \cdots \times a_{n-1})^t x.$$

Also ist $a_1 \times \cdots \times a_{n-1} \neq 0$.

$\Longleftarrow$. Gelte $a_1 \times \cdots \times a_{n-1} \neq 0$. Dann ist $0 \neq (a_1 \times \cdots \times a_{n-1})^t (a_1 \times \cdots \times a_{n-1}) = \det(a_1, \ldots, a_{n-1}, a_1 \times \cdots \times a_{n-1})$. Also sind $a_1, \ldots, a_{n-1}$ linear unabhängig. □

(4) Für $\sigma \in S_{n-1}$ gilt

$$a_{\sigma(1)} \times \cdots \times a_{\sigma(n-1)} = \varepsilon_\sigma a_1 \times \cdots \times a_{n-1}.$$

(5) Es gilt $(a_1 \times \cdots \times a_{n-1})^t a_i = 0$ für $1 \le i < n$, denn det ist alternierend.

(6) Im Spezialfall $n = 3$ kann man das Vektorprodukt explizit angeben. Seien etwa $a = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$ und $b = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} \in K^3$. Dann ist nach (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.

$$\begin{aligned} x &\in H \\ x &\in a_1 + \langle a_2 - a_1, \ldots, a_n - a_1 \rangle && \text{nach 4.1.2} \\ x - a_1 &\in \langle a_2 - a_1, \ldots, a_n - a_1 \rangle \\ 0 &= \det(a_2 - a_1, \ldots, a_n - a_1, x - a_1) \\ \det(a_2 - a_1, \ldots, a_n - a_1, x) &= \det(a_2 - a_1, \ldots, a_n - a_1, a_1) \\ a^t x &= a^t a_1. \end{aligned}$$

□

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) $P(a_1, \ldots, a_n) := \{\, \sum_{i=1}^n \alpha_i a_i \mid \alpha_i \in \mathbb{R}, 0 \le \alpha_i \le 1 \text{ für } 1 \le i \le n \,\}$ 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 $1 \leq i \leq n$. *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

$$\mathrm{Vol}(a_1, a_2) = \|a_1\| \cdot \|h\| \quad \text{(Grundlinie} \cdot \text{Höhe)}$$

$n = 3$: Für linear unabhängige $a_1, a_2, a_3 \in V$ ist

$$\mathrm{Vol}(a_1, a_2, a_3) = \mathrm{Vol}(a_1, a_2) \cdot \|h\| \quad \text{(Grundfläche} \cdot \text{Höhe)}$$

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:

$$\begin{aligned}\sigma(a_i, a_j) &= \sum_{s,t=1}^{n} \alpha_{si}\alpha_{tj}\sigma(e_s, e_t) \\ &= \sum_{s=1}^{n} \alpha_{si}\alpha_{sj} \qquad \text{da } e_1, \ldots, e_n \text{ Orthonormalbasis ist.}\end{aligned}$$

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 $\mathrm{Vol}(a_1, \ldots, a_n)$ wohldefiniert und gleich $|\det(A)|$.

(2) Schreibe $a_n = u + h$ mit $u \in U := \langle a_1, \ldots, a_{n-1}\rangle$, $h \in U^\perp$ (dies ist möglich, da $\langle a_1, \ldots, a_n\rangle \subseteq U \oplus U^\perp$). Dann gilt

$$\begin{aligned}&\mathrm{Vol}(a_1, \ldots, \underbrace{a_n}_{=u+h}) \\ &= |\det(A)| \\ &= \mathrm{Vol}(a_1, \ldots, a_{n-1}, h), \quad \text{da } a_1, \ldots, a_{n-1}, u \text{ linear abhängig sind.}\end{aligned}$$

Also:

$$\begin{aligned}\mathrm{Vol}(a_1, \ldots, a_n)^2 &= \mathrm{Vol}(a_1, \ldots, a_{n-1}, h)^2 \\ &= \begin{vmatrix} \sigma(a_1, a_1) & \ldots & \sigma(a_1, a_{n-1}) & 0 \\ \vdots & & & \vdots \\ \sigma(a_{n-1}, a_1) & \ldots & \sigma(a_{n-1}, a_{n-1}) & 0 \\ 0 & \ldots & 0 & \sigma(h,h) \end{vmatrix} \\ &= \det((\sigma(a_i, a_j))_{1 \leq i,j \leq n-1}) \cdot \sigma(h,h) \\ &= \mathrm{Vol}(a_1, \ldots, a_{n-1})^2 \|h\|^2.\end{aligned}$$

□

KOROLLAR. *Seien $a_1, \ldots, a_{n-1} \in \mathbb{R}^n$ linear unabhängig, $\mathbb{R}^n$ mit dem üblichen Skalarprodukt versehen. Dann gilt*

$$\mathrm{Vol}(a_1, \ldots, a_{n-1}) = \|a_1 \times \cdots \times a_{n-1}\|.$$

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

$$\begin{aligned}
\mathrm{Vol}(a_1, \ldots, a_{n-1}, h) &= \mathrm{Vol}(a_1, \ldots, a_{n-1})\|h\| && \text{nach Teil (2) des Satzes} \\
\mathrm{Vol}(a_1, \ldots, a_{n-1}, h) &= |\det(a_1, \ldots, a_{n-1}, h)| && \text{nach Teil (1) des Satzes} \\
&= |(a_1 \times \cdots \times a_{n-1})^t h| \\
&= \|h\|^2.
\end{aligned}$$

Wegen $h \neq 0$ folgt $\mathrm{Vol}(a_1, \ldots, a_{n-1}) = \|h\|$. □

6.3.3. Determinante eines Endomorphismus, Orientierung. Seien 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

$$\begin{aligned}
A &\sim B \\
B &= CAC^{-1} && \text{für ein } C \in GL(n, K) \\
\det(B) &= \det(CAC^{-1}) \\
&= \det(C)\det(A)\det(C^{-1}) \\
&= \det(C)\det(A)\det(C)^{-1} \\
&= \det(A).
\end{aligned}$$

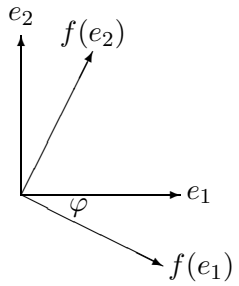
Wir können also die Determinante $\det(f)$ des Endomorphismus f definieren als $\det(A)$, wobei A eine beliebige darstellende Matrix von f ist.

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.

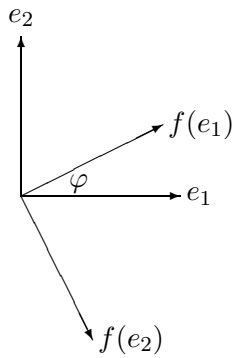
BEMERKUNG. In $\mathbb{R}^n$ nennt man *willkürlich* die Standardbasis $e_1, \ldots, e_n$ positiv orientiert.

BEISPIELE. (1) Seien $V = \mathbb{R}^2$ 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, denn es gilt

$$\det(a_1, \ldots, a_{n-1}, h) = (a_1 \times \cdots \times a_{n-1})^t h = \|h\|^2 > 0.$$

Wir haben also folgendes gezeigt: In $\mathbb{R}^n$ (mit dem üblichen Skalarprodukt) ist das Vektorprodukt aus $n-1$ linear unabhängigen Vektoren $a_1, \ldots, a_{n-1}$ stets 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).

KAPITEL 7

Eigenwerte und Eigenvektoren, Diagonalisierung

In Abschnitt 5.4 hatten wir im Kontext der euklidischen Geometrie zu einer gegebenen Matrix $A \in \mathbb{R}^{n\times n}$ nach einer *orthogonalen* Matrix $C \in \mathrm{O}(n)$ gesucht, so daß C^tAC Diagonalgestalt hat. Das Ergebnis war, daß zu einer symmetrischen Matrix A ein solches C stets gefunden werden kann. Die Hauptachsentransformation nach Jacobi lieferte ein Verfahren, ein solches $C \in \mathrm{O}(n)$ zu konstruieren. (Vorher hatten wir in 4.2.2 mit einem einfacheren Argument ein im Fall $K = \mathbb{R}$ schwächeres Resultat bewiesen, nämlich daß man zu jeder symmetrischen Matrix $A \in K^{n\times n}$ ein $C \in \mathrm{GL}(n, K)$ konstruieren kann, so daß C^tAC Diagonalgestalt hat). Offenbar kann ein solches orthogonales C auch nur für symmetrisches A existieren.

Wir kommen jetzt wieder auf die affine Geometrie zurück und wollen uns mit der Frage befassen, ob und gegebenenfalls wie man eine nicht notwendig symmetrische Matrix $A \in K^{n\times n}$ mittels einer *invertierbaren* Matrix $C \in \mathrm{GL}(n, K)$ durch Bilden von CAC^{-1} in Diagonalgestalt bringen kann. Es wird sich zeigen, daß es dafür wichtig ist, die „Richtungen“ zu identifizieren, in denen der Endomorphismus f_A als Streckung wirkt (sie werden durch die sogenannten Eigenvektoren angegeben), und ferner die Streckungsfaktoren (sogenannte Eigenwerte) zu bestimmen. Die Eigenwerte kann man als Nullstellen eines gewissen, für den Endomorphismus f_A charakteristischen Polynoms ausrechnen. Es ist deshalb nützlich, mit einigen elementaren Betrachtungen über den Polynombegriff zu beginnen.

7.1. Polynome

In 2.1.2 hatten wir $\mathbb{R}[t]$, also die Menge aller Polynome in einer Variablen t mit Koeffizienten aus $\mathbb{R}$, als Beispiel eines Vektorraums eingeführt. Damals hatten wir Polynome $\alpha_n t^n + \cdots + \alpha_1 t + \alpha_0$ als formale Ausdrücke aufgefaßt, und Addition und Skalarmultiplikation definiert durch

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

Wir wollen jetzt den Polynombegriff etwas genauer fassen, und gleichzeitig auch Polynome (in einer Veränderlichen) über einem beliebigen Ring einführen. Es wird sich zeigen, daß man wieder einen Ring erhält, und zwar den sogenannten Polynomring $K[t]$.

Jedem Polynom $f = \sum_{i=0}^{n} \alpha_i t^i \in K[t]$ kann man eine Abbildung von K nach K zuordnen durch $\gamma \mapsto \sum_{i=0}^{n} \alpha_i \gamma^i \in K$. Für allgemeine Körper bestimmt eine solche Abbildung jedoch *nicht* die Koeffizienten. Zum Beispiel

für $\mathbb{F}_2$ und $f = t^2+t$ ist $f(0) = f(1) = 0$, also die zugeordnete Abbildung die Nullabbildung, obwohl die Koeffizienten nicht Null sind. Dies ist ein Grund, Polynome als formale Ausdrücke einzuführen.

7.1.1. Der Polynomring $K[t]$. Unser Ansatz zu einer genauen Fassung des Polynombegriffs ist es, daß ein formaler Ausdruck $\alpha_n t^n + \cdots + \alpha_1 t + \alpha_0$ eindeutig festgelegt ist durch die Folge $(\alpha_0, \alpha_1, \dots, \alpha_n, 0, 0, \dots)$ der Koeffizienten.

DEFINITION. Seien K ein Ring und

$$K^{(\mathbb{N})} := \{\, (\alpha_0, \alpha_1, \alpha_2, \dots) \in K^{\mathbb{N}} \mid \alpha_i \in K, \text{ fast alle } \alpha_i \text{ sind } 0 \,\}.$$

„Fast alle“ bedeutet „alle bis auf endlich viele“. Die Elemente $f \in K^{(\mathbb{N})}$ nennt man *Polynome* über K. Man schreibt auch

$$K[t] := \{\, f \mid f \text{ Polynom über } K \,\} = K^{(\mathbb{N})}.$$

Für $f, g \in K[t]$, etwa $f = (\alpha_0, \alpha_1, \dots)$ und $g = (\beta_0, \beta_1, \dots)$ erklären wir Addition $f + g$ und Multiplikation $f \cdot g$ durch

$$\begin{aligned} f + g &= (\alpha_0 + \beta_0, \alpha_1 + \beta_1, \dots), \\ f \cdot g &= (\gamma_0, \gamma_1, \dots) \qquad\qquad \text{mit } \gamma_i = \textstyle\sum_{k+l=i} \alpha_k \cdot \beta_l. \end{aligned}$$

Offenbar sind $f + g, f \cdot g \in K[t]$.

SATZ. *Sei K ein (kommutativer) Ring.*

(1) $K[t]$ wird mit obiger Addition und Multiplikation zu einem (kommutativen) Ring mit Null $(0, 0, 0, \dots)$ und Eins $(1, 0, 0, \dots)$. Man nennt $K[t]$ den Polynomring *über K in einer Veränderlichen.*

(2) Definiert man $t := (0, 1, 0, 0, \dots) \in K[t]$ und identifiziert man $\alpha \in K$ mit $(\alpha, 0, 0, \dots) \in K[t]$, so läßt sich jedes $f \in K[t]$ schreiben als

$$f = \sum_{i=0}^{n} \alpha_i t^i \quad \text{mit } \alpha_i \in K.$$

(Addition und Multiplikation sind hier in $K[t]$ zu verstehen).

(3) Diese Darstellung ist im folgenden Sinn eindeutig. Gilt $\sum_{i=0}^{n} \alpha_i t^i = \sum_{j=0}^{m} \beta_j t^j$ mit (etwa) $n \le m$, so folgt $\alpha_k = \beta_k$ für $0 \le k \le n$ und $\beta_k = 0$ für $n < k \le m$.

BEWEIS. (1) a. $(K[t], +)$ ist abelsche Gruppe, da die Addition komponentenweise definiert ist.

b. Ist K kommutativ, so ist die Multiplikation in $K[t]$ kommutativ, denn für $f = (\alpha_i)_{i\in\mathbb{N}} \in K^{(\mathbb{N})}$ und $g = (\beta_i)_{i\in\mathbb{N}} \in K^{(\mathbb{N})}$ gilt

$$\begin{aligned} f \cdot g &=: (\gamma_i) \quad \text{mit } \gamma_i = \textstyle\sum_{k+l=i} \alpha_k \beta_l \\ g \cdot f &=: (\gamma'_i) \quad \text{mit } \gamma'_i = \textstyle\sum_{k+l=i} \beta_k \alpha_l = \sum_{k+l=i} \alpha_l \beta_k = \sum_{l+k=i} \alpha_k \beta_l = \gamma_i. \end{aligned}$$

c. Die Multiplikation ist assoziativ, denn ist noch $h = (\gamma_i)_{i\in\mathbb{N}} \in K^{(\mathbb{N})}$, so gilt ($[c]_i$ bezeichne die i-te Komponente von $c \in K^{(\mathbb{N})}$)

$$\begin{aligned}
[(f \cdot g) \cdot h]_i &= \sum_{k+l=i} [f \cdot g]_k \cdot \gamma_l \\
&= \sum_{k+l=i} \sum_{r+s=k} \alpha_r \beta_s \gamma_l \\
&= \sum_{r+s+l=i} \alpha_r \beta_s \gamma_l \\
&= \sum_{r+q=i} \alpha_r [g \cdot h]_q \\
&= [f \cdot (g \cdot h)]_i.
\end{aligned}$$

d. Das Distributivgesetz ergibt sich aus

$$\begin{aligned}
[f \cdot (g + h)]_i &= \sum_{k+l=i} \alpha_k (\beta_l + \gamma_l) \\
&= \sum_{k+l=i} \alpha_k \beta_l + \sum_{k+l=i} \alpha_k \gamma_l \\
&= [f \cdot g]_i + [f \cdot h]_i \\
&= [f \cdot g + f \cdot h]_i.
\end{aligned}$$

Im Fall eines nicht kommutativen Rings K beweist man $(f + g)h = fh + gh$ genauso.

e. $(1, 0, 0, \dots)$ ist Einselement, denn für alle i gilt

$$[f \cdot (1, 0, 0, \dots)]_i = \alpha_i \cdot 1 + \alpha_{i-1} \cdot 0 + \cdots + \alpha_0 \cdot 0 = \alpha_i.$$

Also ist $f \cdot (1, 0, 0, \dots) = f$.

(2) Setze $t^0 := (1, 0, 0, \dots)$. Durch Induktion über i sieht man dann leicht, daß

$$t^i = (\underbrace{0, \dots, 0}_{i}, 1, 0, 0, \dots).$$

Es folgt

$$\alpha \cdot t^i = (\alpha, 0, 0, \dots) \cdot t^i = (\underbrace{0, \dots, 0}_{i}, \alpha, 0, 0, \dots) = t^i \cdot \alpha.$$

Sei jetzt $f = (\alpha_i)_{i\in\mathbb{N}}$ und $\alpha_i = 0$ für $i > n$. Dann gilt

$$\begin{aligned}
&(\alpha_0, \alpha_1, \dots) \\
&= (\alpha_0, 0, 0, \dots) + (0, \alpha_1, 0, 0, \dots) + (0, \dots, 0, \alpha_n, 0, 0, \dots) = \sum_{i=0}^{n} \alpha_i t^i.
\end{aligned}$$

Damit ist die Existenz der im Satz angegebenen Darstellung von $f \in K[t]$ gezeigt.

(3) Die Eindeutigkeit dieser Darstellung ergibt sich mit demselben Argument. ∎

7.1.2. Der Grad eines Polynoms. Sei $f \in K[t]$, etwa $f = \sum_{i=0}^{n} \alpha_i t^i$ mit $\alpha_n \neq 0$. Dann nennt man die (nach dem eben bewiesenen Satz eindeutig bestimmte) Zahl n den *Grad* des Polynoms f und bezeichnet ihn mit $\deg(f)$. Für das Nullpolynom $0 \in K[t]$ setzt man $\deg(0) := -1$.

LEMMA. *Sei K ein Körper.*

(1) Für $f, g \in K[t]$ mit $f, g \neq 0$ gilt $\deg(fg) = \deg(f) + \deg(g)$.

(2) Für $f, g, h \in K[t]$ mit $f \neq 0$ folgt aus $fg = fh$ stets $g = h$ (Kürzungsregel).

BEWEIS. (1) Seien $f = \sum_{i=0}^{n} \alpha_i t^i$ mit $\alpha_n \neq 0$ und $g = \sum_{j=0}^{m} \beta_j t^j$ mit $\beta_m \neq 0$. Dann ist $fg = \sum_{k=0}^{n+m} \gamma_k t^k$ mit $\gamma_k = \sum_{i+j=k} \alpha_i \beta_j$, also $\gamma_{n+m} = \alpha_n \beta_m \neq 0$ (hier wird verwendet, daß wir uns in einem Körper befinden).

(2) Seien $f, g, h \in K[t]$ mit $f \neq 0$ und nehmen wir an, daß $fg = fh$. Dann ist $f(g - h) = 0$. Wegen $f \neq 0$ folgt $g - h = 0$ aus (1). □

7.1.3. Nullstellen von Polynomen. Seien K ein Ring, $\gamma \in K$ und $f = \sum_{i=0}^{n} \alpha_i t^i \in K[t]$. Man definiert dann

$$f(\gamma) := \sum_{i=0}^{n} \alpha_i \gamma^i \in K$$

und nennt dieses Ringelement das Ergebnis der *Einsetzung* von γ in f. (Hierbei ist wie üblich $\gamma^0 := 1$ gesetzt für alle $\gamma \in K$). γ heißt *Nullstelle* von f, wenn $f(\gamma) = 0$. Schließlich nennt man das Polynom f *konstant*, wenn $f = 0$ oder $\deg(f) = 0$ ist.

BEMERKUNGEN. (1) Für $f, g \in K[t]$ und $\gamma \in K$ gilt offenbar

$$(f + g)(\gamma) = f(\gamma) + g(\gamma), \quad (f \cdot g)(\gamma) = f(\gamma) \cdot g(\gamma).$$

(2) Ist $f \in K[t]$ konstant, so hat f genau dann eine Nullstelle, wenn $f = 0$ ist. Ferner hat zum Beispiel $t^2 + 1 \in \mathbb{R}[t]$ keine Nullstelle.

LEMMA. *Seien K ein Ring, $f \in K[t]$ und $\gamma \in K$. Dann ist $f(\gamma) = 0$ genau dann, wenn es ein Polynom $g \in K[t]$ gibt mit $f = (t - \gamma)g$.*

BEWEIS. Gibt es ein Polynom $g \in K[t]$ mit $f = (t - \gamma)g$, so ist $f(\gamma) = (\gamma - \gamma) \cdot g(\gamma) = 0$. Seien nun umgekehrt $f = \sum_{i=0}^{n} \alpha_i t^i \in K[t]$ und $\gamma \in K$ mit $f(\gamma) = \sum_{i=0}^{n} \alpha_i \gamma^i = 0$ gegeben. Dann gilt

$$\begin{aligned} f &= f - \sum_{i=0}^{n} \alpha_i \gamma^i \\ &= \sum_{i=0}^{n} \alpha_i (t^i - \gamma^i) \\ &= \sum_{i=1}^{n} \alpha_i (t - \gamma)(t^{i-1} + \gamma t^{i-2} + \cdots + \gamma^{i-1}) + \alpha_0 (t^0 - \gamma^0) \\ &= (t - \gamma)g + \alpha_0 (1 - 1) \\ &= (t - \gamma)g \end{aligned}$$

für $g = \sum_{i=1}^{n} \alpha_i (t^{i-1} + \gamma t^{i-2} + \cdots + \gamma^{i-1}) \in K[t]$. □

KOROLLAR. *Seien K ein Körper, $f \in K[t]$ und* $\deg(f) = n \geq 0$. *Dann hat f höchstens n Nullstellen.*

BEWEIS. Induktion nach n. *Basis* $n = 0$. Dann ist f ein konstantes, von 0 verschiedenes Polynom und hat also keine Nullstelle. *Schritt* $n-1 \Longrightarrow n$ mit $n \geq 1$. Nehmen wir an, daß f verschiedene Nullstellen $\alpha_1, \alpha_2, \ldots, \alpha_{n+1}$ besitzt. Nach dem Lemma ist $f = (t-\alpha_1)g$ für ein $g \in K[t]$ mit $\deg(g) = n-1$. Wegen $\alpha_i - \alpha_1 \neq 0$ für $i = 2, \ldots, n+1$ sind $\alpha_2, \ldots, \alpha_{n+1}$ verschiedene Nullstellen von g. Dies ist nach Induktionsvoraussetzung nicht möglich. □

DEFINITION. Seien K ein Körper, $\lambda \in K$, $f \in K[t]$ und e eine natürliche Zahl ≥ 1. λ heißt *e-fache Nullstelle* von f, wenn sich f schreiben läßt in der Form $f = (t-\lambda)^e g$ mit $g \in K[t]$, $g(\lambda) \neq 0$.

BEMERKUNG. e ist eindeutig bestimmt.

BEWEIS. Nehmen wir an $f = (t-\lambda)^e g = (t-\lambda)^{e'} g'$ mit $g, g' \in K[t]$ und $g(\lambda) \neq 0 \neq g'(\lambda)$. Wäre etwa $e < e'$, so hätte man nach der Kürzungsregel $g = (t-\lambda)^{e'-e} g'$, was wegen $g(\lambda) \neq 0$ nicht sein kann. □

7.2. Eigenwerte und Eigenvektoren

Gegeben sei ein endlichdimensionaler K-Vektorraum V und eine lineare Abbildung $f\colon V \to V$. Gesucht ist eine Basis $(x_1, \ldots, x_n)$ von V so daß f bezüglich $x_1, \ldots, x_n$ durch eine Diagonalmatrix dargestellt wird, also

$$f \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

DEFINITION. Seien V ein K-Vektorraum und $f\colon V \to V$ linear.

(1) Ein $x \in V$ mit $x \neq 0$ heißt *Eigenvektor* von f, wenn es ein $\lambda \in K$ gibt mit
$$f(x) = \lambda x.$$
(2) Ein $\lambda \in K$ heißt *Eigenwert* von f, wenn es ein $x \in V$ mit $x \neq 0$ gibt so daß $f(x) = \lambda x$. Man nennt dann x einen Eigenvektor zum Eigenwert λ.
(3) Sei λ Eigenwert von f.
$$\mathrm{Eig}(f, \lambda) := \{\, x \in V \mid f(x) = \lambda x \,\}$$
heißt *Eigenraum* von f zum Eigenwert λ.

BEMERKUNGEN. Seien V ein K-Vektorraum und $f\colon V \to V$ linear.

(1) Man beachte, daß $\mathrm{Eig}(f, \lambda) = \ker(f - \lambda\mathrm{id})$ Unterraum von V ist, und nach Definition von Eigenwerten auch $\neq 0$ ist.
(2) Hat V eine Basis aus Eigenvektoren $x_1, \ldots, x_n$ von f, so wird f bezüglich der Basis $x_1, \ldots, x_n$ durch eine Diagonalmatrix dargestellt, und zwar
$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

falls x_i Eigenvektor zum Eigenwert λ_i ist. Es gilt nämlich $f(x_i) = \lambda_i x_i$. Man nennt einen solchen Endomorphismus f *diagonalisierbar* (siehe auch 7.4).

(3) Im allgemeinen ist eine solche Darstellung eines beliebigen linearen $f\colon V \to V$ nicht möglich; man kommt nur „beinahe“ auf eine Diagonalmatrix. Wir werden dies im nächsten Abschnitt 7.3 genauer klären, wo wir die Jordansche Normalform behandeln werden.

7.2.1. Summen von Eigenräumen. Wir wollen uns mit dem Problem befassen, wie man die Eigenwerte und die zugehörigen Eigenräume eines gegebenen Endomorphismus f berechnen kann. Zunächst können wir feststellen, daß eine Summe von Eigenräumen immer eine direkte Summe ist.

SATZ. *Seien V ein K-Vektorraum und $f\colon V \to V$ linear. Sind $\lambda_1, \dots, \lambda_n$ verschiedene Eigenwerte von f und ist $V_i := \operatorname{Eig}(f, \lambda_i)$ der Eigenraum zu λ_i (also $V_i \neq 0$), so gilt*

$$V_1 + \cdots + V_n = V_1 \oplus \cdots \oplus V_n \subseteq V,$$

insbesondere also $n \leq \dim V$.

BEWEIS. Zu zeigen ist, daß für beliebige $x_1 \in V_1, \dots, x_n \in V_n$ aus $\sum_{i=1}^n x_i = 0$ stets folgt $x_1 = \cdots = x_n = 0$. Wir zeigen dies durch Induktion über n. *Basis* $n = 1$. Hier ist nichts zu zeigen. *Schritt* $n-1 \Longrightarrow n$. Gelte $\sum_{i=1}^n x_i = 0$ mit $x_1 \in V_1, \dots, x_n \in V_n$. Dann ist

$$0 = f(\sum_{i=1}^n x_i) = \sum_{i=1}^n f(x_i) = \sum_{i=1}^n \lambda_i x_i.$$

Andererseits hat man

$$0 = \lambda_n \cdot \sum_{i=1}^n x_i = \sum_{i=1}^n \lambda_n x_i.$$

Subtraktion ergibt

$$\begin{aligned}
0 &= \sum_{i=1}^{n-1} (\lambda_n - \lambda_i) x_i \\
0 &= (\lambda_n - \lambda_i) x_i \qquad \text{für } 1 \leq i \leq n-1 \text{ (Induktionsvoraussetzung)} \\
0 &= x_i \qquad\qquad\quad \text{für } 1 \leq i \leq n-1, \text{ da } \lambda_n \neq \lambda_i \\
0 &= x_n.
\end{aligned}$$

□

Insbesondere sind also Eigenvektoren $x_1, \dots, x_n$ zu verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_n$ immer linear unabhängig. Im Fall eines endlichdimensionalen Vektorraums erhalten wir damit ein einfaches hinreichendes Kriterium für die Diagonalisierbarkeit eines Endomorphismus.

KOROLLAR. *Seien V ein n-dimensionaler K-Vektorraum und $f\colon V \to V$ linear. Hat f n verschiedene Eigenwerte, so ist f diagonalisierbar.*

Um dieses Kriterium anwenden zu können, braucht man eine Methode, die Eigenwerte eines Endomorphismus f zu berechnen. Wir werden zeigen, daß die Eigenwerte von f gerade die Nullstellen eines gewissen, durch f bestimmten Polynoms sind.

7.2.2. Das charakteristische Polynom. Seien V ein n-dimensionaler K-Vektorraum und f ein Endomorphismus, also $f\colon V \to V$ linear. Sei ferner $A \in K^{n\times n}$ eine darstellende Matrix von f. Wir gehen jetzt zum Polynomring $K[t]$ über und betrachten $A - tE$ als Element von $K[t]^{n\times n}$. Man nennt

$$p_A := \det(A - tE) \in K[t]$$

das *charakteristische Polynom* von A. Ist noch $B \in K^{n\times n}$ eine weitere darstellende Matrix von f, so gilt $A \sim B$ nach dem Satz in 3.4.3, d.h. $B = CAC^{-1}$ für ein $C \in \mathrm{GL}(n, K)$. Also ergibt sich für das charakteristische Polynom p_B von B (im Ring $K[t]^{n\times n}$)

$$\begin{aligned}\det(B - tE) &= \det(CAC^{-1} - tCEC^{-1})\\ &= \det(C(A - tE)C^{-1})\\ &= \det(C)\det(A - tE)\det(C)^{-1}\\ &= \det(A - tE),\end{aligned}$$

d.h. $p_A = p_B$. Wir können also das charakteristische Polynom p_f des Endomorphismus f definieren als p_A, wobei A eine beliebige darstellende Matrix von f ist. Im Grenzfall eines Endomorphismus f des Nullraums $V = 0$ setzen wir $p_f := 1$.

Satz. *Seien V ein endlichdimensionaler K-Vektorraum, $f\colon V \to V$ linear und $\lambda \in K$. Folgende Aussagen sind äquivalent.*

(1) λ ist Eigenwert von f.
(2) $g := f - \lambda\mathrm{id}$ ist kein Monomorphismus.
(3) $g := f - \lambda\mathrm{id}$ ist kein Epimorphismus.
(4) λ ist Nullstelle des charakteristischen Polynoms von f.

Beweis. Äquivalent sind die Aussagen

λ ist Eigenwert von f
$f(x) = \lambda x$ für ein $x \in V$, $x \neq 0$
$(f - \lambda\mathrm{id})(x) = 0$ für ein $x \in V$, $x \neq 0$
$\ker(f - \lambda\mathrm{id}) \neq 0$
$f - \lambda\mathrm{id}$ ist kein Monomorphismus
$f - \lambda\mathrm{id}$ ist kein Epimorphismus — nach 3.2.3
$f - \lambda\mathrm{id}$ ist kein Isomorphismus — nach 3.2.3
$\det(f - \lambda\mathrm{id}) = 0$ — siehe 6.3
λ ist Nullstelle des charakteristischen Polynoms von f.

□

Bemerkungen. Es sei V ein endlichdimensionaler K-Vektorraum und $f\colon V \to V$ linear.

(1) Die Eigenwerte von f lassen sich also als Nullstellen des charakteristischen Polynoms von f ausrechnen.
(2) Sei λ Eigenwert von f. Den zugehörigen Eigenraum kann man wie folgt berechnen. Man fixiere eine Basis $x_1, \ldots, x_n$ von V. Sei $A = (\alpha_{ij})$ die darstellende Matrix von f bezüglich $x_1, \ldots, x_n$, also $f(x_j) = \sum_{i=1}^{n} \alpha_{ij} x_i$ für $1 \leq j \leq n$, oder in Matrizenschreibweise

$$f \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Dann sind folgende Aussagen äquivalent.

$$\begin{aligned}
x := \sum_{i=1}^{n} \xi_i x_i &\in \mathrm{Eig}(f, \lambda) \\
f(x) &= \lambda x \\
f((\xi_1 \quad \ldots \quad \xi_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) &= \lambda (\xi_1 \quad \ldots \quad \xi_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\
(\xi_1 \quad \ldots \quad \xi_n) A^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &= \lambda (\xi_1 \quad \ldots \quad \xi_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\
(\xi_1 \quad \ldots \quad \xi_n) A^t &= \lambda (\xi_1 \quad \ldots \quad \xi_n) \\
A \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} &= \lambda E \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \\
(A - \lambda E) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} &= 0.
\end{aligned}$$

Die Berechnung des zugehörigen Eigenraums läuft also auf die Lösung eines linearen Gleichungssystems hinaus.
(3) Wir zeigen noch eine Eigenschaft des charakteristischen Polynoms, die uns später bei der Herstellung der Jordanschen Normalform nützlich sein wird. Seien $U, W \subseteq V$ Unterräume mit $V = U \oplus W$. Wir nehmen ferner an, daß U und W f-invariant sind, d.h. daß $f(U) \subseteq U$ und $f(W) \subseteq W$. (Man sagt in diesem Fall, daß der Endomorphismus f durch das Paar (U, W) *reduziert* ist; siehe [**3**, p.72]). Dann läßt sich das charakteristische Polynom von f zerlegen in das Produkt der charakteristischen Polynome der Einschränkungen von f auf die beiden Unterräume, also

$$p_f = p_{f \restriction U} \cdot p_{f \restriction W}.$$

Beweis. Man wähle Basen $x_1, \ldots, x_n$ von U und $y_1, \ldots, y_m$ von W. Dann gilt für geeignete Matrizen $A \in K^{n \times n}$ und $B \in$

$K^{m\times m}$

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Also

$$\begin{aligned} p_f &= \begin{vmatrix} A - tE & 0 \\ 0 & B - tE \end{vmatrix} \\ &= \det(A - tE)\det(B - tE) \quad \text{nach 6.2.5} \\ &= p_{f\restriction U} \cdot p_{f\restriction W}. \end{aligned}$$

□

Beispiel. Wir berechnen das charakteristische Polynom $\det(A-tE)$ für

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

Es ist

$$A - tE = \begin{pmatrix} -t & -1 & 1 \\ -3 & -2-t & 3 \\ -2 & -2 & 3-t \end{pmatrix},$$

also durch Entwicklung nach der ersten Zeile

$$\begin{aligned} &\begin{vmatrix} -t & -1 & 1 \\ -3 & -2-t & 3 \\ -2 & -2 & 3-t \end{vmatrix} \\ &= -t[(-2-t)(3-t)+6] - (-1)[(-3)(3-t)+6] + [6-2(2+t)] \\ &= -t[-6-t+t^2+6] - (-1)[-9+3t+6] + [2-2t] \\ &= -t^3 + t^2 + t - 1 \end{aligned}$$

Man sieht sofort, daß 1 eine Nullstelle ist. Es gibt also ein g mit $-t^3+t^2+t-1 = (t-1)g$; durch Division findet man $g = -t^2+1$, d.h. $-t^3+t^2+t-1 = -(t-1)^2(t+1)$. Es gibt also zwei Eigenwerte $+1$ und -1, und der Eigenwert $+1$ tritt doppelt auf.

7.2.3. Algebraische und geometrische Vielfachheit. Es sei V ein endlichdimensionaler K-Vektorraum und $f\colon V \to V$ linear. Jedem Eigenwert λ von f sind in natürlicher Weise zwei Zahlen zugeordnet.

(1) Die Vielfachheit $\mu(p_f, \lambda)$ von λ als Nullstelle des charakteristischen Polynoms p_f von f; diese Zahl nennt man die *algebraische Vielfachheit* des Eigenwerts λ.

(2) Die Dimension des Eigenraums $\mathrm{Eig}(f, \lambda)$, der dem Eigenwert λ zugeordnet ist; diese Zahl nennt man die *geometrische Vielfachheit* des Eigenwerts λ.

Wir wollen uns überlegen, daß die geometrische Vielfachheit stets kleiner oder gleich der algebraischen Vielfachheit ist.

LEMMA. *Seien V ein endlichdimensionaler K-Vektorraum, $f\colon V \to V$ linear und λ Eigenwert von f. Dann gilt* $\dim \operatorname{Eig}(f,\lambda) \le \mu(p_f,\lambda)$.

BEWEIS. Sei $x_1,\dots,x_r$ eine beliebige Basis von $\operatorname{Eig}(f,\lambda)$. Man ergänze sie zu einer Basis $x_1,\dots,x_r,x_{r+1},\dots,x_n$ von V. Dann gilt $f(x_i)=\lambda x_i$ für $i=1,\dots,r$, also

$$f\begin{pmatrix} x_1\\ \vdots\\ x_r\\ x_{r+1}\\ \vdots\\ x_n\end{pmatrix} = \begin{pmatrix} \lambda & & & & & \\ & \ddots & & & & \\ & & \lambda & & & \\ * & \dots & * & * & \dots & * \\ \vdots & & & & & \\ * & \dots & * & * & \dots & * \end{pmatrix}\begin{pmatrix} x_1\\ \vdots\\ x_r\\ x_{r+1}\\ \vdots\\ x_n\end{pmatrix}.$$

Also ist $\dim \operatorname{Eig}(f,\lambda) \le \mu(p_f,\lambda)$. □

BEMERKUNG. In dem Fall, daß das charakteristische Polynom von f in Linearfaktoren zerfällt (und das ist etwa über dem Körper $\mathbb{C}$ der komplexen Zahlen stets so), werden wir in 7.4 zeigen können, daß f genau dann diagonalisierbar ist, wenn für jeden Eigenwert λ von f die geometrische Vielfachheit gleich der algebraischen Vielfachheit ist, also wenn gilt $\dim \operatorname{Eig}(f,\lambda) = \mu(p_f,\lambda)$. Wir wollen uns jetzt mit dem allgemeinen Fall befassen, daß nämlich die Dimension eines Eigenraums $\operatorname{Eig}(f,\lambda)$ eventuell zu klein ist. Es ist $\operatorname{Eig}(f,\lambda) = \ker g$ mit $g := f - \lambda\mathrm{id}$. Die Idee hier ist, die eventuell zu kleine Dimension zu vergrößern und statt des Eigenraums $\ker g$ den sogenannten Hauptraum $U_{f,\lambda} := \bigcup_{m=0}^{\infty} \ker g^m$ zu betrachten. Dies wird uns auf die Jordansche Normalform führen.

7.3. Jordansche Normalform

In diesem Abschnitt sei K stets ein Körper.

7.3.1. Jordan-Matrizen. Für jedes $\lambda \in K$ und $s \ge 1$ sei die *Jordan-Matrix* $J_{s,\lambda}$ definiert durch

$$J_{s,\lambda} := \begin{pmatrix} \lambda & 1 & & & 0\\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1\\ 0 & & & & \lambda \end{pmatrix} \in K^{s\times s}$$

und insbesondere $J_{1,\lambda} := (\lambda)$.

SATZ (Jordansche Normalform). *Gegeben sei ein n-dimensionaler K-Vektorraum V und $f\colon V \to V$ linear. Ferner sei p_f Produkt linearer Faktoren. Dann existiert eine Basis $x_1,\dots,x_n$ von V mit*

$$f\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix} = \begin{pmatrix} A_1 & & 0\\ & \ddots & \\ 0 & & A_r \end{pmatrix}\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix},$$

wobei jedes A_i eine Jordan-Matrix $J_{s,\lambda}$ ist zu einem Eigenwert λ von f.

BEMERKUNGEN. (1) Es können eventuell mehrere Jordan-Matrizen A_i zu einem Eigenwert λ gehören.

(2) Die A_i sind bis auf die Reihenfolge eindeutig bestimmt, und zwar durch f, also unabhängig von der Basis; dies wird hier nicht bewiesen.

BEWEIS. Sei etwa $p_f = (-1)^n \prod_{i=1}^{m} (t - \lambda_i)^{e_i}$ mit verschiedenen λ_i und $e_i \geq 1$. Wir führen den Beweis durch Induktion über die Anzahl m der verschiedenen Eigenwerte von f.

Schritt 1: Zerlegung von V. Betrachte $\lambda := \lambda_1$. Setze $g := f - \lambda\mathrm{id}$. Dann ist $\ker g$ der Eigenraum von f bezüglich λ. Wir untersuchen den sogenannten *Hauptraum* von f bezüglich λ, definiert als

$$U := U_{f,\lambda} := \bigcup_{m=0}^{\infty} \ker g^m;$$

hierbei ist wie üblich $g^m := g \circ g \circ \cdots \circ g$ mit m Vorkommen von g.

Wir zeigen zunächst

(7.1)

U ist f-invarianter Unterraum, d.h. U ist Unterraum mit $f(U) \subseteq U$.

Zum Beweis von (7.1) stellen wir zunächst fest, daß U ein Unterraum ist. Es genügt zu zeigen, daß aus $x, y \in U$ stets folgt $x + y \in U$. Ist nämlich $x \in \ker g^m$ und $y \in \ker g^k$ mit etwa $m \leq k$, so folgt $x \in \ker g^k$ und damit $x + y \in \ker g^k \subseteq U$. $f(U) \subseteq U$ ergibt sich wie folgt. Offenbar genügt $g(U) \subseteq U$, da $f = g + \lambda\mathrm{id}$. Sei also $x \in U$, etwa $g^m(x) = 0$. Es genügt zu zeigen $g(x) \in U$. Es ist $0 = g(g^m(x)) = g^m(g(x))$, also $g(x) \in \ker g^m \subseteq U$.

Wir setzen jetzt

$$W := \bigcap_{m=0}^{\infty} \operatorname{im} g^m$$

und zeigen

(7.2)

W ist f-invarianter Unterraum, d.h. W ist Unterraum mit $f(W) \subseteq W$.

Zum Beweis von (7.2) beachte man, daß W als Durchschnitt von Unterräumen wieder ein Unterraum ist. $f(W) \subseteq W$ ergibt sich wie folgt. Offenbar genügt wieder $g(W) \subseteq W$, da $f = g + \lambda\mathrm{id}$. Sei also $y \in W$, d.h. $y = g^m(x_m)$ für alle m, mit geeigneten $x_m \in V$. Es genügt zu zeigen $g(y) \in W$. Man erhält $g(y) = g^{m+1}(x_m) \in \operatorname{im} g^{m+1} \subseteq \operatorname{im} g^m$ für alle m.

Unser nächstes Ziel ist $V = U \oplus W$. Dazu werden wir von der Voraussetzung $\dim V < \infty$ Gebrauch machen. Wir zeigen zunächst, daß es ein $k \leq \dim V$ gibt mit $1 \leq k$ und

$$\begin{aligned} &V \supsetneq \operatorname{im} g \supsetneq \operatorname{im} g^2 \supsetneq \cdots \supsetneq \operatorname{im} g^k = \operatorname{im} g^{k+1} = \ldots, \text{ also } \operatorname{im} g^k = W. \\ &0 \subsetneq \ker g \subsetneq \ker g^2 \subsetneq \cdots \subsetneq \ker g^k = \ker g^{k+1} = \ldots, \text{ also } \ker g^k = U. \end{aligned} \tag{7.3}$$

Zum Beweis von (7.3) beachte man zunächst, daß es aus Dimensionsgründen stets ein k mit $\operatorname{im} g^k = \operatorname{im} g^{k+1}$ geben muß. Man wähle k minimal mit $\operatorname{im} g^k = \operatorname{im} g^{k+1}$. Dann gilt

$$g^k[V] = g^{k+1}[V]$$
$$g(g^k[V]) = g(g^{k+1}[V])$$

$$g^{k+1}[V] = g^{k+2}[V] \qquad \text{usw.}$$

Die zweite Aussage in (7.3) ergibt sich daraus wie folgt. Offenbar ist $0 \subseteq \ker g \subseteq \ker g^2 \subseteq \dots$. Wegen $\dim \ker g^m + \dim \operatorname{im} g^m = n$ (nach 3.2.3) folgt die Behauptung aus der ersten Aussage. $1 \leq k$ ergibt sich aus $\ker g \neq 0$ (denn $\ker g$ ist Eigenraum zu λ).

Jetzt erhält man $V = U \oplus W$ wie folgt. $\ker g^k \cap \operatorname{im} g^k = 0$: Sei $x \in \ker g^k \cap \operatorname{im} g^k$. Es ist $g^k \colon \operatorname{im} g^k \to \operatorname{im} g^{2k}$ wohldefiniert und surjektiv, also ein Isomorphismus. Wegen $x \in \operatorname{im} g^k$ und $g^k(x) = 0$ folgt $x = 0$. $\ker g^k + \operatorname{im} g^k = V$ folgt aus

$$\begin{aligned} \dim \ker g^k + \operatorname{im} g^k &= \dim \ker g^k + \dim \operatorname{im} g^k \quad \text{dritte Folgerung in 3.2.3} \\ &= n \end{aligned}$$

Schritt 2: Konstruktion einer Basis von U, bezüglich derer sich g (und damit auch $f = g + \lambda \mathrm{id}$) übersichtlich verhält. Dazu ein

LEMMA. *Sei*

$$\ker g^i = \ker g^{i-1} \oplus \langle y_1, \dots, y_m \rangle$$

mit $y_1, \dots, y_m$ linear unabhängig. Dann existieren $z_1, \dots, z_q \in \ker g^{i-1}$ (mit $q \geq 0$) so daß

$$\ker g^{i-1} = \ker g^{i-2} \oplus \langle g(y_1), \dots, g(y_m), z_1, \dots, z_q \rangle,$$

wobei $g(y_1), \dots, g(y_m), z_1, \dots, z_q$ linear unabhängig sind.

