Lineare Algebra
und
Analytische Geometrie

Oswald Riemenschneider

Vorlesungen Hamburg

Sommersemester 1991, Wintersemester 1991/92

Gekürzte, korrigierte und mit Skizzen versehene

Fassung vom 8. März 2005

Vorwort

Diese Noten dokumentieren den Inhalt meiner jeweils zweistündigen Vorlesungen über Lineare Algebra und Analytische Geometrie, deren Grundlage die Bücher von GERD FISCHER[1] waren. Den Vorlesungsteilnehmern wurden nur die Teile ausgehändigt, in denen ich wesentlich von dem Grundtext, der vor allem auch zum Selbststudium benutzt werden sollte, abgewichen bin. Auf allgemeinen Wunsch habe ich mich später entschlossen, den gesamten Text auszuarbeiten, die in der Vorlesung nur zitierten Abschnitte aufzunehmen und somit eine halbwegs kohärente Materialsammlung einem größeren Leserkreis zugänglich zu machen. Der Text enthielt in dieser Fassung als Anhang einige Anwendungen der Linearen Algebra, insbesondere auf Probleme der Analysis, die ich vorwiegend in der Veranstaltung „Anwendungen und Modelle" präsentiert hatte. Ich habe bei der vorliegenden Neuauflage auf die Aufnahme dieser Teile verzichtet, da sie in meine Manuskripte *Analysis I, Analysis II* und *Analysis III* eingeflossen sind. Am Ende befinden sich auch die Übungsaufgaben, wie sie zu der Vorlesung ausgeteilt wurden. Eine Sammlung von Lösungen ist auf Anfrage separat erhältlich.

An Voraussetzungen sollte der Leser/die Leserin eine Kenntnis der Notationen der „elementaren Mengenlehre" und der „Abbildungstheorie" (siehe dazu z. B. meine eben erwähnten Texte zur Analysis) und ein gutes Training in formaler Logik, zudem Rechenfertigkeit in den Bereichen der natürlichen, ganzen, reellen und komplexen Zahlen mitbringen, vor allem aber Neugier auf mathematisches Basiswissen und Freude am mathematischen Denken.

Hamburg, im März 2005
Oswald Riemenschneider

[1] Siehe die Literaturliste auf p. iii).

Inhalt

Literatur

[1] E. Artin: Analytische Geometrie und Algebra I, II. Vorlesungsausarbeitung 1960/61, Hamburg.

[2] B. Artmann: Lineare Algebra. Basel–Boston–Stuttgart: Birkhäuser 1986.

[3] S. Bosch: Lineare Algebra. Berlin: Springer 2001.

[4] E. Brieskorn: Lineare Algebra und Analytische Geometrie I, II, III. Braunschweig: Vieweg 1983, 1985 und $\geq$ 1993.

[5] N. Bourbaki: Algèbre I, Chap. 1 à 3. Paris: Hermann 1970.

[6] G. Fischer: Lineare Algebra. Braunschweig etc.: Vieweg 1983 (7., durchg. Auflage und spätere).

[7] G. Fischer: Analytische Geometrie. Braunschweig etc.: Vieweg 1985 (3. Auflage und spätere).

[8] P. Halmos: Finite–dimensional vector spaces. Princeton, New Jersey, etc.: Van Nostrand 1958 (2nd edition).

[9] K. Hoffmann, R. Kunze: Linear Algebra. Englewood Cliffs, New Jersey: Prentice Hall 1971 (2nd edition).

[