BEWEIS. a. Es ist $g(y_j) \in \ker g^{i-1}$, da $g^{i-1}(g(y_j)) = g^i(y_j) = 0$ wegen $y_j \in \ker g^i$.

b. Wir zeigen gleichzeitig

$$\begin{aligned} &\ker g^{i-2} \cap \langle g(y_1), \dots, g(y_m) \rangle = 0 \\ &g(y_1), \dots, g(y_m) \text{ sind linear unabhängig.} \end{aligned}$$

Zum Beweis genügt es zu zeigen, daß aus $\sum_{j=1}^m \alpha_j g(y_j) \in \ker g^{i-2}$ folgt $\alpha_1 = \dots = \alpha_m = 0$. Sei also $\sum_{j=1}^m \alpha_j g(y_j) \in \ker g^{i-2}$. Man erhält

$$0 = g^{i-2}\Big(\sum_{j=1}^m \alpha_j g(y_j)\Big) = g^{i-1}\Big(\sum_{j=1}^m \alpha_j y_j\Big)$$

$$\sum_{j=1}^m \alpha_j y_j \in \ker g^{i-1}$$

$$\sum_{j=1}^m \alpha_j y_j = 0 \qquad \text{da } \ker g^i = \ker g^{i-1} \oplus \langle y_1, \dots, y_m \rangle$$

$$\alpha_1 = \dots = \alpha_m = 0 \qquad \text{da } y_1, \dots, y_m \text{ linear unabhängig.}$$

c. Jetzt läßt sich das Lemma leicht beweisen: Man wähle eine Basis $u_1, \dots, u_p$ von $\ker g^{i-2}$, und ergänze $u_1, \dots, u_p, g(y_1), \dots, g(y_m)$ zu einer Basis von ganz $\ker g^{i-1}$, etwa $u_1, \dots, u_p, g(y_1), \dots, g(y_m), z_1, \dots, z_q$. □

Man erhält aus dem Lemma

$$\begin{aligned}
\ker g^k &= \ker g^{k-1} \oplus \langle x_1, \dots, x_{m_1} \rangle \\
\ker g^{k-1} &= \ker g^{k-2} \oplus \langle g(x_1), \dots, g(x_{m_1}), \\
&\qquad\qquad x_{m_1+1}, \dots, x_{m_2} \rangle \\
\ker g^{k-2} &= \ker g^{k-3} \oplus \langle g^2(x_1), \dots, g^2(x_{m_1}), \\
&\qquad\qquad g(x_{m_1+1}), \dots, g(x_{m_2}), \\
&\qquad\qquad x_{m_2+1}, \dots, x_{m_3} \rangle
\end{aligned}$$

und schließlich

$$\begin{aligned}
\ker g = 0 \oplus \langle & g^{k-1}(x_1), \dots, g^{k-1}(x_{m_1}), \\
& g^{k-2}(x_{m_1+1}), \dots, g^{k-2}(x_{m_2}), \\
& \vdots \\
& x_{m_{k-1}+1}, \dots, x_{m_k} \rangle
\end{aligned}$$

mit $1 \le m_1 \le m_2 \le \cdots \le m_k$. Einsetzen ergibt folgende Basis von $\ker g^k = U$.

$$\begin{array}{ll}
x_{i_1}, g(x_{i_1}), \dots, g^{k-1}(x_{i_1}) & \text{für } 1 \le i_1 \le m_1 \\
x_{i_2}, g(x_{i_2}), \dots, g^{k-2}(x_{i_2}) & \text{für } m_1 + 1 \le i_2 \le m_2 \\
\vdots & \\
x_{i_k} & \text{für } m_{k-1} + 1 \le i_k \le m_k.
\end{array} \tag{7.4}$$

Außerdem ist $x_{i_1} \in \ker g^k$, $x_{i_2} \in \ker g^{k-1}$ und schließlich $x_{i_k} \in \ker g$, d.h.

$$\begin{aligned}
& g^k(x_{i_1}) = 0 \\
& g^{k-1}(x_{i_2}) = 0 \\
& \vdots \\
& g(x_{i_k}) = 0
\end{aligned}$$

Jede Zeile von (7.4) hat also die Form

$$y, g(y), \dots, g^{p-1}(y) \quad \text{mit } g^p(y) = 0 \text{ und } p \le k.$$

Aus $g^p(y) = 0$ folgt

$$g \begin{pmatrix} y \\ g(y) \\ \vdots \\ g^{p-1}(y) \end{pmatrix} = \begin{pmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ 0 & & & & 0 \end{pmatrix} \begin{pmatrix} y \\ g(y) \\ \vdots \\ g^{p-1}(y) \end{pmatrix}$$

und wegen $f = g + \lambda\mathrm{id}$

$$f\begin{pmatrix} y \\ g(y) \\ \vdots \\ g^{p-1}(y) \end{pmatrix} = \begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix} \begin{pmatrix} y \\ g(y) \\ \vdots \\ g^{p-1}(y) \end{pmatrix}.$$

Insgesamt ergibt sich für die in (7.4) gefundene Basis $z_1, \ldots, z_d$ von U

$$f\begin{pmatrix} z_1 \\ \vdots \\ z_d \end{pmatrix} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_d \end{pmatrix},$$

wobei jedes A_i eine Jordan-Matrix $J_{p,\lambda}$ ist.

Schritt 3. Abschluß. Wir wissen bereits $V = U \oplus W$, $f(U) \subseteq U$ und $f(W) \subseteq W$. Aus der eben gefundenen darstellenden Matrix für $f|U$ ergibt sich

$$p_{f|U} = \pm(t - \lambda_1)^d$$

mit $d = \dim U$. Nach Bemerkung 3 in 7.2.2 gilt

$$p_f = p_{f|U} \cdot p_{f|W}.$$

Ist $W = 0$, so sind wir fertig (dann ist $p_{f|W} = 1$). Sei also W nicht der Nullraum. Dann ist $p_{f|W}(\lambda_1) \neq 0$ (denn sonst gäbe es ein $x \in W$, $x \neq 0$ mit $f(x) = \lambda_1 x$, also $x \in \ker g$ im Widerspruch zu $V = U \oplus W$ und $\ker g \subseteq U$). Daraus ergibt sich, daß in

$$p_f = \pm(t - \lambda_1)^{e_1} \prod_{i=2}^{m} (t - \lambda_i)^{e_i} = \pm(t - \lambda_1)^d p_{f|W}$$

$e_1 = d$ sein muß (denn kürzt man gemeinsame Faktoren $t - \lambda_1$ heraus, so kann rechts und links nichts übrig bleiben). Also ist

$$p_{f|W} = \pm \prod_{i=2}^{m} (t - \lambda_i)^{e_i}.$$

$p_{f|W}$ hat einen Eigenwert weniger als p_f. Nach Induktionsvoraussetzung gibt es also für $f|W$ eine Basis mit der gewünschten Eigenschaft. Für $f|U$ hatten wir dies in Schritt 2 bewiesen. Damit ergibt sich der Satz. □

BEMERKUNGEN. (1) Die Bedingung „p_f ist Produkt linearer Faktoren“ im Satz ist notwendig.

BEWEIS. Sei $x_1, \ldots, x_n$ eine Basis von V mit

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

mit Jordan-Matrizen A_i. Dann gilt

$$p_f = p_{A_1} p_{A_2} \ldots p_{A_r},$$
$$p_{A_i} = \pm(t - \lambda_i)^{n_i}.$$

□

(2) In der Algebra zeigt man: Zu jedem $f \in K[t]$ existiert ein Erweiterungskörper $L \supseteq K$, in dem f in Linearfaktoren zerfällt.
(3) In „algebraisch abgeschlossenen Körpern“ zerfällt jedes Polynom in Linearfaktoren. Z.B. ist $\mathbb{C}$ algebraisch abgeschlossen (Fundamentalsatz der Algebra; Beweis in der Funktionentheorie).

7.3.2. Umformulierung für Matrizen. Aus den Überlegungen des vorangehenden Abschnitts ergibt sich eine nützliche Fassung der Jordanschen Normalform für Matrizen.

Satz (Jordansche Normalform für Matrizen). *Seien K ein Körper und $A \in K^{n\times n}$. Ferner sei p_A Produkt linearer Faktoren. Dann findet man ein $S \in \mathrm{GL}(n,K)$ so daß*

$$SAS^{-1} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix},$$

wobei jedes A_i eine Jordan-Matrix $J_{s,\lambda}$ ist zu einem Eigenwert λ von f_A.

Beweis. Wir betrachten noch einmal den Beweis zum Satz über die Jordansche Normalform, und zwar für $V = K^n$ und $f = f_B$, $B := A^t$. In (7.4) hatten wir für $g = f - \lambda\mathrm{id}$ die folgende Basis von $\ker g^k = U$ erhalten.

$$\begin{array}{ll} x_{i_1}, g(x_{i_1}), \ldots, g^{k-1}(x_{i_1}) & \text{für } 1 \le i_1 \le m_1 \\ x_{i_2}, g(x_{i_2}), \ldots, g^{k-2}(x_{i_2}) & \text{für } m_1 + 1 \le i_2 \le m_2 \\ \vdots & \\ x_{i_k} & \text{für } m_{k-1} + 1 \le i_k \le m_k. \end{array}$$

Zur Vereinfachung nehmen wir $1 = m_1 = m_2 = \cdots = m_k$ an und schreiben x für x_{i_1}. Sei T die Matrix mit den Spalten $x, g(x), \ldots, g^{k-1}(x)$. Dann gilt für $1 \le i \le k$

$$\begin{aligned} Bg^{i-1}(x) = f_B(g^{i-1}(x)) = (\lambda\mathrm{id} + g)g^{i-1}(x) = \lambda g^{i-1}(x) + g^i(x) \\ = i\text{-te Spalte von } T \begin{pmatrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix} \end{aligned}$$

$$BT = T \begin{pmatrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix}$$

also

$$T^{-1}BT = \begin{pmatrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix},$$

$$T^tA(T^t)^{-1} = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix}.$$

□

BEISPIEL. Sei

$$A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}, \text{ also } B = A^t = \begin{pmatrix} 3 & -1 & 1 \\ 4 & 0 & 2 \\ 3 & -1 & 3 \end{pmatrix}.$$

Dann ist

$$\begin{aligned} p_B = \begin{vmatrix} 3-t & -1 & 1 \\ 4 & -t & 2 \\ 3 & -1 & 3-t \end{vmatrix} &= (3-t)[-3t+t^2+2] + \underbrace{[12-4t-6]+[-4+3t]}_{2-t} \\ &= (3-t)(t-1)(t-2) - (t-2) \\ &= (t-2)[-t^2+4t-4] \\ &= -(t-2)^3. \end{aligned}$$

Sei $g = f_B - 2\mathrm{id}$. Dann ist

$$0 \subsetneq \ker g \subsetneq \ker g^2 \subsetneq \ker g^3 = \mathbb{R}^3.$$

Wir folgen jetzt dem Beweis des Satzes und konstruieren eine Basis von $\mathbb{R}^3$, bezüglich derer sich g übersichtlich verhält. Die Zahl k im Satz ist hier 3. Wir müssen Basen der Räume $\ker g$ und $\ker g^2$ berechnen. Es ist $g = f_C$ mit

$$C = \begin{pmatrix} 1 & -1 & 1 \\ 4 & -2 & 2 \\ 3 & -1 & 1 \end{pmatrix}, \quad \text{also} \quad C^2 = \begin{pmatrix} 0 & 0 & 0 \\ 2 & -2 & 2 \\ 2 & -2 & 2 \end{pmatrix}.$$

$\ker g$ ergibt sich als Lösungsraum des linearen Gleichungssystems $Cx = 0$:

$$\begin{cases} \xi_1 - \xi_2 + \xi_3 = 0 \\ 4\xi_1 - 2\xi_2 + 2\xi_3 = 0 \end{cases}$$

$$\begin{cases} \xi_1 - \xi_2 + \xi_3 = 0 \\ 2\xi_2 - 2\xi_3 = 0 \end{cases}$$

$$\begin{cases} \xi_1 - \xi_2 + \xi_3 = 0 \\ \xi_2 - \xi_3 = 0 \end{cases}$$

Also ist

$$\ker g = \langle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \rangle.$$

Ferner ist

$$\ker g^2 = \{ \begin{pmatrix}\xi_1\\ \xi_2\\ \xi_3\end{pmatrix} \mid \xi_1 - \xi_2 + \xi_3 = 0 \} = \langle \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\1\end{pmatrix} \rangle.$$

Man erhält

$$\begin{aligned}
\mathbb{R}^3 = \ker g^3 = \quad & \underbrace{\ker g^2}_{\langle \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\1\end{pmatrix} \rangle} \quad \oplus \langle \underbrace{x_1}_{\begin{pmatrix}1\\0\\0\end{pmatrix}} \rangle, && \text{also } m_1 = 1\\
\ker g^2 = \ & \underbrace{\ker g}_{\langle \begin{pmatrix}0\\1\\1\end{pmatrix} \rangle} \oplus \langle \underbrace{g(x_1)}_{\begin{pmatrix}1\\4\\3\end{pmatrix}} \rangle, && \text{also } m_2 = m_1\\
\ker g = 0 \oplus & \langle \underbrace{g^2(x_1)}_{\begin{pmatrix}0\\2\\2\end{pmatrix}} \rangle, && \text{also } m_3 = m_2.
\end{aligned}$$

Einsetzung ergibt die folgende Basis von $\ker g^3 = \mathbb{R}^3$:

$$x_1, g(x_1), g^2(x_1) = \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\4\\3\end{pmatrix}, \begin{pmatrix}0\\2\\2\end{pmatrix}.$$

Sei T die Matrix mit den Spalten $x_1, g(x_1), g^2(x_1)$, also

$$T = \begin{pmatrix}1 & 1 & 0\\ 0 & 4 & 2\\ 0 & 3 & 2\end{pmatrix}.$$

Dann ist, wie wir aus der Theorie wissen,

$$T^t A (T^t)^{-1} = \begin{pmatrix}2 & 1 & 0\\ 0 & 2 & 1\\ 0 & 0 & 2\end{pmatrix}$$

Man kann dies (etwa mit MAPLE) leicht kontrollieren.

BEISPIEL. Sei

$$A = \begin{pmatrix}2 & & & & \\ -1 & 2 & & & \\ -1 & -1 & 2 & & \\ -1 & -1 & -1 & 2 & \\ -1 & -1 & -1 & -1 & 2\end{pmatrix}, \text{ also } B = \begin{pmatrix}2 & -1 & -1 & -1 & -1\\ & 2 & -1 & -1 & -1\\ & & 2 & -1 & -1\\ & & & 2 & -1\\ & & & & 2\end{pmatrix}$$

und damit $p_B = (2-t)^5$. Sei wieder $g = f_B - 2\mathrm{id}$. Dann ist

$$0 \subsetneq \ker g \subsetneq \ker g^2 \subsetneq \ker g^3 \subsetneq \ker g^4 \subsetneq \ker g^5 = \mathbb{R}^5.$$

Es ist $g = f_C$ mit

$$C = \begin{pmatrix} 0 & -1 & -1 & -1 & -1 \\ & 0 & -1 & -1 & -1 \\ & & 0 & -1 & -1 \\ & & & 0 & -1 \\ & & & & 0 \end{pmatrix},$$

also

$$C^2 = \begin{pmatrix} 0 & 0 & 1 & 2 & 3 \\ & 0 & 0 & 1 & 2 \\ & & 0 & 0 & 1 \\ & & & 0 & 0 \\ & & & & 0 \end{pmatrix}, \; C^3 = \begin{pmatrix} 0 & 0 & 0 & -1 & -3 \\ & 0 & 0 & 0 & -1 \\ & & 0 & 0 & 0 \\ & & & 0 & 0 \\ & & & & 0 \end{pmatrix}$$

und

$$C^4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 \\ & & & 0 & 0 \\ & & & & 0 \end{pmatrix}.$$

Man erhält

$$\mathbb{R}^5 = \ker g^5 = \underbrace{\ker g^4}_{\langle e_1,e_2,e_3,e_4\rangle} \oplus \langle \underbrace{x_1}_{e_5} \rangle, \quad \text{also } m_1 = 1$$

$$\ker g^4 = \underbrace{\ker g^3}_{\langle e_1,e_2,e_3\rangle} \oplus \langle \underbrace{g(x_1)}_{\begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \\ 0 \end{pmatrix}} \rangle, \quad \text{also } m_2 = m_1$$

$$\ker g^3 = \underbrace{\ker g^2}_{\langle e_1,e_2\rangle} \oplus \langle \underbrace{g^2(x_1)}_{\begin{pmatrix} 3 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}} \rangle, \quad \text{also } m_3 = m_2$$

$$\ker g^2 = \underbrace{\ker g}_{\langle e_1\rangle} \oplus \langle \underbrace{g^3(x_1)}_{\begin{pmatrix} -3 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}} \rangle, \quad \text{also } m_4 = m_3$$

$$\ker g = 0 \oplus \langle \underbrace{g^4(x_1)}_{\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}} \rangle, \qquad \text{also } m_5 = m_4.$$

Einsetzung ergibt die folgende Basis von $\ker g^5 = \mathbb{R}^5$:

$$x_1, g(x_1), g^2(x_1), g^3(x_1), g^4(x_1) = \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}-1\\-1\\-1\\-1\\0\end{pmatrix}, \begin{pmatrix}3\\2\\1\\0\\0\end{pmatrix}, \begin{pmatrix}-3\\-1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}.$$

Sei T die Matrix mit diesen Spalten, also

$$T^t = \begin{pmatrix}0 & 0 & 0 & 0 & 1\\ -1 & -1 & -1 & -1 & 0\\ 3 & 2 & 1 & 0 & 0\\ -3 & -1 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0\end{pmatrix}.$$

Dann ist wieder, wie wir aus der Theorie wissen, $T^t A (T^t)^{-1}$ die Jordan-Matrix $J_{5,2}$. Man kann dies mit MAPLE leicht kontrollieren. Die Eingaben lauten:

```
with(linalg):
T:=matrix([[0,-1,3,-3,1],[0,-1,2,-1,0],[0,-1,1,0,0],[0,-1,0,0,0],[1,0,0,0,0]]):
A:=matrix([[2,0,0,0,0],[-1,2,0,0,0],[-1,-1,2,0,0],[-1,-1,-1,2,0],[-1,-1,-1,-1,2]]):
evalm(transpose(T)*A*inverse(transpose(T)));
```

und die Ausgabe ist

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

7.4. Diagonalisierung

Seien V ein endlichdimensionaler K-Vektorraum und $f\colon V \to V$ linear. f hatten wir *diagonalisierbar* genannt, wenn es eine Basis $x_1, \ldots, x_n$ von V und $\lambda_1, \ldots, \lambda_n \in K$ gibt so daß

$$f \begin{pmatrix}x_1\\ \vdots \\ x_n\end{pmatrix} = \begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix} \begin{pmatrix}x_1\\ \vdots \\ x_n\end{pmatrix}.$$

7.4.1. Diagonalisierbarkeit von Endomorphismen. Wir beweisen ein einfaches Kriterium für die Diagonalisierbarkeit eines Endomorphismus.

SATZ. *Seien V ein n-dimensionaler K-Vektorraum und $f\colon V \to V$ linear. Ferner sei das charakteristische Polynom p_f Produkt linearer Faktoren, etwa $p_f = (-1)^n \prod_{i=1}^m (t-\lambda_i)^{e_i}$ mit verschiedenen λ_i und $e_i \geq 1$. f ist genau dann diagonalisierbar, wenn für alle Eigenwerte λ_i die geometrische Vielfachheit $\dim \mathrm{Eig}(f,\lambda_i)$ gleich der algebraischen Vielfachheit e_i von λ_i ist.*

BEWEIS. 1. Sei f diagonalisierbar, also $V = V_1 \oplus \cdots \oplus V_m$ mit $V_i = \mathrm{Eig}(f,\lambda_i)$. Sei $(x_{ij})_{1\leq j\leq e_i}$ Basis von V_i. Dann gilt

$$f\begin{pmatrix} x_{11} \\ \vdots \\ x_{1e_1} \\ \vdots \\ x_{m1} \\ \vdots \\ x_{me_m} \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_m & & \\ & & & & & \ddots & \\ & & & & & & \lambda_m \end{pmatrix} \begin{pmatrix} x_{11} \\ \vdots \\ x_{1e_1} \\ \vdots \\ x_{m1} \\ \vdots \\ x_{me_m} \end{pmatrix} \tag{7.5}$$

mit e_i Vorkommen von λ_i. Man erhält für das charakteristische Polynom

$$p_f = \det(A - tE) = \prod_{i=1}^m (\lambda_i - t)^{e_i}$$

Also ist e_i sowohl die Vielfachheit von λ_i als auch die Dimension von V_i.

2. Für jeden Eigenwert λ_i von f gelte $\dim V_i = e_i$ mit $V_i = \mathrm{Eig}(f,\lambda_i)$. Sei $(x_{ij})_{1\leq j\leq e_i}$ Basis von V_i. Wegen $\sum_{i=1}^m e_i = n$ ist dann $(x_{ij})_{1\leq i\leq m, 1\leq j\leq e_i}$ eine Basis von V, und wir haben wieder (7.5). □

BEMERKUNG. Man kann auch dem obigen Beweis des Satzes von der Jordanschen Normalform entnehmen, daß f diagonalisierbar ist, falls für jeden Eigenwert λ_i von f gilt $\dim \mathrm{Eig}(f,\lambda_i) = e_i$. Setze $g_i := f - \lambda_i \mathrm{id}$. Führt man die im Beweis beschriebene Konstruktion für λ_i durch, so ergibt sich für den Hauptraum $U = U_{f,\lambda_i} = \ker g_i^k$, daß $\dim U = e_i$ ist. Da auch $\dim \mathrm{Eig}(f,\lambda_i) = e_i$ ist für den Eigenraum $\mathrm{Eig}(f,\lambda_i) = \ker g_i$, haben wir $U = \mathrm{Eig}(f,\lambda_i)$ und $k = 1$. Alle Jordan-Matrizen in der durchgeführten Konstruktion für den Eigenwert λ_i haben also die Form (λ_i). Da dies für alle λ_i gilt, ist f diagonalisierbar.

7.4.2. Diagonalisierung für Matrizen. Wir wollen wieder eine Fassung der gerade durchgeführten Überlegungen für Matrizen angeben.

SATZ (Diagonalisierung von Matrizen). *Seien K ein Körper und $A \in K^{n\times n}$. Ferner sei das charakteristische Polynom p_A Produkt linearer Faktoren, und für jeden Eigenwert sei die geometrische Vielfachheit gleich der algebraischen Vielfachheit. Dann findet man ein $S \in \mathrm{GL}(n,K)$ so daß*

$$SAS^{-1} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

BEWEIS. Sei $B := A^t$. Aufgrund der Voraussetzung gibt es eine Basis $x_1, \ldots, x_n$ des K^n aus Eigenvektoren von f_B. Sei x_i Eigenvektor zum Eigenwert λ_i, und T die Matrix mit den Spalten $x_1, \ldots, x_n$. Dann gilt

$$Bx_i = f_B(x_i) = \lambda_i x_i = i\text{-te Spalte von } T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$$BT = T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$$T^{-1}BT = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$$T^t A (T^t)^{-1} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

□

KOROLLAR (Praktisches Verfahren zur Diagonalisierung). *Seien $A \in K^{n\times n}$ und $B := A^t$.*

1. *Man bestimme p_B und versuche, p_B in Linearfaktoren zu zerlegen. Wenn dies unmöglich ist, so ist f_B nicht diagonalisierbar. Andernfalls:*
2. *Für jeden Eigenwert λ bestimme man $\dim \mathrm{Eig}(f_B, \lambda)$. Wenn für einen Eigenwert λ gilt $\dim \mathrm{Eig}(f_B, \lambda) < \mu(p_B, \lambda)$, so ist f_B nicht diagonalisierbar. Andernfalls:*
3. *Für jeden Eigenwert λ bestimme man eine Basis von $\mathrm{Eig}(f_B, \lambda)$ (durch Lösen des linearen Gleichungssystems $(B - \lambda E)x = 0$). Dann ist A diagonalisierbar mit Transformationsmatrix gemäß dem Beweis des Satzes.*

BEISPIEL. Sei

$$A = \begin{pmatrix} 0 & -3 & -2 \\ -1 & -2 & -2 \\ 1 & 3 & 3 \end{pmatrix}, \text{ also } B = A^t = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}$$

In Abschnitt 7.2 hatten wir das charakteristische Polynom $p_A = p_B = -(t-1)^2(t+1)$ bestimmt. Zu den Eigenwerten $+1$ und -1 berechnen wir jetzt die Eigenvektoren.

Für den Eigenwert -1 ergeben sich die Eigenvektoren als Lösungen des linearen Gleichungssystems

$$\begin{pmatrix} 0+1 & -1 & 1 \\ -3 & -2+1 & 3 \\ -2 & -2 & 3+1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = 0.$$

Wir bringen diese Matrix durch elementare Zeilenumformungen in die Form einer Zeilenstufenmatrix:

$$\begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix} \mapsto \begin{matrix} \\ \text{zweite Zeile } +3\cdot \text{ erste Zeile} \\ \text{dritte Zeile } +2\cdot \text{ erste Zeile} \end{matrix} \begin{pmatrix} 1 & -1 & 1 \\ 0 & -4 & 6 \\ 0 & -4 & 6 \end{pmatrix}$$

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

Zu lösen ist also das Gleichungssystem

$$\begin{aligned} \xi_1 - \xi_2 + \xi_3 &= 0 \\ -2\xi_2 + 3\xi_3 &= 0. \end{aligned}$$

Nach dem Verfahren aus 1.4 erhält man als Lösungsraum alle skalaren Vielfachen von $\begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \\ 1 \end{pmatrix}$. Insbesondere ist also $\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$ Eigenvektor zum Eigenwert -1.

Für den Eigenwert 1 ergeben sich die Eigenvektoren als Lösungen des linearen Gleichungssystems

$$\begin{pmatrix} 0-1 & -1 & 1 \\ -3 & -2-1 & 3 \\ -2 & -2 & 3-1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = 0.$$

Wir bringen diese Matrix durch elementare Zeilenumformungen in die Form einer Zeilenstufenmatrix:

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

Das zugehörige Gleichungssystem besteht nur aus der Gleichung $-\xi_1 - \xi_2 + \xi_3 = 0$. Nach dem Verfahren aus Abschnitt 1.4 hat der Lösungsraum (also $\operatorname{Eig}(f_B, 1)$) die Basis $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

Es ist $\dim \operatorname{Eig}(f_B, -1) = 1 = \mu(p_B, -1)$ und $\dim \operatorname{Eig}(f_B, 1) = 3 - 1 = 2 = \mu(p_B, 1)$. Daher ist f_B diagonalisierbar.

$\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ ist eine Basis des $\mathbb{R}^3$ aus Eigenvektoren von f_B. Wir setzen gemäß dem Beweis des Satzes

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

Dann gilt, wie man leicht nachrechnet,

$$T^t A (T^t)^{-1} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

KAPITEL 8

Komplexifizierung und unitäre Vektorräume

Im Satz über die Jordan-Normalform mußten wir voraussetzen, daß das charakteristische Polynom p_f des gegebenen Endomorphismus $f\colon V \to V$ in Linearfaktoren zerfällt. Dies ist stets für den Körper $\mathbb{C}$ der komplexen Zahlen der Fall, aber nicht für allgemeine Körper und insbesondere nicht für den Körper $\mathbb{R}$ der reellen Zahlen. Wir zeigen jetzt, wie man diesen „Defekt“ des Körpers der reellen Zahlen dadurch umgehen kann, daß man einen gegebenen $\mathbb{R}$-Vektorraum in kanonischer Weise zu einem $\mathbb{C}$-Vektorraum macht. Als Anwendung erhalten wir eine Variante des Satzes über die Jordan-Normalform für $\mathbb{R}$-Vektorräume. Ferner führen wir eine ebenfalls kanonische Erweiterung der Theorie der euklidischen Vektorräume (die ja die reellen Zahlen als Grundkörper voraussetzten) auf den wichtigen Fall der komplexen Zahlen als Grundkörper durch. Man spricht dann von unitären Vektorräumen.

8.1. Komplexifizierung

8.1.1. Komplexe Zahlen. In 1.2.4 hatten wir den Körper $\mathbb{C}$ der komplexen Zahlen als Spezialfall einer Körpererweiterung $K(\sqrt{\xi})$ für $\xi = -1$ eingeführt. Mit einer geeigneten Identifikation galt $\mathbb{R} \subseteq \mathbb{C}$. Offenbar kann man $\mathbb{C}$ als $\mathbb{R}$-Vektorraum mit Basis $1, i$ auffassen. Ferner gilt $(\alpha + i\beta)(\gamma + i\delta) = \alpha\gamma - \beta\delta + i(\alpha\delta + \beta\gamma)$.

Bezeichnungen. Sei $\xi = \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$.

$\mathbb{R}e(\xi) := \alpha$	*Realteil*
$\Im(\xi) := \beta$	*Imaginärteil*
$\overline{\xi} := \alpha - i\beta$	*konjugiert komplexe* Zahl
$\lvert\xi\rvert := \sqrt{\alpha^2 + \beta^2}$	*Betrag*.

Es gilt:

$$\begin{aligned} \mathbb{R}e(\xi) &= \frac{1}{2}(\xi + \overline{\xi}) \\ \Im(\xi) &= \frac{1}{2i}(\xi - \overline{\xi}) \\ |\xi|^2 &= \xi\overline{\xi} \\ \frac{1}{\xi} &= \frac{\overline{\xi}}{\xi\overline{\xi}} = \frac{1}{|\xi|^2}\overline{\xi} \end{aligned}$$

Geometrische Deutung (Gaußsche Zahlenebene): $\xi = \alpha + i\beta$, $r = |\xi|$, $\cos\varphi = \frac{\alpha}{r}$, $\sin\varphi = \frac{\beta}{r}$. Also: $\xi = \alpha + i\beta = r(\cos\varphi + i\sin\varphi)$. Addition in $\mathbb{C}$: Vektoraddition in $\mathbb{R}^2$. Multiplikation in $\mathbb{C}$: Beträge multiplizieren, Winkel addieren,

denn

$$\begin{aligned} &r(\cos\varphi + i\sin\varphi)\cdot r'(\cos\varphi' + i\sin\varphi') \\ &= rr'(\underbrace{\cos\varphi\cos\varphi' - \sin\varphi\sin\varphi'}_{\cos(\varphi+\varphi')} + i(\underbrace{\cos\varphi\sin\varphi' + \sin\varphi\cos\varphi'}_{\sin(\varphi+\varphi')})). \end{aligned}$$

Formale Definition (Erinnerung): Es war $\mathbb{C} := \mathbb{R}\times\mathbb{R}$ als additive Gruppe definiert. Die Multiplikation hatten wir definiert durch

$$(\alpha,\beta)\cdot(\gamma,\delta) = (\alpha\gamma - \beta\delta, \alpha\delta + \beta\gamma)$$

für $\alpha,\beta,\gamma,\delta \in \mathbb{R}$.

BEMERKUNGEN. (1) Für $\xi,\xi' \in \mathbb{C}$ gilt:

$$\begin{aligned} &|\xi\xi'| = |\xi||\xi'| \\ &|\xi+\xi'| \le |\xi| + |\xi'| \\ &|\xi| = 0 \quad \text{genau dann, wenn} \quad \xi = 0 \\ &\overline{\xi\xi'} = \overline{\xi}\,\overline{\xi'} \\ &\overline{\xi+\xi'} = \overline{\xi} + \overline{\xi'} \\ &\overline{\overline{\xi}} = \xi \\ &\overline{\xi} = \xi \quad \text{genau dann, wenn} \quad \xi \in \mathbb{R}. \end{aligned}$$

BEWEIS. Durch Ausrechnen. □

(2) Sei $\sigma\colon \mathbb{C}\to\mathbb{C}$ ein $\mathbb{R}$-*Algebra-Automorphismus*, d.h.

$$\begin{aligned} &\sigma(\alpha) = \alpha && \text{für alle } \alpha\in\mathbb{R} \\ &\sigma(\xi+\xi') = \sigma(\xi)+\sigma(\xi') && \text{für alle } \xi,\xi'\in\mathbb{C} \\ &\sigma(\xi\cdot\xi') = \sigma(\xi)\cdot\sigma(\xi') && \text{für alle } \xi,\xi'\in\mathbb{C} \\ &\sigma \text{ bijektiv.} \end{aligned}$$

(Die letzten drei Eigenschaften besagen, daß σ ein *Ringisomorphismus* ist). Dann ist σ die Identität, d.h. $\sigma(\xi) = \xi$ für alle $\xi\in\mathbb{C}$, oder σ ist die Konjugation, d.h. $\sigma(\xi) = \overline{\xi}$ für alle $\xi\in\mathbb{C}$.

BEWEIS. Sei $\xi = \alpha + i\beta$ mit $\alpha,\beta\in\mathbb{R}$. Dann ist

$$\sigma(\xi) = \underbrace{\sigma(\alpha)}_{\alpha} + \sigma(i)\underbrace{\sigma(\beta)}_{\beta} = \alpha + \sigma(i)\beta$$

Es genügt zu zeigen $\sigma(i) = i$ (dann ist σ die Identität) oder $\sigma(i) = -i$ (dann ist σ die Konjugation). Dies folgt aus:

$$\begin{aligned} &i^2 = -1 \\ &-1 = \sigma(-1) = \sigma(i^2) = \sigma(i)\cdot\sigma(i) \\ &\sigma(i)^2 + 1 = 0 \\ &(\sigma(i)+i)(\sigma(i)-i) = 0 \\ &\sigma(i) = i \quad \text{oder} \quad \sigma(i) = -i. \end{aligned}$$

□

(3) Fortsetzung der Konjugation auf den Polynomring $\mathbb{C}[t]$. Für

$$p = \sum_{j=0}^{n} \alpha_j t^j \in \mathbb{C}[t] \text{ mit } \alpha_j \in \mathbb{C} \text{ sei}$$
$$\overline{p} := \sum_{j=0}^{n} \overline{\alpha_j} t^j.$$

Wieder gilt für $p, q \in \mathbb{C}[t]$

$$\overline{p \cdot q} = \overline{p} \cdot \overline{q},$$
$$\overline{p + q} = \overline{p} + \overline{q},$$
$$\overline{\overline{p}} = p,$$
$$\overline{p} = p \quad \text{genau dann, wenn} \quad p \in \mathbb{R}[t].$$

BEWEIS. Durch Ausrechnen. Zum Beispiel für $p = \sum_{j=0}^{n} \alpha_j t^j$ und $q = \sum_{l=0}^{m} \beta_l t^l$ mit $\alpha_j, \beta_l \in \mathbb{C}$ ergibt sich

$$p \cdot q = \sum_{k=0}^{n+m} \Big(\sum_{l+j=k} \alpha_j \beta_l \Big) t^k,$$
$$\overline{p \cdot q} = \sum_{k=0}^{n+m} \underbrace{\overline{\Big(\sum_{l+j=k} \alpha_j \beta_l \Big)}}_{= \sum_{l+j=k} \overline{\alpha_j} \overline{\beta_l}} t^k = \overline{p} \cdot \overline{q}.$$

□

8.1.2. Fundamentalsatz der Algebra. Eine wesentliche Eigenschaft des Körpers $\mathbb{C}$ der komplexen Zahlen ist seine *algebraische Abgeschlossenheit*, d.h. daß jedes nichtkonstante Polynom in $\mathbb{C}[t]$ Produkt linearer Faktoren ist. Man nennt diese Aussage auch den Fundamentalsatz der Algebra. Einen Beweis werden wir im Rahmen dieser Vorlesung nicht geben können; er wird in Vorlesungen über Funktionentheorie geführt. Jedes Polynom $p \in \mathbb{C}[t]$ hat also eine Darstellung

$$p = \alpha \prod_{j=1}^{n} (t - \alpha_j) \quad \text{mit } \alpha, \alpha_j \in \mathbb{C}.$$

$\alpha_1, \ldots, \alpha_n$ sind die Nullstellen in $\mathbb{C}$ von p.

KOROLLAR. *Jedes nicht konstante Polynom $p \in \mathbb{R}[t]$ ist Produkt von linearen und quadratischen Faktoren über $\mathbb{R}$. Genauer gilt*

$$p = \alpha \prod_{j=1}^{r} (t - \alpha_j)^{e_j} \prod_{j=r+1}^{n} \Big(\underbrace{(t - \alpha_j)(t - \overline{\alpha_j})}_{\in \mathbb{R}[t] \text{ vom Grad 2}} \Big)^{e_j},$$

wobei $e_j \geq 1$, $\alpha \in \mathbb{R}$, $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$, $\alpha_{r+1}, \ldots, \alpha_n \in \mathbb{C} \setminus \mathbb{R}$, $\alpha_j \neq \alpha_l$ für $j \neq l$. ($\alpha_1, \ldots, \alpha_r$ sind die reellen, $\alpha_{r+1}, \overline{\alpha_{r+1}}, \ldots, \alpha_n, \overline{\alpha_n}$ die komplexen Nullstellen von p).

BEWEIS. Nach dem Fundamentalsatz der Algebra gibt es ein $\lambda \in \mathbb{C}$ mit

$$p = (t-\lambda)^e q \quad \text{mit } q \in \mathbb{C}[t],\ e \geq 1 \text{ und } q(\lambda) \neq 0.$$

Nach Voraussetzung ist

$$p = \overline{p} = (t-\overline{\lambda})^e \overline{q} \quad \text{mit } \overline{q}(\overline{\lambda}) = \overline{q(\lambda)} \neq 0, \text{ da } q(\lambda) \neq 0.$$

Damit ist gezeigt: ist λ Nullstelle von p mit Vielfachheit e, so ist $\overline{\lambda}$ Nullstelle von p ebenfalls mit Vielfachheit e. Außerdem gilt

$$(t-\alpha_j)(t-\overline{\alpha_j}) = t^2 - t\underbrace{(\alpha_j + \overline{\alpha_j})}_{2\cdot\mathbb{R}e(\alpha_j)} + \underbrace{\alpha_j\overline{\alpha_j}}_{|\alpha_j|^2} \in \mathbb{R}[t].$$

Fortsetzung dieses Verfahrens ergibt die Behauptung. □

KOROLLAR. *Sei $p \in \mathbb{R}[t]$, p nicht konstant und $\deg(p)$ ungerade. Dann hat p eine reelle Nullstelle.*

BEWEIS. Nach Voraussetzung ist $\deg(p) = \sum_{j=1}^{r} e_j + 2\sum_{j=r+1}^{n} e_j$ ungerade. Es folgt $r \geq 1$ und damit die Behauptung. □

8.1.3. Komplexifizierung. Gegeben sei ein $\mathbb{R}$-Vektorraum V. Aus V konstruieren wir einen $\mathbb{C}$-Vektorraum $\tilde{V}$, seine *Komplexifizierung*. Dazu setzen wir $\tilde{V} := V \times V$ als additive Gruppe. Die Skalarmultiplikation $\mathbb{C}\times\tilde{V} \to \tilde{V}$ wird definiert durch

$$(\alpha + i\beta)(v_1, v_2) := (\alpha v_1 - \beta v_2, \alpha v_2 + \beta v_1).$$

BEMERKUNGEN. (1) $\tilde{V}$ mit der angegebenen Struktur ist ein $\mathbb{C}$-Vektorraum.

BEWEIS. Durch Nachrechnen, wie für $V = \mathbb{R}$. □

(2) Die Abbildung

$$\begin{aligned} V &\to \tilde{V} \\ v &\mapsto (v, 0) \end{aligned}$$

ist ein $\mathbb{R}$-*Monomorphismus*, d.h. es gilt

$$\begin{aligned} v + v' \mapsto &(v + v', 0) = (v, 0) + (v', 0), \\ \alpha \cdot v \mapsto &(\alpha v, 0) = \alpha \cdot (v, 0) \quad \text{für } \alpha \in \mathbb{R}. \end{aligned}$$

Weiter gilt

$$i \cdot (v, 0) = (0, v),$$

also

$$(v_1, v_2) = (v_1, 0) + i(v_2, 0) \quad \text{in } \tilde{V}.$$

Wir identifizieren jetzt v mit $(v, 0)$; in diesem Sinn haben wir dann $V \subseteq \tilde{V}$. Jedes $x \in \tilde{V}$ läßt sich also eindeutig schreiben in der Form

$$x = v_1 + iv_2 \quad \text{mit } v_1, v_2 \in V,$$

und es gilt

$$(\alpha + i\beta)(v_1 + iv_2) = \alpha v_1 - \beta v_2 + i(\alpha v_2 + \beta v_1).$$

8.1.4. Komplexifizierung eines Homomorphismus. Seien V, W $\mathbb{R}$-Vektorräume und $f\colon V \to W$ linear, also $f \in \mathrm{Hom}_{\mathbb{R}}(V, W)$. Wir definieren die Komplexifizierung $\tilde{f}$ von f durch

$$\begin{aligned} \tilde{f}&\colon \tilde{V} \to \tilde{W} \\ &\tilde{f}(v_1 + iv_2) = f(v_1) + if(v_2) \quad \text{für } v_1, v_2 \in V. \end{aligned}$$

Bemerkungen. (1) $\tilde{f}$ ist ein $\mathbb{C}$-Homomorphismus.

Beweis. Durch Nachrechnen, z.B.

$$\begin{aligned} &\tilde{f}((\alpha + i\beta)(v_1 + iv_2)) \\ &= f(\alpha v_1 - \beta v_2) + if(\alpha v_2 + \beta v_1) && \text{nach Definition} \\ &= \alpha f(v_1) - \beta f(v_2) + i(\alpha f(v_2) + \beta f(v_1)) \\ &= (\alpha + i\beta)\tilde{f}(v_1 + iv_2). \end{aligned}$$

□

(2) $\tilde{f}$ ist eindeutig bestimmte Fortsetzung von f auf $\tilde{V}$ („Fortsetzung“ meint $\tilde{f}\restriction V = f$).

Beweis. Sei $f'\colon \tilde{V} \to \tilde{W}$ ein $\mathbb{C}$-Homomorphismus, $f'\restriction V = f$. Dann gilt

$$\begin{aligned} f'(v_1 + iv_2) &= f'(v_1) + if'(v_2) \\ &= f(v_1) + if(v_2) \\ &= \tilde{f}(v_1 + iv_2). \end{aligned}$$

□

(3) Seien V, W, U $\mathbb{R}$-Vektorräume und $V \xrightarrow{f} W \xrightarrow{g} U$ $\mathbb{R}$-Homomorphismen. Dann gilt offenbar $\tilde{V} \xrightarrow{\tilde{f}} \tilde{W} \xrightarrow{\tilde{g}} \tilde{U}$. Wir zeigen
   a. $\widetilde{g \circ f} = \tilde{g} \circ \tilde{f}$.
   b. $\tilde{\mathrm{id}} = \mathrm{id}$.

   Man sagt in diesem Fall, daß $\tilde{}$ ein „*Funktor*“ ist.

Beweis. a.

$$\begin{aligned} \tilde{g}(\tilde{f}(v_1 + iv_2)) &= \tilde{g}(f(v_1) + if(v_2)) \\ &= g(f(v_1)) + ig(f(v_2)) \\ &= (g \circ f)(v_1) + i(g \circ f)(v_2) \\ &= \widetilde{(g \circ f)}(v_1 + iv_2) \end{aligned}$$

b. $\widetilde{\mathrm{id}}(v_1 + iv_2) = v_1 + iv_2$. □

Lemma. *Seien V n-dimensionaler $\mathbb{R}$-Vektorraum und $f \in \mathrm{Hom}_{\mathbb{R}}(V, V)$.*

(1) Ist $x_1, \ldots, x_n$ eine Basis von V über $\mathbb{R}$, so ist $x_1, \ldots, x_n$ Basis von $\tilde{V}$ über $\mathbb{C}$.

(2) Es gilt $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{C}}(\tilde{V})$ und $\dim_{\mathbb{R}}(\tilde{V}) = 2\dim_{\mathbb{R}}(V)$.

(3) Das charakteristische Polynom p_f von f ist gleich dem charakteristischen Polynom $p_{\tilde{f}}$ von $\tilde{f}$.

BEWEIS. (1) Offenbar ist $V \subseteq \tilde{V}$. $x_1, \ldots, x_n$ sind linear unabhängig über $\mathbb{C}$: Seien $\alpha_j, \beta_j \in \mathbb{R}$ und gelte $\sum_{j=1}^n (\alpha_j + i\beta_j)x_j = 0$. Dann folgt

$$\begin{aligned}
&\sum_{j=1}^n \alpha_j x_j + i \sum_{j=1}^n \beta_j x_j = 0 \\
&\sum_{j=1}^n \alpha_j x_j = 0 \quad \text{und} \quad \sum_{j=1}^n \beta_j x_j = 0 \\
&\alpha_1 = \cdots = \alpha_n = 0 \quad \text{und} \quad \beta_1 = \cdots = \beta_n = 0, \\
&\quad \text{da } x_1, \ldots, x_n \ \mathbb{R}\text{-linear unabhängig sind.}
\end{aligned}$$

$x_1, \ldots, x_n$ bilden ein Erzeugendensystem über $\mathbb{C}$ von $\tilde{V}$: Sei $v + iw \in \tilde{V}$ mit $v, w \in V$. Dann gilt $v = \sum_{j=1}^n \alpha_j x_j$ und $w = \sum_{j=1}^n \beta_j x_j$ mit $\alpha_j, \beta_j \in \mathbb{R}$, also $\sum_{j=1}^n (\alpha_j + i\beta_j)x_j = v + iw$.

(2) $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{C}}(\tilde{V})$ gilt nach Teil (1). Es ist $\dim_{\mathbb{R}}(\tilde{V}) = 2n$, da $x_1, \ldots, x_n, ix_1, \ldots, ix_n$ eine $\mathbb{R}$-Basis von $\tilde{V}$ bildet.

(3) Seien $x_1, \ldots, x_n$ Basis von V über $\mathbb{R}$ und A darstellende Matrix von f bezüglich $x_1, \ldots, x_n$.

$$f \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Also gilt $p_f = \det(A - tE)$. Nach Teil (1) ist wegen

$$\tilde{f} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

A auch darstellende Matrix von $\tilde{f}$. Also hat man $p_{\tilde{f}} = \det(A - tE) = p_f$. □

KOROLLAR. *Seien V ein endlichdimensionaler $\mathbb{R}$-Vektorraum und $f \in \mathrm{Hom}_{\mathbb{R}}(V, V)$. Dann gilt:*

(1) Ist $\dim V$ ungerade, so besitzt f mindestens einen Eigenvektor, d.h. es gibt einen Unterraum $U \subseteq V$ der Dimension 1, welcher f-invariant ist (also $f(U) \subseteq U$).
(2) Ist $\dim V$ beliebig (aber ≥ 1), so gibt es einen f-invarianten Unterraum $U \subseteq V$ der Dimension 1 oder 2.

BEWEIS. (2) Betrachte $\tilde{f} \in \mathrm{Hom}_{\mathbb{C}}(\tilde{V}, \tilde{V})$. Wegen $\dim V \geq 1$ hat $p_{\tilde{f}}$ eine Nullstelle λ über $\mathbb{C}$. Sei etwa $\lambda =: \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$. λ ist dann Eigenwert von $\tilde{f}$. Sei $x = v + iw \in \tilde{V}$ Eigenvektor von $\tilde{f}$ zum Eigenwert λ (also $x \neq 0$

und $v, w \in V$). Man erhält

$$\tilde{f}(x) = \lambda x$$
$$\tilde{f}(v + iw) = (\alpha + i\beta)(v + iw)$$
$$f(v) + if(w) = \alpha v - \beta w + i(\alpha w + \beta v)$$
$$\begin{cases} f(v) = \alpha v - \beta w \\ f(w) = \alpha w + \beta v \end{cases}$$
$$f\begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}\begin{pmatrix} v \\ w \end{pmatrix}$$

Der Unterraum $U := \mathbb{R}v + \mathbb{R}w$ leistet also das Verlangte.

(1) p_f hat eine reelle Nullstelle $\alpha \in \mathbb{R}$ (nach einer Folgerung in 8.1.2). Also ist α ein Eigenwert von f. Sei $v \neq 0$, $v \in V$ Eigenvektor von f zum Eigenwert α. Dann ist $f(v) = \alpha v$. Der Unterraum $U := \mathbb{R}v$ leistet also das Verlangte. □

8.1.5. Konjugation. Sei V ein $\mathbb{R}$-Vektorraum. Die Abbildung

$$\tilde{V} \to \tilde{V}$$
$$v + iw \mapsto v - iw \quad \text{für } v, w \in V$$

heißt *Konjugation* auf $\tilde{V}$.

Bemerkungen. (1) Für alle $x, y \in \tilde{V}$ und $\alpha \in \mathbb{C}$ gilt

$$\overline{x + y} = \overline{x} + \overline{y}$$
$$\overline{\alpha x} = \overline{\alpha}\,\overline{x}$$
$$\overline{\overline{x}} = x$$
$$\overline{x} = x \quad \text{genau dann, wenn} \quad x \in V.$$

Dies zeigt man wie für $\mathbb{C} = \tilde{\mathbb{R}}$. Zum Beispiel zeigt man die letzte Äquivalenz wie folgt. Sei $x = v + iw$ mit $v, w \in V$, also $\overline{x} = v - iw$. Dann sind folgende Aussagen äquivalent.

$$x = \overline{x}$$
$$v + iw = v - iw$$
$$2iw = 0$$
$$w = 0$$
$$x \in V.$$

(2) Sei wieder $x = v + iw$ mit $v, w \in V$. Dann gilt

$$v = \frac{1}{2}(x + \overline{x}), \quad w = \frac{1}{2i}(x - \overline{x}).$$

8.1.6. Abstieg von $\mathbb{C}$ nach $\mathbb{R}$. Wir zeigen, wann man aus dem Vorhandensein von Homomorphismen oder Unterräumen in der Komplexifizierung $\tilde{V}$ schließen kann auf das Vorhandensein entsprechender Objekte in V.

Satz. *Seien V, W $\mathbb{R}$-Vektorräume.*

(1) Für alle $x \in \tilde{V}$ gilt $x \in V$ genau dann, wenn $x = \overline{x}$.

(2) *Ein beliebiges* $g \in \mathrm{Hom}_{\mathbb{C}}(\tilde{V}, \tilde{W})$ *hat die Form* $g = \tilde{f}$ *für ein* $f \in \mathrm{Hom}_{\mathbb{R}}(V, W)$ *genau dann, wenn* $\overline{g(x)} = g(\overline{x})$ *für alle* $x \in \tilde{V}$. *In diesem Fall ist* $f = g{\restriction}V$.

(3) *Für einen beliebigen* $\mathbb{C}$*-Unterraum* U' *von* $\tilde{V}$ *gilt: Es gibt einen* $\mathbb{R}$*-Unterraum* U *von* V *mit* $U' = \tilde{U}$ *genau dann, wenn für alle* $x \in U'$ *gilt* $\overline{x} \in U'$, *d.h. wenn* U' *invariant ist unter Konjugation. In diesem Fall ist* $U = \{\, x \in U' \mid x = \overline{x} \,\}$.

Beweis. (1) Dies war oben als Bemerkung (1) bewiesen.

(2) Sei $g \in \mathrm{Hom}_{\mathbb{C}}(\tilde{V}, \tilde{W})$. Nehmen wir zunächst an, daß $g = \tilde{f}$ für ein $f \in \mathrm{Hom}_{\mathbb{R}}(V, W)$. Dann gilt für ein beliebiges $x = v_1 + iv_2$ mit $v_1, v_2 \in V$

$$\begin{aligned}\overline{\tilde{f}(x)} &= \overline{f(v_1) + if(v_2)} \\ &= f(v_1) - if(v_2) \\ &= f(v_1) + if(-v_2) \\ &= \tilde{f}(v_1 - iv_2) \\ &= \tilde{f}(\overline{x}).\end{aligned}$$

Sei nun umgekehrt $\overline{g(x)} = g(\overline{x})$ für alle $x \in \tilde{V}$. Für alle $x \in V$ gilt dann $\overline{g(x)} = g(\overline{x}) = g(x)$ (da $x = \overline{x}$), also $g(x) \in W$ nach Teil 1. Daher ist $g{\restriction}V =: f \in \mathrm{Hom}_{\mathbb{R}}(V, W)$. Weiter gilt $\tilde{f} = g$, da die Fortsetzung von f zu einem $\tilde{f} \in \mathrm{Hom}_{\mathbb{C}}(\tilde{V}, \tilde{W})$ eindeutig bestimmt ist.

(3) Sei $U' \subseteq \tilde{V}$ ein $\mathbb{C}$-Unterraum. Nehmen wir zunächst an, daß $U' = \tilde{U}$ für einen $\mathbb{R}$-Unterraum U von V. Dann läßt sich jedes $x \in \tilde{U}$ schreiben in der Form $x = u_1 + iu_2$ mit $u_1, u_2 \in U$. Also ist auch $\overline{x} = u_1 - iu_2 \in \tilde{U}$. Sei nun umgekehrt U' invariant unter Konjugation. Gesucht ist ein $\mathbb{R}$-Unterraum U von V mit $U' = \tilde{U}$. Wir setzen $U := \{\, x \in U' \mid x = \overline{x} \,\}$.

Zu $U' \subseteq \tilde{U}$. Sei $x = v_1 + iv_2 \in U'$ mit $v_1, v_2 \in V$. Dann ist $v_1 = \frac{1}{2}(x + \overline{x})$ und $v_2 = \frac{1}{2i}(x - \overline{x})$. Beide Vektoren sind in U', da mit x auch $\overline{x} \in U'$ ist. Wegen $v_1 = \overline{v_1}$ und $v_2 = \overline{v_2}$ haben wir $v_1, v_2 \in U$ und damit $x \in \tilde{U}$.

Zu $\tilde{U} \subseteq U'$. Sei $x \in \tilde{U}$, also $x = u_1 + iu_2$ mit $u_1, u_2 \in U$. Dann sind $u_1, u_2 \in U'$, also auch $x \in U'$.

Wir müssen noch zeigen, daß U eindeutig bestimmt ist. Sei also $U_0 \subseteq V$ ein weiterer $\mathbb{R}$-Unterraum mit $U' = \tilde{U_0}$. Gezeigt wird $U_0 = U := \{\, x \in U' \mid x = \overline{x} \,\}$. Sei zunächst $x \in U_0$. Dann ist $x = \overline{x}$ und $x \in \tilde{U_0} = U'$, also $x \in U$. Sei nun umgekehrt $x \in U$. Dann ist $x \in U'$ mit $x = \overline{x}$, und ferner $x = v_1 + iv_2$ mit $v_1, v_2 \in U_0$ (da $x \in U' = \tilde{U_0}$). Wegen $x = \overline{x}$ folgt $v_2 = 0$, also $x = v_1 \in U_0$. ∎

Bemerkungen. (1) Wir haben die Komplexifizierung bezüglich $\mathbb{R} \subseteq \mathbb{C}$ durchgeführt. Allgemeiner kann man eine entsprechende Konstruktion für eine beliebige Körpererweiterung $K \subseteq L$ und einen K-Vektorraum V durchführen. Man kann dann $L \otimes_K V$ als L-Vektorraum auffassen (siehe Abschnitt 10.4 über Tensorprodukte).

(2) Ist eine Körpererweiterung $K \subseteq L$ eine sogenannte *Galoiserweiterung* (siehe Algebra-Vorlesung), so läßt sich der Abstieg von L nach K wie eben der von $\mathbb{C}$ nach $\mathbb{R}$ durchführen.

8.2. Anwendung auf die reelle Jordansche Normalform

Zunächst sei noch einmal an den Beweis des Satzes über die Jordansche Normalform erinnert (wir schreiben jetzt V^λ anstelle von $U_{f,\lambda}$). Für jeden Eigenwert λ von f hatten wir $g := f - \lambda\mathrm{id}$ gesetzt und den Hauptraum V^λ zum Eigenwert λ definiert durch

$$V^\lambda := \bigcup_{m=0}^{\infty} \ker g^m.$$

Es war

$$0 \subsetneq \ker g \subsetneq \cdots \subsetneq \underbrace{\ker g^k}_{V^\lambda} = \ker g^{k+1} = \ldots.$$

In V^λ wurde dann eine geeignete Basis gewählt. Die Dimension von V^λ war die Vielfachheit von λ als Nullstelle von p_f. Ferner hatten wir eine Zerlegung

$$V = \underbrace{\ker g^k}_{V^\lambda} \oplus \operatorname{im} g^k$$

in f-invariante Unterräume erreicht. Das volle Resultat ergab sich durch Iteration. Aus dem damaligen Überlegungen erhält man insbesondere:

8.2.1. Hauptraumzerlegung.

Satz. *Seien V ein n-dimensionaler K-Vektorraum und $f\colon V \to V$ linear. Ferner sei p_f Produkt linearer Faktoren, etwa $p_f = (-1)^n \prod_{j=1}^s (t-\lambda_j)^{e_j}$ mit verschiedenen λ_j und $e_j \geq 1$. Für jedes $\lambda \in K$ setzen wir $V^\lambda := \bigcup_{m=0}^\infty \ker g^m$. Dann gilt*

$$V = \bigoplus_{\lambda \in K} V^\lambda.$$

Beweis. Offenbar ist $V^\lambda \neq 0$ genau dann, wenn λ Eigenwert von f ist. Es seien $\lambda_1, \ldots, \lambda_s$ die verschiedenen Eigenwerte von f. Wir betrachten die Abbildung

$$\begin{aligned} h\colon V^{\lambda_1} \times \cdots \times V^{\lambda_s} &\to V \\ (x_1, \ldots, x_s) &\mapsto x_1 + \cdots + x_s, \end{aligned}$$

Man sieht sofort, daß h linear ist. h ist surjektiv aufgrund der Konstruktion der Basis im Beweis des Satzes über die Jordansche Normalform (Basiselemente, die zum Eigenwert λ gehören, liegen in V^λ). Ferner gilt nach dem Dimensionssatz

$$\begin{aligned} \dim \ker h + \dim V &= \dim V^{\lambda_1} \times \cdots \times V^{\lambda_s} \\ &= \dim V^{\lambda_1} + \cdots + \dim V^{\lambda_s} \\ &= e_1 + \cdots + e_s \\ &= \dim V. \end{aligned}$$

Demnach ist $\ker h = 0$, also h ein Isomorphismus, also $V = \bigoplus_{j=1}^s V^{\lambda_j}$. ∎

8.2.2. Reelle Jordansche Normalform.

SATZ. *Seien V ein n-dimensionaler $\mathbb{R}$-Vektorraum und $f\colon V \to V$ linear. Dann gibt es eine Basis $v_1, \dots, v_n$ von V mit*

$$f\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

wobei jedes $A_j \in \mathbb{R}^{n_j \times n_j}$ die Form

$$\begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix}$$

(λ reeller Eigenwert von f) oder die Form

$$\begin{pmatrix} \boxed{\begin{matrix} \alpha & -\beta \\ \beta & \alpha \end{matrix}} & \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & & 0 \\ & \boxed{\begin{matrix} \alpha & -\beta \\ \beta & \alpha \end{matrix}} & \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & \\ & & \ddots & \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} \\ 0 & & & \boxed{\begin{matrix} \alpha & -\beta \\ \beta & \alpha \end{matrix}} \end{pmatrix}$$

($\lambda := \alpha + i\beta$ ($\beta \neq 0$) komplexer Eigenwert von $\tilde{f}$, mit $\alpha, \beta \in \mathbb{R}$) hat.

BEWEIS. Wir betrachten die eindeutige Fortsetzung $\tilde{f}\colon \tilde{V} \to \tilde{V}$ von f auf $\tilde{V}$. Nach 8.2.1 gilt $\tilde{V} = \bigoplus_{\lambda \in \mathbb{C}} \tilde{V}^\lambda$. Für $\tilde{V}$ können wir jetzt den Satz über die Jordansche Normalform anwenden, da $\mathbb{C}$ algebraisch abgeschlossen ist. Er liefert eine Basis, so daß die darstellende Matrix von f Jordan-Normalform hat. Unser Ansatz ist die Konstruktion einer solchen Basis, und anschließend ein Abstieg von $\mathbb{C}$ nach $\mathbb{R}$.

1. Sei $\lambda \in \mathbb{C}$, λ Eigenwert von $\tilde{f}$. Wir setzen $g := \tilde{f} - \lambda\mathrm{id}$ und betrachten $\overline{g} := \tilde{f} - \overline{\lambda}\mathrm{id} \in \mathrm{Hom}_{\mathbb{C}}(\tilde{V}, \tilde{V})$. Dann gilt

$$\overline{g^l(x)} = \overline{g}^l(\overline{x}) \quad \text{für alle } x \in \tilde{V}, l \geq 0.$$

BEWEIS. Es ist $\overline{g(x)} = \overline{\tilde{f}(x) - \lambda x} = \tilde{f}(\overline{x}) - \overline{\lambda}\overline{x} = \overline{g}(\overline{x})$. Wir führen den Beweis durch Induktion über l. *Basis* $l = 0$. Wegen $g^0 = \overline{g}^0 = \mathrm{id}$ ist hier nichts zu zeigen. *Schritt* $l \mapsto l + 1$.

$$\begin{aligned} \overline{g^{l+1}(x)} &= \overline{g(g^l(x))} \\ &= \overline{g}(\overline{g^l(x)}) && \text{nach der Vorbemerkung} \\ &= \overline{g}(\overline{g}^l(\overline{x})) && \text{nach Induktionsvoraussetzung} \\ &= \overline{g}^{l+1}(\overline{x}). \end{aligned}$$

□

2. Sei λ komplexer Eigenwert von $\tilde{f}$. Dann ist $\tilde{V}^\lambda \to \tilde{V}^{\overline{\lambda}}$, $x \mapsto \overline{x}$ bijektiv.

BEWEIS. Sei $x \in \tilde{V}^\lambda$, also $g^l(x) = 0$ für ein l. Nach (1) ist $\overline{g^l(x)} = \overline{g}^l(\overline{x}) = 0$, also $\overline{x} \in \tilde{V}^{\overline{\lambda}}$. Genauso zeigt man, daß aus $y \in \tilde{V}^{\overline{\lambda}}$ folgt $\overline{y} \in \tilde{V}^\lambda$. □

3. Sei λ reeller Eigenwert von $\tilde{f}$ (also auch von f). Nach (2) ist $\tilde{V}^\lambda$ invariant bei Konjugation. Nach Teil 3 des Satzes in 8.1.6 ist dann $\tilde{V}^\lambda = \tilde{U}$ für einen $\mathbb{R}$-Unterraum $U \subseteq V$, und es gilt

$$U = \{\, x \in \tilde{V}^\lambda \mid x = \overline{x} \,\} = \{\, x \in V \mid \exists l \underbrace{(\tilde{f} - \lambda\mathrm{id})^l(x)}_{=(f-\lambda\mathrm{id})^l(x)} = 0 \,\} = V^\lambda.$$

Also ist $\widetilde{V^\lambda} = \tilde{V}^\lambda$. Man wähle man nun wie im Beweis der Jordan-Normalform eine „gute“ Basis von V^λ bezüglich f. Sie ist dann auch Basis von $\tilde{V}^\lambda$ bezüglich $\tilde{f}$.

4. Sei λ komplexer Eigenwert von $\tilde{f}$ mit $\lambda \neq \overline{\lambda}$. Sei $v_1 + iw, \ldots, v_m + iw_m$ eine „gute“ Basis von $\tilde{V}^\lambda$ mit $v_j, w_j \in V$, so daß $\tilde{f}$ bezüglich dieser Basis durch eine Matrix in Jordanscher Normalform dargestellt wird. Da $\tilde{V}^\lambda \to \tilde{V}^{\overline{\lambda}}$, $x \mapsto \overline{x}$ ein Isomorphismus ist, muß $v_1 - iw_1, \ldots, v_m - iw_m$ Basis von $\tilde{V}^{\overline{\lambda}}$ sein. Also ist $v_1, w_1, \ldots, v_m, w_m$ eine $\mathbb{C}$-Basis von $\tilde{V}^\lambda \oplus \tilde{V}^{\overline{\lambda}}$ (da es sich um ein Erzeugendensystem aus $\dim \tilde{V}^\lambda \oplus \tilde{V}^{\overline{\lambda}} = 2m$ Elementen handelt). OBdA sei

$$\tilde{f}\begin{pmatrix} v_1 + iw_1 \\ \vdots \\ v_m + iw_m \end{pmatrix} = \begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix} \begin{pmatrix} v_1 + iw_1 \\ \vdots \\ v_m + iw_m \end{pmatrix}$$

(man betrachte jede Jordan-Matrix einzeln). Also gilt für $j < m$ mit $\lambda =: \alpha + i\beta$

$$\underbrace{\tilde{f}(v_j + iw_j)}_{f(v_j)+if(w_j)} = \underbrace{(\alpha + i\beta)(v_j + iw_j)}_{\alpha v_j - \beta w_j + i(\beta v_j + \alpha w_j)} + v_{j+1} + iw_{j+1}$$

$$\begin{cases} f(v_j) = \alpha v_j - \beta w_j + v_{j+1} \\ f(w_j) = \beta v_j + \alpha w_j + w_{j+1} \end{cases}$$

$$f\begin{pmatrix} v_j \\ w_j \end{pmatrix} = \begin{pmatrix} \alpha & -\beta & 1 & 0 \\ \beta & \alpha & 0 & 1 \end{pmatrix} \begin{pmatrix} v_j \\ w_j \\ v_{j+1} \\ w_{j+1} \end{pmatrix}.$$

Ebenso erhält man für $j = m$

$$f\begin{pmatrix} v_m \\ w_m \end{pmatrix} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} v_m \\ w_m \end{pmatrix}.$$

5. Konstruiert wurde eine $\mathbb{C}$-Basis von $\tilde{V} = \bigoplus_{\lambda \in \mathbb{C}} \tilde{V}^\lambda$, bestehend aus Vektoren aus V. Da $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{C}}(\tilde{V})$, und die Familie der $\mathbb{C}$-Basis auch $\mathbb{R}$-linear unabhängig ist, hat man damit eine $\mathbb{R}$-Basis von V konstruiert. f hat bezüglich dieser Basis die gewünschte Form. □

8.3. Unitäre Vektorräume

Die komplexen Zahlen kann man als den euklidischen Vektorraum $\mathbb{R}^2$ auffassen, indem man $\xi = \alpha + i\beta$ mit dem Paar (α, β) identifiziert (in 1.2.4 waren die komplexen Zahlen so definiert worden). Für den Betrag gilt dann

$$\|(\alpha,\beta)\| = \sqrt{\alpha^2+\beta^2} = |\xi| = \sqrt{\xi\overline{\xi}}.$$

Wir betrachten jetzt die folgende Abbildung

$$\begin{aligned}\sigma\colon \mathbb{C}\times\mathbb{C} &\to \mathbb{C}\\ (\xi,\xi') &\mapsto \xi\overline{\xi'}.\end{aligned}$$

Dann gilt offenbar

$$\begin{aligned}&\sigma(\xi,\xi) \geq 0,\\ &+\sqrt{\sigma(\xi,\xi)} = |\xi|,\\ &\sigma(\xi,\xi') = \xi\overline{\xi'} = \overline{\overline{\xi}\,\xi'} = \overline{\sigma(\xi',\xi)}.\end{aligned}$$

Entsprechend erweitern wir jetzt den Begriff eines euklidischen Vektorraums (über $\mathbb{R}$) zu dem eines unitären Vektorraums (über $\mathbb{C}$).

8.3.1. Hermitesches Skalarprodukt.

Definition. Seien V ein $\mathbb{C}$-Vektorraum und $\sigma\colon V\times V\to\mathbb{C}$ eine Abbildung. σ heißt *hermitesches Skalarprodukt*, wenn σ linear im ersten Argument ist, d.h. wenn für alle $x,y,z\in V$ und alle $\alpha,\beta\in\mathbb{C}$ gilt

$$\sigma(\alpha x+\beta y,z) = \alpha\sigma(x,z)+\beta\sigma(y,z),$$

ferner σ eine *hermitesche Abbildung* ist, d.h. für alle $x,y\in V$ gilt

$$\sigma(x,y) = \overline{\sigma(y,x)},$$

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 *unitärer Vektorraum*, wenn σ ein hermitesches Skalarprodukt auf V ist.

Bemerkungen. Sei (V,σ) ein unitärer Vektorraum.

(1) Es ist $\sigma(x,x)\in\mathbb{R}$ für alle $x\in V$ (wegen $\sigma(x,x) = \overline{\sigma(x,x)}$). Deshalb ist die Forderung $\sigma(x,x)>0$ in der Definition sinnvoll.

(2) Offenbar gilt $\sigma(x,x)=0$ genau dann, wenn $x=0$. Beweis: Aus $\sigma(x,x)=0$ folgt $x=0$ wegen der positiven Definitheit. Umgekehrt ist $\sigma(0,y)=0$ für alle y, da $\sigma(x,y)$ linear im ersten Argument ist.

(3) Es gilt $\sigma(x,\alpha y+\beta z) = \overline{\alpha}\sigma(x,y)+\overline{\beta}\sigma(x,z)$, denn

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

Man sagt deshalb, daß σ *semilinear* im zweiten Argument ist.

(4) Man nennt Abbildungen $\sigma\colon V\times V\to\mathbb{C}$, die linear im ersten und semilinear im zweiten Argument sind, auch *Sesquilinearformen* (lat. „anderthalb“).

BEISPIELE. (1) Das Standardbeispiel ist $V = \mathbb{C}^n$, wobei für $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ und $y = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}$ gesetzt wird

$$\sigma(x, y) := x^t \overline{y} = \xi_1 \overline{\eta_1} + \cdots + \xi_n \overline{\eta_n}.$$

Man zeigt leicht, daß damit ein hermitesches Skalarprodukt definiert ist: Offenbar ist σ linear im ersten Argument. σ ist hermitesche Abbildung, denn

$$\overline{\sigma(x, y)} = \overline{x^t \overline{y}} = \overline{x}^t y = y^t \overline{x} = \sigma(y, x).$$

σ ist positiv definit, denn

$$\sigma(x, x) = x^t \overline{x} = \xi_1 \overline{\xi_1} + \cdots + \xi_n \overline{\xi_n} = |\xi_1|^2 + \cdots + |\xi_n|^2 > 0,$$

falls ein $|\xi_j| \neq 0$. Man nennt σ das *kanonische Skalarprodukt* im $\mathbb{C}^n$.

(2) Seien (V, σ) ein unitärer Vektorraum, W ein $\mathbb{C}$-Vektorraum und $f\colon W \to V$ ein $\mathbb{C}$-Monomorphismus (z.B. $W \subseteq V$ ein Unterraum und f die Einbettung). Setze

$$\sigma_f(x, y) := \sigma(f(x), f(y))$$

für alle $x, y \in W$. Dann ist (W, σ_f) ein unitärer Vektorraum.

BEWEIS. Offenbar ist σ_f linear im ersten Argument. σ_f ist hermitesche Abbildung, denn

$$\overline{\sigma_f(x, y)} = \overline{\sigma(f(x), f(y))} = \sigma(f(y), f(x)) = \sigma_f(y, x).$$

σ_f ist positiv definit, denn für $x \in W$ mit $x \neq 0$ ist $f(x) \neq 0$ (da f ein Monomorphismus ist), also $0 < \sigma(f(x), f(x)) = \sigma_f(x, x)$. □

Viele Definitionen und Sätze mit ihren Beweisen lassen sich leicht von euklidischen auf unitäre Vektorräume übertragen. Wir wollen dies im folgenden durchführen.

8.3.2. Orthogonale Projektion auf einen Vektor. Sei (V, σ) ein unitärer 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) $p \in \mathbb{C}y$ und
(2) $x - p \perp y$.

$h := x - p$ heißt *Lot* von x auf y.

SATZ. *Sei (V, σ) ein unitärer 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. Wie in Abschnitt 5.1. □

8.3.3. Schwarzsche Ungleichung.

SATZ. *Seien (V,σ) ein unitärer Vektorraum und $x,y\in V$. Dann gilt*

(1) $\|x\|=0$ *genau dann, wenn* $x=0$.
(2) $\|\alpha x\|=|\alpha|\cdot\|x\|$ *für alle* $\alpha\in\mathbb{C}$.
(3) (Schwarzsche Ungleichung). $|\sigma(x,y)|\le\|x\|\cdot\|y\|$. *Die Gleichheit gilt genau dann, wenn x und y linear abhängig sind.*
(4) (Dreiecksungleichung). $\|x+y\|\le\|x\|+\|y\|$.

BEWEIS. Wie in Abschnitt 5.1. Nur für die Schwarzsche Ungleichung und die Dreiecksungleichung muß die frühere Argumentation leicht verändert werden.

(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

$$\begin{aligned}
0&\le\sigma(h,h)\\
&=\sigma(x-p,x-p)\\
&=\sigma(x,x)-\sigma(p,x)-\sigma(x,p)+\sigma(p,p)\\
&=\sigma(x,x)-\frac{\sigma(x,y)}{\sigma(y,y)}\sigma(y,x)-\frac{\overline{\sigma(x,y)}}{\overline{\sigma(y,y)}}\sigma(x,y)+\frac{\sigma(x,y)\overline{\sigma(x,y)}}{\sigma(y,y)\overline{\sigma(y,y)}}\sigma(y,y)\\
&=\frac{\sigma(x,x)\sigma(y,y)-\sigma(x,y)\overline{\sigma(x,y)}}{\sigma(y,y)},
\end{aligned}$$

also $|\sigma(x,y)|^2\le\sigma(x,x)\sigma(y,y)$. Für den Zusatz (Gleichheit gilt genau dann, wenn x und y linear abhängig sind) kann man den Beweis in 5.1 unverändert übernehmen.

(4) Man erhält aufgrund der Schwarzschen Ungleichung

$$\begin{aligned}
\sigma(x+y,x+y)&=\sigma(x,x)+\sigma(x,y)+\sigma(y,x)++\sigma(y,y)\\
&=\sigma(x,x)+\sigma(x,y)+\overline{\sigma(x,y)}+\sigma(y,y)\\
&=\sigma(x,x)+2\mathbb{R}e(\sigma(x,y))+\sigma(y,y)\\
&\le\sigma(x,x)+2|\sigma(x,y)|+\sigma(y,y)\\
&\le\|x\|^2+2\|x\|\cdot\|y\|+\|y\|^2\\
&=(\|x\|+\|y\|)^2.
\end{aligned}$$

Das war zu zeigen. □

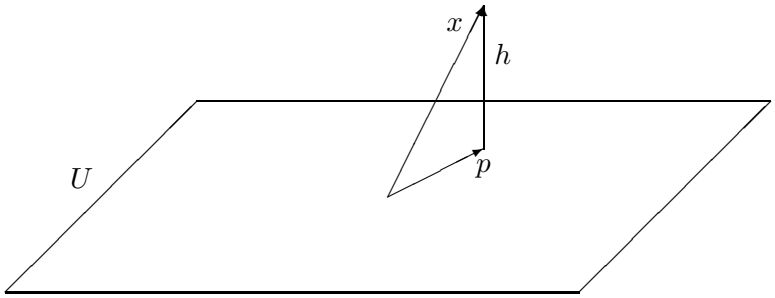
8.3.4. Orthonormalsystem und Orthonormalbasis. Eine Familie $(e_i)_{i\in I}$ von Vektoren aus einem unitären Vektorraum heißt *Orthonormalsystem* (ONS), wenn gilt

$$\begin{aligned}
\sigma(e_i,e_j)&=0\quad\text{für } i,j\in I,\ i\neq j,\\
\sigma(e_i,e_i)&=1\quad\text{für } i\in I.
\end{aligned}$$

$(e_i)_{i\in I}$ heißt *Orthonormalbasis* (ONB), wenn zusätzlich $(e_i)_{i\in I}$ Basis ist. Im Standardbeispiel $(\mathbb{C}^n,\sigma)$ mit $\sigma(x,y)=x^t\overline{y}$ ist $(e_1,\dots,e_n)$ Orthonormalbasis. Ferner gilt wieder: Ist $(e_i)_{i\in I}$ Orthonormalsystem, so ist $(e_i)_{i\in I}$ linear unabhängig. Ist $x\in V$ und $U\subseteq V$ Unterraum, so heißt $p\in V$ *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 $U = \mathbb{C}y$ ist das die alte Definition).



8.3.5. Orthogonale Projektion auf einen Unterraum.

SATZ. *Seien (V, σ) ein unitärer Vektorraum, $x \in V$, ferner $U \subseteq V$ ein 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. Wie in Abschnitt 5.1. □

8.3.6. Schmidtsches Orthonormalisierungsverfahren.

SATZ. *Seien (V, σ) ein unitärer 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. Wie in Abschnitt 5.1. □

Jeder endlichdimensionale unitäre Vektorraum besitzt also eine Orthonormalbasis.

8.3.7. Orthogonales Komplement.

SATZ. *Seien V ein unitärer Vektorraum und $U \subseteq V$ ein endlichdimensionaler Unterraum. Dann gilt*

(1) $V = U \oplus U^\perp$.
(2) $U = U^{\perp\perp}$.

BEWEIS. (1) $V = U + U^\perp$. Seien $x \in V$ und p die orthogonale Projektion von x auf U. Dann ist

$$x = \underbrace{p}_{\in U} + \underbrace{(x - p)}_{\in U^\perp} \in U + U^\perp.$$

$U \cap U^\perp = 0$. Sei $x \in U \cap U^\perp$. Dann ist $x \perp x$, also $\sigma(x, x) = 0$ und damit $x = 0$.

(2) $U \subseteq U^{\perp\perp}$ gilt stets (auch für eine Menge M statt U). $U^{\perp\perp} \subseteq U$: Sei $x \in U^{\perp\perp}$. Dann gilt

$$x = \underbrace{p}_{\in U} + \underbrace{(x - p)}_{\in U^\perp} \qquad \text{nach Teil (1)}$$

$$\begin{cases} x-p \perp x & \text{da } x-p \in U^\perp \text{ und } x \in U^{\perp\perp} \\ x-p \perp p & \text{da } x-p \in U^\perp \text{ und } p \in U \end{cases}$$
$$\sigma(x-p, x-p) = 0$$
$$x - p = 0$$
$$x = p \in U.$$

□

8.3.8. Komplexifizierung. Einen euklidischen Vektorraum kann man stets in eindeutiger Weise zu einem unitären Vektorraum machen.

SATZ. *Seien (V, σ) euklidischer Vektorraum und $\tilde{V}$ die Komplexifizierung von V. Dann gibt es genau ein hermitesches Skalarprodukt $\tilde{\sigma}$ auf $\tilde{V}$ mit $\tilde{\sigma}\restriction V = \sigma$.*

BEWEIS. Es war $V \subseteq \tilde{V} = \{\, x + iy \mid x, y \in V \,\}$.

Eindeutigkeit von $\tilde{\sigma}$. Sei $\tilde{\sigma}$ hermitesches Skalarprodukt mit $\tilde{\sigma}\restriction V = \sigma$. Dann gilt für alle $x, y, u, v \in V$

$$\begin{aligned} \tilde{\sigma}(x+iy, u+iv) &= \tilde{\sigma}(x,u) + i\tilde{\sigma}(y,u) - i\tilde{\sigma}(x,v) + \tilde{\sigma}(y,v) \\ &= \sigma(x,u) + \sigma(y,v) + i(\sigma(y,u) - \sigma(x,v)). \end{aligned}$$

Also ist $\tilde{\sigma}$ nur von σ abhängig.

Existenz von $\tilde{\sigma}$. Für $x, y, u, v \in V$ setze

$$\tilde{\sigma}(x+iy, u+iv) := \sigma(x,u) + \sigma(y,v) + i(\sigma(y,u) - \sigma(x,v)).$$

Durch Nachrechnen zeigt man, daß dann $\tilde{\sigma}$ ein hermitesches Skalarprodukt ist:

1. $\tilde{\sigma}$ ist linear im ersten Argument: Die Additivität im ersten Argument ist klar. Ferner gilt

$$\begin{aligned} &\tilde{\sigma}((\alpha+i\beta)(x+iy), u+iv) \\ &= \tilde{\sigma}(\alpha x - \beta y + i(\alpha y + \beta x), u+iv) \\ &= \sigma(\alpha x - \beta y, u) + \sigma(\alpha y + \beta x, v) + i(\sigma(\alpha y + \beta x, u) - \sigma(\alpha x - \beta y, v)). \end{aligned}$$

Andererseits ist

$$\begin{aligned} &(\alpha+i\beta)\tilde{\sigma}(x+iy, u+iv) \\ &= \alpha\sigma(x,u) + \alpha\sigma(y,v) - \beta\sigma(y,u) + \beta\sigma(x,v) \\ &\quad + i(\beta\sigma(x,u) + \beta\sigma(y,v) + \alpha\sigma(y,u) - \alpha\sigma(x,v)). \end{aligned}$$

2. $\tilde{\sigma}$ ist hermitesch: Es gilt

$$\begin{aligned} \tilde{\sigma}(u+iv, x+iy) &= \sigma(u,x) + \sigma(v,y) + i(\sigma(v,x) - \sigma(u,y)) \\ &= \sigma(x,u) + \sigma(y,v) + i(\sigma(x,v) - \sigma(y,u)) \\ &= \overline{\tilde{\sigma}(x+iy, u+iv)}. \end{aligned}$$

3. $\tilde{\sigma}$ ist positiv definit: Es gilt

$$\tilde{\sigma}(x+iy, x+iy) = \sigma(x,x) + \sigma(y,y) + i(\underbrace{\sigma(y,x) - \sigma(x,y)}_{=0}) > 0,$$

falls $x \neq 0$ oder $y \neq 0$. □

BEMERKUNG. Unter den Voraussetzungen des Satzes gilt: Ist $(e_1, \dots, e_n)$ eine Orthonormalbasis von V, so ist $(e_1, \dots, e_n)$ auch eine Orthonormalbasis von $\tilde{V}$. Beweis. Nach Lemma 8.1.4 ist $(e_1, \dots, e_n)$ eine $\mathbb{C}$-Basis von $\tilde{V}$, und außerdem gilt $\tilde{\sigma}(e_r, e_s) = \sigma(e_r, e_s) = \delta_{rs}$.

BEISPIEL. Auf $V = \mathbb{R}^n$ sei $\sigma(x, y) = x^t y$ für $x, y \in \mathbb{R}^n$. Auf $\tilde{V} = \mathbb{C}^n$ gilt dann $\tilde{\sigma}(a, b) = a^t \overline{b}$ für $a, b \in \mathbb{C}^n$. Dies kann man wie früher durch Verifikation beweisen, indem man zeigt, daß $\tilde{\sigma}$ ein hermitesches Skalarprodukt ist. Eine andere Möglichkeit ist die folgende. Nach dem Satz existiert $\tilde{\sigma}$ so daß $\tilde{\sigma} \restriction V = \sigma$. Für $a = \sum_{j=1}^n \alpha_j e_j$ und $b = \sum_l \beta_l e_l$ ist dann

$$\tilde{\sigma}(a, b) = \sum_{j,l} \alpha_j \overline{\beta_l} \sigma(e_j, e_l) = a^t \overline{b}.$$

8.3.9. Darstellungssatz.

SATZ. *Seien (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und $g \in V^* = \mathrm{Hom}_{\mathbb{C}}(V, \mathbb{C})$ (bzw. $= \mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R})$). Dann gibt es genau ein $y \in V$ so daß für alle $x \in V$ gilt*

$$g(x) = \sigma(x, y).$$

Ist $(e_1, \dots, e_n)$ eine Orthonormalbasis von V, so gilt $y = \sum_{j=1}^n \overline{g(e_j)} e_j$.

BEWEIS. Sei (V, σ) ein unitärer Vektorraum (ebenso geht es im Fall eines euklidischen Vektorraums).

Existenz von y. Seien $(e_1, \dots, e_n)$ eine Orthonormalbasis von V und $x \in V$. Dann gilt

$$x = \sum_{j=1}^n \sigma(x, e_j) e_j$$

$$g(x) = \sum_{j=1}^n \sigma(x, e_j) g(e_j) = \sigma(x, \underbrace{\sum_{j=1}^n \overline{g(e_j)} e_j}_{=:y}).$$

Eindeutigkeit von y. Seien y, y' mit den im Satz formulierten Eigenschaften gegeben. Dann gilt

$$\begin{aligned} &\sigma(x, y) = \sigma(x, y') && \text{für alle } x \in V \\ &\sigma(x, y - y') = 0 && \text{für alle } x \in V \\ &\sigma(y - y', y - y') = 0 \\ &y - y' = 0. \end{aligned}$$

□

8.3.10. Adjungierter Endomorphismus. Für $A = (\alpha_{ij}) \in \mathbb{C}^{n \times n}$ definieren wir

$$\overline{A} := (\overline{\alpha_{ij}}) \quad \text{und} \quad A^* := \overline{A}^t = (\overline{\alpha_{ji}})_{ij}.$$

$\overline{A}$ heißt die *konjugierte* Matrix und A^* die *adjungierte* Matrix zu A.

SATZ. *Seien (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$).*

(1) Es gibt genau ein $f^* \in \mathrm{Hom}_{\mathbb{C}}(V,V)$ *(bzw.* $\in \mathrm{Hom}_{\mathbb{R}}(V,V)$*) so daß für alle* $x, y \in V$ *gilt*

$$\sigma(f(x), y) = \sigma(x, f^*(y)).$$

f^* *heißt der zu* f adjungierte *Endomorphismus. Ist* $(e_1, \ldots, e_n)$ *eine Orthonormalbasis von* V*, so gilt*

$$f^*(y) = \sum_{j=1}^{n} \sigma(y, f(e_j))e_j.$$

(2) Ist $A \in \mathbb{C}^{n\times n}$ *(bzw.* $\in \mathbb{R}^{n\times n}$*) die darstellende Matrix von* f *bzgl. einer Orthonormalbasis* $\mathcal{A} = (e_1, \ldots, e_n)$ *von* V*, so ist* A^* *die darstellende Matrix von* f^* *bzgl.* $(e_1, \ldots, e_n)$*, also*

$$\mathcal{M}_{\mathcal{A}}(f^*) = \mathcal{M}_{\mathcal{A}}(f)^*.$$

Beweis. (1) Existenz von f^*. Sei $y \in V$ fest. Betrachte

$$g\colon V \to \mathbb{C} \quad (\text{bzw. } \mathbb{R})$$
$$x \mapsto \sigma(f(x), y).$$

Es ist $g \in V^*$. Sei $(e_1, \ldots, e_n)$ eine Orthonormalbasis von V. Nach dem Darstellungssatz gilt

$$\sigma(f(x), y) = g(x) = \sigma(x, \sum_{j=1}^{n} \overline{g(e_j)}e_j) = \sigma(x, \sum_{j=1}^{n} \overline{\sigma(f(e_j), y)}e_j).$$

Setze

$$f^*(y) := \sum_{j=1}^{n} \overline{\sigma(f(e_j), y)}e_j = \sum_{j=1}^{n} \sigma(y, f(e_j))e_j.$$

Dann ist offenbar $f^* \in \mathrm{Hom}_{\mathbb{C}}(V,V)$, und für alle $x, y \in V$ gilt $\sigma(f(x), y) = \sigma(x, f^*(y))$.

Eindeutigkeit von f^*. $f^*(y)$ ist eindeutig bestimmt nach dem Darstellungssatz.

(2) Sei $A = (\alpha_{ij}) \in \mathbb{C}^{n\times n}$ (bzw. $\in \mathbb{R}^{n\times n}$) darstellende Matrix von f bzgl. der Orthonormalbasis $(e_1, \ldots, e_n)$, also

$$f(e_j) = \sum_{i=1}^{n} \alpha_{ij}e_i.$$

Dann gilt

$$f^*(e_j) = \sum_s \sigma(e_j, \underbrace{f(e_s)}_{\sum_i \alpha_{is}e_i})e_s = \sum_{s,i} \overline{\alpha_{is}}\, \underbrace{\sigma(e_j, e_i)}_{\delta_{ji}} e_s = \sum_s \overline{\alpha_{js}}e_s.$$

Das war zu zeigen. □

Bemerkung. Seien (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum, $f, g \in \mathrm{Hom}_{\mathbb{C}}(V,V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V,V)$) und $\alpha \in \mathbb{C}$ (bzw. $\in \mathbb{R}$). Dann gilt

(1) $(f+g)^* = f^* + g^*$.
(2) $(f \circ g)^* = g^* \circ f^*$.
(3) $(\alpha f)^* = \overline{\alpha} f^*$.
(4) $f^{**} = f$.

(5) Ist f ein Isomorphismus, so gilt $(f^{-1})^* = (f^*)^{-1}$.

Die Abbildung $f \mapsto f^*$ für $f \in \mathrm{Hom}_{\mathbb{C}}(V,V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V,V)$) hat also ähnliche Eigenschaften wie die Konjugation $\alpha \mapsto \overline{\alpha}$ für $\alpha \in \mathbb{C}$.

BEWEIS. (1) Für alle $x, y \in V$ gilt

$$\begin{aligned}\sigma((f+g)(x), y) &= \sigma(f(x) + g(x), y)\\ &= \sigma(f(x), y) + \sigma(g(x), y)\\ &= \sigma(x, f^*(y)) + \sigma(x, g^*(y))\\ &= \sigma(x, f^*(y) + g^*(y))\\ &= \sigma(x, (f^* + g^*)(y)).\end{aligned}$$

Also ist $(f+g)^* = f^* + g^*$.

(2) Für alle $x, y \in V$ gilt

$$\sigma((f \circ g)(x), y) = \sigma(g(x), f^*(y)) = \sigma(x, (g^* \circ f^*)(y)).$$

Also ist $(f \circ g)^* = g^* \circ f^*$.

(3) Für alle $x, y \in V$ gilt

$$\sigma(\alpha f(x), y) = \alpha\sigma(f(x), y) = \alpha\sigma(x, f^*(y)) = \sigma(x, \overline{\alpha} f^*(y)).$$

Also ist $(\alpha f)^* = \overline{\alpha} f^*$.

(4) Zu zeigen ist

$$\sigma(f^*(x), y) = \sigma(x, f(y))$$

für alle $x, y \in V$, denn nach der Definition adjungierter Abbildungen folgt daraus die Behauptung. Es gilt aber

$$\sigma(f^*(x), y) = \overline{\sigma(y, f^*(x))} = \overline{\sigma(f(y), x)} = \sigma(x, f(y)).$$

(5) Sei f ein Isomorphismus. Dann ist $f \circ f^{-1} = \mathrm{id}$, also wegen der Eindeutigkeit des adjungierten Endomorphismus auch $(f \circ f^{-1})^* = \mathrm{id}^* = \mathrm{id}$. Nach (2) ist aber $(f \circ f^{-1})^* = (f^{-1})^* \circ f^*$. Daraus folgt die Behauptung $(f^{-1})^* = (f^*)^{-1}$. □

BEMERKUNG. Teil (4) kann man auch leicht über darstellende Matrizen beweisen: Ist $A \in \mathbb{C}^{n\times n}$ darstellende Matrix von f bzgl. einer Orthonormalbasis $(e_1, \ldots, e_n)$, so ist nach Teil (2) des Satzes 8.3.10 über adjungierte Endomorphismen A^* die darstellende Matrix von f^* bzgl. derselben Orthonormalbasis $(e_1, \ldots, e_n)$, und auch $A^{**} = A$ die darstellende Matrix von f^{**} bzgl. $(e_1, \ldots, e_n)$. Also muß $f^{**} = f$ sein.

8.3.11. Darstellende Matrizen von Endomorphismen. Wir stellen einige Eigenschaften darstellender Matrizen von Endomorphismen zusammen, die wir im folgenden verwenden. Seien V ein n-dimensionaler K-Vektorraum und $\mathcal{A} = (x_1, \ldots, x_n)$ eine Basis von V. In 3.4.3 hatten wir gesehen, daß

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

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

ein Vektorraum-Isomorphismus ist. Nun trägt $\mathrm{Hom}_K(V,V) = \mathrm{End}_K(V)$ wie $K^{n\times n}$ sogar eine Ringstruktur: die (nicht kommutative) Multiplikation ist die Komposition $\circ$ zweier Endomorphismen, und die Identität id das Einselement. Wir wollen uns überlegen, daß der Vektorraum-Isomorphismus $\mathcal{M}_\mathcal{A}$ auch mit der Ringstruktur verträglich ist. Seien also $f, g \in \mathrm{End}_K(V)$ mit darstellenden Matrizen $A = (\alpha_{ij})$ und $B = (\beta_{js})$ bzgl. derselben Basis $\mathcal{A} = (x_1, \ldots, x_n)$ gegeben. Wir zeigen

$$\begin{aligned} \mathcal{M}_\mathcal{A}(f \circ g) &= \mathcal{M}_\mathcal{A}(f) \circ \mathcal{M}_\mathcal{A}(g), \\ \mathcal{M}_\mathcal{A}(\mathrm{id}) &= E. \end{aligned}$$

Beweis. Die zweite Geichung gilt offensichtlich. Für die erste haben wir zunächst

$$f(x_j) = \sum_i \alpha_{ij} x_i \quad \text{und} \quad g(x_s) = \sum_j \beta_{js} x_j,$$

also

$$\begin{aligned} f(g(x_s)) &= f\Big(\sum_j \beta_{js} x_j\Big) = \sum_j \beta_{js} f(x_j) = \sum_j \beta_{js} \sum_i \alpha_{ij} x_i \\ &= \sum_i \Big(\sum_j \alpha_{ij}\beta_{js}\Big) x_i. \end{aligned}$$

□

Ist also der Endomorphismus f oder die Matrix $\mathcal{M}_\mathcal{A}(f)$ invertierbar, so gilt

$$\mathcal{M}_\mathcal{A}(f^{-1}) = (\mathcal{M}_\mathcal{A}(f))^{-1}.$$

8.3.12. Unitäre Endomorphismen.

Satz. *Sei (V,σ) ein unitärer (bzw. euklidischer) Vektorraum, ferner $\mathcal{A} = (e_1, \ldots, e_n)$ eine Orthonormalbasis von V und $A \in \mathbb{C}^{n\times n}$ (bzw. $\in \mathbb{R}^{n\times n}$) die darstellende Matrix von f bzgl. $(e_1, \ldots, e_n)$. Dann sind folgende Aussagen äquivalent.*

(1) $f^* = f^{-1}$ (insbesondere also: f ist Isomorphismus).
(2) $A^* = A^{-1}$ (insbesondere also: A ist invertierbar).
(3) $\sigma(x,y) = \sigma(f(x), f(y))$ für alle $x, y \in V$.
(4) $\|x\| = \|f(x)\|$ für alle $x \in V$.
(5) $(f(e_1), \ldots, f(e_n))$ ist eine Orthonormalbasis von V.

Beweis. (1) $\Rightarrow$ (2). Sei f ein Isomorphismus und es gelte $f^* = f^{-1}$. Dann ist

$$\begin{aligned} \mathcal{M}_\mathcal{A}(f)^* &= \mathcal{M}_\mathcal{A}(f^*) && \text{nach 8.3.10} \\ &= \mathcal{M}_\mathcal{A}(f^{-1}) && \text{da } f^* = f^{-1} \\ &= \mathcal{M}_\mathcal{A}(f)^{-1} && \text{nach 8.3.11.} \end{aligned}$$

(2) $\Rightarrow$ (1). Sei $A = \mathcal{M}_\mathcal{A}(f)$ invertierbar und es gelte $A^* = A^{-1}$. Dann ist

$$\begin{aligned} \mathcal{M}_\mathcal{A}(f^{-1}) &= \mathcal{M}_\mathcal{A}(f)^{-1} && \text{nach 8.3.11} \\ &= \mathcal{M}_\mathcal{A}(f)^* && \text{nach Voraussetzung} \\ &= \mathcal{M}_\mathcal{A}(f^*) && \text{nach 8.3.10,} \end{aligned}$$

also $f^* = f^{-1}$.

(1) $\Leftrightarrow$ (3). (3) ist äquivalent zu

$$\sigma(x,y) = \sigma(f(x),f(y)) = \sigma(x,(f^*\circ f)(y)) \quad \text{für alle } x,y \in V.$$

Nach dem Eindeutigkeitsteil des Darstellungssatzes 8.3.9 ist dies äquivalent zu $f^* \circ f = \mathrm{id}$, also zu (1).

(3) $\Leftrightarrow$ (4). Dies folgt aus der Linearität von f und

$$\|x\| = \sqrt{\sigma(x,x)}$$
$$\sigma(x,y) = \frac{1}{4}\|x+y\|^2 - \frac{1}{4}\|x-y\|^2 + \frac{i}{4}\|x+iy\|^2 - \frac{i}{4}\|x-iy\|^2,$$

was man durch Nachrechnen leicht verifiziert.

(3) $\Leftrightarrow$ (5). Folgende Aussagen sind äquivalent.

$$\forall x,y\in V\colon (\sigma(x,y) = \sigma(f(x),f(y)))$$
$$\forall s,t\colon (\underbrace{\sigma(e_s,e_t)}_{\delta_{st}} = \sigma(f(e_s),f(e_t)))$$
$$(f(e_1),\dots,f(e_n)) \text{ ist Orthonormalsystem}$$
$$(f(e_1),\dots,f(e_n)) \text{ ist Orthonormalbasis.}$$

□

DEFINITION. Seien (V,σ) ein unitärer (bzw. euklidischer) Vektorraum und $f \in \mathrm{Hom}_{\mathbb{C}}(V,V)$ (bzw. $f \in \mathrm{Hom}_{\mathbb{R}}(V,V)$). f heißt *unitär* (bzw. *orthogonal*), wenn f eine Bedingung (und damit alle) des vorangehenden Satzes erfüllt, etwa $f^* = f^{-1}$. Eine Matrix $A \in \mathbb{C}^{n\times n}$ heißt *unitär* (bzw. *orthogonal*), wenn $A^* = A^{-1}$. Folgende Bezeichnungen sind gebräuchlich.

$$\mathrm{U}(n) := \{\, A \in \mathrm{GL}(n,\mathbb{C}) \mid \overline{A}^t A = E \,\} \quad \textit{unitäre Gruppe}$$
$$\mathrm{SU}(n) := \{\, A \in \mathrm{U}(n) \mid \det A = +1 \,\} \quad \textit{spezielle unitäre Gruppe.}$$
$$\mathrm{O}(n) := \{\, A \in \mathrm{GL}(n,\mathbb{R}) \mid A^t A = E \,\} \quad \textit{orthogonale Gruppe}$$
$$\mathrm{SO}(n) := \{\, A \in \mathrm{O}(n) \mid \det A = +1 \,\} \quad \textit{spezielle orthogonale Gruppe}$$

BEMERKUNG. Es gelten die Untergruppenbeziehungen

$$\mathrm{SO}(n) \subseteq \mathrm{O}(n) \subseteq \mathrm{GL}(n,\mathbb{R}),$$
$$\mathrm{SU}(n) \subseteq \mathrm{U}(n) \subseteq \mathrm{GL}(n,\mathbb{C}).$$

8.3.13. Selbstadjungierte Endomorphismen. Seien (V,σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und ferner $f \in \mathrm{Hom}_{\mathbb{C}}(V,V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V,V)$). f heißt *selbstadjungiert*, wenn $f = f^*$ ist. Eine Matrix $A \in \mathbb{C}^{n\times n}$ (bzw. $\in \mathbb{R}^{n\times n}$) heißt *hermitesch* (bzw. *symmetrisch*), wenn $A^* = A$ ist, also $\overline{A}^t = A$ (bzw. $A^t = A$).

BEISPIEL. $\left(\begin{smallmatrix} 1 & i \\ -i & 0 \end{smallmatrix}\right)$ ist hermitesch.

SATZ. *Seien (V,σ) ein endlichdimensionaler unitärer Vektorraum, $f \in \mathrm{Hom}_{\mathbb{C}}(V,V)$ und $\lambda \in \mathbb{C}$ ein Eigenwert von f. Dann gilt*

(1) Ist f selbstadjungiert, so ist $\lambda \in \mathbb{R}$.

(2) Ist f unitär, so ist $|\lambda| = 1$.

BEWEIS. Seien $x \in V$, $x \neq 0$ und $f(x) = \lambda x$. Dann erhält man folgendes.
(1)

$$\begin{aligned} & f = f^* \\ & \sigma(f(x), x) = \sigma(x, f(x)) \\ & \sigma(\lambda x, x) = \sigma(x, \lambda x) \\ & \lambda\sigma(x, x) = \overline{\lambda}\sigma(x, x) \\ & \lambda = \overline{\lambda}, \qquad\qquad \text{da } x \neq 0. \end{aligned}$$

(2)

$$\begin{aligned} & \sigma(f(x), f(x)) = \sigma(x, x) \\ & \sigma(\lambda x, \lambda x) = \sigma(x, x) \\ & \lambda\overline{\lambda}\sigma(x, x) = \sigma(x, x) \\ & \lambda\overline{\lambda} = 1, \qquad\qquad \text{da } x \neq 0 \\ & |\lambda| = 1. \end{aligned}$$

□

8.3.14. Diagonalisierung. Im nächsten Kapitel 9 wollen wir uns mit dem Problem befassen, wann ein Endomorphismus f eines gegebenen endlichdimensionalen unitären (bzw. euklidischen) Vektorraums (V, σ) diagonalisierbar ist, d.h. eine Orthonormalbasis aus Eigenvektoren von f besitzt. Dafür zunächst einige vorbereitende Anmerkungen.

BEMERKUNGEN. Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum.

(1) Seien $(e_1, \ldots, e_n)$ und $(e'_1, \ldots, e'_n)$ Orthonormalbasen von V. Nach dem Satz in 8.3.12 über orthogonale Endomorphismen (insbesondere der Äquivalenz von (2) und (5)) gilt

$$\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} \quad \text{für ein } C \in \mathrm{U}(n) \text{ (bzw. } \in \mathrm{O}(n)).$$

(2) Seien $A, B \in \mathbb{C}^{n\times n}$ (bzw. $\in \mathbb{R}^{n\times n}$). Dann sind folgende Aussagen äquivalent.
- A und B sind darstellende Matrizen desselben $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$) f bezüglich evtl. verschiedener Orthonormalbasen.
- $B = C^{-1}AC$ für ein $C \in \mathrm{U}(n)$ (bzw. $\in \mathrm{O}(n)$), d.h. A und B sind *unitär* (bzw. *orthogonal*) *äquivalent*.

BEWEIS. Seien A, B darstellende Matrizen desselben Endomorphismus $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$) bezüglich der Orthonormalbasen $(e_1, \ldots, e_n)$ und $(e'_1, \ldots, e'_n)$. Dann gilt

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = A^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \text{und} \quad f \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = B^t \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}.$$

Nach (1) hat man ein $C \in U(n)$ mit

$$\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}.$$

Es folgt

$$A^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = Cf \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = CB^t \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = CB^tC^{-1} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Also ist

$$A^t = CB^tC^{-1}$$
$$B^t = C^{-1}A^tC$$
$$B = C^tA(C^{-1})^t = D^{-1}AD \quad \text{mit } D = (C^t)^{-1} \in U(n) \text{ (bzw. } \in \mathrm{O}(n)).$$

Sei umgekehrt $B = C^{-1}AC$ für ein $C \in \mathrm{U}(n)$ (bzw. $\in \mathrm{O}(n)$). Man wähle eine Orthonormalbasis $(e'_1, \ldots, e'_n)$ und setze

$$\begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = C^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Dann gilt für den durch

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = A^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

definierten Endomorphismus

$$f \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = C^t f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^tA^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^tA^t(C^t)^{-1} \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = (C^{-1}AC)^t \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}.$$

Also sind A und B darstellende Matrizen von f bezüglich verschiedener Orthonormalbasen. □

(3) Sei A darstellende Matrix von f bezüglich einer Orthonormalbasis. Dann sind folgende Aussagen äquivalent.

- V hat eine Orthonormalbasis aus Eigenvektoren von f.
- $$C^{-1}AC = \begin{pmatrix} \gamma_1 & & 0 \\ & \ddots & \\ 0 & & \gamma_n \end{pmatrix}$$
  für ein $C \in \mathrm{U}(n)$ (bzw. $\in \mathrm{O}(n)$), mit $\gamma_j \in \mathbb{C}$ (bzw. $\in \mathbb{R}$).

Beweis. Dies folgt aus (1) und (2), da die darstellende Matrix bezüglich einer Orthonormalbasis aus Eigenvektoren eine Diagonalmatrix ist. □

(4) Nach der Hauptachsentransformation in 5.4 gilt für jeden Endomorphismus f eines endlichdimensionalen euklidischen Vektorraums V:

Ist f selbstadjungiert, so hat V eine Orthonormalbasis aus Eigenvektoren.

Für Matrizen bedeutet dies: Ist $A \in \mathbb{R}^{n\times n}$ symmetrisch, so gibt es $C \in \mathrm{O}(n)$, $\gamma_j \in \mathbb{R}$ mit

$$C^{-1}AC = \begin{pmatrix} \gamma_1 & & 0 \\ & \ddots & \\ 0 & & \gamma_n \end{pmatrix}.$$

Der damalige Beweis verwendete Jacobische Rotationsmatrizen und etwas Analysis. Im folgenden werden wir einen neuen Beweis dieses Satzes geben, der die Komplexifizierung und den Fundamentalsatz der Algebra benutzt. Gleichzeitig wird sich dabei eine Verallgemeinerung auf unitäre Räume ergeben.

KAPITEL 9

Endomorphismen

Wir wollen in diesem Kapitel die Endomorphismen von endlichdimensionalen unitären (bzw. euklidischen) Vektorräumen genauer studieren. Als erstes erhalten wir eine Charakterisierung derjenigen Endomorphismen, die eine Orthonormalbasis aus Eigenvektoren besitzen. In Abschnitt 9.2 befassen wir uns mit der Zerlegung eines Endomorphismus in geometrisch einfach zu beschreibende Bestandteile (Drehungen oder Spiegelungen, Streckungen und orthogonale Projektionen). Es wird sich zeigen, daß eine solche Zerlegung immer eindeutig möglich ist. Im letzten Abschnitt 9.3 verschaffen wir uns eine Übersicht über die orthogonalen Endomorphismen eines endlichdimensionalen euklidischen Vektorraums.

9.1. Normale Endomorphismen

Wir geben jetzt eine Charakterisierung der diagonalisierbaren Endomorphismen f eines endlichdimensionalen unitären (bzw. euklidischen) Vektorraums (V, σ); ein Endomorphismus f heißt diagonalisierbar, wenn (V, σ) eine Orthonormalbasis aus Eigenvektoren von f besitzt. Dazu benötigen wir einen etwas technischen Begriff, nämlich den eines normalen Endomorphismus. Wir werden zeigen:

- In einem unitären Vektorraum ist ein Endomorphismus genau dann diagonalisierbar, wenn er *normal* ist.
- In einem euklidischen Vektorraum ist ein Endomorphismus genau dann diagonalisierbar, wenn er *selbstadjungiert* ist.

9.1.1. Normale Endomorphismen. Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$). f heißt *normal*, wenn

$$f^* \circ f = f \circ f^*.$$

Eine Matrix $A \in \mathbb{C}^{n\times n}$ (bzw. $\in \mathbb{R}^{n\times n}$) heißt *normal*, wenn $A^*A = AA^*$ ist.

Bemerkungen. Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$).

(1) Sei A darstellende Matrix von f bzgl. einer Orthonormalbasis. Dann gilt: f ist normal (selbstadjungiert) genau dann, wenn A normal (hermitesch bzw. symmetrisch) ist. Dies folgt aus dem Satz (in 8.3.10) über adjungierte Endomorphismen.
(2) Ist f selbstadjungiert, so ist f auch normal, denn aus $f = f^*$ folgt offenbar $f^* \circ f = f \circ f^*$.
(3) Ist f unitär (bzw. orthogonal), so ist f auch normal, denn aus $f^* = f^{-1}$ folgt wieder $f^* \circ f = f \circ f^*$.

(4) Seien $U_i \subseteq V$ Unterräume mit $U_1 \perp U_2$ und $V = U_1 \oplus U_2$. Wir setzen voraus, daß $f[U_i] \subseteq U_i$ für $i = 1, 2$. Dann gilt:

f normal $\iff$ $f\restriction U_i \colon U_i \to U_i$ $(i = 1, 2)$ normal;

f selbstadjungiert $\iff$ $f\restriction U_i \colon U_i \to U_i$ $(i = 1, 2)$ selbstadjungiert.

Dies sieht man wie folgt. Sei $e_1, \ldots, e_m$ eine Orthonormalbasis von U_1 und $e_{m+1}, \ldots, e_n$ Orthonormalbasis von U_2. Dann ist $e_1, \ldots, e_n$ Orthonormalbasis von $V = U_1 \oplus U_2$, da $U_1 \perp U_2$. Sei A_i die transponierte der darstellenden Matrix von $f\restriction U_i$ bzgl. obiger Orthonormalbasis von U_i. Dann ist

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

und damit

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

die transponierte der darstellenden Matrix von f bzgl. $e_1, \ldots, e_n$. Weiter gilt

$$A^* = \begin{pmatrix} A_1^* & 0 \\ 0 & A_2^* \end{pmatrix}.$$

Dann sind folgende Aussagen äquivalent.

f ist normal	
A ist normal	nach Bemerkung (1)
$AA^* = A^*A$	
$A_iA_i^* = A_i^*A_i$	für $i = 1, 2$
$f\restriction U_i$ sind normal	für $i = 1, 2$.

Entsprechend schließt man für selbstadjungierte Endomorphismen.

9.1.2. Charakterisierung normaler Endomorphismen.

SATZ. *Sei (V, σ) ein endlichdimensionaler K-Vektorraum ($K \in \{\mathbb{R}, \mathbb{C}\}$) und $\sigma \colon V \times V \to K$ ein Skalarprodukt. Sei $f \in \mathrm{Hom}_K(V, V)$. Das charakteristische Polynom f zerfalle über K in Linearfaktoren. Dann sind folgende Aussagen äquivalent.*

(1) f ist normal.
(2) V hat eine Orthonormalbasis aus Eigenvektoren von f.

BEWEIS. (1) $\Rightarrow$ (2). Für $\lambda \in K$ sei

$$V_\lambda := \{\, x \in V \mid f(x) = \lambda x \,\}$$

der Eigenraum von f zu λ.

Schritt (a). Ist $g \in \mathrm{Hom}_K(V, V)$ und $U \subseteq V$ ein Unterraum mit $g[U] \subseteq U$, so ist $g^*[U^\perp] \subseteq U^\perp$; dies gilt, da für alle $x \in U$ und $y \in U^\perp$

$$\sigma(\underbrace{g(x)}_{\in U}, y) = \sigma(x, g^*(y)) = 0, \quad \text{also } g^*(y) \in U^\perp \text{ für } y \in U^\perp.$$

Schritt (b). Ist $\lambda \in K$ und $g \in \mathrm{Hom}_K(V,V)$ mit $f \circ g = g \circ f$, so ist $g[V_\lambda] \subseteq V_\lambda$; dies gilt, da man für alle $x \in V_\lambda$ hat

$$\begin{aligned} &f(x) = \lambda x \\ &f(g(x)) = g(f(x)) = g(\lambda x) = \lambda g(x) \quad \text{wegen } f \circ g = g \circ f \\ &g(x) \in V_\lambda. \end{aligned}$$

Schritt (c). Ist f normal, so gilt $f[V_\lambda^\perp] \subseteq V_\lambda^\perp$ für alle $\lambda \in K$, denn man hat

$$\begin{aligned} f^*[V_\lambda] &\subseteq V_\lambda && \text{nach Schritt (b)} \\ f^{**}[V_\lambda^\perp] &\subseteq V_\lambda^\perp && \text{nach Schritt (a)} \\ f[V_\lambda^\perp] &\subseteq V_\lambda^\perp && \text{wegen } f = f^{**}. \end{aligned}$$

Schritt (d). Wir zeigen (1) $\Rightarrow$ (2) durch Induktion über $\dim V$.

Fall $\dim V = 1$. Man wähle ein $e \in V$ mit $\|e\| = 1$. Dann ist (e) Basis von V, und e ist Eigenvektor von f wegen $f(e) \in Ke = V$.

Fall $\dim V > 1$. Da das charakteristische Polynom von f über K in Linearfaktoren zerfällt, hat f einen Eigenwert, etwa $\lambda \in K$. Es gilt $V = V_\lambda \oplus V_\lambda^\perp$, und wegen $\dim V_\lambda > 0$ folgt $\dim V_\lambda^\perp < \dim V$. V_λ hat trivialerweise ein Orthonormalbasis aus Eigenvektoren von f, da ja V_λ Eigenraum ist.

Weiter hat $V_\lambda^\perp$ eine Orthonormalbasis aus Eigenvektoren von f; dies sieht man wie folgt. Offenbar ist $f[V_\lambda] \subseteq V_\lambda$ und nach Schritt (c) auch $f[V_\lambda^\perp] \subseteq V_\lambda^\perp$. Nach der obigen Bemerkung (4) ist dann auch $f \restriction V_\lambda^\perp$ normal. Ferner teilt nach Bemerkung (3) in 7.2.2 das charakteristische Polynom von $f \restriction V_\lambda^\perp$ dasjenige von f. Da nach Voraussetzung p_f in Linearfaktoren zerfällt, muß dies also auch für $p_{f \restriction V_\lambda^\perp}$ gelten (dies zeigt man leicht allgemein mit Hilfe des Lemmas in 7.1.3, und zwar durch Induktion über den Grad). Nach Induktionsvoraussetzung hat also $V_\lambda^\perp$ eine Orthonormalbasis aus Eigenvektoren von $f \restriction V_\lambda^\perp$ (und damit von f). Wegen $V = V_\lambda \oplus V_\lambda^\perp$ hat also auch V eine Orthonormalbasis aus Eigenvektoren von f.

(2) $\Rightarrow$ (1). Sei $e_1, \ldots, e_n$ eine Orthonormalbasis aus Eigenvektoren von f. Dann gilt

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = A \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \text{mit } A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

A ist normal, da

$$A^* = \begin{pmatrix} \overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda_n} \end{pmatrix}$$

und deshalb $AA^* = A^*A$. Also ist f normal nach Bemerkung (1). ∎

KOROLLAR. *Sei (V,σ) ein endlichdimensionaler unitärer Vektorraum und $f \in \mathrm{Hom}_{\mathbb{C}}(V,V)$. Dann sind folgende Aussagen äquivalent.*

(1) f ist normal.

(2) V hat eine Orthonormalbasis aus Eigenvektoren von f.

BEWEIS. Nach dem Fundamentalsatz der Algebra zerfällt das charakteristische Polynom f über $\mathbb{C}$ in Linearfaktoren. Also ergibt sich aus dem obigen Satz die Behauptung. □

9.1.3. Hauptachsentransformation. Als Folgerung aus dem gerade bewiesenen Satz ergibt sich jetzt leicht ein neuer Beweis der Hauptachsentransformation für euklidische Vektorräume. Er beruht auf dem folgenden einfachen Lemma.

LEMMA. *Sei (V,σ) endlichdimensionaler euklidischer Vektorraum und $f \in \mathrm{Hom}_{\mathbb{R}}(V,V)$ selbstadjungiert. Dann zerfällt das charakteristische Polynom von f über $\mathbb{R}$ in Linearfaktoren.*

BEWEIS. Sei A die darstellende Matrix von f bzgl. irgendeiner Orthonormalbasis von V. Dann ist A symmetrisch, nach Bemerkung (1). Faßt man A als $\mathbb{C}^{n\times n}$-Matrix auf, so ist A hermitesch und damit hat (nach dem Satz in 8.3.13) das charakteristische Polynom $\det(A - tE)$ nur reelle Nullstellen über $\mathbb{C}$ (denn A ist darstellende Matrix eines selbstadjungierten Endomorphismus). Andererseits zerfällt p_f nach dem Fundamentalsatz der Algebra über $\mathbb{C}$ in Linearfaktoren. Insgesamt hat man, daß p_f über $\mathbb{R}$ in Linearfaktoren zerfällt. □

KOROLLAR (Hauptachsentransformation). *Sei (V,σ) endlichdimensionaler euklidischer Vektorraum und $f \in \mathrm{Hom}_{\mathbb{R}}(V,V)$. Dann sind folgende Aussagen äquivalent.*

(1) f ist selbstadjungiert.
(2) V hat eine Orthonormalbasis aus Eigenvektoren von f.

BEWEIS. (1) $\Rightarrow$ (2). Die Behauptung folgt sofort aus dem Satz in 9.1.2 zusammen mit obigem Lemma; man braucht noch, daß jeder selbstadjungierte Endomorphismus normal ist.

(2) $\Rightarrow$ (1). Sei $e_1,\dots,e_n$ eine Orthonormalbasis aus Eigenvektoren von f. Dann gilt

$$f\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = A\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \text{mit } A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ und } \lambda_j \in \mathbb{R}.$$

Also ist $A = A^t$. □

9.2. Zerlegung von Endomorphismen; geometrische Deutung

9.2.1. Charakterisierung orthogonaler Projektionen. Sei V ein K-Vektorraum und f ein Endomorphismus. f heißt *Projektion*, wenn es Unterräume $U, W \subseteq V$ gibt mit $V = U \oplus W$, so daß für alle $u \in U$ und $w \in W$ gilt: $f(u+w) = u$. Man nennt f *Parallelprojektion von U in Richtung W*.

Sei (V,σ) endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und f ein Endomorphismus. f heißt *orthogonale Projektion*, wenn es einen Unterraum $U \subseteq V$ gibt, so daß für alle $p \in U$ und $h \in U^{\perp}$ gilt $f(p+h) = p$.

LEMMA. *Sei (V,σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und f ein Endomorphismus. Folgende Aussagen sind äquivalent.*

(1) f ist orthogonale Projektion.
(2) f ist selbstadjungiert und es gilt $f^2 = f$.

BEWEIS. (1) $\Rightarrow$ (2). Sei $V = U \oplus U^\perp$ und f orthogonale Projektion auf U. Sei $e_1, \ldots, e_m$ Orthonormalbasis von U und $e_{m+1}, \ldots, e_n$ Orthonormalbasis von $U^\perp$. Dann ist $e_1, \ldots, e_n$ Orthonormalbasis von V, und es gilt

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}.$$

Daraus folgt $f^2 = f$ sowie die Selbstadjungiertheit von f.

(2) $\Rightarrow$ (1). Sei f selbstadjungiert mit $f^2 = f$. Nach der Folgerung in 9.1.2 (bzw. in 9.1.3) hat V eine Orthonormalbasis $e_1, \ldots, e_n$ aus Eigenvektoren von f. Sei $x \neq 0$ Eigenvektor von f zum Eigenwert λ. Dann gilt

$$\begin{aligned} &f(x) = \lambda x \\ &f(x) = f^2(x) = \lambda f(x) = \lambda^2 x, \quad \text{wegen } f^2 = f \end{aligned}$$

also

$$\begin{aligned} &\lambda = \lambda^2 && \text{da } x \neq 0 \\ &0 = \lambda(\lambda - 1) \\ &\lambda = 0 \text{ oder } \lambda = 1. \end{aligned}$$

Seien $e_1, \ldots, e_m$ die e_j mit $\lambda_j = 1$. Dann

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}.$$

Also ist f orthogonale Projektion auf $\langle e_1, \ldots, e_m \rangle$ (insbesondere $f = 0$, falls alle $\lambda_j = 0$). □

9.2.2. Streckungen, Drehungen und Spiegelungen. Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und f ein Endomorphismus von V. f heißt *Streckung* (oder *Dehnung*) genau dann, wenn f selbstadjungiert ist und alle Eigenwerte von f größer als 0 sind.

Sei f eine Streckung. Dann gibt es nach der Folgerung in 9.1.2 (bzw. in 9.1.3) eine Orthonormalbasis $e_1, \ldots, e_n$ von V aus Eigenvektoren von f und $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit $\lambda_i > 0$ für $1 \leq i \leq n$, so daß

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

f bewirkt also in der Richtung e_j eine Streckung, falls $\lambda_j \geq 1$, oder eine Stauchung, falls $\lambda_j \leq 1$ ist.

Definition. Sei (V, σ) ein endlichdimensionaler euklidischer Vektorraum und f ein Endomorphismus.

(1) f heißt *Drehung*, falls f orthogonal und $\det f = 1$ ist.
(2) f heißt *Spiegelung*, falls f orthogonal und $\det f = -1$ ist.

Bemerkungen. (1) Ist f orthogonal, so gilt stets $\det f = +1$ oder $\det f = -1$. Dies sieht man wie folgt. Sei A die darstellende Matrix von f bzgl. der Orthonormalbasis $e_1, \ldots, e_n$. Dann gilt:

$$A^* = A^t = A^{-1} \qquad \text{da } f \text{ orthogonal}$$
$$AA^t = E$$
$$\det(AA^t) = \det A \cdot \underbrace{\det(A^t)}_{=\det A} = \det(E) = 1$$

$\det A = 1$ oder $\det A = -1$.

(2) Sei (V, σ) endlichdimensionaler euklidischer Vektorraum, $U \subseteq V$ Untervektorraum und $V = U \oplus U^\perp$. Dann definiere

$$f\colon V \to V$$
$$x = p + h \mapsto p - h \quad \text{mit } p \in U \text{ und } h \in U^\perp.$$

Sei etwa $e_1, \ldots, e_m$ eine Orthonormalbasis von U und $e_{m+1}, \ldots, e_n$ eine Orthonormalbasis von $U^\perp$. Dann gilt

$$f\begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} = \underbrace{\begin{pmatrix} \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{m} & \underbrace{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}}_{n-m} \end{pmatrix}}_{=:A} \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix},$$

also $\det A = (-1)^{n-m}$. Demnach ist f Drehung, falls $n - m$ gerade ist, und Spiegelung, falls $n - m$ ungerade ist.

Ist zum Beispiel $V = \mathbb{R}^3$ und $U = \mathbb{R}e_1$, so folgt $U^\perp = \langle e_2, e_3 \rangle$ und wir haben eine Drehung um $\mathbb{R}e_1$. Ist $U = \langle e_1, e_2 \rangle$, so folgt $U^\perp = \mathbb{R}e_3$. und wir haben eine Spiegelung an der Ebene $\langle e_1, e_2 \rangle$.

Im Spezialfall, daß U eine Hyperebene ist (d.h. $m := \dim U = n - 1$), heißt f *Hyperebenenspiegelung*.

(3) Normalformen für Drehungen und Spiegelungen folgen im nächsten Abschnitt. Beispiel: Sei V zweidimensional, e_1, e_2 eine Orthonormalbasis von V und

$$f\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}}_{A} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \quad \text{mit } 0 \leq \varphi < 2\pi.$$

Dann gilt $AA^t = E$, also ist f orthogonal. Es gilt $\det A = \cos^2\varphi + \sin^2\varphi = 1$, also ist f eine Drehung. Es handelt sich um eine Drehung gegen den Uhrzeigersinn um den Winkel φ.

9.2.3. Zerlegung von Endomorphismen. Wir wollen uns jetzt überlegen, daß man in einem endlichdimensionalen euklidischen Vektorraum jeden Endomorphismus darstellen kann

- als Hintereinanderausführung einer orthogonalen Projektion, einer Streckung und einer anschließenden Drehung oder Spiegelung, oder auch
- als Hintereinanderausführung einer Drehung oder Spiegelung, einer Streckung und einer anschließenden orthogonalen Projektion.

Für unitäre Vektorräume gilt eine entsprechende Aussage.

SATZ. *Sei (V,σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und f ein Endomorphismus von V. Dann gibt es Endomorphismen p, p', s, s', d, d' mit p, p' orthogonale Projektionen, s, s' selbstadjungiert mit reellen, positiven Eigenwerten und d, d' unitär (bzw. orthogonal), so daß*

$$\begin{aligned} f &= d \circ s \circ p \\ &= p' \circ s' \circ d'. \end{aligned}$$

BEWEIS. $f^* \circ f$ ist selbstadjungiert, da $(f^* \circ f)^* = f^* \circ f^{**} = f^* \circ f$. Mit der Folgerung in 9.1.2 (bzw. in 9.1.3) erhält man eine Orthonormalbasis $e_1, \ldots, e_n$ von V aus Eigenvektoren von $f^* \circ f$. Es gilt also

$$(f^* \circ f) \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \text{mit } \lambda_i \in \mathbb{R}.$$

OBdA seien $\lambda_1, \ldots, \lambda_m \neq 0$ und $\lambda_{m+1}, \ldots, \lambda_n = 0$ für $0 \leq m \leq n$. Für $1 \leq i, j \leq m$ gilt

$$\sigma(f(e_i), f(e_j)) = \sigma(e_i, \underbrace{(f^* \circ f)(e_j)}_{=\lambda_j e_j}) = \lambda_j \delta_{ij},$$

also

$$\begin{aligned} &f(e_i) \neq 0 && \text{für } 1 \leq i \leq m \\ &f(e_i) \perp f(e_j) && \text{für } 1 \leq i, j \leq m,\ i \neq j. \end{aligned}$$

Nun normiere man die $f(e_i)$ für $1 \leq i \leq m$; man erhält

$$e_i' := \frac{f(e_i)}{\|f(e_i)\|} \qquad \text{für } 1 \leq i \leq m.$$

Also ist $e_1', \ldots, e_m'$ ein Orthonormalsystem. Nach dem Schmidtschen Orthonormalisierungsverfahren existiert eine Orthonormalbasis $e_1', \ldots, e_m', e_{m+1}', \ldots e_n'$ von V. Damit erhält man eine Matrix $C \in \mathbb{C}^{n\times n}$ (bzw. $\in \mathbb{R}^{n\times n}$) mit $C^* = C^{-1}$ und

$$\begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix} = C \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Es gilt $f(e_i) = \gamma_i e_i'$ mit $\gamma_i := \|f(e_i)\| \in \mathbb{R}$ sogar für $1 \leq i \leq n$ (mit $\gamma_{m+1} = \cdots = \gamma_n = 0$). Also

$$\begin{aligned} f\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} &= \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix} \begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix} \\ &= \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix} C \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} \gamma_1 & & & & & \\ & \ddots & & & & \\ & & \gamma_m & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} C}_{=:A} \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} \end{aligned}$$

Es seien d, s, p wie folgt definiert.

$$\begin{aligned} d\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} &= C \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \\ s\begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} &= \begin{pmatrix} \gamma_1 & & & & & \\ & \ddots & & & & \\ & & \gamma_m & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} \\ p\begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} &= \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} \end{aligned}$$

Also ist d unitär (orthogonal), s selbstadjungiert mit positiven Eigenwerten und p die orthogonale Projektion auf $\langle e_1, \ldots, e_m \rangle$. Es folgt $f = d \circ s \circ p$,

denn es gilt:

$$
\begin{aligned}
&(d \circ s \circ p) \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \\
&= (d \circ s)(\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}) \\
&= \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} d(s \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}) \\
&= \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} d(\begin{pmatrix} \gamma_1 & & & & & \\ & \ddots & & & & \\ & & \gamma_m & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}) \\
&= \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} \gamma_1 & & & & & \\ & \ddots & & & & \\ & & \gamma_m & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \underbrace{d \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}}_{C \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}} \\
&= f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.
\end{aligned}
$$

Die zweite Zerlegung ergibt sich durch Anwendung der ersten auf f^* anstelle von f.

$$
\begin{aligned}
f^* &= d \circ s \circ p \\
f = f^{**} &= \underbrace{p^*}_{=p} \circ \underbrace{s^*}_{=s} \circ d^*.
\end{aligned}
$$

Es genügt also zu zeigen, daß d^* unitär (bzw. orthogonal) ist; dies folgt aber aus $d^{**} = (d^{-1})^* = (d^*)^{-1}$. □

BEMERKUNGEN. (1) Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und f ein Endomorphismus von V. Dann gibt es

g selbstadjungierter Endomorphismus mit Eigenwerten ≥ 0

d unitärer (bzw. orthogonaler) Endomorphismus

so daß

$f = d \circ g$ (siehe vorheriger Satz; setze $g := s \circ p$)

Dabei ist g *eindeutig bestimmt.* (Ebenso für $f = g' \circ d'$ mit g' selbstadjungiert und alle Eigenwerte von $g' \geq 0$, sowie d' unitär bzw. orthogonal).

BEWEIS. Seien g_1, g selbstadjungiert, alle ihre Eigenwerte ≥ 0 und d_1, d unitär (bzw. orthogonal), so daß $f = d \circ g = d_1 \circ g_1$ ist. Man erhält:

$$f^* \circ f = \underbrace{g^*}_{=g} \circ \underbrace{d^*}_{=d^{-1}} \circ d \circ g = g^2$$
$$= g_1^* \circ d_1^* \circ d_1 \circ g_1 = g_1^2.$$

Da g_1 selbstadjungiert ist, existiert nach der Folgerung in 9.1.2 (bzw. in 9.1.3) eine Orthonormalbasis $e_1, \ldots, e_n$ von V aus Eigenvektoren von g_1.

$$g_1 \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \text{mit } \lambda_j \geq 0 \text{ für } 1 \leq j \leq n.$$

Also gilt

$g_1^2(e_j) = \lambda_j^2 e_j = g^2(e_j)$

Nach (9.1) (siehe unten) folgt somit

$g_1(e_j) = \lambda_j e_j = g(e_j)$ für $1 \leq j \leq n$,

$g = g_1$

Zu zeigen bleibt, daß für jeden selbstadjungierten Endomorphismus g, dessen sämtliche Eigenwerte ≥ 0 sind, jedes $x \in V$ mit $x \neq 0$ und jedes $\lambda \in \mathbb{R}$ mit $\lambda \geq 0$ gilt

(9.1) $$g(x) = \lambda x \quad \text{genau dann, wenn} \quad g^2(x) = \lambda^2 x.$$

Dies beweisen wir wie folgt. $\Rightarrow$. Klar. $\Leftarrow$. Sei $e_1, \ldots, e_n$ eine Orthonormalbasis von V aus Eigenvektoren von g, also

$$g \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \text{mit } \lambda_j \geq 0 \text{ für } 1 \leq j \leq n.$$

Sei $x = \sum_{j=1}^{n} \alpha_j e_j$. Dann gilt

$$
\begin{aligned}
& g^2(x) = \lambda^2 x \\
& \sum_{j=1}^{n} \alpha_j \lambda_j^2 e_j = \sum_{j=1}^{n} \alpha_j \lambda^2 e_j \\
& \alpha_j \lambda_j^2 = \alpha_j \lambda^2 && \text{für } 1 \le j \le n \\
& \lambda_j = \lambda && \text{für } 1 \le j \le n \text{ mit } \alpha_j \neq 0, \text{ da } \lambda_j, \lambda \ge 0.
\end{aligned}
$$

Also ergibt sich $\sum_{j=1}^{n} \alpha_j \lambda_j e_j = \sum_{j=1}^{n} \alpha_j \lambda e_j$ und damit $g(x) = \lambda x$. □

(2) Ist in (1) speziell f ein Isomorphismus, so sind die nach (1) existierenden Endomorphismen d, g mit $f = d \circ g$ offenbar eindeutig bestimmt.

9.2.4. Praktische Durchführung der Hauptachsentransformation und der Zerlegung von Endomorphismen. Sei (V, σ) ein endlichdimensionaler euklidischer Vektorraum, f ein selbstadjungierter Endomorphismus, $e_1, \ldots, e_n$ eine Orthonormalbasis von V und $A \in \mathbb{R}^{n \times n}$ mit

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = A \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Wir betrachten das folgende Problem.

(9.2)

Gesucht: Orthonormalbasis $e_1', \ldots, e_n'$ von V aus Eigenvektoren von f.

Ein äquivalentes Problem ist

(9.3)

Gesucht: $C \in O(n)$ und $\gamma_1, \ldots, \gamma_n \in \mathbb{R}$ mit $C^t A C = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}$.

Wir erinnern noch einmal an den Beweis der Äquivalenz: (9.2)⇒(9.3). Sei $e_1', \ldots, e_n'$ eine Orthonormalbasis von V aus Eigenvektoren von f, also

$$f \begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix} = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix} \begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix}.$$

Es gibt ein $C \in O(n)$ mit

$$\begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix} = C^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Also:

$$f\begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix} = C^t f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^t A \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^t AC \begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix}$$

$$C^t AC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

(9.3)⇒(9.2): Sei $C \in O(n)$ gegeben mit $C^t AC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}$. Setze $\begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix} := C^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$. Dann gilt

$$f\begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix} = C^t f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^t A \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^t AC \begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix}$$

$$= \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix} \begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix}.$$

Nun wollen wir eine Lösung für das Problem (9.2) bzw. (9.3) konstruieren:

Schritt 1. Bestimme alle Eigenwerte von A als Nullstellen von $\det(A - tE)$. (Dies ist insbesondere für $n = 2, 3$ möglich).

Schritt 2. Bestimme zu jedem Eigenwert λ von A eine Basis des Eigenraums

$$U_\lambda = \{\, x \in \mathbb{R}^n \mid Ax = \lambda x \,\}.$$

Dies entspricht der Lösung eines homogenen Gleichungssystems mit der Matrix $A - \lambda E$.

Schritt 3. Bestimme nach dem Schmidtschen Orthonormalisierungsverfahren eine Orthonormalbasis in jedem Eigenraum U_λ. Sei C die Matrix aus $\mathbb{R}^{n\times n}$, deren Spalten die Orthonormalbasen der U_λ sind, λ Eigenwert von A. Dann gilt:

(a) $C \in O(n)$

(b) $C^t AC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}$ mit γ_i Eigenwert von A (wobei mehrfache Eigenwerte entsprechend mehrfach aufgeschrieben sind).

Beweis. (a) Wir zeigen $C \in O(n)$. Nach Bemerkung (1) (s.u.) gilt $\mathbb{R}^n = \bigoplus U_\lambda$, λ Eigenwert von A, und $U_\lambda \perp U_\gamma$ für $\lambda \neq \gamma$. (Denn U_λ ist Eigenraum zum Eigenwert λ der Abbildung $\mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto Ax$). Also bilden die Spalten von C eine Orthonormalbasis von $\mathbb{R}^n$ und somit ist $C \in O(n)$.

(b) Wir zeigen, daß C^tAC Diagonalmatrix ist. Folgende Aussagen sind äquivalent.

$$C^tAC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}$$

$$AC = C \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}$$

$Ac_i = c_i\gamma_i$ (c_i i-te Spalte von C) für $1 \leq i \leq n$.

Solche γ_i existieren nach Konstruktion von C. □

BEMERKUNGEN. (1) Umformulierung der Folgerung in 9.1.2 (bzw. in 9.1.3). Seien (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum, f ein Endomorphismus von V und $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f. (Also $\lambda_i \neq \lambda_j$ für $i \neq j$). Sei

$$V_\lambda = \{\, x \in V \mid f(x) = \lambda x \,\}$$

der Eigenraum von f zum Eigenwert λ. f ist genau dann normal (bzw. selbstadjungiert), wenn $V = \bigoplus_{i=1}^m V_{\lambda_i}$ und $V_{\lambda_i} \perp V_{\lambda_j}$ $(i \neq j)$.

Dies beweist man wie folgt. $\Leftarrow$. V hat dann eine Orthonormalbasis aus Eigenvektoren von f. $\Rightarrow$. Sei $e_1, \ldots, e_n$ eine Orthonormalbasis von V aus Eigenvektoren von f und seien Räume V_i definiert durch

$$V_i := \langle \{\, e_j \mid e_j \text{ Eigenvektor zum Eigenwert } \lambda_i \,\} \rangle.$$

Dann ist $V = \bigoplus_{i=1}^m V_i$ und $V_i \perp V_j$ für $i \neq j$. Außerdem gilt:

$V_i \subseteq V_{\lambda_i}$

$\dim V_i \leq \dim V_{\lambda_i}$

$$\dim V = \sum_{i=1}^m \dim V_i \leq \sum_{i=1}^m \dim V_{\lambda_i} \leq \dim V \quad \text{nach dem Satz in 7.2.1}$$

$\dim V_i = \dim V_{\lambda_i}$

$V_i = V_{\lambda_i}$.

(2) Entsprechend berechnet man eine Orthonormalbasis aus Eigenvektoren eines normalen Endomorphismus f eines unitären Vektorraums (V, σ).

(3) Zur Folgerung über die Zerlegung von Endomorphismen: Die Zerlegung läßt sich explizit angeben, falls eine Orthonormalbasis aus Eigenvektoren von $f^* \circ f$ (selbstadjungiert) bekannt ist.

9.2.5. Simultanes Diagonalisieren. Es folgt nun eine für die Anwendungen wichtige Verallgemeinerung der Folgerung in 9.1.2 (bzw. in 9.1.3).

SATZ. *Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und $M \subseteq \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $M \subseteq \mathrm{Hom}_{\mathbb{R}}(V, V)$) eine Teilmenge. Dann sind äquivalent:*

(1) (a) Für alle $f \in M$ gilt: f ist normal (bzw. selbstadjungiert).

(b) Für alle $f, g \in M$ gilt: $f \circ g = g \circ f$.

(2) Es existiert eine Orthonormalbasis $e_1, \ldots, e_n$ von V, so daß für alle $f \in M$ gilt: $e_1, \ldots, e_n$ sind Eigenvektoren von f. ($e_1, \ldots, e_n$ ist also eine „simultane" Orthonormalbasis).

BEWEIS. (2) $\Rightarrow$ (1). (a) gilt nach der Folgerung 9.1.2 (bzw. in 9.1.3). (b) Seien $f, g \in M$. Dann existieren γ_i, λ_i für $1 \leq i \leq n$ mit:

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \text{ und } g \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Also folgt $f \circ g = g \circ f$.

(1) $\Rightarrow$ (2). Sei (V, σ) unitär. (Den euklidischen Fall behandelt man analog).

Schritt 1. Sei U der von M erzeugte Unterraum von $\mathrm{Hom}_{\mathbb{C}}(V, V)$. Es ist $\dim U < \infty$, denn es gilt $\dim \mathrm{Hom}_{\mathbb{C}}(V, V) = (\dim V)^2 < \infty$. Sei $f_1, \ldots, f_m$ eine Basis von U. Es genügt, eine simultane Orthonormalbasis $e_1, \ldots, e_n$ aus Eigenvektoren für $f_1, \ldots, f_m$ zu finden. Ist nämlich $e_1, \ldots, e_n$ eine solche Orthonormalbasis und ist $f \in M$, so ist $f = \sum_{j=1}^{m} \alpha_j f_j$ mit $\alpha_j \in \mathbb{C}$ und damit folgt für $1 \leq i \leq n$:

$$\begin{aligned} f(e_i) &= \sum_{j=1}^{m} \alpha_j f_j(e_i) \\ &= \sum_{j=1}^{m} \alpha_j \gamma_{ij} e_i \qquad \gamma_{ij} \text{ Eigenwert von } f_j \text{ zum Eigenvektor } e_i \\ &= \underbrace{\Big(\sum_{j=1}^{m} \alpha_j \gamma_{ij}\Big)}_{\in \mathbb{C}} e_i. \end{aligned}$$

Also ist e_i Eigenvektor von f.

Schritt 2. Nach Schritt 1 können wir also ohne Einschränkung $M = \{f_1, \ldots, f_m\}$ setzen. Wir beweisen nun die Aussage des Satzes durch Induktion nach m.

$m = 1$. Siehe Folgerung 9.1.2.

$m-1 \Rightarrow m$. Seien $\lambda_1, \ldots, \lambda_t$ die verschiedenen Eigenwerte von f_1 und V_{λ_i} der Eigenraum zum Eigenwert λ_i. Nach der vorangegangenen Bemerkung (1) (Umformulierung der Folgerung in 9.1.2) gilt $V = \bigoplus_{i=1}^{t} V_{\lambda_i}$ und $V_{\lambda_i} \perp V_{\lambda_j}$ für $i \neq j$. Wegen $f_1 \circ f_j = f_j \circ f_1$ gilt $f_j[V_\lambda] \subseteq V_\lambda$ für jeden Eigenwert λ von f_1, denn für $x \in V_\lambda$ haben wir

$$\begin{aligned} &f_1(x) = \lambda x \\ &f_1(f_j(x)) = f_j(f_1(x)) = f_j(\lambda x) = \lambda f_j(x) \\ &f_j(x) \in V_\lambda. \end{aligned}$$

Sei nun λ ein fest gewählter Eigenwert von f_1 und $2 \leq j \leq t$. Setze

$$\begin{aligned} g_j := f_j \restriction V_\lambda \colon V_\lambda &\to V_\lambda \\ x &\mapsto f_j(x). \end{aligned}$$

Nach den Bemerkungen in 9.1.2 gilt:

- g_j ist normal, da $V = \bigoplus_{i=1}^{t} V_{\lambda_i}$ und $V_{\lambda_i} \perp V_{\lambda_j}$ für $i \neq j$, und
- $g_s \circ g_t = g_t \circ g_s$ für $2 \leq s, t \leq n$, da $f_s \circ f_t = f_t \circ f_s$.

Nach Induktionsvoraussetzung hat V_λ eine simultane Orthonormalbasis aus Eigenvektoren von g_j, $j = 2, \ldots, m$. Die Vereinigung dieser Orthonormalbasen für alle V_λ, λ Eigenwert von f_1, liefert die Behauptung. □

Bemerkungen. (1) Genauso zeigt man: Sei V ein endlichdimensionaler K-Vektorraum und $M \subseteq \mathrm{Hom}_K(V, V)$ eine Teilmenge. Dann sind äquivalent:
   (a) (i) Für alle $f \in M$ gilt: f ist diagonalisierbar.
       (ii) Für alle $f, g \in M$ gilt: $f \circ g = g \circ f$.
   (b) Es existiert eine simultane Basis von V aus Eigenvektoren aller $f \in M$.

(2) Matrizenformulierung des Satzes: Sei $M \subseteq \mathbb{C}^{n\times n}$ (bzw. $M \subseteq \mathbb{R}^{n\times n}$) eine Teilmenge. Dann sind äquivalent:
   (a) (i) Für alle $A \in M$ gilt: A ist normal (bzw. selbstadjungiert).
       (ii) Für alle $A, B \in M$ gilt: $AB = BA$.
   (b) Es existiert ein $C \in U(n)$ (bzw. $\in O(n)$), so daß für alle $A \in M$ gilt: C^*AC ist eine Diagonalmatrix.

9.3. Orthogonale Endomorphismen

Wir wollen uns eine Übersicht über alle orthogonalen Endomorphismen eines endlichdimensionalen euklidischen Vektorraums verschaffen. Zunächst zeigen wir, daß jeder orthogonale Endomorphismus als Hintereinanderausführung von Hyperebenenspiegelungen dargestellt werden kann. Wir betrachten dann den zwei- und dreidimensionalen Fall und können dort leicht die orthogonalen Endomorphismen klassifizieren. Mit der reellen Jordanschen Normalform ergibt sich, daß damit auch schon der allgemeine Fall einer beliebigen endlichen Dimension erledigt ist. Abschließend verschaffen wir uns noch eine Parameterdarstellung der $\mathrm{SO}(n)$, unter Verwendung der schon in 5.4 benutzten Jacobischen Rotationsmatrizen.

9.3.1. Zerlegung in Hyperebenenspiegelungen. Sei (V, σ) ein endlichdimensionaler euklidischer Vektorraum. Wir definieren

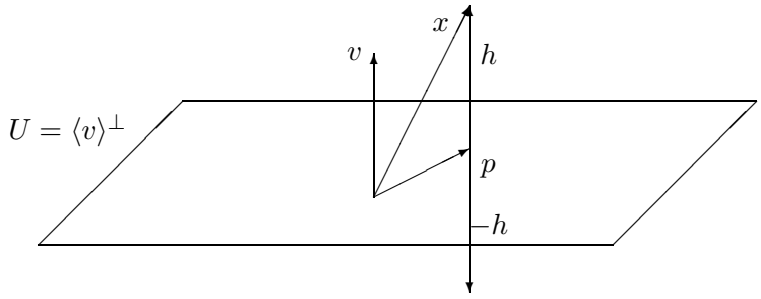
$$\mathrm{O}(V) := \mathrm{O}(V, \sigma) := \{\, f \in \mathrm{Hom}_{\mathbb{R}}(V, V) \mid f \text{ orthogonal} \,\},$$
$$\mathrm{SO}(V) := \mathrm{SO}(V, \sigma) := \{\, f \in \mathrm{O}(V) \mid \det f = 1 \,\}.$$

Wir wollen uns zunächst überlegen, daß sich jede Hyperebenenspiegelung auf einem Unterraum von V zu einer Hyperebenenspiegelung auf ganz V fortsetzen läßt.

Beweis. Sei etwa $\dim V = n$. Unter einer Hyperebene verstehen wir einen Unterraum $U \subseteq V$ mit $\dim U = n - 1$. Sei $U^\perp = \langle v \rangle$, also $U = \langle v \rangle^\perp$ und $V = \langle v \rangle^\perp \oplus \langle v \rangle$. Als Hyperebenenspiegelung hatten in wir in 9.2.2 Endomorphismen der Gestalt

$$f_v \colon V \to V$$
$$x = p + h \mapsto p - h \quad \text{mit } p \in \langle v \rangle^\perp \text{ und } h \in \langle v \rangle$$

bezeichnet.



Es gilt

$$f_v(x) = x - 2\frac{\sigma(x,v)}{\sigma(v,v)}v, \tag{9.4}$$

denn für $x = p + \alpha v$ mit $p \in \langle v\rangle^\perp$ und $\alpha \in \mathbb{R}$ ist

$$\begin{aligned} \sigma(x,v) &= \alpha\sigma(v,v) && \text{da } p \perp v,\\ f_v(x) &= p - \alpha v\\ &= x - 2\alpha v\\ &= x - 2\frac{\sigma(x,v)}{\sigma(v,v)}v. \end{aligned}$$

Aus (9.4) folgt offenbar unsere Behauptung, daß sich nämlich jede Hyperebenenspiegelung auf einem Unterraum von V zu einer Hyperebenenspiegelung auf ganz V fortsetzen läßt. □

Satz (Zerlegung in Hyperebenenspiegelungen). *Sei (V,σ) ein euklidischer Vektorraum mit $\dim V = n < \infty$, sei $f \in \mathrm{O}(V)$ und $f \neq \mathrm{id}$. Dann existieren ein $r \leq n$ und Hyperebenenspiegelungen $f_1, \ldots, f_r \in \mathrm{O}(V)$ mit $f = f_1 \circ \cdots \circ f_r$.*

Beweis. Durch Induktion nach n.
Basis $n = 1$. Sei e eine Orthonormalbasis von V. Dann

$$\begin{aligned} f(e) &= \pm e && \text{da } f \text{ orthogonal},\\ f(e) &= -e && \text{da } f \neq \mathrm{id},\\ f(\alpha e) &= -\alpha e && \text{für alle } \alpha \in \mathbb{R}. \end{aligned}$$

Also ist $f = f_e$.
Schritt $n-1 \Rightarrow n$ (für $n > 1$).

(1) Annahme: Es existiert ein $v \in V$ mit $f(v) = v$ und $v \neq 0$. Es ist

$$V = \underbrace{\langle v\rangle^\perp}_{=:U} \oplus \langle v\rangle.$$

Zunächst zeigen wir: $f[\langle v\rangle^\perp] \subseteq \langle v\rangle^\perp$. Sei dazu $u \in \langle v\rangle^\perp$. Dann gilt:

$\sigma(f(u), v) = \sigma(f(u), f(v)) = \sigma(u,v) = 0$ und deshalb

$f(u) \in \langle v\rangle^\perp$.

Also ist $f\restriction U \in \mathrm{O}(U)$. Nach Induktionsvoraussetzung existieren $v_1, \dots, v_r \in U$, $r \leq n-1$ mit

$$f\restriction U = (f_{v_1}\restriction U) \circ \cdots \circ (f_{v_r}\restriction U) \quad (\text{oder } f\restriction U = \mathrm{id},\ r = 0)$$

(siehe vorige Bemerkung). Damit folgt $f = f_{v_1} \circ \cdots \circ f_{v_r}$, denn für ein $u \in U$ gilt nach Konstruktion $f(u) = (f_{v_1} \circ \cdots \circ f_{v_r})(u)$, für v gilt $f(v) = v = f_{v_1}(\dots f_{v_r}(v)\dots)$, da $v \perp v_i$ für alle $1 \leq i \leq r$, und $V = U \oplus \langle v \rangle$.

(2) Reduktion des allgemeinen Falls auf (1). Es ist $f \neq \mathrm{id}$, also existiert ein $v \in V$ mit $f(v) \neq v$. Setze $w := f(v) - v$. Wir zeigen $(f_w \circ f)(v) = v$.

Zum Beweis beachte man, daß $2f(v) = f(v) + v + \underbrace{f(v) - v}_{w}$. Es gilt $(f(v) + v) \perp w$, denn

$$\begin{aligned}\sigma(f(v)+v, f(v)) &= \underbrace{\sigma(f(v), f(v))}_{=\sigma(v,v)} + \sigma(v, f(v)) \\ &= \sigma(v, v + f(v)) \\ &= \sigma(v + f(v), v).\end{aligned}$$

Also folgt nach der Definition von f_w

$$f_w(2f(v)) = f(v) + v + \underbrace{-w}_{v - f(v)} = 2v$$

$$f_w(f(v)) = v.$$

Nach (1) folgt nun wegen $f_w \circ f \in \mathrm{O}(V)$, daß es Hyperebenenspiegelungen $f_1, \dots, f_r \in \mathrm{O}(V)$ gibt ($r \leq n-1$) mit

$$f_w \circ f = f_1 \circ \cdots \circ f_r$$

$$f = f_w \circ f_1 \circ \cdots \circ f_r, \quad \text{da } f_w^2 = \mathrm{id}.$$

□

9.3.2. Fixräume.

DEFINITION. Sei V ein K-Vektorraum und f ein Endomorphismus. Dann heißt $\{\, x \in V \mid f(x) = x \,\}$ *Fixraum* von f.

KOROLLAR. *Es sei (V, σ) ein endlichdimensionaler euklidischer Vektorraum und $f \in \mathrm{O}(V)$.*

(1) Sind $f_1, \dots, f_r$ ($r \leq n$) Hyperebenenspiegelungen in $\mathrm{O}(V)$ mit $f = f_1 \circ \cdots \circ f_r$, so hat der Fixraum von f eine Dimension $\geq n - r$.

(2) Sei $\dim V = n$ ungerade.

- *Gilt $\det f = +1$, so ist $+1$ Eigenwert von f.*
- *Gilt $\det f = -1$, so ist -1 Eigenwert von f.*

BEWEIS. (1) Sei $f_i = f_{v_i}$ mit $v_i \in V$ für $1 \leq i \leq r$. Es ist $V = \langle v_1 \dots, v_r \rangle \oplus \langle v_1, \dots, v_r \rangle^\perp$, also $\dim \langle v_1, \dots, v_r \rangle \leq r$ und $\dim \langle v_1, \dots, v_r \rangle^\perp \geq n - r$. Ferner gilt

$$\langle v_1, \dots, v_r \rangle^\perp \subseteq \{\, x \in V \mid f(x) = x \,\},$$

denn aus $x \perp v_i$ für alle $1 \leq i \leq r$ folgt $(f_{v_1} \circ \cdots \circ f_{v_r})(x) = x$.

(2) Sei $\dim V = n$ ungerade. Nach dem Satz existieren Hyperebenenspiegelungen $f_1, \ldots, f_r$, $r \leq n$ mit $f = f_1 \circ \cdots \circ f_r$. Also ist $\det f = \det(f_1) \ldots \det(f_r) = (-1)^r$.

Fall $\det f = +1$. Dann ist r gerade, also $r < n$. Nach (1) hat f einen Fixpunkt $\neq 0$. Also ist 1 Eigenwert von f.

Fall $\det f = -1$. Dann ist $\det(-f) = (-1)^n \det f = (-1)(-1) = 1$. Also hat $-f$ einen Fixpunkt $x \neq 0$, d.h. $(-f)(x) = x$, d.h. $f(x) = -x$. Also ist -1 ein Eigenwert von f. ∎

9.3.3. Geometrische Veranschaulichung im Fall $\dim V = 2$ oder $= 3$. Sei (V, σ) ein euklidischer Vektorraum mit $\dim V = 2$. Es seien e_1, e_2 eine Orthonormalbasis von V und $v = \alpha_1 e_1 + \alpha_2 e_2 \in V$ $(\alpha_1, \alpha_2 \in \mathbb{R})$. Es gelte $\|v\| = 1$, d. h. $\alpha_1^2 + \alpha_2^2 = 1$. Wir betrachten die Geradenspiegelung

$$\begin{aligned} &f_v \colon V \to V \\ &f_v(x) = x - 2\sigma(x, v)v \quad (\text{da } \sigma(v, v) = 1). \end{aligned}$$

Man erhält

$$\begin{aligned} f_v(e_1) &= e_1 - 2\alpha_1(\alpha_1 e_1 + \alpha_2 e_2) \\ f_v(e_2) &= e_2 - 2\alpha_2(\alpha_1 e_1 + \alpha_2 e_2) \end{aligned}$$

und damit

$$f_v \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} 1 - 2\alpha_1^2 & -2\alpha_1\alpha_2 \\ -2\alpha_1\alpha_2 & 1 - 2\alpha_2^2 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}.$$

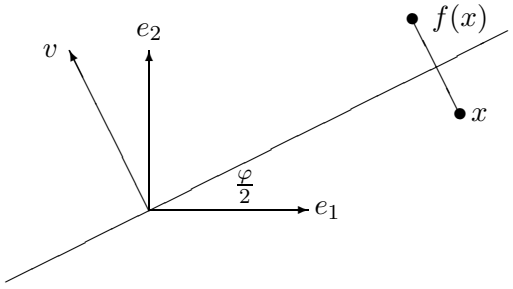
Es gilt $1 = \alpha_1^2 + \alpha_2^2$, also $1 - 2\alpha_1^2 = \alpha_2^2 - \alpha_1^2$. Wegen $\alpha_1^2 + \alpha_2^2 = 1$ kann man ein $\varphi \in \mathbb{R}$ wählen mit $\cos\frac{\varphi}{2} = \alpha_2$ und $\sin\frac{\varphi}{2} = -\alpha_1$. Nach dem Additionstheorem für sin und cos erhält man dann:

$$\begin{aligned} 1 - 2\alpha_1^2 = \alpha_2^2 - \alpha_1^2 &= \cos^2\frac{\varphi}{2} - \sin^2\frac{\varphi}{2} = \cos\varphi \\ -2\alpha_1\alpha_2 &= 2\cos\frac{\varphi}{2}\sin\frac{\varphi}{2} = \sin\varphi \\ 1 - 2\alpha_2^2 = \alpha_1^2 - \alpha_2^2 &= -\cos\varphi. \end{aligned}$$

Also ergibt sich

$$f_v \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}.$$

Geometrisch handelt es sich um eine Spiegelung an $\langle v \rangle^\perp$, $v = -\sin\frac{\varphi}{2}\, e_1 + \cos\frac{\varphi}{2}\, e_2$.



Spiegelachse

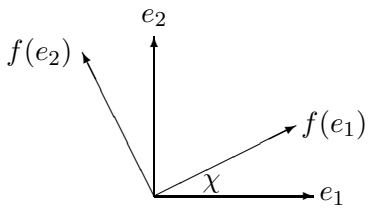
Komposition von Geradenspiegelungen: Seien $\varphi, \psi \in \mathbb{R}$.

$$\begin{aligned}
&\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \begin{pmatrix} \cos\psi & \sin\psi \\ \sin\psi & -\cos\psi \end{pmatrix} \\
&= \begin{pmatrix} \cos\varphi\cos\psi + \sin\varphi\sin\psi & \cos\varphi\sin\psi - \sin\varphi\cos\psi \\ \sin\varphi\cos\psi - \cos\varphi\sin\psi & \sin\varphi\sin\psi + \cos\varphi\cos\psi \end{pmatrix} \\
&= \begin{pmatrix} \cos(\psi-\varphi) & \sin(\psi-\varphi) \\ -\sin(\psi-\varphi) & \cos(\psi-\varphi) \end{pmatrix}
\end{aligned}$$

bedeutet eine Drehung um den Winkel $\psi - \varphi$ gegen den Uhrzeigersinn, denn

$$f \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \cos\chi & \sin\chi \\ -\sin\chi & \cos\chi \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

bedeutet geometrisch



Ergebnis: Sei e_1, e_2 eine Orthonormalbasis von V und $f \in \mathrm{O}(V)$. Dann gilt:

$\det f = -1 \quad \Rightarrow \quad f$ *Geradenspiegelung*, es existiert ein $\varphi \in \mathbb{R}$ mit

$$f \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \quad (-\pi < \varphi \leq \pi).$$

$\det f = +1 \quad \Rightarrow \quad f$ *Drehung*, es existiert ein $\varphi \in \mathbb{R}$ mit

$$f \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \quad (-\pi < \varphi \leq \pi).$$

f ist Komposition zweier Geradenspiegelungen.

Bemerkung. Diese Gestalt der darstellenden Matrizen kann man auch leicht direkt einsehen.

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{O}(2)$$

$$\Rightarrow A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \text{ oder } A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \text{ für ein } \varphi \in \mathbb{R}.$$

Beweis. A ist orthogonal, also gilt

$$\begin{aligned}
\alpha^2 + \beta^2 &= 1 \\
\gamma^2 + \delta^2 &= 1 \\
\alpha\gamma + \beta\delta &= 0.
\end{aligned}$$

Annahme: $\alpha, \beta \neq 0$.

$$\begin{aligned}
&\frac{\gamma}{\beta} = -\frac{\delta}{\alpha} =: \chi \\
&\gamma = \beta\chi \quad , \quad \delta = -\alpha\chi \\
&1 = \gamma^2 + \delta^2 = \chi^2(\beta^2 + \alpha^2) = \chi^2 \\
&\chi = \pm 1.
\end{aligned}$$

Damit folgt

$$A = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} \text{ oder } A = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Wegen $\alpha^2 + \beta^2 = 1$ gibt es ein $\varphi \in \mathbb{R}$ mit $\alpha = \cos\varphi$ und $\beta = \sin\varphi$.

Falls $\alpha = 0$: $\beta = \pm 1$, $\delta = 0$, $\gamma = \pm 1$.

Falls $\beta = 0$: $\alpha = \pm 1$, $\gamma = 0$, $\delta = \pm 1$. □

Wir behandeln jetzt den dreidimensionalen Fall. Sei also (V, σ) ein euklidischer Vektorraum, $\dim V = 3$ und $f \in \mathrm{O}(V)$.

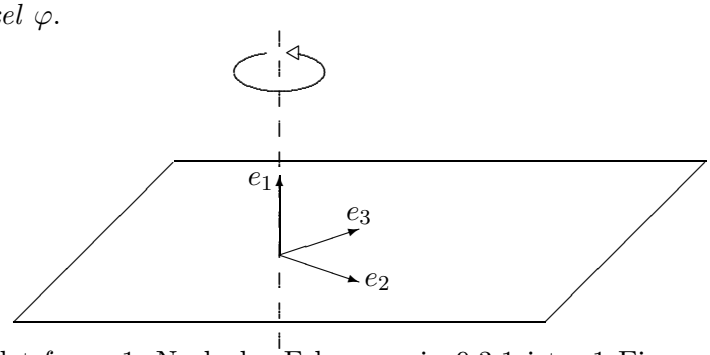
Fall $\det f = 1$. Nach der Folgerung in 9.3.1 ist 1 Eigenwert von f, da $\dim V = 3$ ungerade ist. Sei e_1 Eigenvektor zum Eigenwert 1 mit $\|e_1\| = 1$, und e_2, e_3 eine Orthonormalbasis von $\langle e_1\rangle^\perp$. Dann ist

$$f \restriction \langle e_2, e_3\rangle \colon \langle e_2, e_3\rangle \to \langle e_2, e_3\rangle$$

eine orthogonale Abbildung mit $\det(f \restriction \langle e_2, e_3\rangle) = 1$. Nach der vorangehenden Überlegung im zweidimensionalen Fall gibt es ein $\varphi \in \mathbb{R}$ mit

$$f \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}, \quad -\pi < \varphi \leq \pi.$$

f ist eine *Drehung* mit *Drehachse* $\mathbb{R}e_1$, *Drehebene* $\mathbb{R}e_2 \oplus \mathbb{R}e_3 = \langle e_1\rangle^\perp$ und *Drehwinkel* φ.



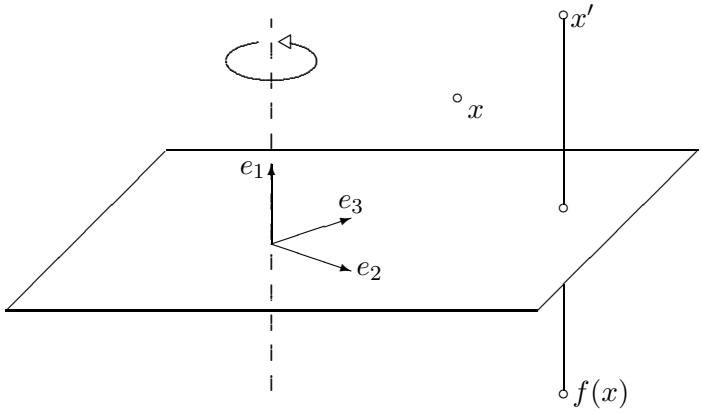
Fall $\det f = -1$. Nach der Folgerung in 9.3.1 ist -1 Eigenwert von f. Sei e_1 Eigenvektor von f zum Eigenwert -1 mit $\|e_1\| = 1$, und e_2, e_3 eine Orthonormalbasis von $\langle e_1\rangle^\perp$. Dann ist

$$f \restriction \langle e_2, e_3\rangle \colon \langle e_2, e_3\rangle \to \langle e_2, e_3\rangle$$

eine orthogonale Abbildung mit $\det(f \restriction \langle e_2, e_3\rangle) = 1$, da $\det f = -1$. Nach der vorangehenden Überlegung im zweidimensionalen Fall existiert ein $\varphi \in \mathbb{R}$ mit

$$f \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}, \quad -\pi < \varphi \leq \pi.$$

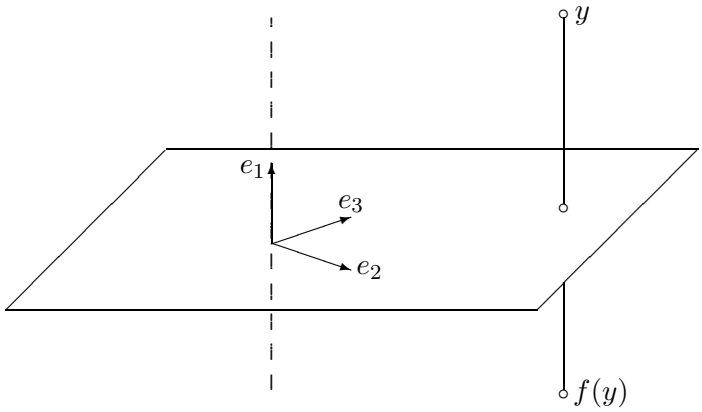
f ist eine *Drehspiegelung*, d.h. eine Drehung um die Achse $\mathbb{R}e_1$ um den Winkel φ und eine anschließende Spiegelung an der Drehebene $\mathbb{R}e_2 \oplus \mathbb{R}e_3$.



Grenzfälle: $\varphi = 0$.

$$f\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

Es handelt sich um eine Spiegelung an der Ebene $\langle e_2, e_3\rangle$, und zwar um die Hyperebenenspiegelung f_{e_1}.



$\varphi = \pi$.

$$f\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

Hier hat man eine Spiegelung am Ursprung.

Wir zeigen jetzt, daß der Drehwinkel φ nur von der Orientierung der gewählten Orthonormalbasis abhängt.

KOROLLAR. *Es sei (V, σ) ein euklidischer, orientierter Vektorraum mit $\dim V = 2$ und $f \in \mathrm{SO}(V)$. Dann existiert genau ein $\varphi \in \mathbb{R}$, $-\pi < \varphi \leq \pi$, so daß für alle positiv orientierten Orthonormalbasen e_1, e_2 von V gilt*

$$f\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} e_1 \\ e_2 \end{pmatrix}.$$

φ heißt orientierter Drehwinkel *und $|\varphi|$* absoluter Drehwinkel.

BEWEIS. Sei e_1, e_2 eine fest gewählte, positiv orientierte Orthonormalbasis von V. Dann existiert nach obiger Bemerkung genau ein $\varphi \in \mathbb{R}$, $-\pi < \varphi \leq \pi$, mit

$$f\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

(φ ist eindeutig bestimmt nach dem Verlauf von sin und cos). Sei e_1', e_2' eine beliebige, positiv orientierte Orthonormalbasis von V. Dann gibt es ein $A \in \mathrm{SO}(2)$ mit

$$\begin{pmatrix} e_1' \\ e_2' \end{pmatrix} = A\begin{pmatrix} e_1 \\ e_2 \end{pmatrix}.$$

A vertauscht mit $\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$, da $A = \begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix}$ für ein $\psi \in \mathbb{R}$ ist, d.h.

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} A = \begin{pmatrix} \cos(\varphi+\psi) & \sin(\varphi+\psi) \\ -\sin(\varphi+\psi) & \cos(\varphi+\psi) \end{pmatrix} = A\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}.$$

Daraus folgt

$$f\begin{pmatrix} e_1' \\ e_2' \end{pmatrix} = \underbrace{A\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}A^t}_{\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}}\begin{pmatrix} e_1' \\ e_2' \end{pmatrix}.$$

□

DEFINITION. Für $A = (\alpha_{ij}) \in K^{n\times n}$ ist die *Spur* von A definiert durch $\mathrm{Spur}A := \sum_{i=1}^n \alpha_{ii}$.

BEMERKUNG. Sei (V, σ) ein euklidischer, orientierter Vektorraum mit $\dim V = 2$ und $f \in \mathrm{SO}(V)$. Sei v_1, v_2 eine beliebige Basis von V und $A \in \mathbb{R}^{2\times 2}$ mit

$$f\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = A\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

Dann ist $\mathrm{Spur}A = 2\cos\varphi$, also ist der Drehwinkel $|\varphi|$ durch $\mathrm{Spur}A$ festgelegt.

BEWEIS. A ist ähnlich zu $\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$. Es genügt also zu zeigen, daß ähnliche Matrizen dieselbe Spur haben.

$\mathrm{Spur}CAC^{-1} = \mathrm{Spur}AC^{-1}C = \mathrm{Spur}A \quad$ für $A, C \in K^{n\times n}$ und $C \in \mathrm{GL}(n, K)$,

denn allgemein gilt für $A = (\alpha_{ij}), B = (\beta_{ij}) \in K^{n\times n}$:

$$\mathrm{Spur}AB = \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij}\beta_{ji} = \mathrm{Spur}BA.$$

□

KOROLLAR. *Sei (V, σ) ein euklidischer Vektorraum mit $\dim V = 3$ und $f \in \mathrm{SO}(V)$. Sei e_1 ein Eigenvektor von f zum Eigenwert 1 mit $\|e_1\| = 1$ (e_1 existiert nach der Folgerung in 9.3.1) und sei $\langle e_1\rangle^\perp$ orientiert. Dann existiert*

genau ein $\varphi \in \mathbb{R}$ *mit* $-\pi < \varphi \leq \pi$, *so daß für alle positiv orientierten Orthonormalbasen* e_2, e_3 *von* $\mathbb{R}e_1^{\perp}$ *gilt:*

$$f\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}.$$

BEWEIS. f induziert eine Drehung in $\langle e_1\rangle^{\perp}$. Also folgt die Behauptung unmittelbar aus dem obigen Korollar. □

BEMERKUNG. Sei $A \in \mathbb{R}^{3\times 3}$ beliebige darstellende Matrix von f bzgl. einer beliebigen Basis. Dann ist $\mathrm{Spur}A = 1 + 2\cos\varphi$, da A ähnlich ist zu

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{pmatrix}$$

(Beweis wie zuvor). Also ist der absolute Drehwinkel $|\varphi|$ durch $\mathrm{Spur}A$ festgelegt.

9.3.4. Geometrische Veranschaulichung im allgemeinen Fall. Aus den bisher behandelten Spezialfällen ergibt sich jetzt mit Hilfe des Satzes 8.2.2 von der reellen Jordanschen Normalform eine entsprechende Aussage für den allgemeinen Fall.

SATZ. *Sei* (V,σ) *ein euklidischer Vektorraum mit* $1 \leq \dim V < \infty$ *und sei* $f \in \mathrm{O}(V)$. *Dann existieren eine Orthonormalbasis* $e_1, \ldots, e_n$ *von* V, $r, s, t \in \mathbb{N}$, *und* $\varphi_1, \ldots, \varphi_t \in \mathbb{R}$ *mit*

$$f\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \underbrace{\begin{smallmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{smallmatrix}}_{r-mal} & & & & \\ & \underbrace{\begin{smallmatrix} -1 & & \\ & \ddots & \\ & & -1 \end{smallmatrix}}_{s-mal} & & & \\ & & \boxed{\begin{smallmatrix} \cos\varphi_1 & \sin\varphi_1 \\ -\sin\varphi_1 & \cos\varphi_1 \end{smallmatrix}} & & \\ & & & \ddots & \\ & & & & \boxed{\begin{smallmatrix} \cos\varphi_t & \sin\varphi_t \\ -\sin\varphi_t & \cos\varphi_t \end{smallmatrix}} \end{pmatrix}\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

BEWEIS. (1) Gezeigt wird: Es existieren Unterräume $U_i \subseteq V$, $1 \leq i \leq m$ mit $V = \bigoplus_{i=1}^{m} U_i$, $f[U_i] \subseteq U_i$, $\dim U_i = 1, 2$ und $U_i \perp U_j$ für alle $i \neq j$. Dies beweisen wir durch Induktion nach $\dim V$.

$\dim V = 1$. Klar.

Induktionsschritt. Nach der reellen Jordanschen Normalform 8.2.2 gibt es zu jedem $f \in \mathrm{Hom}_{\mathbb{R}}(V,V)$ einen Unterraum $U \subseteq V$

mit $\dim U = 1, 2$ und $f[U] \subseteq U$. Denn ist etwa

$$J = \begin{pmatrix} \boxed{\begin{smallmatrix} \alpha & -\beta \\ \beta & \alpha \end{smallmatrix}} & & & & \\ & \boxed{\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}} & & & \\ & & \ddots & & \\ & & & \boxed{\begin{smallmatrix} \alpha & -\beta \\ \beta & \alpha \end{smallmatrix}} & \\ & & & & \boxed{\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}} \\ & & & & \boxed{\begin{smallmatrix} \alpha & -\beta \\ \beta & \alpha \end{smallmatrix}} \end{pmatrix},$$

die Matrix aus 8.2.2 und ist U der von den beiden Basisvektoren, die zu dem letzten Kästchen $\boxed{\begin{smallmatrix} \alpha & -\beta \\ \beta & \alpha \end{smallmatrix}}$ gehören, erzeugte Unterraum, so ist $\dim U = 2$ und es gilt $f[U] \subseteq U$. Es folgt $f[U^\perp] \subseteq U^\perp$, denn $f[U] = U$ da f ein Isomorphismus ist. Also ist $f^{-1}\restriction U\colon U \to U$ orthogonal. Sei $v \in U^\perp$, $u \in U$. Dann ist

$$\sigma(f(v), u) = \sigma(f^{-1}(f(v)), f^{-1}(u)) = \sigma(v, f^{-1}(u)) = 0,$$

also $f(v) \in U^\perp$. Daraus folgt $V = U \oplus U^\perp$, $f\restriction U^\perp\colon U^\perp \to U^\perp$, $\dim U = 1$ oder $\dim U = 2$, $\dim U^\perp < \dim V$ und $U \perp U^\perp$. Auf $U^\perp$ läßt sich also die Induktionsvoraussetzung anwenden, woraus die Behauptung folgt.

(2) Nach (1) genügt es wegen 9.3.3, den Satz für $\dim V = 1, 2$ zu beweisen. Für $\dim V = 1$ gilt er offensichtlich, und für $\dim V = 2$ beachte man, daß die Matrix $\left(\begin{smallmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{smallmatrix}\right)$ nach 9.3.3 eine Spiegelung an $\langle v\rangle^\perp$, $v = -\sin\frac{\varphi}{2}\, e_1 + \cos\frac{\varphi}{2}\, e_2$ darstellt und deshalb durch Wahl einer geeigneten Orthonormalbasis in die Form $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ gebracht werden kann.

□

9.3.5. Parameterdarstellung der $\mathrm{SO}(n)$**.** Nach 9.3.3 ist $\mathbb{R} \to \mathrm{SO}(2)$, $\varphi \mapsto \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$ surjektiv; diese Abbildung heißt *Parameterdarstellung* von $\mathrm{SO}(2)$.

Gesucht ist eine ähnliche Parameterdarstellung für $\mathrm{SO}(n)$. Heuristisch betrachtet erkennt man: „Für $\mathrm{SO}(n)$ braucht man $\frac{n(n-1)}{2}$ Parameter“. Denn für $A = (\alpha_{ij}) \in \mathbb{R}^{n\times n}$ sind äquivalent

$A \in \mathrm{SO}(n)$

die Spalten von A bilden eine Orthonormalbasis von $\mathbb{R}^n$

$\sum_{i=1}^n \alpha_{ij}\alpha_{il} = \delta_{jl}$ für $j \le l$.

Das sind $\frac{n(n-1)}{2} + n$ Relationen, und es bleiben noch $\frac{n(n-1)}{2}$ „freie Parameter“ übrig. Dies sieht man wie folgt. Sei x die Anzahl der Elemente von A über der Diagonalen und n die Anzahl der Elemente auf der Diagonalen. Dann gilt: $n + 2x = n^2$, also $x = \frac{n^2-n}{2} = \frac{n(n-1)}{2}$.

Sei $1 \leq i < j \leq n$ und $\varphi \in \mathbb{R}$. In Abschnitt 5.4 hatten wir die die Jacobimatrix $R_{ij}(\varphi)$ definiert durch

$$\begin{pmatrix}
1 & & & & & & & & & \\
& \ddots & & & & & & & & \\
& & 1 & & & & & & & \\
& & & \cos\varphi & & & & \sin\varphi & & \\
& & & & 1 & & & & & \\
& & & & & \ddots & & & & \\
& & & & & & 1 & & & \\
& & & -\sin\varphi & & & & \cos\varphi & & \\
& & & & & & & & 1 & \\
& & & & & & & & & \ddots \\
& & & & & & & & & & 1
\end{pmatrix} \begin{matrix} \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{matrix}$$

$$\begin{matrix} & & i & & & & j & & & \end{matrix}$$

$$= (\alpha_{st}) \in \mathbb{R}^{n\times n} \text{ mit } \alpha_{st} = \begin{cases} \delta_{st} & \text{falls } (s,t) \neq (i,i),(i,j),(j,i),(j,j) \\ \cos\varphi & \text{falls } (s,t) = (i,i),(j,j) \\ \sin\varphi & \text{falls } (s,t) = (i,j) \\ -\sin\varphi & \text{falls } (s,t) = (j,i). \end{cases}$$

Bemerkungen. (1) Jacobimatrizen wurden in Abschnitt 5.4 beim Beweis der Hauptachsentransformation nach dem Jacobischen Verfahren verwendet.

(2) $R_{ij}(\varphi) \in \mathrm{SO}(n)$ beschreibt eine „ebene Drehung“. Sei (V,σ) ein euklidischer Vektorraum und $e_1,\ldots,e_n$ eine Orthonormalbasis von V,

$$f\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = R_{ij}(\varphi)\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Dann ist $V = \langle e_i, e_j\rangle \oplus \langle e_i, e_j\rangle^\perp$, $f\restriction\langle e_i,e_j\rangle$ ist eine Drehung und $f\restriction\langle e_i,e_j\rangle^\perp$ ist die Identität.

(3) $R_{ij}(\varphi)R_{ij}(-\varphi) = E$, d.h. $R_{ij}(\varphi)^t = R_{ij}(\varphi)^{-1} = R_{ij}(-\varphi)$.

Satz. *Sei $n \in \mathbb{N}$ und $A \in \mathrm{SO}(n)$. Dann ist A Produkt von $\frac{n(n-1)}{2}$ Jacobimatrizen.*

BEWEIS. 1. Seien $A = (\alpha_{rt})$, $1 \leq i < j \leq n$, $\varphi \in \mathbb{R}$, $s := \sin\varphi$ und $c := \cos\varphi$. Betrachte

$$AR_{ij}(\varphi) = (\alpha_{rt}) \begin{pmatrix} & & & i & & & j & & \\ 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & c & & & s & & i \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & \ddots & & \\ & & & & & & & 1 & \\ & & & -s & & & c & & j \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ & & & & & & & & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \dots & \alpha_{1i-1} & c\alpha_{1i}-s\alpha_{1j} & \alpha_{1i+1} & \dots & \alpha_{1j-1} & s\alpha_{1i}+c\alpha_{1j} & \alpha_{1i+1} & \dots \\ & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \\ \dots & \alpha_{ni-1} & c\alpha_{ni}-s\alpha_{nj} & \alpha_{ni+1} & \dots & \alpha_{nj-1} & s\alpha_{ni}+c\alpha_{nj} & \alpha_{ni+1} & \dots \end{pmatrix}$$

$$= \begin{pmatrix} \dots & a_{i-1} & ca_i-sa_j & a_{i+1} & \cdots & a_{j-1} & sa_i+ca_j & a_{j+1} & \dots \end{pmatrix},$$

wobei a_t die t-te Spalte von A ist. Zum Beweis des Satzes wird nun φ passend gewählt. Dazu benötigen wir folgende Aussage: Sind $\alpha, \beta \in \mathbb{R}$, so existiert ein $\varphi \in \mathbb{R}$ derart daß mit $s := \sin\varphi$ und $c := \cos\varphi$ gilt

$$\begin{aligned} c\alpha - s\beta &= 0 \\ s\alpha + c\beta &\geq 0. \end{aligned} \tag{9.5}$$

Beweis von (9.5). *Fall* $\alpha \neq 0$ oder $\beta \neq 0$. Wähle φ mit

$$\begin{aligned} \sin\varphi = s &= \frac{\alpha}{\sqrt{\alpha^2+\beta^2}} \\ \cos\varphi = c &= \frac{\beta}{\sqrt{\alpha^2+\beta^2}}. \end{aligned}$$

Dies ist möglich, da $s^2 + c^2 = 1$ ist. Dann gilt

$$\begin{aligned} c\alpha - s\beta &= 0 \\ s\alpha + c\beta &= \sqrt{\alpha^2+\beta^2} > 0. \end{aligned}$$

Fall $\alpha = \beta = 0$. Dann gilt für beliebige φ

$$\begin{aligned} c\alpha - s\beta &= 0 \\ s\alpha + c\beta &= 0. \end{aligned}$$

2. Nach (1) existiert ein $\varphi_{12} \in \mathbb{R}$, so daß mit $R_{12} := R_{12}(\varphi_{12})$ gilt:

$$AR_{12} = \begin{pmatrix} * & \cdots & * \\ \vdots & & \vdots \\ * & \cdots & * \\ 0 & \underbrace{(\sin\varphi)\alpha_{n1} + (\cos\varphi)\alpha_{n2}}_{\geq 0} & \alpha_{n3} \quad \ldots \quad \alpha_{nn} \end{pmatrix}$$

(nach (9.5) mit $\alpha = \alpha_{n1}, \beta = \alpha_{n2}$). Ebenso gibt es ein $\varphi_{23} \in \mathbb{R}$ so daß mit $R_{23} := R_{23}(\varphi_{23})$ gilt:

$$AR_{12}R_{23} = \begin{pmatrix} * & \cdots & * \\ \vdots & & \vdots \\ * & \cdots & * \\ 0 & 0 & \underbrace{\alpha'_{n3}}_{\geq 0} \quad \alpha_{n4} \quad \ldots \quad \alpha_{nn} \end{pmatrix}.$$

Durch Iteration erhält man $\varphi_{12}, \varphi_{23}, \varphi_{34}, \ldots, \varphi_{n-1,n}$ mit

$$AR_{12}R_{23}R_{34}\ldots R_{n-1n} = \begin{pmatrix} * & \cdots & * \\ \vdots & & \vdots \\ * & \cdots & * \\ 0 & \ldots \quad \ldots \quad 0 & \underbrace{\tilde{\alpha}_{nn}}_{\geq 0} \end{pmatrix} =: (\tilde{\alpha}_{rt}).$$

Wegen $A, R_{12}, \ldots, R_{n-1n} \in \mathrm{SO}(n)$ folgt $(\tilde{\alpha}_{rt}) \in \mathrm{SO}(n)$. Also folgt

$$1 = \sum_{j=1}^{n} \tilde{\alpha}_{nj}^2 = \tilde{\alpha}_{nn}^2 \quad \text{und damit} \quad \tilde{\alpha}_{nn} = 1, \quad \text{da } \tilde{\alpha}_{nn} \geq 0.$$

Außerdem gilt

$$\begin{aligned} 1 = \sum_{i=1}^{n} \tilde{\alpha}_{in}^2 = 1 + \sum_{i=1}^{n-1} \tilde{\alpha}_{in}^2 \quad & \text{da } \tilde{\alpha}_{nn} = 1, \text{ also} \\ \tilde{\alpha}_{in} = 0 \qquad\qquad\qquad\quad & \text{für } 1 \leq i \leq n-1. \end{aligned}$$

Als Ergebnis erhalten wir

$$AR_{12}R_{23}\ldots R_{n-1n} = \begin{pmatrix} \boxed{\quad * \quad} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \begin{matrix} 0 \ldots 0 \end{matrix} & 1 \end{pmatrix}.$$

Wendet man nun das gleiche Verfahren auf $\boxed{*} \in \mathrm{SO}(n-1)$ an, so folgt durch Iteration

$$A \underbrace{R_{12}R_{23}\ldots R_{n-1n}}\underbrace{R_{12}\ldots R_{n-2n-1}}\cdots\underbrace{R_{12}R_{23}}\underbrace{R_{12}} = \begin{pmatrix} \alpha & 0 & \ldots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{pmatrix},$$

$\alpha \in \mathbb{R}$, und es ist $\alpha = 1$, da $A, R_{st} \in \mathrm{SO}(n)$ sind. Das abschließende Ergebnis lautet also

$$\underbrace{A \underbrace{R_{12}R_{23}\dots R_{n-1n}}_{n-1}\underbrace{R_{12}\dots R_{n-2n-1}}_{n-2}\dots\underbrace{R_{12}R_{23}}_{2}\underbrace{R_{12}}_{1}}_{\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2} \text{ Jacobimatrizen}} = E.$$

Damit folgt die Behauptung, da $R_{ij}(\varphi)^{-1} = R_{ij}(-\varphi)$ ist. □

KOROLLAR. *Ist $A \in \mathrm{SO}(3)$, so existieren $\varphi_1, \varphi_2, \varphi_3 \in \mathbb{R}$ mit*

$$A = \begin{pmatrix} \cos\varphi_1 & \sin\varphi_1 & 0 \\ -\sin\varphi_1 & \cos\varphi_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi_2 & \sin\varphi_2 \\ 0 & -\sin\varphi_2 & \cos\varphi_2 \end{pmatrix} \begin{pmatrix} \cos\varphi_3 & \sin\varphi_3 & 0 \\ -\sin\varphi_3 & \cos\varphi_3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

BEWEIS. Nach dem Satz gibt es $\varphi_1, \varphi_2, \varphi_3 \in \mathbb{R}$ mit

$$A = R_{12}(\varphi_1)R_{23}(\varphi_2)R_{12}(\varphi_3).$$

□

Die Zahlen $\varphi_1, \varphi_2, \varphi_3$ heißen *Eulersche Winkel*. Sie werden z.B. bei der Beschreibung einer Kreiselbewegung verwendet. Geometrisch kann man dies wie folgt deuten. Seien $\varphi_1, \varphi_2, \varphi_3 \in \mathbb{R}$ und $A = R_{12}(\varphi_1)R_{23}(\varphi_2)R_{12}(\varphi_3)$. Sei (V, σ) ein euklidischer Vektorraum mit einer Orthonormalbasis e_1, e_2, e_3 und

$$f \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = A \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}.$$

Wir können dann f schreiben als Hintereinanderausführung der Drehungen f_1, f_2, f_3, also in der Form $f = f_1 \circ f_2 \circ f_3$ mit

$$f_3 \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} := R_{12}(\varphi_3) \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} =: \begin{pmatrix} e_1' \\ e_2' \\ e_3' \end{pmatrix} \quad \text{Orthonormalbasis von } V,$$

$$f_2 \begin{pmatrix} e_1' \\ e_2' \\ e_3' \end{pmatrix} := R_{23}(\varphi_2) \begin{pmatrix} e_1' \\ e_2' \\ e_3' \end{pmatrix} =: \begin{pmatrix} e_1'' \\ e_2'' \\ e_3'' \end{pmatrix} \quad \text{Orthonormalbasis von } V,$$

$$f_1 \begin{pmatrix} e_1'' \\ e_2'' \\ e_3'' \end{pmatrix} := R_{12}(\varphi_1) \begin{pmatrix} e_1'' \\ e_2'' \\ e_3'' \end{pmatrix} = \underbrace{R_{12}(\varphi_1)R_{23}(\varphi_2)R_{12}(\varphi_3)}_{A} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

f_3 ist eine Drehung mit Drehachse $\mathbb{R}e_3$ und Drehwinkel φ_3, f_2 eine Drehung mit Drehachse $\mathbb{R}e_1'$ und Drehwinkel φ_2, und f_1 eine Drehung mit Drehachse $\mathbb{R}e_3''$ und Drehwinkel φ_1.

BEMERKUNGEN. Es gibt hier einen Zusammenhang zur Topologie: man fasse $\mathrm{SO}(n)$, $\mathrm{O}(n)$ und $\mathrm{GL}(n, \mathbb{R})$ als Teilmengen von $\mathbb{R}^{n^2}$ mit der induzierten Topologie auf, also

$$\mathrm{SO}(n) \subseteq \mathrm{O}(n) \subseteq \mathrm{GL}(n, \mathbb{R}) \subseteq \mathbb{R}^{n \times n} \cong \mathbb{R}^{n^2}.$$

(1) $\mathrm{O}(n)$ ist beschränkt und abgeschlossen in $\mathbb{R}^{n^2}$, also *kompakt*, und ebenso $\mathrm{SO}(n)$.

BEWEIS. Sei $A = (\alpha_{ij}) \in \mathbb{R}^{n^2}$. Dann gilt

$$A \in \mathrm{O}(n) \equiv \sum_{j=1}^{n} \alpha_{ij}\alpha_{lj} = \delta_{il} \quad \text{für alle } i, l$$
$$A \in \mathrm{SO}(n) \equiv A \in \mathrm{O}(n) \text{ und } \det A = 1.$$

Also sind $\mathrm{O}(n)$ (und $\mathrm{SO}(n)$) abgeschlossen als Nullstellengebilde von Gleichungen. Außerdem folgt aus $A \in \mathrm{O}(n)$, daß $\sum_{j=1}^{n} \alpha_{ij}^2 = 1$ für alle i, also $|\alpha_{ij}| \leq 1$ für alle i, j. Also ist $\mathrm{O}(n)$ beschränkt. □

(2) Nach dem Satz ist

$$\begin{aligned} \Phi\colon \mathbb{R}^{\frac{n(n-1)}{2}} &\to \mathrm{SO}(n) \\ (\varphi_1, \ldots, \varphi_{\frac{n(n-1)}{2}}) &\mapsto R_{12}(\varphi_1) \ldots R_{n-1n}(\varphi_{n-1}) \\ &\quad R_{12}(\varphi_n) \ldots R_{n-2n-1}(\varphi_{2n-2}) \\ &\quad \ldots \\ &\quad R_{12}(\varphi_{\frac{n(n-1)}{2}}) \end{aligned}$$

surjektiv. Φ ist stetig, da $\varphi \mapsto R_{ij}(\varphi)$ und die Matrizenmultiplikation stetig sind. Da $\mathbb{R}^{\frac{n(n-1)}{2}}$ zusammenhängend ist, ist also auch $\mathrm{SO}(n)$ zusammenhängend. Aber es gibt keine Bijektion

$$\Psi\colon \mathbb{R}^{\frac{n(n-1)}{2}} \xrightarrow{\cong} \mathrm{SO}(n),$$

so daß Ψ und Ψ^{-1} stetig sind, denn $\mathrm{SO}(n)$ ist kompakt und $\mathbb{R}^{\frac{n(n-1)}{2}}$ ist nicht kompakt.

Es gibt jedoch eine „lokale Isomorphie“, die Cayley-Transformation (siehe Übungen).

$$\begin{aligned} \mathbb{R}^{\frac{n(n-1)}{2}} \cong \{\, A \in \mathbb{R}^{n\times n} \mid A^t = -A \,\} &\xrightarrow{\cong} \{\, B \in \mathrm{SO}(n) \mid \det(B+E) \neq 0 \,\} \\ A &\mapsto (E-A)(E+A)^{-1} \\ (E-B)(E+B)^{-1} &\leftarrow\!\shortmid B \end{aligned}$$

d.h. $\mathrm{SO}(n)$ (und auch $\mathrm{O}(n)$, s.u.) sind „Mannigfaltigkeiten“ der Dimension $\frac{n(n-1)}{2}$.

(3) Sei $A \in \mathrm{O}(n)$, $\det A = -1$, z.B.

$$A = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}.$$

Dann ist $\mathrm{O}(n)$ disjunkte Vereinigung von $\mathrm{SO}(n)$ und $A \cdot \mathrm{SO}(n)$ (also $\mathrm{O}(n) = \mathrm{SO}(n) \cup A \cdot \mathrm{SO}(n)$ und $\mathrm{SO}(n) \cap A \cdot \mathrm{SO}(n) = \emptyset$). Dies ist eine Zerlegung in *Zusammenhangskomponenten*. (Hierbei ist $A \cdot \mathrm{SO}(n) := \{\, AB \mid B \in \mathrm{SO}(n) \,\}$).

BEWEIS. Sei $B \in \mathrm{O}(n)$ mit $B \notin \mathrm{SO}(n)$. Dann ist $\det(B) = -1$, also $\det(A^t B) = 1$, damit $B = A \underbrace{A^t B}_{\in \mathrm{SO}(n)} \in A \cdot \mathrm{SO}(n)$. Außerdem ist

$\mathrm{SO}(n) \cap A \cdot \mathrm{SO}(n) = \emptyset$ und nach (b) ist $\mathrm{SO}(n)$ zusammenhängend, also ist auch $A \cdot \mathrm{SO}(n)$ zusammenhängend. □

BEMERKUNG. Aus dem Beweis des Satzes folgt allgemein für ein $A \in \mathbb{R}^{n\times n}$ (also nicht notwendig $A \in \mathrm{SO}(n)$): Es existiert ein $B \in \mathrm{SO}(n)$, welches Produkt von $\frac{n(n-1)}{2}$ Jacobimatrizen ist, mit der Eigenschaft, daß

$$AB = C = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1n} \\ & \ddots & \vdots \\ & & \gamma_{nn} \end{pmatrix}$$

eine obere Dreiecksmatrix ist. Außerdem gilt $A = CB^t$, also $A^t = BC^t$, wobei C^t eine untere Dreiecksmatrix ist. Damit erhalten wir: Es existiert ein $Q \in \mathrm{SO}(n)$ und eine untere Dreiecksmatrix R mit $A = QR$. Diese so genannte *QR-Zerlegung* spielt in der numerischen Mathematik eine wichtige Rolle.

KAPITEL 10

Gruppen, Moduln, Tensorprodukt

In diesem Kapitel wollen wir die grundlegenden algebraischen Hilfsmittel ausbauen. Als Anwendung ergibt sich in Abschnitt 10.1 eine Klassifikation der endlichen Untergruppen der SO(3). In Abschnitt 10.2 definieren wir Faktorgruppen und beweisen ihre universelle Eigenschaft; Anwendungen sind einige einfache zahlentheoretische Aussagen, z.B. $x^p = x \mod p$ für $x \in \mathbb{Z}$ und Primzahlen p. Der Begriff eines Moduls (eine Art Vektorraum über einem Ring statt eines Körpers) wird in Abschnitt 10.3 eingeführt; mit seiner Hilfe erhalten wir einen einfachen Beweis des Satzes von Cayley/Hamilton, welcher aussagt, daß das „Einsetzen“ einer Matrix in ihr eigenes charakteristisches Polynom stets die Nullmatrix ergibt. In Abschnitt 10.4 konstruieren wir schließlich die in der Physik und auch in der Differentialgeometrie wichtigen Tensorprodukte von Moduln; auf sie wird man in natürlicher Weise durch die Betrachtung von mehrstelligen linearen Abbildungen (sog. multilinearen Abbildungen) geführt.

10.1. Gruppen

In Abschnitt 1.2.2 hatten wir den Gruppenbegriff eingeführt und seine einfachsten Eigenschaften bewiesen. Wir wollen uns jetzt etwas mehr Kenntnisse über Gruppen und insbesondere endliche Gruppen verschaffen; Ziel ist eine Klassifikation der endlichen Untergruppen der SO(3). Im folgenden verwenden wir x, y, z (statt α, β, γ) für die Elemente einer Gruppe und schreiben xy statt $x \circ y$.

Sei G eine endliche Gruppe. Dann heißt die Anzahl der Elemente von G die *Ordnung* von G, geschrieben $\mathrm{Ord}(G)$. Für eine endliche Gruppe $G = \{g_1, \ldots, g_n\}$ kann man die Multiplikation in einer Gruppentafel explizit aufschreiben:

$\cdot$	g_1	g_2	$\cdots$	g_n
g_1	$g_1\ g_1$	$g_1\ g_2$	$\cdots$	$g_1\ g_n$
g_2	$g_2\ g_1$	$g_2\ g_2$	$\cdots$	$g_2\ g_n$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
g_n	$g_n\ g_1$	$g_n\ g_2$	$\cdots$	$g_n\ g_n$

In jeder Zeile und in jeder Spalte kommt jedes Element von G genau einmal vor.

10.1.1. Gruppenhomomorphismen, Untergruppen, Normalteiler. Seien G, H Gruppen und $f\colon G \to H$ eine Abbildung. f heißt *Gruppenhomomorphismus*, falls für alle $x, y \in G$ gilt

$$f(xy) = f(x)f(y).$$

f heißt *Gruppenisomorphismus*, falls f ein bijektiver Gruppenhomomorphismus ist. G ist *isomorph* zu H (geschrieben $G \cong H$), falls ein Gruppenisomorphismus $g\colon G \to H$ existiert.

BEMERKUNG. Seien G, H Gruppen und $f\colon G \to H$ ein Gruppenhomomorphismus. Dann gilt

$$f(e) = e$$
$$f(x^{-1}) = f(x)^{-1} \quad \text{für alle } x \in G.$$

BEWEIS.

$$f(e) = f(ee) = f(e)f(e), \qquad \text{also } f(e) = e.$$
$$f(e) = f(xx^{-1}) = f(x)f(x^{-1}) \quad \text{also } f(x)^{-1} = f(x^{-1}).$$

□

In 5.3.1 hatten wir bereits den Begriff einer Untergruppe wie folgt definiert. Eine Teilmenge U einer Gruppe G heißt *Untergruppe*, wenn die Gruppenstruktur auf G eine Gruppenstruktur auf U induziert, d.h. wenn gilt (1) $xy \in U$ für alle $x, y \in U$, (2) $e \in U$ und (3) $x^{-1} \in U$ für alle $x \in U$. Eine Untergruppe N einer Gruppe G heißt *Normalteiler*, wenn für alle $x \in G$ gilt

$$xN = Nx,$$

wobei $xN := \{\, xy \mid y \in N \,\}$ und $Nx := \{\, yx \mid y \in N \,\}$.

BEMERKUNG. Sei G eine Gruppe und $U \subseteq G$ eine Teilmenge. Dann ist U ist genau dann Normalteiler von G, wenn U eine Untergruppe von G ist und für alle $x \in G$ und $y \in U$ auch $xyx^{-1} \in U$ ist.

BEWEIS. $\Longrightarrow$. Aus $xU = Ux$ folgt $xyx^{-1} \in U$. $\Longleftarrow$. Es ist $xU \subseteq Ux$ und aus $x^{-1}U \subseteq Ux^{-1}$ folgt $Ux \subseteq xU$. Also $xU = Ux$. □

LEMMA. *Sei G eine Gruppe und $U \subseteq G$ eine Teilmenge. Dann sind äquivalent:*

(1) U ist Untergruppe von G.
(2) (a) $U \neq \emptyset$.
(b) Für alle $x, y \in U$ ist $xy^{-1} \in U$.

BEWEIS. (1) $\Rightarrow$ (2). klar. (2) $\Rightarrow$ (1). Da $U \neq \emptyset$ ist, gibt es ein $x_0 \in U$. Wegen (2b) ist dann $e = x_0 x_0^{-1} \in U$. Für ein beliebiges $y \in U$ ist wieder mit (2b) auch $y^{-1} = ey^{-1} \in U$. Schließlich folgt für beliebige $x, y \in U$ wegen $y^{-1} \in U$ und wegen (2b) auch $xy = x(y^{-1})^{-1} \in U$. □

LEMMA. *Seien G, H Gruppen, $f\colon G \to H$ ein Gruppenhomomorphismus, U eine Untergruppe von G und V eine Untergruppe von H. Dann gilt:*

(1) $f[U]$ ist Untergruppe von H.
(2) $f^{-1}[V]$ ist Untergruppe von G.
(3) Falls V Normalteiler von H ist, so ist $f^{-1}[V]$ Normalteiler von G.

BEWEIS. (1) und (2) sind klar. (3) Sei $x \in G$ und $y \in f^{-1}[V]$. Dann gilt

$$f(xyx^{-1}) = f(x)f(y)f(x)^{-1} \in V,$$

da $f(y) \in V$ und V Normalteiler ist. Also ist $xyx^{-1} \in f^{-1}[V]$. □

BEMERKUNG. Sei U Normalteiler in G und $f\colon G \to H$ ein Gruppenhomomorphismus. Dann muß $f[U]$ nicht Normalteiler in H sein. Dies sieht man durch folgendes Beispiel: Wähle $G \subseteq H$ Untergruppe, die kein Normalteiler von H ist und $U = G$, sowie $f\colon G \to H$ die Einbettung von G nach H.

DEFINITION. Sei G eine Gruppe und $M \subseteq G$ eine Teilmenge.

(1) $\langle M\rangle := \{\, x \in G \mid \exists n{\in}\mathbb{N}, x_1, \ldots, x_n{\in}M, \varepsilon_1, \ldots, \varepsilon_n{\in}\{1,-1\}\colon x = x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n} \,\}$ heißt *von M erzeugte Untergruppe*. (Dabei sei $\langle\emptyset\rangle := \{e\}$).

(2) G heißt *zyklisch*, falls es ein $x \in G$ gibt mit $G = \langle x\rangle$.

BEMERKUNGEN. (1) $\langle M\rangle$ ist die kleinste Untergruppe von G, die M enthält. Dies beweist man wie die analoge Aussage für Vektorräume.

(2) Ist $G = \langle x\rangle$ für ein $x \in G$, so ist $G = \{\, x^s \mid s \in \mathbb{Z} \,\}$.

BEISPIELE. (1) Ist V ein K-Vektorraum, so ist V mit der Addition eine Gruppe.

(2) $\mathbb{Z}$ ist mit der Addition eine Gruppe. Es gilt $\mathbb{Z} = \langle 1\rangle$. In additiven Gruppen schreibt man nx mit $n \in \mathbb{Z}$ statt x^n und $-x$ statt x^{-1}.

(3) Sei R ein Ring und $\mathrm{U}(R) := \{\, r \in R \mid \exists s{\in}R : rs = sr = 1 \,\}$. Dann ist $\mathrm{U}(R)$ mit der Multiplikation von R eine Gruppe, sie heißt *Einheitengruppe* von R. Im Spezialfall eines Körpers K ist $\mathrm{U}(K) = K^* = \{\, x \in K \mid x \neq 0 \,\}$.

(4) Sei M eine Menge und $S(M) := \{\, f\colon M \to M \mid f \text{ bijektiv} \,\}$. Dann ist $S(M)$ mit der Komposition von Abbildungen eine Gruppe. Sie heißt *symmetrische Gruppe* auf M.

Im Spezialfall $M = \{1, 2, \ldots, n\}$ für ein $n \in \mathbb{N}$ ist $S_n := S(\{1, \ldots, n\})$ die *Permutationsgruppe* von n Elementen, die wir bereits in 6.1.5 betrachtet hatten. Definiere den Gruppenhomomorphismus

$$\begin{aligned} \varepsilon_n\colon S_n &\to \{1,-1\} \\ \sigma &\mapsto \text{Vorzeichen der Permutation } \sigma \end{aligned}$$

und setze $A_n := \ker \varepsilon_n$. A_n ist die Menge der geraden Permutationen und heißt *alternierende Gruppe* von n Elementen. Leicht zu erkennen ist (siehe 6.1.5)

$$\mathrm{Ord}(S_n) = n!, \qquad \mathrm{Ord}(A_n) = \frac{n!}{2}.$$

Sei $N \subseteq M$ eine Teilmenge. Dann sind

$$\begin{aligned} G &:= \{\, f \in S(M) \mid f[N] = N \,\} \\ H &:= \{\, f \in S(M) \mid \forall x{\in}N\colon f(x) = x \,\} \end{aligned}$$

Untergruppen von $S(M)$, denn etwa für G gilt: 1. $\mathrm{id}_M \in G$. 2. Sind $f, g \in G$, so ist $(fg)[N] = f(g[N]) = f[N] = N$, also ist $fg \in G$. 3. Sei $f \in G$. Dann ist $f{\restriction}N\colon N \to N$ bijektiv, also gilt für alle $x \in N$: $f^{-1}(f(x)) = x$. Damit erhält man $f^{-1}[N] = N$, also ist auch $f^{-1} \in G$. Ebenso beweist man, daß H Untergruppe von S_n ist.

Insbesondere ist $S_n \subseteq S_{n+1}$ Untergruppe bei der Identifikation

$$S_n = \{\, \sigma \in S_{n+1} \mid \sigma(n+1) = n+1 \,\}.$$

(5) Sei V ein Vektorraum über einem Körper K. Dann ist $\mathrm{Aut}_K(V) := \{\, f \in \mathrm{Hom}_K(V,V) \mid f \text{ Isomorphismus} \,\}$ mit der Komposition von Abbildungen eine Gruppe. Sie heißt *Automorphismengruppe* von V.

(6) $\mathrm{GL}(n,K)$ heißt *allgemeine lineare Gruppe*. Es ist $\det\colon \mathrm{GL}(n,K) \to K^*$ ein Gruppenhomomorphismus und $\mathrm{SL}(n,K) := \ker\det$ ist die *spezielle lineare Gruppe*. Für $K = \mathbb{R}$ ist $\mathrm{O}(n)$ die *orthogonale Gruppe* und $\mathrm{SO}(n) = \ker\det\colon \mathrm{O}(n) \to \{1,-1\}$ ist die *spezielle orthogonale Gruppe*.

(7) Sei V ein endlichdimensionaler K-Vektorraum. Dann ist $\{\, g\colon V \to V \mid g \text{ Affinität} \,\}$ die *affine Gruppe* von V. (Erinnerung: Eine Abbildung $g\colon V \to V$ hieß Affinität, wenn es ein $f \in \mathrm{Aut}_K(V)$ und ein $a \in V$ gibt, so daß für alle $x \in V$ gilt $g(x) = f(x) + a$).

(8) Sei (V,σ) ein endlichdimensionaler euklidischer Vektorraum. Dann heißt $\{\, g\colon V \to V \mid g \text{ Bewegung} \,\}$ *euklidische Gruppe* von V. (Erinnerung: Eine Abbildung $g\colon V \to V$ ist genau dann eine Bewegung (oder Kongruenz), falls es ein $f \in \mathrm{O}(V)$ und ein $a \in V$ gibt, so daß für alle $x \in V$ gilt $g(x) = f(x) + a$).

DEFINITION. Sei (V,σ) ein endlichdimensionaler euklidischer Vektorraum und sei $F \subseteq V$ eine Teilmenge („geometrische Figur“). Dann heißt

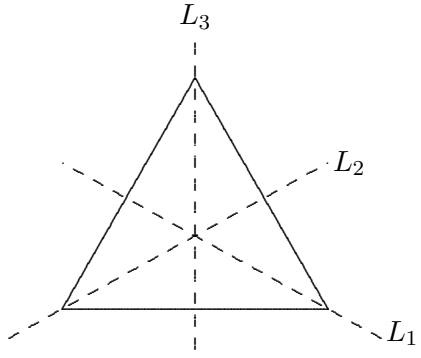
$$\{\, g\colon V \to V \mid g \text{ Bewegung von } V,\ g(F) = F \,\}$$

Symmetriegruppe von F.

BEMERKUNG. Die Symmetriegruppe von F ist tatsächlich eine Gruppe, nämlich eine Untergruppe der euklidischen Gruppe von V. Dies zeigt man wie im obigen Beispiel bei $S(M)$.

BEISPIELE. (Anschaulich).

(1) Sei $F \subseteq \mathbb{R}^2$ ein gleichseitiges Dreieck.



Es sei oBdA 0 im Schnittpunkt von L_1, L_1, L_3, und G sei die Symmetriegruppe von F. Leicht einzusehen ist $G \subseteq \mathrm{O}(\mathbb{R}^2)$. Sei f die Drehung um den Winkel $\frac{2\pi}{3}$ und g_i die Spiegelung an L_i für $i = 1,2,3$. Man beachte, daß $\mathrm{id}, f, f^2, g_1, g_2, g_3$ verschieden sind.

Behauptung. $G = \{\mathrm{id}, f, f^2, g_1, g_2, g_3\}$.

BEWEIS. Jedes $g \in G$ bewirkt eine Permutation der Eckpunkte (g ist durch die Bilder der Eckpunkte festgelegt). Also gilt

$$\{\mathrm{id}, f, f^2, g_1, g_2, g_3\} \subseteq G \hookrightarrow S_3,$$

($\hookrightarrow$ soll bedeuten: injektive Abbildung, also Einbettung) und S_3 hat die Ordnung 6. □

Aus dem obigen Beweis folgt sogar $G \cong S_3$. Es ist

$$G = \{\mathrm{id}, f, f^2, g, gf, gf^2\},$$

wobei g eine der Spiegelungen g_i ist (Beweis einfach). Ferner gelten folgende Relationen:

$$\begin{aligned} f^3 &= \mathrm{id} \\ g^2 &= \mathrm{id} \\ gfgf &= \mathrm{id} \quad \text{(leicht nachzurechnen)}, \end{aligned}$$

also auch $fg = gf^{-1} = gf^2$.

Allgemein kann man zeigen (dies wird hier aber nicht durchgeführt): Sei G die Symmetriegruppe des regulären n-Ecks mit $n \geq 3$, f die Drehung um den Winkel $\frac{2\pi}{n}$ und g eine beliebige Spiegelung in G. Dann ist

$$G = \{\mathrm{id}, f, \ldots, f^{n-1}, g, gf, \ldots, gf^{n-1}\} =: D_{2n}.$$

$D_{2n} = \langle f, g \rangle$ heißt *Diedergruppe* der Ordnung $2n$. (Oft wird auch die Bezeichnung D_n statt D_{2n} verwendet). Es gelten folgende Relationen:

$$\begin{aligned} f^n &= \mathrm{id} \\ f^t &\neq \mathrm{id} \qquad \text{für } 1 \leq t < n \\ fg &= gf^{n-1} \end{aligned}$$

(2) $\{ f \in \mathrm{SO}(\mathbb{R}^2) \mid f \text{ Symmetrie des regulären } n\text{-Ecks} \}$ wird von der Drehung um den Winkel $\frac{2\pi}{n}$ erzeugt, ist also die zyklische Gruppe der Ordnung n.

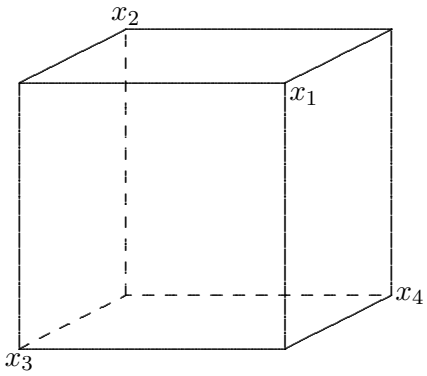
(3) Für ein $F \subseteq \mathbb{R}^3$ nennt man $\{ f \in \mathrm{SO}(\mathbb{R}^3) \mid f \text{ Symmetrie von } F \}$ die *Drehgruppe* von F.

Die Bestimmung der Drehgruppen für die so genannten *platonischen Körper* (dies sind die fünf konvexen, regulären Polyeder im $\mathbb{R}^3$) liefert folgendes Ergebnis.

Polyeder	Beschreibung	Drehgruppe	mit Ordnung
Tetraeder	4 Dreiecke	A_4	$4 \cdot 3 = 12$
Würfel	6 Quadrate	S_4	$6 \cdot 4 = 24$
Oktaeder	8 Dreiecke	S_4	$8 \cdot 3 = 24$
Dodekaeder	12 Fünfecke	A_5	$12 \cdot 5 = 60$
Ikosaeder	20 Dreiecke	A_5	$20 \cdot 3 = 60$

Wir wollen dies im folgenden begründen.

(a) Sei G die Drehgruppe des Würfels



Durch jeden Eckpunkt geht genau eine Raumdiagonale. Bei Symmetrien werden Raumdiagonalen in Raumdiagonale abgebildet. Sei L_i die Raumdiagonale durch den Punkt x_i. Betrachte

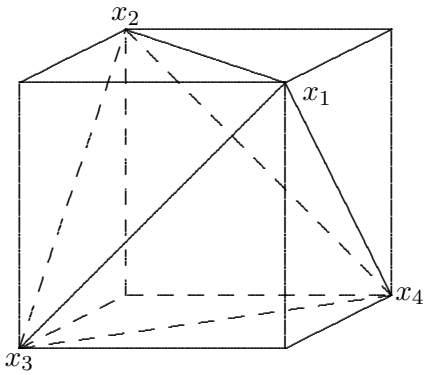
$$\varphi\colon G \to S_4$$
$$f \mapsto (i \mapsto f'(i)) \quad \text{wobei } f' \text{ definiert ist durch } f(L_i) = L_{f'(i)}.$$

φ ist ein injektiver Gruppenhomomorphismus. Zählt man die Elemente von G auf, so erhält man als Symmetrieachsen (Drehachsen von Elementen in G):

(i) 4 Raumdiagonalen: Je 2 Drehungen $\neq$ id.
(ii) 3 Geraden durch die Mittelpunkte gegenüberliegender Seiten: Je 3 Drehungen $\neq$ id.
(iii) 6 Geraden durch die Mittelpunkte gegenüberliegender Kanten: Je 1 Drehung $\neq$ id.

Also hat G mindestens $8 + 9 + 6 + 1 = 24$ Elemente. (Die letzte 1 gehört zur Identität). Da φ ein Monomorphismus ist, $\mathrm{Ord}(G) \geq 24$ und $\mathrm{Ord}(S_4) = 4! = 24$ ist, erhält man als Ergebnis $G \cong S_4$.

(b) Sei H die Drehgruppe des Tetraeders



Sei f die Drehung des Würfels mit $f(x_1) = x_1$, d.h. $f(L_1) = L_1$. Jedes solche f ist eine Drehung des Dreiecks x_2, x_3, x_4. Für $f \neq \mathrm{id}$ ergeben sich zwei Möglichkeiten, nämlich

$$\varphi(f_1) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix} \quad \text{oder} \quad \varphi(f_2) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}.$$

Es sind $\varphi(f_1)$ und $\varphi(f_2)$ gerade Permutationen. Sei g_i die Drehung des Würfels, die den Tetraeder in sich selbst abbildet mit $g_i(x_1) = x_i$ für $i = 2, 3, 4$. Wieder ist $\varphi(g_i)$ eine gerade Permutation. Sei $g \in H$ beliebig. Dann gilt $g(x_1) = x_i$ für ein $1 \leq i \leq 4$.
Fall 1. $g(x_1) = x_1$. Dann ist $g = \mathrm{id}$, $g = f_1$ oder $g = f_2$.
Fall 2. $g(x_1) = x_i$ mit $i \neq 1$. Dann ergibt sich

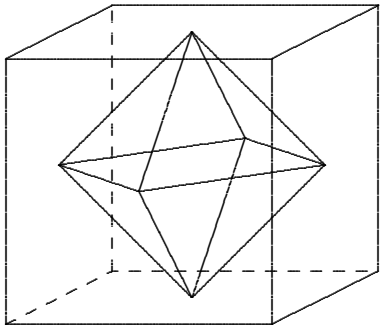
$$g_i^{-1} g(x_1) = x_1$$
$$g_i^{-1} g = \mathrm{id},\ g_i^{-1} g = f_1 \text{ oder } g_i^{-1} g = f_2$$
$$g = g_i,\ g = g_i f_1 \text{ oder } g = g_i f_2.$$

Also ist $\varphi(g)$ wieder eine gerade Permutation.
Außerdem hat H mindestens 12 Elemente, nämlich

$$f_0 := \mathrm{id}, f_1, f_2,$$
$$g_i f_0, g_i f_1, g_i f_2 \quad \text{für } i = 2, 3, 4.$$

Ergebnis: $H \cong A_4$, denn $\mathrm{Ord}(H) \geq 12$ und $\mathrm{Ord}(A_4) = \frac{4!}{2} = 12$.

(c) Drehgruppe des Oktaeders ist die Drehgruppe des Würfels, denn verbindet man die Mitten der Seiten des Würfels, so ergibt sich ein Oktaeder.



(d) Durch ähnliche Überlegungen folgt: Der Dodekaeder und der Ikosaeder haben die gleiche Drehgruppe, und diese ist isomorph zu A_5.

10.1.2. Operationen einer Gruppe auf einer Menge. Unser Ziel ist eine Klassifikation der endlichen Untergruppen der SO(3). Vorbereitend dazu sollen einige nützliche Begriffe eingeführt werden.

Definition. Sei G eine Gruppe, S eine Menge und $m\colon G \times S \to S$ eine Abbildung. Wir schreiben $gx := m(g, x)$ für $g \in G$ und $x \in S$. m heißt *Operation* von G auf S, falls für alle $x \in S$ und $g, h \in G$ gilt:

$$(gh)x = g(hx)$$
$$ex = x.$$

Ferner heißen

$Gx := \{\, gx \mid g \in G \,\}$ *Orbit* (oder *Bahn*) von x,
$G_x := \{\, g \in G \mid gx = x \,\}$ *Isotropiegruppe* (oder *Stabilisator*) von x.

BEMERKUNG. Sei G eine Gruppe, S eine Menge und $m\colon G\times S\to S$ eine Operation von G auf S sowie $x\in S$. Dann ist G_x eine Untergruppe von G.

BEWEIS. 1. Wegen $ex=x$ ist $e\in G_x$. 2. Seien $g,h\in G_x$. Dann ist $(gh)x=g(hx)=gx=x$, also auch $gh\in G_x$. 3. Sei $g\in G_x$, also $gx=x$. Es ist $x=ex=(g^{-1}g)x=g^{-1}(gx)=g^{-1}x$, also auch $g^{-1}\in G_x$. □

BEISPIELE. (1) Sei H eine Gruppe und $G\subseteq H$ eine Untergruppe. Dann ist

$$\begin{aligned} G\times H &\to H\\ (g,x) &\mapsto gx \quad \text{(Produkt in } H\text{)}\end{aligned}$$

eine Operation von G auf H.

(2) Sei G eine Gruppe. Dann ist

$$\begin{aligned} G\times G &\to G\\ (g,x) &\mapsto gxg^{-1}\end{aligned}$$

eine Operation von G auf G, sie wird auch *Konjugation* genannt.

BEWEIS. Seien $g,h\in G$ und $x\in G$. Es gilt $(e,x)\mapsto exe^{-1}=x$ und $((gh),x)\mapsto (gh)x(gh)^{-1}=ghxh^{-1}g^{-1}=g(hxh^{-1})g^{-1}$. □

(3) Sei V ein endlichdimensionaler euklidischer Vektorraum und G die euklidische Gruppe von V (oder eine Untergruppe davon). Dann ist

$$\begin{aligned} G\times V &\to V\\ (g,x) &\mapsto g(x)\end{aligned}$$

eine Operation von G auf V. Ist speziell $G=\mathrm{O}(V)$, so ist für ein $x\in V$ der Orbit von x

$$Gx := \{\, gx \mid g\in G\,\} = \{\, y\in V \mid \|y\|=\|x\|\,\}.$$

Für $\dim V=2$ ist also der Orbit Gx der Kreis um den Nullpunkt mit Radius $\|x\|$, und für $\dim V=3$ ist der Orbit Gx die Oberfläche der Kugel um den Nullpunkt mit Radius $\|x\|$.

SATZ. *Sei G eine Gruppe, S eine Menge und $G\times S\to S$ eine Operation. Dann gilt:*

(1) S ist disjunkte Vereinigung von Orbits.

(2) Für endliches G,S und beliebiges $x\in S$ gilt

$$\mathrm{Ord}(G)=\mathrm{Ord}(G_x)\cdot|Gx|.$$

Hierbei ist $|Gx|$ die Anzahl der Elemente in Gx.

BEWEIS. (1) (a) Es ist $S=\bigcup_{x\in S}Gx$, denn für $x\in S$ gilt $x=ex$.

(b) Es genügt, noch zu zeigen, daß aus $Gx\cap Gy\neq\emptyset$ folgt $Gx=Gy$. Seien also Gx,Gy gegeben mit $Gx\cap Gy\neq\emptyset$. Dann existieren $g,h\in G$ mit $gx=hy$, also $x=g^{-1}hy\in Gy$, woraus $Gx\subseteq Gy$ folgt. Analog zeigt man $Gy\subseteq Gx$. Also ist $Gx=Gy$.

(2) Betrachte die surjektive Abbildung

$$\begin{aligned}\varphi\colon G &\to Gx\\ g &\mapsto gx.\end{aligned}$$

(a) Sei $r := |Gx|$, etwa $Gx = \{f_1x, \ldots, f_rx\}$ mit $f_i \in G$. Wir haben dann

$$G = \bigcup_{i=1}^{r} \varphi^{-1}(f_ix),$$
$$\varphi^{-1}(f_ix) \cap \varphi^{-1}(f_jx) = \emptyset \quad \text{für } i \neq j$$

(dies gilt für jede Abbildung; $\varphi^{-1}(y)$ heißt *Faser* von $y \in Gx$).

(b) Berechnung der Fasern. Sei $f \in G$. Dann ist

$$\varphi^{-1}(fx) = fG_x,$$

denn für $g \in G$ sind folgende Aussagen äquivalent.

$$g \in \varphi^{-1}(fx)$$
$$gx = fx$$
$$f^{-1}gx = x$$
$$f^{-1}g \in G_x$$
$$g \in fG_x.$$

(c) Für ein $f \in G$ ist die Abbildung

$$G_x \to fG_x$$
$$g \mapsto fg$$

bijektiv, denn sie hat die Umkehrabbildung $h \mapsto f^{-1}h$. Also

$$\begin{aligned}
\operatorname{Ord}(G) &= \sum_{i=1}^{r} \operatorname{Ord}(\varphi^{-1}(f_ix)) && \text{nach (a)} \\
&= \sum_{i=1}^{r} \operatorname{Ord}(f_iG_x) && \text{nach (b)} \\
&= r \cdot \operatorname{Ord}(G_x) \\
&= |Gx| \cdot \operatorname{Ord}(G_x) && \text{nach Definition von } r \text{ in (a).}
\end{aligned}$$

□

10.1.3. Satz von Lagrange.

Satz. *Sei H eine endliche Gruppe und $G \subseteq H$ eine Untergruppe. Es sei $[H : G]$ (der* Index *von G in H) die Anzahl der verschiedenen Teilmengen $Gx = \{\, gx \mid g \in G \,\}$, $x \in H$, von H. Dann gilt:*

$$\operatorname{Ord}(H) = \operatorname{Ord}(G) \cdot [H : G].$$

Beweis. Betrachte die Operation

$$G \times H \to H$$
$$(g, x) \mapsto gx \quad \text{(Produkt in } H\text{)}.$$

Seien $Gx_1, \ldots, Gx_r$ die paarweise verschiedenen Orbits. Nach obigem Satz, Teil (1), gilt

$$H = \bigcup_{i=1}^{r} Gx_i \quad \text{und} \quad Gx_i \cap Gx_j = \emptyset \quad \text{für } i \neq j.$$

Die Abbildung $G \to Gx$, $g \mapsto gx$ ist bijektiv für alle $x \in H$, denn $y \mapsto yx^{-1}$ ist die Umkehrabbildung. Also ist $\mathrm{Ord}(H) = r \cdot \mathrm{Ord}(G)$. □

KOROLLAR. *Sei G eine endliche Gruppe mit $\mathrm{Ord}(G) = p$ und p prim. Dann ist G zyklisch.*

BEWEIS. Es ist $\mathrm{Ord}(G) = p \geq 2$, also existiert ein $x \in G$ mit $x \neq e$. Also ist $\langle x \rangle$ Untergruppe von G mit $\mathrm{Ord}(\langle x \rangle) \neq 1$. Nach dem Satz von Lagrange teilt $\mathrm{Ord}(\langle x \rangle)$ die Ordnung von G, und wegen p prim folgt $\mathrm{Ord}(\langle x \rangle) = \mathrm{Ord}(G)$. Also ist $\langle x \rangle = G$. □

10.1.4. Klassifikation der endlichen Untergruppen der $\mathrm{SO}(3)$. Es folgt die Anwendung auf endliche Drehgruppen. Zunächst wird der Begriff der Diedergruppe eingeführt.

DEFINITION. Seien $t \in \mathbb{N}$, $t \geq 1$ und G eine Gruppe. G heißt *Diedergruppe* der Ordnung $2t$, wenn gilt:

(1) $\mathrm{Ord}(G) = 2t$.
(2) Es existieren $f, g \in G$ mit

$$f^t = e$$
$$g^2 = e$$
$$fg = gf^{-1}$$
$$G = \{e, f, \ldots, f^{t-1}, g, gf, \ldots, gf^{t-1}\}.$$

BEMERKUNGEN. (1) D_{2t} ist die Symmetriegruppe des regulären t-Ecks in der Ebene.

(2) Sei G eine Diedergruppe der Ordnung $2t$, etwa

$$G = \{e, f, \ldots, f^{t-1}, g, gf, \ldots, gf^{t-1}\}$$

mit g, f wie in der Definition. Dann hat jedes $h \in G$ eine *eindeutige* Darstellung

$$h = g^i f^j \quad \text{mit } 0 \leq i \leq 1,\ 0 \leq j \leq t-1.$$

Die Multiplikation in G hängt nur von den Paaren (i, j) ab, denn für $0 \leq i, k \leq 1$ und $0 \leq j, l \leq t-1$ gilt

$$g^i f^j g^k f^l = \begin{cases} g^i f^{j+l} & \text{falls } k = 0 \\ g^{i+1} f^{-j+l} & \text{falls } k = 1. \end{cases}$$

Also sind je zwei Diedergruppen der Ordnung $2t$ zueinander isomorph.

SATZ. *(1) Sei $G \subseteq \mathrm{SO}(2)$ eine endliche Untergruppe. Dann ist G zyklisch.*

(2) Sei $G \subseteq \mathrm{SO}(3)$ eine endliche Untergruppe. Dann gilt:
G ist zyklisch
oder G ist Diedergruppe, d.h. $G \cong D_{2n}$ (Symmetriegruppe des regulären n-Ecks),
oder $\mathrm{Ord}(G) \in \{12, 24, 60\}$.

BEMERKUNG. In (2) gilt sogar: Ist G nicht zyklisch und keine Diedergruppe, also $\mathrm{Ord}(G) = 12$, 24 bzw. 60, so ist $G \cong A_4$, S_4 bzw. A_5. (Dies wird hier nicht bewiesen.)

BEWEIS. (1) Sei $f \in \mathrm{SO}(\mathbb{R}^2)$. Dann existiert genau ein φ, $0 \leq \varphi < 2\pi$ mit

$$f = f_\varphi := \text{Drehung um } \varphi \text{ gegen den Uhrzeigersinn.}$$

Sei $\mathrm{Ord}(G) = n$, etwa $G = \{f_{\varphi_1}, \ldots, f_{\varphi_n}\}$ mit $0 \leq \varphi_i < 2\pi$. Wir können oBdA $\varphi_1 = 0$ annehmen. Sei $\varphi := \min\{\, \varphi_i \mid 2 \leq i \leq n \,\}$. Dann existiert für jedes i mit $2 \leq i \leq n$ ein $m \in \mathbb{N}$ und ein ψ_i mit $0 \leq \psi_i < \varphi$, so daß gilt:

$$\begin{aligned} &\varphi_i = m\varphi + \psi_i \\ &f_{\varphi_i} = f_{m\varphi+\psi_i} = f_\varphi^m f_{\psi_i} \\ &f_{\psi_i} \in G && \text{da } G \text{ Gruppe} \\ &\psi_i = 0 && \text{da } \psi_i < \varphi. \end{aligned}$$

Also ist $G = \langle f_\varphi \rangle = \{\, f_\varphi^0, f_\varphi, \ldots, f_\varphi^{n-1} \mid \}$.

(2) Jedes $f \in \mathrm{SO}(\mathbb{R}^3)$, $f \neq \mathrm{id}$, bildet die Oberfläche der Einheitskugel in sich selbst ab und hat dort genau zwei Fixpunkte (nämlich wo die Drehachse die Einheitskugel durchstößt). Sei

$$S := \{\, x \in \mathbb{R}^3 \mid \|x\| = 1 \text{ und es gibt ein } f \in G,\ f \neq \mathrm{id} \text{ mit } f(x) = x \,\}$$

die Menge der *Pole* von G. Wir können oBdA $\mathrm{Ord}(G) > 1$ annehmen, also ist $S \neq \emptyset$.

a. Wir zeigen, daß

$$\begin{aligned} G \times S &\to S \\ (g, x) &\mapsto g(x) \end{aligned}$$

eine wohldefinierte Operation ist. Zu zeigen ist, daß aus $g \in G$ und $x \in S$ folgt $g(x) \in S$. Wegen $x \in S$ gibt es ein $f \in G$, $f \neq \mathrm{id}$ mit $f(x) = x$. Also ist $gfg^{-1}g(x) = g(x)$, und es gilt $gfg^{-1} \in G$, $gfg^{-1} \neq \mathrm{id}$. Daraus folgt $g(x) \in S$.

b. Behauptung: Aus $x \in S$ und $g \in G$ folgt $G_{g(x)} = gG_xg^{-1}$, insbesondere also $\mathrm{Ord}(G_{g(x)}) = \mathrm{Ord}(G_x)$.

Beweis: Folgende Aussagen sind äquivalent.

$$\begin{aligned} &f \in G_{g(x)} \\ &fg(x) = g(x) \\ &g^{-1}fg(x) = x \\ &g^{-1}fg \in G_x \\ &f \in gG_xg^{-1}. \end{aligned}$$

Weiter gilt $\mathrm{Ord}(G_{g(x)}) = \mathrm{Ord}(G_x)$, da die Abbildung

$$\begin{aligned} G &\to G \\ h &\mapsto ghg^{-1} \end{aligned}$$

bijektiv ist und gG_xg^{-1} das Bild von G_x unter dieser Abbildung ist.

c. Seien $Gx_1, \ldots, Gx_r$ die paarweise verschiedenen Orbits bezüglich dieser Operation (S ist also die disjunkte Vereinigung von $Gx_1, \ldots, Gx_r$). Nach dem Satz gilt für jedes i mit $1 \leq i \leq r$

$$\underbrace{\mathrm{Ord}(G)}_{=:n} = \underbrace{|Gx_i|}_{=:m_i} \cdot \underbrace{\mathrm{Ord}(G_{x_i})}_{=:n_i}.$$

Sei etwa

$$Gx_i = \{f_1(x_i), \ldots, f_{m_i}(x_i)\},$$

Die Faser von jedem $f_j(x_i)$ ist dann eine Teilmenge von G mit $n_i = \mathrm{Ord}(G_{x_i})$-vielen Elementen.

Unser Ansatz besteht nun darin, alle Paare (f, x) mit $f \in G$, $f \neq \mathrm{id}$ und x Pol von f zu zählen.

(a) Einerseits gibt es $n-1$ Möglichkeiten für f mit jeweils 2 Polen.

(b) Andererseits: Sei i fest gewählt. Betrachte $f_j(x_i) \in Gx_i$. Wieviele $h \in G$ mit dem Pol $f_j(x_i)$ gibt es? Soviele wie

$$\mathrm{Ord}(\underbrace{\{\, h \in G \mid hf_j(x_i) = f_j(x_i) \,\}}_{G_{f_j(x_i)}} = \mathrm{Ord}(G_{x_i}) = n_i \quad \text{nach (b)}.$$

Bei festem i haben wir also m_i Möglichkeiten für x, mit jeweils $n_i - 1$ zugehörigen f.

Dies ergibt

$$2(n-1) = \sum_{i=1}^{r} m_i(n_i - 1)$$

$$2 - \frac{2}{n} = \sum_{i=1}^{r} \frac{m_i(n_i-1)}{n} = \sum_{i=1}^{r}(1 - \frac{1}{n_i}) \quad \text{wegen } n = m_i n_i.$$

Wir diskutieren nun die Gleichung

$$\boxed{2 - \frac{2}{n} = \sum_{i=1}^{r}(1 - \frac{1}{n_i})} \quad (\text{oBdA } n_i \geq 2). \tag{10.1}$$

d. Behauptung: $r = 2$ oder $r = 3$.

Beweis: Unmöglich ist der Fall $r = 1$. Annahme: $r \geq 4$. Dann folgt aus (10.1) der Widerspruch:

$$2 - \frac{2}{n} = \sum_{i=1}^{r}(1 - \frac{1}{n_i}) \geq 4 - \frac{4}{2} = 2.$$

Fall $r = 2$. Aus (10.1) folgt

$$2 - \frac{2}{n} = 2 - \frac{1}{n_1} - \frac{1}{n_2}$$

$$\frac{2}{n} = \frac{1}{n_1} + \frac{1}{n_2} \qquad (n_1, n_2 \leq n)$$

$$n_1 = n_2 = n \quad \text{und damit} \quad m_1 = m_2 = 1.$$

Also gibt es genau 2 Pole und jeder Pol kommt n-mal vor. Sei $x \in S$, also ein Pol. Man erhält

$$G = \{\, g \in \mathrm{SO}(\mathbb{R}^3) \mid g(x) = x \,\}$$

$\cong$ endliche Untergruppe von $\mathrm{SO}(\mathbb{R}^2)$ (Ebene: $(\mathbb{R}x)^\perp$)

G zyklisch nach (1).

Fall $r = 3$. OBdA gelte $2 \le n_1 \le n_2 \le n_3 \le n$. Nach (10.1) haben wir

$$2 - \frac{2}{n} = 3 - \sum_{i=1}^{3} \frac{1}{n_i}.$$

Annahme $n_1 \ge 3$. Dann erhalten wir den Widerspruch

$$2 - \frac{2}{n} = 3 - \underbrace{\sum_{i=1}^{3} \frac{1}{n_i}}_{\le 1} \ge 2.$$

Also ist $n_1 = 2$ und damit

$$2 - \frac{2}{n} = 3 - \frac{1}{2} - \frac{1}{n_2} - \frac{1}{n_3},$$
$$\frac{1}{n_2} + \frac{1}{n_3} = \frac{1}{2} + \frac{2}{n}.$$

Annahme: $n_2 \ge 4$. Dann haben wir $\frac{1}{n_2} + \frac{1}{n_3} \le \frac{2}{4} = \frac{1}{2}$ und damit einen Widerspruch. Also ist $n_2 = 2$ oder $n_2 = 3$. Es ergeben sich die folgenden Möglichkeiten:

i. $n_1 = 2$ und $n_2 = 2$. Dann ist $n_3 = \frac{n}{2}$ und n gerade.

ii. $n_1 = 2$, $n_2 = 3$ und $n_3 = 3$. Dann gilt $\frac{2}{3} = \frac{1}{2} + \frac{2}{n}$, also $\frac{4}{6} - \frac{3}{6} = \frac{2}{n}$, also $n = 12$.

iii. $n_1 = 2$, $n_2 = 3$ und $n_3 = 4$. Dann gilt $\frac{1}{3} + \frac{1}{4} = \frac{1}{2} + \frac{2}{n}$, also $\frac{4}{12} - \frac{3}{12} = \frac{2}{n}$, also $n = 24$.

iv. $n_1 = 2$, $n_2 = 3$ und $n_3 = 5$. Dann folgt $\frac{1}{3} + \frac{1}{5} = \frac{1}{2} + \frac{2}{n}$, also $\frac{10}{30} + \frac{6}{30} - \frac{15}{30} = \frac{2}{n}$, also $n = 60$.

Andere Fälle sind nicht möglich, denn es gilt

$$\frac{1}{3} + \frac{1}{n_3} = \frac{1}{2} + \frac{2}{n}$$
$$\frac{1}{n_3} = \frac{1}{6} + \frac{2}{n} > \frac{1}{6}$$
$$n_3 < 6 \qquad \text{(außerdem } n_2 = 3 \le n_3\text{)}.$$

Es bleibt noch zu zeigen, daß im Fall (i) G eine Diedergruppe ist.

e. Behauptung: Aus $r = 3$, $n_1 = 2$, $n_2 = 2$, $n_3 = \frac{n}{2}$ folgt $G \cong D_{2\cdot\frac{n}{2}}$.

Beweis: Es sei $S = Gx_1 \cup Gx_2 \cup Gx_3$ (disjunkte Vereinigung) und $m_1 = \frac{n}{2}$, $m_2 = \frac{n}{2}$ und $m_3 = \frac{n}{\frac{n}{2}} = 2$. Es ist $|Gx_3| = 2$, sei also $Gx_3 =: \{x, y\}$. Damit erhält man

$G_x = \{\, f \in G \mid f(x) = x \,\}$

$\cong$ endliche Untergruppe von $\mathrm{SO}(\mathbb{R}^2)$ (Ebene: $(\mathbb{R}x)^\perp$).

Also ist nach (1) $G_x = \langle f \rangle$ für ein $f \in G$. Wähle ein $g \in G$ mit $g \notin G_x$ (möglich wegen $\mathrm{Ord}(G_x) = \frac{n}{2}$). Dann folgt mit $t := \frac{n}{2}$:

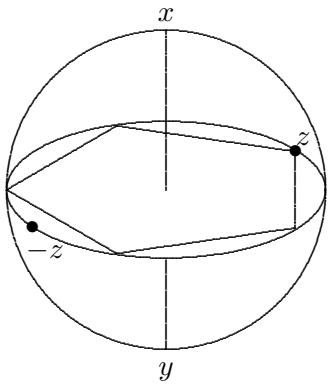
$$G = \{f^0 = \mathrm{id}, f, \ldots, f^{t-1}, g, gf, \ldots, gf^{t-1}\}, \quad G_x = \{f^0, f, \ldots, f^{t-1}\}.$$

Außerdem folgt für ein $g \in G$ mit $g \notin G_x$ (d.h. $g(x) \neq x$), daß $g^2 = \mathrm{id}$. Dies sieht man wie folgt. Zunächst gilt für jedes $g \in G$

$$g(Gx_3) = gGx_3 = Gx_3,$$

d.h. $g(\{x, y\}) = \{x, y\}$, d.h. $g(x) = x$ und $g(y) = y$, oder $g(x) = y$ und $g(y) = x$. Daraus folgt im Fall $g(x) \neq x$, daß $g(x) = y$ und damit $g^2(x) = g(y) = x$. Ebenso folgt im Fall $g(y) \neq y$, daß $g^2(y) = y$. Da g eine Drehung mit $g(x) \neq x$ und $g(y) \neq y$ ist, besitzt g einen Pol z mit $g(z) = z$ und $z \neq x, y$. Dann ist auch $g^2(z) = z$, also ist $g^2 = \mathrm{id}$, da g^2 drei Fixpunkte hat.

Ergebnis: $G = \langle f, g\rangle$, $f^t = \mathrm{id}$, $f^s \neq \mathrm{id}$ für $s < t$, $g^2 = \mathrm{id}$ und $fg = gf^{-1}$ da $(gf)^2 = gfgf = \mathrm{id}$. Also ist $G \cong D_{2t}$ mit $t \geq 1$. □



10.2. Faktorgruppen

Sei G eine Gruppe und $H \subseteq G$ eine Untergruppe. Die Faktorgruppe G/H ist dann wie der Quotientenvektorraum V/U in 2.3.5 definiert. Wir konstruieren die Faktorgruppe G/N für einen Normalteiler $N \subseteq G$, zusammen mit dem natürlichen Gruppenhomomorphismus $\mathrm{kan}\colon G \to G/N$. Es zeigt sich, daß hierfür die Eigenschaft von N, Normalteiler zu sein, notwendig ist. Als einfache Folgerung erhalten wir eine Charakterisierung von Normalteilern als Kerne von Gruppenhomomorphismen. Wir beweisen dann die universelle Eigenschaft von Faktorgruppen und als Folgerung den Homomorphiesatz (vgl. den Homomorphiesatz für Vektorräume in 3.2.2). Als einfache Anwendung erhalten wir eine Klassifikation der zyklischen Gruppen und für endliche Gruppen den Satz von Fermat, daß $x^{\mathrm{Ord}(G)} = e$ ist. Ferner ergibt sich $x^p \equiv x \mod p$ für $x \in \mathbb{Z}$ und Primzahlen p.

10.2.1. Nebenklassen, Konstruktion der Faktorgruppe. Sei G eine Gruppe und $H \subseteq G$ eine Untergruppe. Für $x \in G$ heißt

$$xH := \{\, xh \mid h \in H \,\}$$

die von x erzeugte *Linksnebenklasse* bzgl. H. (Entsprechend heißt $Hx := \{\, hx \mid h \in H \,\}$ *Rechtsnebenklasse*). Die Menge der Linksnebenklassen von G bzgl. H, also

$$G/H := \{\, xH \mid x \in G \,\},$$

wird G *modulo* H genannt.

BEMERKUNGEN. (1) Sei G eine Gruppe, $H \subseteq G$ eine Untergruppe und $x, y \in G$. Dann gilt

$$xH = yH \text{ genau dann, wenn } y^{-1}x \in H.$$

BEWEIS. $\Rightarrow$. Gelte $xH = yH$. Dann existiert ein $h \in H$ mit $x = xe = yh$. Also $y^{-1}x = h \in H$.

$\Longleftarrow$. Sei $y^{-1}x \in H$. Dann gibt es ein $h \in H$ mit $y^{-1}x = h$, also $x = yh$. Damit gilt für alle $h' \in H$, daß $xh' = yhh' \in yH$, also $xH \subseteq yH$. Wegen $(y^{-1}x)^{-1} = x^{-1}y$ folgt genauso $yH \subseteq xH$. □

(2) Sei G eine Gruppe und $H \subseteq G$ eine Untergruppe. Für $x, y \in G$ definieren wir

$$x \sim_H y :\equiv y^{-1}x \in H.$$

Nach (1) ist $\sim_H$ eine Äquivalenzrelation auf G. Wir zeigen

$$xH = \{\, y \mid y \sim_H x \,\} =: [x] \quad \text{Äquivalenzklasse von } x.$$

BEWEIS. $\subseteq$. Sei $h \in H$. Zu zeigen ist $xh \sim_H x$, d.h. $x^{-1}xh \in H$. $\supseteq$. Sei $y \sim_H x$, also $y^{-1}x \in H$. Zu zeigen ist $y \in xH$. Es existiert ein $h \in H$ mit

$$\begin{aligned} y^{-1}x &= h \\ x^{-1}y &= h^{-1} \\ y &= xh^{-1} \in xH. \end{aligned}$$

□

Nach Abschnitt 1.2 (den Bemerkungen zu Äquivalenzklassen) ist also G die disjunkte Vereinigung von Linksnebenklassen.

(3) Sei G eine Gruppe und $H \subseteq G$ eine Untergruppe. Wähle in jeder Linksnebenklasse ein Element (Repräsentant) x_i aus. Dann ist $G = \bigcup_{i \in I} x_i H$ (disjunkte Vereinigung) wobei $(x_i)_{i \in I}$ die Familie der Repräsentanten ist. Es ist dann auch $G = \bigcup_{i \in I} H x_i^{-1}$ disjunkte Vereinigung von Rechtsnebenklassen, denn $\varphi\colon G \to G$, $x \mapsto x^{-1}$ ist bijektiv und es gilt $\varphi(xH) = Hx^{-1}$.

BEISPIELE. (1) Sei V ein Vektorraum über einen Körper K und $U \subseteq V$ ein Untervektorraum. Faßt man V als additive Gruppe und U als Untergruppe auf, so ist

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

(2) Sei $n \in \mathbb{N}$. Faßt man $\mathbb{Z}$ auf als additive Gruppe und definiert man

$$\begin{aligned} n\mathbb{Z} &= \{\, nz \mid z \in \mathbb{Z} \,\} \\ &= \{0, \pm n, \pm 2n, \pm 3n, \dots\} \\ &= \text{die von } n \text{ erzeugte Untergruppe,} \end{aligned}$$

so gilt: $\mathbb{Z}/n\mathbb{Z} = \{[0], [1], \dots, [n-1]\}$ hat n Elemente.

BEWEIS. a. Für alle $x \in \mathbb{Z}$ existieren $z, r \in \mathbb{Z}$ mit $x = zn + r$ und $0 \leq r < n$ (Division mit Rest). Also ist $[x] = [r]$.

b. Noch zu zeigen: Für alle $0 \leq r, s < n$ mit $r \neq s$ ist $[r] \neq [s]$. Sei dazu oBdA $r > s$. Annahme: $[r] = [s]$. Dann existiert ein $z \in \mathbb{Z}$

mit $r - s = n > 0$ und $r - s < n$. Also folgt $0 < r - s < nz$, was ein Widerspruch zu $z \in \mathbb{Z}$ ist. □

BEZEICHNUNGEN. Statt $x - y \in n\mathbb{Z}$ schreibt man $x \equiv y \mod n$ und sagt „x ist kongruent y modulo n“. Rechnet man mit $\equiv$ statt mit $=$, so bedeutet dies das Rechnen in $\mathbb{Z}/n\mathbb{Z}$ statt in $\mathbb{Z}$.

10.2.2. Konstruktion der Faktorgruppe.

SATZ. *Sei G eine Gruppe und $N \subseteq G$ Normalteiler. Definiere*

$$\begin{aligned} \operatorname{kan}\colon G &\to G/N \\ x &\mapsto [x]. \end{aligned}$$

Dann gibt es auf G/N genau eine Gruppenstruktur, so daß $\operatorname{kan}\colon G \to G/N$ *ein Gruppenhomomorphismus wird. In diesem Fall gilt* $\ker \operatorname{kan} = N$.

BEWEIS. *Existenz.* Für alle $x, y \in G$ setzen wir

$$[x] \cdot [y] := [xy] \quad \text{Produkt in } G/N.$$

Wir zeigen zunächst, daß die Multiplikation wohldefiniert ist. Seien also $x, x', y, y' \in G$ mit $[x] = [x']$ und $[y] = [y']$ gegeben. Zu zeigen ist $[xy] = [x'y']$, d.h. die obige Definition des Produkts in G/N ist unabhängig vom Repräsentanten von $[x]$, $[y]$. Dies folgt aus

$$\begin{aligned} (x'y')^{-1}(xy) &= y'^{-1}\underbrace{x'^{-1}x}_{\in N}y \\ &= y'^{-1}yn && \text{für ein } n \in N \text{ wegen } Ny = yN, \\ & && \text{denn } N \text{ ist Normalteiler} \\ &\in Nn && \text{wegen } [y] = [y'] \\ &= N. \end{aligned}$$

Also ist $[x'y'] = [xy]$.

Wir zeigen jetzt, daß die Multiplikation assoziativ ist. Dies folgt aus

$$[x]([y][z]) = [x]([yz]) = [x(yz)] = [(xy)z] = [xy][z] = ([x][y])[z].$$

Ferner ist $[e]$ das neutrale Element, denn es gilt

$$[e][x] = [ex] = [x] \quad \text{für alle } x \in G.$$

Schließlich ist $[x^{-1}]$ das Inverse zu $[x]$, denn es gilt

$$[x^{-1}][x] = [x^{-1}x] = [e].$$

Demnach ist G/N mit obiger Multiplikation eine Gruppe. kan ist offenbar Gruppenhomomorphismus und es gilt $\ker \operatorname{kan} = N$.

Eindeutigkeit. Eine Verknüpfung $\circ$ auf G/N, so daß kan Gruppenhomomorphismus ist, muß die Bedingung $xN \circ yN = xyN = xyNN = xN \cdot yN$ erfüllen. □

BEMERKUNGEN. (1) Sei G eine Gruppe und $N \subseteq G$ Untergruppe. Sei noch

$$\begin{aligned} \operatorname{kan}\colon G &\to G/N \\ x &\mapsto [x] \end{aligned}$$

Gruppenhomomorphismus bzgl. irgendeiner (fest gewählten) Gruppenstruktur auf G/N. Dann ist N Normalteiler.

BEWEIS. Es ist $\mathrm{kan}(e) = [e]$ neutrales Element in G/N, da kan Gruppenhomomorphismus ist. Ferner ist

$$\ker \mathrm{kan} = \{\, x \in G \mid [x] = [e] \,\} = N.$$

Also ist N Normalteiler, denn nach der folgenden Bemerkung gilt dies stets für Kerne von Gruppenhomomorphismen. □

(2) Sei G eine Gruppe und $N \subseteq G$ Untergruppe. Dann ist N Normalteiler von G genau dann, wenn es eine Gruppe G' und einen Gruppenhomomorphismus $f\colon G \to G'$ gibt mit $N = \ker f$.

BEWEIS. $\Longleftarrow$. Es ist $N = \ker f = f^{-1}(\{e\})$, und $\{e\}$ ist Normalteiler. $\Rightarrow$ folgt aus dem eben bewiesenen Satz. □

3. Sei $G = \mathbb{Z}$ und $N = n\mathbb{Z}$ für ein $n \in N$. Dann gilt für alle $x, y, z \in \mathbb{Z}$

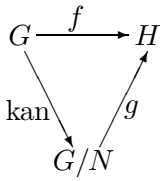
$$[x] + [y] = [z] \quad \text{genau dann, wenn} \quad x + y = z + nu \quad \text{für ein } u \in \mathbb{Z}$$
$$[y] = -[x] \quad \text{genau dann, wenn} \quad x + y \in n\mathbb{Z}.$$

Ist zum Beispiel $n = 4$, so hat man

$$\mathbb{Z}/4\mathbb{Z} = \{[0], [1], [2], [3]\},$$
$$[1] + [2] = [3],$$
$$[2] + [3] = [5] = [1],$$
$$-[1] = [3].$$

10.2.3. Universelle Eigenschaft der Faktorgruppe.

SATZ. *Seien G, H Gruppen, $N \subseteq G$ Normalteiler und $f\colon G \to H$ Gruppenhomomorphismus mit $N \subseteq \ker f$. Dann gibt es genau einen Gruppenhomomorphismus $g\colon G/N \to H$ mit $g \circ \mathrm{kan} = f$.*



BEWEIS. *Existenz.* Für $x \in G$ definieren wir $g([x]) := f(x)$.
a. g ist wohldefiniert, denn für $x, y \in G$ mit $[x] = [y]$ gilt

$$y^{-1}x \in N$$
$$x = yn \quad \text{für ein } n \in N$$
$$f(x) = f(yn) = f(y)\underbrace{f(n)}_{=e} = f(y).$$

Also ist $g([x]) := f(x)$ unabhängig von der Wahl des Repräsentanten x.
b. g ist Gruppenhomomorphismus, denn für $x, y \in G$ gilt

$$g([x][y]) = g([xy]) = f(xy) = f(x)f(y) = g([x])g([y]).$$

Eindeutigkeit. Für eine Abbildung g mit den gewünschten Eigenschaften muß gelten $g([x]) = (g \circ \mathrm{kan})(x) = f(x)$ für alle $x \in G$. □

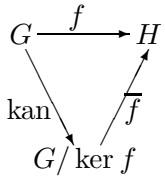
BEMERKUNG. Seien G, H Gruppen und $N \subseteq G$ Normalteiler. Dann ist die Abbildung

$$\begin{aligned}\{\, g\colon G/N \to H \mid g \text{ Homom.}\,\} &\to \{\, f\colon G \to H \mid f \text{ Homom. mit } f[N] = \{e\}\,\}\\ g &\mapsto g \circ \mathrm{kan}\end{aligned}$$

offenbar eine Bijektion.

10.2.4. Homomorphiesatz.

KOROLLAR (Homomorphiesatz). *Seien G, H Gruppen, $f\colon G \to H$ surjektiver Gruppenhomomorphismus und* $\mathrm{kan}\colon G \to G/\ker f$ *wie in Satz 10.2.2 definiert. Dann gibt es genau eine Abbildung $\overline{f}\colon G/\ker f \to H$ mit $\overline{f} \circ \mathrm{kan} = f$. Dieses eindeutig bestimmte $\overline{f}$ ist ein Isomorphismus.*



BEWEIS. Nach dem Satz ist $\overline{f}$ ein wohldefinierter Gruppenhomomorphismus. $\overline{f}$ ist surjektiv, denn zu jedem $y \in H$ gibt es ein $x \in G$ mit $y = f(x) = \overline{f}([x])$. $\overline{f}$ ist injektiv, denn für $x \in G$ mit $\overline{f}([x]) = e$ gilt $f(x) = \overline{f}([x]) = e$, also $x \in \ker f$ und damit $[x] = [e]$. □

10.2.5. Anwendungen: zyklische Gruppen, Satz von Fermat.

Als erste einfache Anwendung erhalten wir eine Übersicht über alle zyklischen Gruppen.

KOROLLAR. *Sei G eine zyklische Gruppe. Dann gilt*

$$G \cong \begin{cases} \mathbb{Z}, & \text{falls } \mathrm{Ord}(G) = \infty \\ \mathbb{Z}/m\mathbb{Z}, & \text{falls } \mathrm{Ord}(G) = m < \infty. \end{cases}$$

BEWEIS. 1. Sei G zyklisch, etwa $G = \langle x \rangle$ für ein $x \in G$. Dann ist

$$\begin{aligned} f\colon \mathbb{Z} &\to G\\ i &\mapsto x^i\end{aligned}$$

ein (offenbar surjektiver) Gruppenhomomorphismus, denn es gilt

$$f(i+j) = x^{i+j} = x^i x^j = f(i)f(j).$$

2. Noch zu zeigen ist

$$\ker f = \begin{cases} 0 & \text{falls } \mathrm{Ord}(G) = \infty \\ n\mathbb{Z} & \text{falls } \mathrm{Ord}(G) = n < \infty. \end{cases}$$

Zum Beweis beachte man, daß $\ker f$ Untergruppe von $\mathbb{Z}$ ist, also nach der folgenden Bemerkung von der Form $m\mathbb{Z}$ für ein $m \in \mathbb{Z}$ mit $m \geq 0$ ist. Im Fall $m \neq 0$ ist $G \cong \mathbb{Z}/\ker f = \mathbb{Z}/m\mathbb{Z}$, und im Fall $m = 0$ ist $f\colon \mathbb{Z} \to G$ ein Isomorphismus. □

BEMERKUNG. Wir haben noch nachzutragen, daß jede Untergruppe $N \subseteq \mathbb{Z}$ von der Form $N = m\mathbb{Z}$ ist für ein $m \in \mathbb{Z}$ mit $m \geq 0$. Zum Beweis können wir annehmen, daß $N \neq \{0\}$ ist. Sei m das kleinste positive Element von N. Für ein beliebiges $z \in N$ gibt es dann $r, s \in \mathbb{Z}$ mit $0 \leq r < m$ so daß $z = ms + r$, also $r = z - ms \in N$ und damit $r = 0$, denn m war minimal gewählt. Also ist $z \in m\mathbb{Z}$ und damit $N = m\mathbb{Z}$.

Als weitere Folgerung ergibt sich der Satz von Fermat.

KOROLLAR (Fermat). *Sei G eine endliche Gruppe der Ordnung n und $x \in G$. Dann ist $x^n = e$.*

BEWEIS. $\langle x \rangle \subseteq G$ ist eine Untergruppe. Nach dem Satz 10.1.3 von Lagrange gilt

$$n = \mathrm{Ord}(G) = \underbrace{\mathrm{Ord}\langle x \rangle}_{=:m} \cdot \underbrace{[G : \langle x \rangle]}_{=:r}.$$

Nach der vorangehenden Folgerung ist $\langle x \rangle \cong \mathbb{Z}/m\mathbb{Z}$ für ein $m \in \mathbb{Z}$ mit $m > 0$, also $x^m = e$, denn in $\mathbb{Z}/m\mathbb{Z}$ gilt

$$m \cdot \overline{a} = \underbrace{\overline{a} + \overline{a} + \cdots + \overline{a}}_{m\text{-mal}} = \overline{m \cdot a} = 0,$$

also $x^n = x^{mr} = (x^m)^r = e$. □

Sei G eine Gruppe und $x \in G$. Dann nennt man $\mathrm{Ord}(x) := \mathrm{Ord}(\langle x \rangle)$ die *Ordnung* von x in G.

BEMERKUNGEN. (1) Sei G eine Gruppe und $x \in G$ mit $\mathrm{Ord}(x) = n$. Dann gilt

a. $\langle x \rangle = \{x^0, x^1, \ldots, x^{n-1}\}$.
b. $n = \min\{\, i \mid i > 0, x^i = e \,\}$.
c. Für alle $m \in \mathbb{N}$ gilt $x^m = e$ genau dann, wenn n (also die Ordnung von x) die Zahl m teilt.

BEWEIS. Nach der Folgerung ist $\langle x \rangle \cong \mathbb{Z}/n\mathbb{Z}$ vermöge $x \mapsto [1]$. Wegen

$$\mathbb{Z}/n\mathbb{Z} = \{[0], [1], \ldots, [n-1]\}$$

gelten (a) und (b). Für (c) ist die Richtung $\Longleftarrow$ klar. $\Rightarrow$. Sei $x^m = e$. Dann gibt es $q, z \in \mathbb{Z}$ mit $0 \leq r < n$ so daß $m = qn + r$, also $e = x^m = x^{qn+r} = (x^n)^q \cdot x^r = e \cdot x^r$. Daraus folgt $r = 0$, denn n ist die kleinste positive Zahl mit $x^n = e$. □

(2) Sei p Primzahl. Dann ist (wie unten bewiesen wird) $\mathbb{Z}/p\mathbb{Z}$ ein Körper mit der folgenden Struktur.

$$[x] + [y] = [x + y],$$
$$[x] \cdot [y] = [x \cdot y].$$

Also ist $G = \{\, [x] \in \mathbb{Z}/p\mathbb{Z} \mid [x] \neq 0 \,\}$ eine multiplikative Gruppe der Ordnung $p - 1$. Nach dem Satz von Fermat gilt dann

$$[x]^{p-1} = [1] \qquad \text{für alle } x \in \mathbb{Z} \text{ mit } [x] \neq 0$$
$$[x]^p = [x] \qquad \text{für alle } x \in \mathbb{Z}$$
$$x^p \equiv x \mod p \quad \text{für alle } x \in \mathbb{Z}.$$

Nachzutragen bleibt, daß für Primzahlen p stets $\mathbb{Z}/p\mathbb{Z}$ ein Körper ist. Offenbar genügt es zu zeigen, daß es zu jedem $x \in \mathbb{Z}$ mit $[x] \neq 0$ ein $y \in \mathbb{Z}$ mit $[y] \neq 0$ gibt so daß $[x][y] = [1]$ ist. Sei also $x \in \mathbb{Z}$ mit $[x] \neq 0$ gegeben. Dann ist der größte gemeinsame Teiler (x, p) von x und p gleich 1. Man betrachte

$$\{\, ax + bp \mid a, b \in \mathbb{Z} \,\} =: G.$$

Offenbar ist $G \subseteq \mathbb{Z}$ eine Untergruppe, also $G = m\mathbb{Z}$ für ein $m > 0$. Es ist $m = 1$, denn aus $x = mz_1$ und $p = mz_2$ für $z_1, z_2 \in \mathbb{Z}$ folgt $m|x$ und $m|p$, also $m = 1$ wegen $(x, p) = 1$. Also ist $ax + bp = 1$ für geeignete $a, b \in \mathbb{Z}$, also $[a][x] + [0] = [1]$. Es folgt $[a] \neq 0$.

10.3. Moduln, Satz von Cayley-Hamilton

Wir definieren den Begriff eines Moduls, als eine Art Vektorraum über einem Ring statt einem Körper. Als Anwendung beweisen wir den Satz von Cayley/Hamilton, welcher aussagt, daß das „Einsetzen“ einer Matrix in ihr eigenes charakteristisches Polynom stets die Nullmatrix ergibt. Ferner führen wir den Begriff des Minimalpolynoms ein. Schliesslich stellen wir einige Hilfsmittel bereit (Faktormoduln, frei erzeugte Moduln, Basis und lineare Forsetzung für Moduln), die zur Einführung des Tensorprodukts im folgenden Abschnitt gebraucht werden. – In diesem Abschnitt sei R ein nicht notwendig kommutativer Ring mit Einselement.

10.3.1. Moduln, Modulhomomorphismen.

Definition. Sei M eine abelsche Gruppe und $m\colon R \times M \to M$ eine Abbildung; wir schreiben rx für $m(r, x)$. (M, m) oder kurz M heißt *R-Linksmodul*, wenn für alle $r, s \in R$ und alle $x, y \in M$ gilt

(1) $r(sx) = (rs)x$ (Assoziativgesetz).
(2) $1x = x$ (1 ist linksneutrales Element).
(3) $(r + s)x = rx + sx$ und $r(x + y) = rx + ry$ (Distributivgesetze).

Bemerkungen. (1) Eine analoge Definition trifft man für *Rechtsmoduln*. Im folgenden verstehen wir unter einem *Modul* stets einen Linksmodul.

(2) Einen Modul kann man ansehen als eine Art Vektorraum über einem Ring (statt eines Körpers). Einfache Rechenregeln beweist man genauso wie bei Vektorräumen, etwa

$$0x = 0,$$
$$r0 = 0,$$
$$(-1)x = -x.$$

Definition. Seien M, M' Moduln über R und $f\colon M \to M'$ eine Abbildung. f heißt *R-Modulhomomorphismus*, wenn für alle $x, y \in M$ und $r \in R$ gilt

$$f(x + y) = f(x) + f(y),$$
$$f(rx) = rf(x).$$

Beispiele. (1) Jeder Vektorraum V über einem Körper K ist ein K-Modul.

(2) Sei R der Ring $\mathbb{Z}$ der ganzen Zahlen und M eine abelsche Gruppe. Dann kann man M als $\mathbb{Z}$-Modul auffassen, und zwar mit folgendem Skalarprodukt. Sei $n \in Z$ und $x \in M$. Im Fall $n \geq 0$ setzen wir

$$nx = \underbrace{x + x + \cdots + x}_{n\text{-mal}},$$

und im Fall $n < 0$ sei $nx := -(-n)x$. Die Eigenschaften verifiziert man durch einfaches Nachrechnen (mit Fallunterscheidungen).

(3) Sei V Vektorraum über einem Körper K und R der Endomorphismenring $\mathrm{End}_K(V)$. Dann ist V ein R-(Links)modul mittels

$$f \cdot x = f(x).$$

(4) Sei K ein Ring. Dann ist $M := K^n = \{ \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \mid \alpha_1, \ldots, \alpha_n \in K \}$ ein K-Modul mit folgender Addition und Skalarmultiplikation.

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} := \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix},$$

$$\alpha \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} := \begin{pmatrix} \alpha\alpha_1 \\ \vdots \\ \alpha\alpha_n \end{pmatrix}.$$

(5) Sei V ein K-Modul über einem kommutativen Ring K und $f\colon V \to V$ ein K-Modulhomomorphismus. Wir definieren V_f wie folgt als $K[t]$-Modul. $V_f := V$ als abelsche Gruppe. Ferner wird das Skalarprodukt

$$K[t] \times V \to V \quad \text{definiert durch}$$
$$\Big(\sum_{i=0}^{n} \alpha_i t^i\Big) \cdot v = \sum_{i=0}^{n} \alpha_i f^i(v).$$

Wir zeigen, daß V_f ein $K[t]$-Modul ist. Seien also $r := \sum_{i=0}^{n} \alpha_i t^i$ und $s := \sum_{j=0}^{n} \beta_j t^j$ aus $K[t]$, ferner $v, w \in V$. Dann gilt

$$\begin{aligned} r(sv) &= \Big(\sum_{i=0}^{n} \alpha_i t^i\Big)\Big(\sum_{j=0}^{m} \beta_j f^j(v)\Big) \\ &= \sum_{i=0}^{n} \alpha_i f^i\Big(\sum_{j=0}^{m} \beta_j f^j(v)\Big) \\ &= \sum_{i,j} \alpha_i \beta_j f^{i+j}(v) \\ &= \sum_{l} \Big(\sum_{i+j=l} \alpha_i \beta_j \Big) f^l(v) \\ &= (rs)v. \end{aligned}$$

Ferner gilt $1v = f^0(v) = v$, und $r(v + w) = \sum_i \alpha_i f^i(v + w) = \sum_i \alpha_i (f^i(v) + f^i(w)) = rv + rw$ sowie (wobei wir oBdA $n = m$

annehmen), und auch

$$(r+s)v \ = \sum_{i=0}^{n}(\alpha_i+\beta_i)f^i(v) = \sum_{i=0}^{n}\alpha_i f^i(v) + \sum_{i=0}^{n}\beta_i f^i(v) = rv+sv.$$

BEMERKUNG. Die Skalarmultiplikation im $K[t]$-Modul V_f, also

$$\Big(\sum_{i=0}^{n}\alpha_i t^i\Big)\cdot v := \sum_{i=0}^{n}\alpha_i f^i(v)$$

kann man auffassen als Abbildung

$$\Phi_f\colon K[t] \to \mathrm{End}_K(V)$$
$$\Phi_f(r)(v) = rv.$$

Dieses Φ_f ist ein K-Vektorraumhomomorphismus und auch ein Ringhomomorphismus, denn es gilt

$$\Phi_f(rs)(v) = (rs)v = r(sv) = \Phi_f(r)\left(\Phi_f(s)(v)\right) = \left(\Phi_f(r)\circ\Phi_f(s)\right)(v),$$

also $\Phi_f(rs) = \Phi_f(r)\circ\Phi_f(s)$.

10.3.2. Satz von Cayley/Hamilton. Als Anwendung des eben konstruierten Moduls V_f oder auch des entsprechenden Ringhomomorphismus $\Phi_f\colon K[t]\to\mathrm{End}_K(V)$ wollen wir jetzt eine wichtige Eigenschaft des charakteristischen Polynoms beweisen.

Sei also K ein Körper, V ein n-dimensionaler K-Vektorraum und f ein Endomorphismus, also $f\colon V\to V$ linear. Das charakteristische Polynom p_f von f war in 7.2.2 definiert als

$$p_f := p_A := \det(A-tE) \in K[t],$$

wobei $A\in K^{n\times n}$ eine (beliebige) darstellende Matrix von f ist.

Wir wollen uns eine Übersicht über die Menge aller Polynome aus $K[t]$ verschaffen, die durch Φ_f auf $0\in\mathrm{End}_K(V)$ abgebildet werden, also über

$$\mathcal{I}_f := \{\,\sum_{i=0}^{n}\alpha_i t^i \in K[t] \mid \sum_{i=0}^{n}\alpha_i f^i = 0 \in \mathrm{End}_K(V)\,\}.$$

Es ist klar, daß $\mathcal{I}_f$ ein von 0 verschiedenes Polynom vom Grad höchstens n^2 enthält, denn $f^0, f^1, \ldots, f^{n^2}$ sind linear abhängig wegen $\dim\mathrm{End}_K(V) = n^2$. Der zu beweisende Satz von Cayley/Hamilton sagt aus, daß sogar schon ein Polynom vom Grad n in $\mathcal{I}_f$ liegen muß, nämlich das charakteristische Polynom p_f.

SATZ (Cayley/Hamilton). *Sei K ein Körper, V ein n-dimensionaler K-Vektorraum, $f\in\mathrm{End}_K(V)$ ein Endomorphismus und p_f das charakteristische Polynom von f. Dann ist $\Phi_f(p_f)=0$. Das heißt: Aus $p_f=\sum_{i=0}^{n}\alpha_i t^i$ folgt $\sum_{i=0}^{n}\alpha_i f^i = 0\in\mathrm{End}_K(V)$.*

BEWEIS. Sei $A\in K^{n\times n}$ darstellende Matrix von f bzgl. einer Basis von V, und V_f definiert wie im Beispiel oben. Nach Definition der Skalarmultiplikation in V_f gilt dann für alle $v\in V$

$$t\cdot v = f(v) = Av.$$

Wir schieben jetzt eine allgemeine Bemerkung ein, die die folgende Argumentation verständlich machen soll. Sei M ein R-Modul. Ist $A \in R^{n\times m}$ und $x \in M$, so wird Ax komponentenweise erklärt. Ist noch $B \in K^{m\times k}$, so gilt $(AB)x = A(Bx)$.

Es ist $1 \cdot v = v$, wenn man 1 als $\in K[t]$, ferner $v \in V = K^n$ und $\cdot$ als das Skalarprodukt im Modul V_f auffaßt. Also ist Ev nach der Bemerkung oben erklärt, und ebenso tEv (was klammerfrei geschrieben werden kann). Gleichfalls kann man in Av die Matrix A als $\in K[t]^{n\times n}$ und die Multiplikation als Skalarmultiplikation in V_f auffassen. Also gilt

$$\begin{aligned} tv = tEv = Av, \\ (A - tE)v = 0. \end{aligned}$$

Setze $B := A - tE \in K[t]^{n\times n}$. Nach 6.2.4 gibt es zu B eine komplementäre Matrix $B^{\#} \in K[t]^{n\times n}$ mit

$$B^{\#}B = \det(B)E.$$

Aus $Bv = 0$ folgt

$$0 = B^{\#}Bv = \det(B)v.$$

Es war $p_f = \det(B) = \sum_{i=0}^{n} \alpha_i t^i$. Für unser beliebiges $v \in V$ ist also

$$0 = \det(B)v = \Big(\sum_{i=0}^{n} \alpha_i t^i\Big)v = \sum_{i=0}^{n} \alpha_i f^i(v) = \Big(\sum_{i=0}^{n} \alpha_i f^i\Big)(v)$$

und damit $\sum_{i=0}^{n} \alpha_i f^i = 0 \in \mathrm{End}_K(V)$. □

Bemerkung. Ist $x_1, \ldots, x_n$ Basis von V und $A \in K^{n\times n}$ eine Matrix mit

$$f \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

so gilt auch

$$f^i \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^i \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Aus $\Phi_f(p_f) = \sum_{i=0}^{n} \alpha_i f^i = 0$ folgt also, daß

$$\Big(\sum_{i=0}^{n} \alpha_i A^i\Big) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \Big(\sum_{i=0}^{n} \alpha_i f^i\Big) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0,$$

also $\sum_{i=0}^{n} \alpha_i A^i = 0$. Das „Einsetzen“ der Matrix A in ihr eigenes charakteristisches Polynom p_A liefert also die Nullmatrix.

10.3.3. Das Minimalpolynom. Das charakteristische Polynom p_A zu A ist i.a. keineswegs ein Polynom kleinsten Grades mit A als „Nullstelle“. Ist zum Beispiel $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \in K^{2\times 2}$, so ist $p_A = (1-t)^2$, aber $\mu_A := t - 1$ besitzt ebenfalls A als Nullstelle, denn es gilt $\mu_A(A) = A - 1 \cdot E = 0$.

Sei $V \neq 0$ ein endlichdimensionaler K-Vektorraum und $f\colon V \to V$ linear. $\mathcal{I}_f$ sei definiert wie in 10.3.2. Offenbar ist $\mathcal{I}_f$ abgeschlossen gegen Addition (d.h. aus $g, h \in \mathcal{I}_f$ folgt $g+h \in \mathcal{I}_f$) und gegen Multiplikation mit beliebigen Elementen des Rings $K[t]$ (d.h. aus $g \in K[t]$ und $h \in \mathcal{I}_f$ folgt $gh \in \mathcal{I}_f$); eine solche Teilmenge eines Ringes nennt man ein *Ideal*. Nach dem Satz von Cayley/Hamilton ist $p_f \in \mathcal{I}_f$, wir kennen also bereits ein Polynom $\neq 0$ in $\mathcal{I}_f$.

Definition. Sei K ein Körper. Unter den Polynomen $\neq 0$ in $\mathcal{I}_f$ gibt es eines kleinsten Grades, das wir (wegen der Abgeschlossenheit von $\mathcal{I}_f$ gegen Multiplikation mit beliebigen Körperelementen) als normiert annehmen können, d.h. der Leitkoeffizient (also der Koeffizient der höchsten t-Potenz) ist 1. Dieses Polynom ist offenbar eindeutig bestimmt; man nennt es das *Minimalpolynom* von f und bezeichnet es mit μ_f.

Wir wollen zeigen, daß das Minimalpolynom Teiler des charakteristischen Polynoms p_f ist. Dazu benötigen wir die Division mit Rest im Polynomring $K[t]$.

Satz (Division mit Rest). *Sei K ein Körper. Zu $f, g \in K[t]$ mit $f \neq 0$ gibt es eindeutig bestimmte $q, r \in K[t]$ mit*

$$g = fq + r \quad \textit{und } \deg(r) < \deg(f).$$

Beweis. *Existenz.* Sei

$$\begin{aligned} f &= a_n t^n + \cdots + a_1 t + a_0 \quad \text{mit } a_n \neq 0 \\ g &= b_m t^m + \cdots + b_1 t + b_0 \quad \text{mit } b_m \neq 0. \end{aligned}$$

Den Beweis führten wir durch Induktion nach m.

Fall $m = 0$, $n > 0$. Setze $q = 0$ und $r = g$.

Fall $m \geq 0$, $n = 0$. Setze $q = \frac{1}{a_0} g$ und $r = 0$.

Fall $m > 0$, $n > 0$. OBdA sei $n \leq m$ (sonst setze $q = 0$ und $r = g$). Es gibt ein $g_1 \in K[t]$ mit

$$g = \frac{b_m}{a_n} t^{m-n} f + g_1 \quad \text{und } \deg(g_1) < m.$$

Nach Induktionsvoraussetzung gibt es $q_1, r_1 \in K[t]$ mit

$$\begin{aligned} g &= \frac{b_m}{a_n} t^{m-n} f + q_1 f + r \quad \text{und } \deg(r) < n \\ &= \underbrace{\left(\frac{b_m}{a_n} t^{m-n} + q_1\right)}_{=:q} f + r \end{aligned}$$

Eindeutigkeit. Gelte $g = q_1 f + r_1 = q_2 f + r_2$ mit $\deg(r_1), \deg(r_2) < \deg(f)$. Dann ist $f(q_1 - q_2) = r_2 - r_1$ und $\deg(r_2 - r_1) < \deg(f)$. Wäre $q_1 - q_2 \neq 0$, so würde gelten $\deg(f(q_1 - q_2)) = \deg(f) + \deg(q_1 - q_2) \geq \deg(f)$ (wegen 7.1.2, da K ein Körper ist), also ein Widerspruch (wegen $\deg(0) = -\infty$). Folglich muß $q_1 = q_2$ sein und somit auch $r_1 = r_2$. ∎

KOROLLAR. *Sei $V \neq 0$ ein endlichdimensionaler K-Vektorraum und $f\colon V \to V$ linear. Dann ist das Minimalpolynom von f ein Teiler des charakteristischen Polynoms von f, also*

$$\mu_f | p_f.$$

BEWEIS. Wegen $\mu_f \neq 0$ gibt es nach dem Satz $q, r \in K[t]$ mit

$$p_f = \mu_f q + r \quad \text{und } \deg(r) < \deg(\mu_f).$$

Wegen $r(f) = p_f(f) - \mu_f(f)q(f) = 0$ und der Minimalität des Grades des Minimalpolynoms muß $r = 0$ sein. □

SATZ. *Sei $V \neq 0$ ein endlichdimensionaler K-Vektorraum und $f\colon V \to V$ linear. Dann haben das charakteristische Polynom p_f und das Minimalpolynom μ_f die gleichen Nullstellen (i.a. haben also die Nullstellen von μ_f kleinere Vielfachheit).*

BEWEIS. Wegen $\mu_f | p_f$ ist jede Nullstelle von μ_f auch Nullstelle von p_f. Sei umgekehrt λ Nullstelle von p_f, also Eigenwert von f, und x ein Eigenvektor zu λ, also $x \neq 0$ mit $f(x) = \lambda x$. Sei

$$\mu_f = \sum_{i=0}^{k} \alpha_i t^i$$

das Minimalpolynom von f. Dann gilt

$$0 = \Big(\sum_{i=0}^{k} \alpha_i f^i\Big)(x) = \sum_{i=0}^{k} \alpha_i f^i(x) = \sum_{i=0}^{k} \alpha_i \lambda^i x = \Big(\sum_{i=0}^{k} \alpha_i \lambda^i\Big) x.$$

Also ist $\mu_f(\lambda) = 0$. □

10.3.4. Berechnung des Minimalpolynoms. Der folgende Satz liefert ein Verfahren, das Minimalpolynom aus der Jordanschen Normalform zu berechnen.

SATZ. *Sei V ein n-dimensionaler K-Vektorraum und $f\colon V \to V$ linear. Ferner sei p_f Produkt linearer Faktoren, etwa*

$$p_f = (-1)^n (t - \lambda_1)^{e_1} \cdot \dots \cdot (t - \lambda_m)^{e_m} \quad \textit{mit verschiedenen } \lambda_i \textit{ und } e_i \geq 1.$$

Sei $k_i = \min\{\, k \mid \ker g_i^k = \ker g_i^{k+1} \,\}$ mit $g_i := f - \lambda_i \mathrm{id}$. Dann ist

$$\mu_f = (t - \lambda_1)^{k_1} \cdot \dots \cdot (t - \lambda_m)^{k_m}$$

das Minimalpolynoms von f.

BEMERKUNG. Die Existenz von k_i wurde im Beweis des Satzes 7.3.1 von der Jordanschen Normalform bewiesen. Aus diesem Beweis geht auch hervor, daß k_i die Zeilenzahl des größten zum Eigenwert λ_i gehörigen Jordan-Kästchens ist, und daß auch $k_i = \min\{\, k \mid (g_i \restriction U_i)^k = 0 \,\text{und}\, (g_i \restriction U_i)^{k-1} \neq 0 \,\}$ ist mit $U_i = \ker g_i^k$. Man nennt k_i den *Index* der nilpotenten Abbildung $g_i \restriction U_i$. Ist insbesondere f diagonalisierbar, so folgt

$$\mu_f = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_m).$$

BEWEIS. Nach der obigen Folgerung ist μ_f ein Teiler von p_f, und nach dem darauf folgenden Satz hat μ_f dieselben Nullstellen wie p_f, also

$$\mu_f = (t-\lambda_1)^{n_1} \cdot \cdots \cdot (t-\lambda_m)^{n_m}$$

mit $n_i \le e_i$ für $i \in \{1,\dots,m\}$. Zu zeigen ist $n_i = k_i$.

Nach dem Beweis von 7.3.1 ist $V = U_1 \oplus \cdots \oplus U_m$ mit $U_i = \ker g_i^{k_i}$ eine Zerlegung von V in f-invariante Unterräume.

Wir zeigen $n_i \le k_i$. Da die n_i Exponenten des Minimalpolynoms sind, genügt

$$(f-\lambda_1\mathrm{id})^{k_1} \circ \cdots \circ (f-\lambda_m\mathrm{id})^{k_m} = 0, \quad \text{also } g_1^{k_1} \circ \cdots \circ g_m^{k_m} = 0.$$

Wegen $V = U_1 \oplus \cdots \oplus U_m$ genügt es dafür zu zeigen, daß jedes $x_i \in U_i$ auf 0 abgebildet wird. Sei etwa $x_1 \in U_1$. Dann gilt

$$(g_1^{k_1} \circ g_2^{k_2} \circ \cdots \circ g_m^{k_m})(x_1) = (g_2^{k_2} \circ \cdots \circ g_m^{k_m})(g_1^{k_1}(x_1)) = 0,$$

da g_i und g_j kommutieren (denn $(f-\lambda_i\mathrm{id}) \circ (f-\lambda_j\mathrm{id}) = f^2 - (\lambda_i+\lambda_j)f + \lambda_i\lambda_j\mathrm{id} = (f-\lambda_j\mathrm{id}) \circ (f-\lambda_i\mathrm{id})$) und $x_1 \in U_1 = \ker g_1^{k_1}$.

Wir zeigen jetzt, daß die Annahme $n_i < k_i$ auf einen Widerspruch führt. Nehmen wir etwa $n_1 < k_1$ an. Dann gibt es ein $x_1 \in U_1$ mit $y := g_1^{n_1}(x_1) \neq 0$. Da U_1 f-invariant ist (und $g_j = f - \lambda_j\mathrm{id}$), ist U_1 auch g_j-invariant, also auch $g_j^{n_j}$-invariant. Für $j \neq 1$ ist

$$g_j^{n_j} \restriction U_1 \colon U_1 \to U_1$$

injektiv, denn aus $g_j^{n_j}(x) = 0$ für ein $x \in U_1$ folgt $x \in \ker g_j^{n_j} \subseteq U_j$ (wegen $n_j \le k_j$), also $x \in U_1 \cap U_j = 0$. Also ist $g_j^{n_j} \restriction U_1 \colon U_1 \to U_1$ ein Isomorphismus und damit auch $g_2^{n_2} \restriction U_1 \circ \cdots \circ g_m^{n_m} \restriction U_1$. Wegen des schon verwendeten Kommutierens von g_i und g_j folgt

$$(g_1^{n_1} \circ g_2^{n_2} \circ \cdots \circ g_m^{n_m})(x_1) = (g_2^{n_2} \circ \cdots \circ g_m^{n_m})(\underbrace{g_1^{n_1}(x_1)}_{=y\neq 0}) \neq 0$$

im Widerspruch zu $\mu_f(f) = 0$. □

10.3.5. Faktormoduln. Sei M ein R-Modul und $N \subseteq M$ eine Teilmenge. N heißt *R-Untermodul* von M, wenn N gegen Addition und gegen Skalarmultiplikation mit beliebigen Ringelementen abgeschlossen ist.

SATZ (Konstruktion des Faktormoduls). *Sei M ein R-(Links)modul, $N \subseteq M$ ein Untermodul und*

$$\mathrm{kan} \colon M \to M/N$$
$$x \mapsto [x].$$

Dann gibt es genau eine R-Modulstruktur auf M/N, so daß kan *ein R-Homomorphismus wird.*

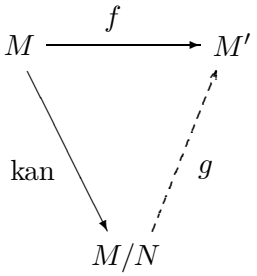
BEWEIS. M/N ist abelsche Gruppe bei $[x]+[y] = [x+y]$. Definiere

$$R \times M/N \to M/N$$
$$(r,[x]) \mapsto [rx].$$

Zu zeigen ist nur noch, daß diese Abbildung wohldefiniert ist (denn es ist klar, daß M/N damit zu einem R-Modul wird). Seien also $x, y \in M$, $r \in R$

und gelte $[x] = [y]$. Dann ist $x-y =: n \in N$, also $rx-ry = r(x-y) = rn \in N$ und damit $[rx] = [ry]$. □

SATZ (Universelle Eigenschaft). *Seien M, M' R-Moduln, $N \subseteq M$ Untermodul und $f\colon M \to M'$ ein R-Modulhomomorphismus mit $f[N] = \{0\}$. Dann gibt es genau einen R-Modulhomomorphismus $g\colon M/N \to M'$ mit $g \circ \mathrm{kan} = f$.*



BEWEIS. Wie für Gruppen. □

KOROLLAR (Homomorphiesatz). *Seien M, M' R-Moduln und $f\colon M \to M'$ ein R-Modulhomomorphismus. Dann gilt*

$$\begin{aligned} \overline{f}\colon M/\ker f &\to \operatorname{im} f \\ [x] &\mapsto f(x) \end{aligned}$$

ist ein wohldefinierter R-Modulisomorphismus (d.h. ein bijektiver R-Modulhomomorphismus; es ist $\ker f := f^{-1}(0)$).

BEWEIS. Wie für Gruppen. □

10.3.6. Basis eines Moduls.

DEFINITION. Sei I eine Menge und R ein *Ring*.

$$\begin{aligned} R^I &:= \{\, f\colon I \to R \mid f \text{ Abbildung} \,\} \\ &= \{\, (r_i)_{i\in I} \mid r_i \in R \text{ für } i \in I \,\} \end{aligned}$$

wird zu einem R-Modul durch

$$\begin{aligned} (r_i) + (s_i) &:= (r_i + s_i) \\ r \cdot (r_i) &:= (rr_i) \end{aligned}$$

für $r_i, s_i, r \in R$. Weiter definieren wir

$$R^{(I)} := \{\, (r_i)_{i\in I} \in R^I \mid \text{für fast alle } i \in I \text{ gilt } r_i = 0 \,\}.$$

Für $j \in I$ sei

$$e_j := (r_i)_{i\in I} \quad \text{mit} \quad r_i := \delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j. \end{cases}$$

Sei M ein R-Modul und I eine Menge. Eine Familie $(x_i)_{i\in I}$ von Elementen aus M heißt *Basis* von M wenn folgendes gilt.

- Für alle $x \in M$ gibt es eine endliche Teilmenge $J \subseteq I$ von I und $r_i \in R$ für $i \in J$ so daß $x = \sum_{i\in J} r_i x_i$.
- Für alle endlichen Teilmengen $J \subseteq I$ und $r_i, s_i \in R$ für $i \in J$

folgt aus $\sum_{i\in J} r_i x_i = \sum i \in J s_i x_i$ stets $r_i = s_i$ für alle $i \in J$.

BEMERKUNGEN. (1) Sei I eine Menge und R ein Ring. Dann ist offenbar $R^{(I)}$ ein R-Modul mit der Basis $(e_i)_{i\in I}$.

(2) Man kann zeigen, daß nicht jeder Modul eine Basis besitzt (ein Beispiel ist $\mathbb{Z}/n\mathbb{Z}$). Ferner ist die Basislänge i.a. nicht eindeutig bestimmt. Diese für Vektorräume gültigen Aussagen lassen sich also nicht auf Moduln übertragen.

(3) Auch für Moduln gilt der Satz von der *linearen Fortsetzung*. Seien M, M' R-Moduln, $(x_i)_{i\in I}$ eine Basis von M und $(y_i)_{i\in I}$ eine beliebige Familie von Elementen aus M'. Dann gibt es genau einen Modulhomomorphismus $f\colon M \to M'$ mit $f(x_i) = y_i$ für alle $i \in I$. Dies beweist man wie in 3.1.4 für Vektorräume.

10.4. Tensorprodukt

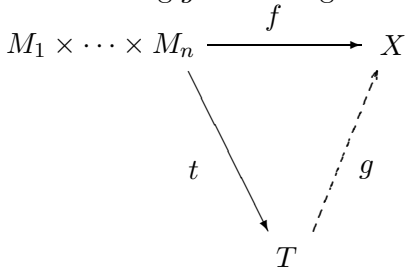
In diesem Abschnitt sei K ein kommutativer Ring, z.B. ein Körper.

10.4.1. Multilineare Abbildungen.

DEFINITION. Sei n eine natürliche Zahl und $M_1, \dots, M_n, M$ R-Moduln. Eine Abbildung $f\colon M_1 \times \cdots \times M_n \to M$ heißt *n-multilinear*, wenn für jedes i und alle $x_j \in M_j$ mit $j \neq i$ die Abbildung

$$\begin{aligned} M_i &\to M \\ x_i &\mapsto f(x_1, \dots, x_n) \end{aligned}$$

R-linear ist. Sei T ein R-Modul und $t\colon M_1 \times \cdots \times M_n \to T$ eine Abbildung. t heißt *universell n-multilinear*, wenn t n-multilinear ist und für jeden R-Modul X und jede n-multilineare Abbildung $f\colon M_1 \times \cdots \times M_n \to X$ es genau eine R-lineare Abbildung $g\colon T \to X$ gibt mit $f = g \circ t$, d.h. so daß

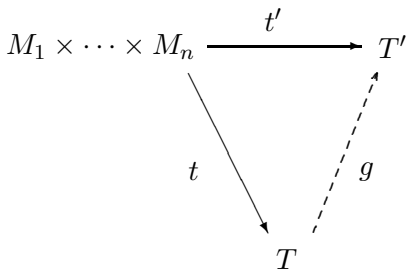


kommutativ ist.

BEMERKUNG. Seien

$$\begin{aligned} t&\colon M_1 \times \cdots \times M_n \to T \\ t'&\colon M_1 \times \cdots \times M_n \to T' \end{aligned}$$

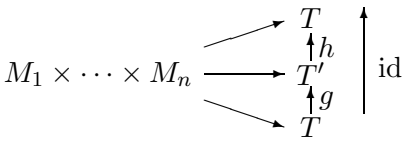
universell n-multilinear. Dann gibt es genau einen Isomorphismus $g\colon T \to T'$ mit $t' = g \circ t$.



BEWEIS. 1. Da t multilinear und t' universell multilinear ist, gibt es ein $h\colon T' \to T$ mit $t = h \circ t'$.

2. Da t' multilinear und t universell multilinear ist, gibt es ein $g\colon T \to T'$ mit $t' = g \circ t$.

3. Aus (1), (2) folgt $t = h \circ t' = h \circ g \circ t$, also $h \circ g = \mathrm{id}$ (da t universell ist) und ebenso $t' = g \circ t = g \circ h \circ t'$, also $g \circ h = \mathrm{id}$ (da t' universell ist). □



10.4.2. Existenz des Tensorprodukts.

SATZ. *Seien $M_1, \ldots, M_n$ R-Moduln. Dann gibt es einen R-Modul T und eine universelle n-multilineare Abbildung $t\colon M_1 \times \cdots \times M_n \to T$.*

BEWEIS. Sei $F := R^{(M_1 \times \cdots \times M_n)}$ der von der Menge $M_1 \times \cdots \times M_n$ frei erzeugte R-Modul, also der R-Modul mit der aus allen $e_{(x_1,\ldots,x_n)}$ für $(x_1, \ldots, x_n) \in M_1 \times \cdots \times M_n$ bestehenden Basis. In F identifizieren wir $e_{(x_1,\ldots,x_n)}$ mit $(x_1, \ldots, x_n)$. Sei $N \subseteq F$ der Untermodul, der erzeugt wird von allen Elementen von F der Form

$$\begin{aligned}(x_1, \ldots, x_{i-1}, x_i + x_i', x_{i+1}, \ldots, x_n) - (x_1, \ldots, x_i, \ldots, x_n)& \\ - (x_1, \ldots, x_i', \ldots, x_n),& \\ (x_1, \ldots, x_{i-1}, rx_i, x_{i+1}, \ldots, x_n) - r(x_1, \ldots, x_i, \ldots, x_n)&\end{aligned}$$

mit $x_1 \in M_1, \ldots, x_i, x_i' \in M_i, \ldots, x_n \in M_n$ und $r \in R$. Definiere

$$\begin{aligned}&T := F/N,\\ &t\colon M_1 \times \cdots \times M_n \to T = F/N,\\ &t(x_1, \ldots, x_n) := [(x_1, \ldots, x_n)].\end{aligned}$$

Wir zeigen zunächst, daß t multilinear ist. Zum Beispiel für $r \in R$ und $x_i \in M_i$ haben wir

$$\begin{aligned}t(rx_1, x_2, \ldots, x_n) &= [(rx_1, x_2, \ldots, x_n)]\\ &= [r(x_1, x_2, \ldots, x_n)]\\ &= r[(x_1, x_2, \ldots, x_n)]\\ &= rt(x_1, x_2, \ldots, x_n),\end{aligned}$$

da $(rx_1, x_2, \ldots, x_n) - r(x_1, \ldots, x_n) \in N$. Wir zeigen jetzt, daß t universell multilinear ist. Sei also $f\colon M_1 \times \cdots \times M_n \to X$ multilinear, X ein R-Modul. Wir definieren

$$\begin{aligned}&f'\colon F \to X \quad \text{durch}\\ &f'(e_{(x_1,\ldots,x_n)}) := f(x_1, \ldots, x_n).\end{aligned}$$

Es gibt genau einen R-Homomorphismus mit dieser Eigenschaft (lineare Fortsetzung). Es gilt $f'[N] = 0$, denn f' ist auf jedem der obigen Erzeugenden von N Null; z.B. ist

$$\begin{aligned} &f'((rx_1, x_2, \ldots, x_n) - r(x_1, \ldots, x_n)) \\ &= f'(rx_1, x_2, \ldots, x_n) - f'(r(x_1, \ldots, x_n)) \\ &= f(rx_1, x_2, \ldots, x_n) - rf(x_1, \ldots, x_n) \\ &= 0, \quad \text{da } f \text{ multilinear ist.} \end{aligned}$$

Nach der universellen Eigenschaft des Faktormoduls gibt es einen R-Homomorphismus

$$\begin{aligned} g\colon F/N = T &\to X \\ g([e_{(x_1,\ldots,x_n)}]) &\mapsto f(x_1, \ldots, x_n). \end{aligned}$$

Es gilt $f = g \circ t$, denn für $x_i \in M_i$ $(1 \leq i \leq n)$ ist

$$g(t(x_1, \ldots, x_n)) = g([e_{(x_1,\ldots,x_n)}]) = f(x_1, \ldots, x_n).$$

g ist eindeutig bestimmt durch $f = g \circ t$: Sei noch $g'\colon F/N = T \to X$ gegeben mit $f = g' \circ t$. Dann gilt

$$g(t(x_1, \ldots, x_n)) = g([e_{(x_1,\ldots,x_n)}]) = g'([e_{(x_1,\ldots,x_n)}]) = g'(t(x_1, \ldots, x_n)).$$

Die Abbildungen g und g' stimmen also auf einem Erzeugendensystem von $T = F/N$ überein (denn $(e_{(x_1,\ldots,x_n)})_{(x_1,\ldots,x_n)\in M_1\times\cdots\times M_n}$ ist Basis von F), also $g = g'$. ∎

10.4.3. Definition des Tensorprodukts.

DEFINITION. Seien $M_1, \ldots, M_n$ R-Moduln.

$$M_1 \otimes_R \cdots \otimes_R M_n = M_1 \otimes \cdots \otimes M_n := T$$

mit T wie im Beweis des Satzes heißt *Tensorprodukt*. Die im Beweis ebenfalls konstruierte Abbildung

$$\text{kan} := t\colon M_1 \times \cdots \times M_n \to M_1 \otimes \cdots \otimes M_n$$

heißt *kanonische Abbildung*. Für $x_i \in M_i$ $(1 \leq i \leq n)$ setzt man

$$x_1 \otimes \cdots \otimes x_n := t(x_1, \ldots, x_n).$$

BEMERKUNGEN. (1) Nach dem Satz ist $t\colon M_1 \times \cdots \times M_n \to M_1 \otimes \cdots \otimes M_n$ universell multilinear und deshalb nach einer obigen Bemerkung bis auf Isomorphie eindeutig bestimmt.

(2) Da $t\colon M_1 \times \cdots \times M_n \to M_1 \otimes \cdots \otimes M_n$ multilinear ist, gilt

$$\begin{aligned} x_1 \otimes \cdots \otimes (x_i + x_i') \otimes \cdots \otimes x_n &= x_1 \otimes \cdots \otimes x_i \otimes \cdots \otimes x_n + \\ &\quad x_1 \otimes \cdots \otimes x_i' \otimes \cdots \otimes x_n, \\ x_1 \otimes \cdots \otimes rx_i \otimes \cdots \otimes x_n &= r(x_1 \otimes \cdots \otimes x_i \otimes \cdots \otimes x_n) \end{aligned}$$

für $x_1 \in M_1, \ldots, x_i, x_i' \in M_i, \ldots, x_n \in M_n$, $r \in R$ und $1 \leq i \leq n$. Insbesondere ist

$$x_1 \otimes \cdots \otimes x_n = 0 \quad \text{falls ein } x_i = 0 \text{ ist.}$$

(3) Die Elemente $x_1 \otimes \cdots \otimes x_n \in M_1 \otimes \cdots \otimes M_n$ bilden ein Erzeugendensystem von $M_1 \otimes \cdots \otimes M_n$ als R-Modul (nach dem Beweis des Satzes) und sogar als abelsche Gruppe, denn es gilt

$$r(x_1 \otimes \cdots \otimes x_n) = rx_1 \otimes \cdots \otimes x_n.$$

BEISPIEL. $R = \mathbb{Z}$, $M_1 = \mathbb{Z}/2\mathbb{Z}$, $M_2 = \mathbb{Z}/3\mathbb{Z}$. Dann gilt

$$M_1 \otimes M_2 = \mathbb{Z}/2\mathbb{Z} \otimes \mathbb{Z}/3\mathbb{Z} = 0,$$

denn für $\alpha, \beta \in \mathbb{Z}$ ist $3 \cdot (\alpha + 2\mathbb{Z}) = 3\alpha + 2\mathbb{Z} = \alpha + 2\mathbb{Z}$ (Fallunterscheidung α gerade/ungerade), also

$$\begin{aligned}(\alpha + 2\mathbb{Z}) \otimes (\beta + 3\mathbb{Z}) &= 3(\alpha + 2\mathbb{Z}) \otimes (\beta + 3\mathbb{Z}) \\ &= (\alpha + 2\mathbb{Z}) \otimes \underbrace{3(\beta + 3\mathbb{Z})}_{=3\mathbb{Z}=[0]} \\ &= 0.\end{aligned}$$

10.4.4. Basis des Tensorprodukts.

SATZ. *Seien $M_1, \ldots, M_n$ R-Moduln, ferner $(x_{i_1})_{i_1 \in I_1}$ Basis von M_1, …, $(x_{i_n})_{i_n \in I_n}$ Basis von M_n. Dann ist*

$$(x_{i_1} \otimes \cdots \otimes x_{i_n})_{i_1 \in I_1, \ldots, i_n \in I_n}$$

Basis von $M_1 \otimes \ldots, \otimes M_n$.

BEWEIS. *Erzeugendensystem.* Sei $m_j \in M_j$ für $1 \leq j \leq n$. Dann ist m_j R-Linearkombination von Elementen aus $(x_{i_j})_{i_j \in I_j}$. Also ist $m_1 \otimes \cdots \otimes m_n$ R-Linearkombination von Elementen $x_{i_1} \otimes \cdots \otimes x_{i_m}$, da kan: $M_1 \times \cdots \times M_n \to M_1 \otimes \cdots \otimes M_n$ multilinear war.

Lineare Unabhängigkeit. Seien $i_1 \in I_1, \ldots, i_n \in I_n$ fest gewählt. Betrachte

$$\begin{aligned}f &\colon M_1 \times \cdots \times M_n \to R \\ f(m_1, \ldots, m_n) &= r_{i_1} \cdots r_{i_n},\end{aligned}$$

wobei r_{i_j} der Koeffizient von x_{i_j} in der Basisdarstellung von m_j ist. Offensichtlich ist f multilinear. Nach der universellen Eigenschaft des Tensorprodukts gibt es einen R-Homomorphismus

$$g \colon M_1 \otimes \cdots \otimes M_n \to R$$

mit $f = g \circ \text{kan}$. Also ist

$$g(x_{l_1} \otimes \cdots \otimes x_{l_n}) = \begin{cases} 1 & \text{falls } l_1 = i_1, \ldots, l_n = i_n \\ 0 & \text{sonst.} \end{cases}$$

Sei jetzt

$$\sum_{l_1, \ldots, l_n} r_{l_1 \cdots l_n} x_{l_1} \otimes \cdots \otimes x_{l_n} = 0 \qquad \text{mit } r_{l_1 \cdots l_n} \in R, \text{ fast alle } r_{l_1 \cdots l_n} = 0.$$

Daraus folgt

$$0 = g\Big(\sum_{l_1, \ldots, l_n} r_{l_1 \cdots l_n} x_{l_1} \otimes \cdots \otimes x_{l_n} \Big) = r_{i_1 \cdots i_n}.$$

Dies gilt für alle $i_1, \ldots, i_n$. ∎

DEFINITION. Sei V Vektorraum über einem Körper K und $p, q \in \mathbb{N}$.

$$V_p^q := \underbrace{V \otimes \cdots \otimes V}_{p\text{-mal}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q\text{-mal}}$$

heißt *Tensorraum* der *p-fach kontravarianten* und *q-fach kovarianten Tensoren* (oder auch Tensorraum der Stufe (p, q)). Hierbei ist $V^* := \mathrm{Hom}_K(V, V)$ der zu V duale Vektorraum.

BEMERKUNG. Sei $v_1, \ldots, v_n$ Basis von V. $f_i \in V^*$ sei definiert durch $f_i(v_j) := \delta_{ij}$ für $1 \le i \le n$. Dann ist $f_1, \ldots, f_n$ Basis von V^*. (f_i) heißt die zu (v_i) *duale Basis*. Im folgenden verwenden wir die Schreibweise $f_i =: v^i \in V^*$.

BEWEIS. *Erzeugendensystem.* Sei $f \in V^*$ und $f(v_i) =: \alpha_i$. Dann gilt

$$\Big(\sum_i \alpha_i f_i\Big)(v_j) = \sum_i \alpha_i f_i(v_j) = \alpha_j,$$

also $\sum_i \alpha_i f_i = f$.

Lineare Unabhängigkeit. Gelte $\sum_i \alpha_i f_i = 0$. Also

$$0 = \Big(\sum_i \alpha_i f_i\Big)(v_j) = \sum_i \alpha_i f_i(v_j) = \alpha_j$$

für alle j mit $1 \le j \le n$. ∎

10.4.5. Transformationsformel.

KOROLLAR. *Sei V ein K-Vektorraum, $v_1, \ldots, v_n$ Basis von V unf $p, q \in \mathbb{N}$. Dann hat*

$$V_p^q = \underbrace{V \otimes \cdots \otimes V}_{p\text{-mal}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q\text{-mal}}$$

die K-Basis

$$(v_{i_1} \otimes \cdots \otimes v_{i_p} \otimes v^{j_1} \otimes \cdots \otimes v^{j_q})_{1 \le i_1, \ldots, i_p, j_1, \ldots, j_q \le n}$$

(also $\dim V_p^q = n^{p+q}$). Jedes $x \in V_p^q$ hat also folgende Basisdarstellung.

$$x = \sum_{\substack{i_1, \ldots, i_p \\ j_1, \ldots, j_q}} \xi_{j_1 \cdots j_q}^{i_1 \cdots i_p} v_{i_1} \otimes \cdots \otimes v_{i_p} \otimes v^{j_1} \otimes \cdots \otimes v^{j_q}$$

mit $\xi_{j_1 \cdots j_q}^{i_1 \cdots i_p} v_{i_1} \in K$. (Die Schreibweise der Indizes ist Konvention. Sie ist so getroffen, daß stets über oben und unten stehende Indizes summiert wird). Sei $\tilde{v}_1, \ldots, \tilde{v}_n$ eine weitere Basis von V, $(\tilde{v}^i)$ die duale Basis zu $(\tilde{v}_i)$. Seien weiter α_i^j, $\beta_i^j \in K$, $1 \le i, j \le n$ mit

$$\boxed{\tilde{v}_i = \sum_j \alpha_i^j v_j, \qquad v_i = \sum_j \beta_i^j \tilde{v}_j}$$

(also $(\alpha_i^j)_{ij}(\beta_i^j)_{ij} = E$). x habe bzgl. der neuen Basis folgende Darstellung:

$$x = \sum_{\substack{s_1, \ldots, s_p \\ t_1, \ldots, t_q}} \tilde{\xi}_{t_1 \cdots t_q}^{s_1 \cdots s_p} \tilde{v}_{s_1} \otimes \cdots \otimes \tilde{v}_{s_p} \otimes \tilde{v}^{t_1} \otimes \cdots \otimes \tilde{v}^{t_q}.$$

Dann gilt die Transformationsformel:

$$\boxed{\tilde{\xi}^{s_1\cdots s_p}_{t_1\cdots t_q} = \sum_{\substack{i_1,\ldots,i_p\\ j_1,\ldots,j_q}} \xi^{i_1\cdots i_p}_{j_1\cdots j_q}\alpha^{j_1}_{t_1}\cdots\alpha^{j_q}_{t_q}\beta^{s_1}_{i_1}\cdots\beta^{s_p}_{i_p}.}$$

Beweis. Die erste Aussage ist klar nach dem vorherigem Satz. Für die zweite Aussage ist zu berechnen

(1) (v_i) aus $(\tilde{v}_i)$.
(2) (v^j) aus $(\tilde{v}^j)$.

Zu (1). Nach Voraussetzung ist

$$v_i = \sum_s \beta^s_i \tilde{v}_s.$$

Zu (2). Gelte $v^j = \sum_t \gamma^j_t \tilde{v}^t$. Man erhält

$$\begin{aligned} v^j(\tilde{v}_i) &= \Big(\sum_t \gamma^j_t \tilde{v}^t\Big)(\tilde{v}_i)\\ &= \sum_t \gamma^j_t \tilde{v}^t(\tilde{v}_i)\\ &= \gamma^j_i \end{aligned}$$

und andererseits wegen $\tilde{v}_i = \sum_s \alpha^s_i v_s$

$$\begin{aligned} v^j(\tilde{v}_i) &= v^j\Big(\sum_s \alpha^s_i v_s\Big)\\ &= \sum_s \alpha^s_i v^j(v_s)\\ &= \alpha^j_i. \end{aligned}$$

Also ist

$$v^j = \sum_t \alpha^j_t \tilde{v}^t.$$

Man erhält

$$\sum_{\substack{i_1,\ldots,i_p,j_1,\ldots,j_q\\ s_1,\ldots,s_p,t_1,\ldots,t_q}} \xi^{i_1\cdots i_p}_{j_1\cdots j_q}\beta^{s_1\cdots s_p}_{i_1\cdots i_p}\alpha^{j_1\cdots j_q}_{t_1\cdots t_q}\tilde{v}_{s_1}\otimes\cdots\otimes\tilde{v}_{s_p}\otimes\tilde{v}^{t_1}\otimes\cdots\otimes\tilde{v}^{t_q}.$$

Also

$$\tilde{\xi}^{s_1\cdots s_p}_{t_1\cdots t_q} = \sum_{\substack{i_1,\ldots,i_p\\ j_1,\ldots,j_q}} \xi^{i_1\cdots i_p}_{j_1\cdots j_q}\beta^{s_1\cdots s_p}_{i_1\cdots i_p}\alpha^{j_1\cdots j_q}_{t_1\cdots t_q}.$$

□

Bemerkungen. (1) Die Transformationsformel für die Koeffizienten von Tensoren kann man sich folgendermaßen merken.

$$\tilde{v}^i = \sum_j \alpha^j_i v_j \qquad \text{Basiswechsel,} \quad A := (\alpha^j_i)_{ij}.$$

Kovariante Indizes $j_1,\ldots,j_q$ transformieren sich mittels A, kontravariante Indizes $i_1,\ldots,i_p$ mittels $(A^{-1})^t$. Für A orthogonal ist

$(A^{-1})^t = A$, d.h. kovariante und kontravariante Indizes transformieren sich gleich.

(2) Häufig werden Tensoren durch das Transformationsverhalten der Koeffizienten definiert: Jeder Basis $v_1, \dots, v_n$ von V ordnen wir eine Familie $\xi(v_1, \dots, v_n) = (\xi^{i_1 \cdots i_p}_{j_1 \cdots j_q})_{1 \le i_\nu, j_\mu \le n}$ $(\in K^{n^{p+q}})$ zu:

$$(v_1, \dots, v_n) \mapsto \xi(v_1, \dots, v_n).$$

Bei einem Basiswechsel von (v_i) zu $(\tilde{v}_i)$ gilt zwischen

$$\xi(v_1, \dots, v_n) =: (\xi^{i_1 \cdots i_p}_{j_1 \cdots j_q}) \quad \text{und}$$
$$\xi(\tilde{v}_1, \dots, \tilde{v}_n) =: (\tilde{\xi}^{s_1 \cdots s_p}_{t_1 \cdots t_q})$$

der im Satz beschriebene Zusammenhang. Dann ist

$$\begin{aligned} x &:= \sum_{i_\nu, j_\mu} \xi^{i_1 \cdots i_p}_{j_1 \cdots j_q} v_{i_1} \otimes \cdots \otimes v_{i_p} \otimes v^{j_1} \otimes \cdots \otimes v^{j_q} \\ &= \sum_{s_\nu, t_\mu} \tilde{\xi}^{s_1 \cdots s_p}_{t_1 \cdots t_q} \tilde{v}_{s_1} \otimes \cdots \otimes \tilde{v}_{s_p} \otimes \tilde{v}^{t_1} \otimes \cdots \otimes \tilde{v}^{t_q} \end{aligned}$$

Tensor in V^q_p. Für jede Basis $v_1, \dots, v_n$ von V ist also $\xi(v_1, \dots, v_n)$ das System der Koeffizienten des Tensors x.

(3) Für $\dim V < \infty$ läßt sich der Tensorraum V^q_p auch anders definieren: Für K-Vektorräume $V_1, \dots, V_n$ sei

$$M := \operatorname{Mult}(V_1 \times \dots V_n, K) := \{\, f \colon V_1 \times \cdots \times V_n \to K \mid f \text{ multilinear} \,\}$$

die Menge der *Multilinearformen*. M ist K-Vektorraum durch

$$(f+g)(v_1, \dots, v_n) := f(v_1, \dots, v_n) + g(v_1, \dots, v_n)$$
$$(\alpha f)(v_1, \dots, v_n) := \alpha f(v_1, \dots, v_n)$$

für $f, g \in M$ und $\alpha \in K$.

Außerdem gilt für $\dim V_1 < \infty, \dots, \dim V_n < \infty$:

$$t \colon V_1 \times \cdots \times V_n \to \operatorname{Mult}(V_1^* \times \cdots \times V_n^*, K)$$
$$(v_1, \dots, v_n) \mapsto ((f_1, \dots, f_n) \mapsto f_1(v_1) \dots f_n(v_n))$$

ist universelle multilineare Abbildung, d.h.

$$V_1 \otimes \cdots \otimes V_n \xrightarrow{\cong} \operatorname{Mult}(V_1^* \times \cdots \times V_n^*, K)$$
$$v_1 \otimes \cdots \otimes v_n \mapsto ((f_1, \dots, f_n) \mapsto f_1(v_1) \cdots f_n(v_n))$$

Insbesondere gilt für $\dim V < \infty$:

$$\begin{aligned} V^q_p &= \underbrace{V \otimes \cdots \otimes V}_{p\text{-mal}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q\text{-mal}} \\ &\cong \operatorname{Mult}(\underbrace{V^* \times \cdots \times V^*}_{p\text{-mal}} \times \underbrace{V^{**} \times \cdots \times V^{**}}_{q\text{-mal}}, K) \end{aligned}$$

Wegen

$$V \xrightarrow{\cong} V^{**}$$
$$v \mapsto (f \mapsto f(v))$$

folgt

$$V_p^q \cong \mathrm{Mult}(\underbrace{V^* \times \cdots \times V^*}_{p\text{-mal}} \times \underbrace{V \times \cdots \times V}_{q\text{-mal}}, K).$$

BEWEIS. t ist offensichtlich multilinear. Wir zeigenjetzt, daß t universell multilinear ist. Sei X ein K-Vektorraum und $f\colon V_1 \times \cdots \times V_n \to X$ multilinear. Sei $(v_{i_\nu})_{i_\nu \in I_\nu}$ Basis von V_ν für $1 \le \nu \le n$. f ist bestimmt durch die Werte auf $\{\, (v_{i_1}, \ldots, v_{i_n}) \mid i_1 \in I_1, \ldots, i_n \in I_n \,\}$. Es genügt nun zu zeigen:

$(t(v_{i_1}, \ldots, v_{i_n}))_{i_1 \in I_1, \ldots, i_n \in I_n}$ ist Basis von $\mathrm{Mult}(V_1^* \times \cdots \times V_n^*, K)$.

Dies sieht man wie folgt. *Erzeugendensystem.* Sei $h \in \mathrm{Mult}(V_1^* \times \cdots \times V_n^*, K)$,

$$h(v^{i_1}, \ldots, v^{i_n}) =: \alpha_{i_1 \cdots i_n} \in K.$$

Dann gilt:

$$\begin{aligned}
&\Big(\sum_{j_1,\ldots,j_n} \alpha_{j_1 \cdots j_n} t(v_{j_1}, \ldots, v_{j_n}) \Big)(v^{i_1}, \ldots, v^{i_n}) \\
&= \sum_{j_1,\ldots,j_n} \alpha_{j_1 \cdots j_n} v^{i_1}(v_{j_1}) \cdots v^{i_n}(v_{j_n}) \\
&= \alpha_{i_1 \cdots i_n},
\end{aligned}$$

d.h. $h = \sum_{j_1,\ldots,j_n} \alpha_{j_1 \cdots j_n} t(v_{j_1}, \ldots, v_{j_n})$.

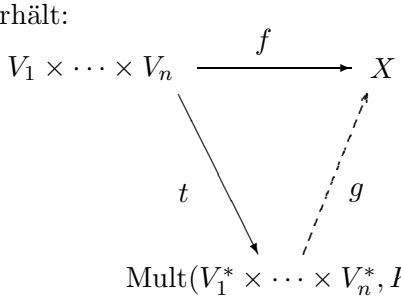
Lineare Unabhängigkeit. Gelte

$$\sum_{j_1,\ldots,j_n} \alpha_{j_1 \cdots j_n} t(v_{j_1}, \ldots, v_{j_n}) = 0.$$

Dann folgt

$$\begin{aligned}
0 &= \Big(\sum_{j_1,\ldots,j_n} \alpha_{j_1 \cdots j_n} t(v_{j_1}, \ldots, v_{j_n}) \Big)(v^{i_1}, \ldots, v^{i_n}) \\
&= \alpha_{i_1 \cdots i_n}.
\end{aligned}$$

Man erhält:



Definiere nun g durch

$$g(t(v_{i_1}, \ldots, v_{i_n})) = f(v_{i_1}, \ldots, v_{i_n}).$$

Offensichtlich ist g wohldefiniert, K-Homomorphismus, und es gilt $f = g \circ t$. ∎

BEMERKUNG. Im allgemeinen ($\dim V = \infty$) ist $V^{**} \ncong V$ und $\mathrm{Mult}(V_1^* \times \cdots \times V_n^*, K) \ncong V_1 \otimes \cdots \otimes V_n$.

Bemerkung. In 8.1.6 haben wir die Komplexifizierung bezüglich $\mathbb{R} \subseteq \mathbb{C}$ durchgeführt. Allgemeiner kann man eine entsprechende Konstruktion für eine beliebige Körpererweiterung $K \subseteq L$ und einen K-Vektorraum V durchführen. Man bildet dann das Tensorprodukt $L \otimes_K V$ und faßt es als L-Vektorraum auf.

Literaturverzeichnis

1. Oliver Deiser, *Einführung in die Mengenlehre*, 2nd ed., Springer Verlag, Berlin, Heidelberg, New York, 2004.
2. Gerd Fischer, *Lineare Algebra*, 14th ed., Vieweg, 2003.
3. Paul R. Halmos, *Finite-dimensional vector spaces*, Springer Verlag, Berlin, Heidelberg, New York, 1974.
4. Hannes Stoppel and Birgit Griese, *Übungsbuch zur Linearen Algebra*, Vieweg, Braunschweig, 1997.

Index

$\mathbb{R}$-Algebra-Automorphismus, 142
Äquivalenzklasse, 7
Äquivalenzrelation, 7

Abbildung, 4
  affine, 69
  alternierende, 104
  bijektive, 5
  Einschränkung einer, 4
  hermitesche, 152
  injektive, 4
  kanonische, 224
  lineare, 47
  multilineare, 104, 222
  natürliche, 45
  normierte, 104
  semilineare, 152
  surjektive, 4
  universell multilineare, 222
Abbildungen
  Komposition von, 5
Abstand, 85, 88
affine Hülle, 66
affiner Unterraum
  Parallelität, 66
affiner Unterraum, 65
  Dimension, 66
Affinität, 69
algebraisch abgeschlossen, 143
Allquantor, 4, 7
Assoziativgesetz, 8
Austauschlemma, 39
Austauschsatz, 39
Auswahlsaxiom, 37
Automorphismengruppe, 198
Automorphismus, 53

Bahn, 201
Basis, 34
  eines Moduls, 221
  negativ orientierte, 116
  positiv orientierte, 116
Basisergänzungssatz, 36
Betrag, 141
  einer komplexen Zahl, 14
Bewegung, 92
Bild, 5, 48, 55
bilinear, 79
Bolzano-Weierstraß
  Satz von, 99

Cauchyfolge, 13
Cayley-Hamilton
  Satz von, 216
Cayley-Transformation, 193
Cosinussatz, 79, 83
Cramersche Regel, 111

Definitionsbereich, 4
Dehnung, 169
Determinante, 104
Diagonalgestalt, 72
Diedergruppe, 199, 204
Differenz, 3
Dimension, 36
Dimensionssatz, 56
Distributivgesetz, 12
Dodekaeder, 199, 201
Doppelsumme, 12
Drehachse, 184
Drehebene, 184
Drehgruppe, 199
Drehspiegelung, 185
Drehung, 170, 183
Drehwinkel, 184
  absoluter, 185
  orientierter, 185
Dreiecksungleichung, 82, 154
Durchschnitt, 3, 6

Ebene, 39, 66
Eigenraum, 123
Eigenvektor, 123
Eigenwert, 123
Einheitengruppe, 197
Einheitsmatrix, 17
Einheitsvektor, 33
Einsetzung, 122
Element
  maximales, 37

Ellipse, 70, 77
Ellipsoid, 77
endlichdimensional, 34
Endomorphismus, 53
  adjungierter, 158
  diagonalisierbarer, 124, 137
  normaler, 165
  orthogonaler, 91, 161
  orthonaler, 160
  selbstadjungierter, 161
  unitärer, 161
Epimorphismus, 53
Erzeugendensystem, 34
euklidisches Skalarprodukt, 79
Eulersche Winkel, 192
Existenzquantor, 4

Familie, 5
Faser, 203
Fixraum, 181
Fläche
  parabolische, 75, 78, 93
Folge, 6
  beschränkte, 98
  konvergente, 98
Fortsetzung
  lineare, 52, 222
Fourierkoeffizient, 85
Fundamentalsatz
  der Algebra, 133
  der Algebra, 168
  der Algebra, 143, 168
Funktion, 4
  charakteristische, 6
  quadratische, 70
Funktor, 145

Galoiserweiterung, 148
Gaußsches Eliminationsverfahren, 20
gemeinsames Lot, 88
Gerade, 40, 66
Geraden
  windschiefe, 88
Geradenspiegelung, 182, 183
gleichorientiert, 116
Gleichungssystem
  homogenes, 50
  inhomogenes, 50
  lineares, 1
Grad eines Polynoms, 122
Gruppe, 1, 8
  abelsche, 9
  affine, 69, 198
  allgemeine lineare, 27, 198
  alternierende, 197
  euklidische, 198
  orthogonale, 92, 161, 198
  spezielle lineare, 198
  spezielle orthogonale, 161, 198
  spezielle unitäre, 161
  symmetrische, 10, 197
  unitäre, 161
  zyklische, 197
Gruppenhomomorphismus, 195
Gruppenisomorphismus, 196

Hauptachsentransformation, 164, 168
Hauptraum, 129
hermitesches Skalarprodukt, 152
Hessesche Normalform, 89
Homomorphiesatz, 212, 221
Homomorphismus, 53
  natürlicher, 56
Hyperbel, 70, 77
Hyperboloid
  einschaliges, 77
  zweischaliges, 77
Hyperebene, 66
Hyperebenenspiegelung, 170
Hyperfläche
  quadratische, 70

Ideal, 218
Identität, 5
Ikosaeder, 199, 201
imaginäre Einheit, 14
Imaginärteil, 141
Index, 203
  einer nilpotenten Abbildung, 219
Interpolation, 112
inverses Element, 9
isomorph, 196
Isomorphismus, 53
Isotropiegruppe, 201

Jacobi, 95
Jacobimatrix, 189
Jacobische Rotationsmatrix, 96
Jordan-Matrix, 128
Jordan-Normalform, 63

Körper, 1, 11
  endlicher, 214
Kürzungsregel, 122
kartesisches Produkt, 4
Kegel, 77
Kegelfläche, 75, 77, 93
Kern, 48, 55
Klasse, 3
  echte, 3
Koeffizient, 15
Kommutativgesetz, 9
kompakt, 192
Komplexaddition, 44
Komplexifizierung, 144
Komprehensionsprinzip, 2

Konjugation, 147, 202
Koordinatensystem, 29
Kreis, 70
Kruskal-Algorithmus, 39
kumulative Typenstruktur, 2

Länge, 81, 153
Lagrange
  Satz von, 203
Laplacescher Entwicklungssatz, 108
linear unabhängig, 32
linear abhängig, 32
lineare Varietät, 67
Linearkombination, 34
Linksmodul, 214
Linksnebenklasse, 208
Lot, 81, 84, 153, 155

Matrix, 15
  adjungierte, 157
  darstellende, 60, 63
  hermitesche, 161
  inverse, 26
  invertierbare, 25
  komplementäre, 109
  konjugierte, 157
  normale, 165
  orthogonale, 92, 161
  symmetrische, 161
  transponierte, 25
  unitäre, 161
Matrizen
  ähnliche, 63
  orthogonal äquivalente, 162
  unitär äquivalente, 162
Menge, 1, 3
  reine, 2
Mengen
  disjunkte, 4
Minimalpolynom, 218
Mittelpunktsfläche, 75, 93
Modul, 214
Modulhomomorphismus, 214
modulo, 208
Monomorphismus, 53
Multilinearform, 228
Multiplikationssatz, 109

neutrales Element, 9
Normalenvektor, 89
Normalteiler, 196
Nullmatrix, 16
Nullstelle, 122
  e-fache, 123

Oktaeder, 199, 201
Operation, 201
Orbit, 201
Ordnung, 195
  eines Elements, 213
  lineare, 37
  partielle, 37
Orientierung, 116
orthogonal, 81, 153
orthogonale Projektion, 81, 84, 153, 154
orthogonales Komplement, 87
Orthonormalbasis, 84, 154
Orthonormalisierungsverfahren, 85, 155
Orthonormalsystem, 84, 154

Parabel, 71, 77
Paraboloid
  elliptisches, 78
  hyperbolisches, 78
Parallelotop, 114
Parallelprojektion, 168
Parameterdarstellung
  von SO(2), 188
Permutation, 11
  gerade, 107
  ungerade, 107
Permutationsgruppe, 197
Pivot, 19
plantonische Körper, 199
Pole, 205
Polynom, 120
  charakteristisches, 125
  konstantes, 122
Polynomring, 120
positiv definit, 80, 152
Potenzmenge, 6
Projektion, 168
  orthogonale, 153, 168
Punkt, 66
Pythagoras
  Satz des, 83

QR-Zerlegung, 194
quadratische Hyperfläche, 70
quadratische Funktion, 70
Quadrik, 70
Quotientenraum, 44

Rang, 58
  einer linearen Abbildung, 62
Realteil, 141
Rechtsmodul, 214
Rechtsnebenklasse, 208
Reduzierbarkeit, 126
Regel von Sarrus, 110
Relation, 4
  antisymmetrische, 37
  auf M, 4
  rechtseindeutige, 4
  reflexive, 7, 37
  symmetrische, 7

transitive, 7, 37
Richtung, 66
Ring, 103
kommutativer, 103
Ringisomorphismus, 142
Russellsche Antinomie, 2

Schranke
kleinste obere, 37
obere, 37
Schwarzsche Ungleichung, 82, 154
senkrecht, 87
Sesquilinearform, 152
Shoenfield-Prinzip, 2
Skalar, 30
Skalarprodukt
hermitesches, 152
kanonisches, 80, 153
Spaltenrang, 57
Spaltenraum, 57
Spaltenumformung, 19
Spiegelung, 170
Spur, 186
Stützstellen, 112
Stabilisator, 201
Streckung, 169
Summe
direkte, 42
von Unterräumen, 41
Supremum, 37
Sylvestersches Trägheitsgesetz, 78
Symmetriegruppe, 198
symmetrisch, 80

Teilmenge, 3
Teilverhältnis, 70
Tensor
p-fach kontravarianter und q-fach kovarianter, 226
Tensorprodukt, 224
Tensorraum, 226
der Stufe (p, q), 226
Tetraeder, 199
Trägheitsgesetz
Sylvestersches, 78
Translation, 69
Transposition, 107
Typ
einfacher, 3

Umkehrabbildung, 5
Unabhängigkeit vom Repräsentanten, 8
Untergruppe, 92, 196
von M erzeugt, 197
Untermodul, 220
Unterraum, 40
von M erzeugter, 41
Untervektorraum, 40

Urbild, 5

Vandermondesche Determinante, 112
Vektor, 30
Vektorprodukt, 113
Vektorraum, 30
euklidischer, 80
Vektorraumhomomorphismus, 53
Vereinigung, 3, 6
Vielfachheit
algebraische, 127, 138
geometrische, 127, 138
Volumen
eines Parallelotops, 114
Vorzeichen, 107

Würfel, 199
Wertebereich, 4
Winkel, 83

Zahl
ganze, 6, 9
komplexe, 14, 141
konjugiert komplexe, 14, 141
rationale, 11, 12
reelle, 13
Zeilenrang, 57
Zeilenraum, 57
Zeilenstufenmatrix, 19
Zeilenumformung, 19
Zornsches Lemma, 37
Zusammenhangskomponente, 193
Zylinder
elliptischer, 77
hyperbolischer, 77
parabolischer, 78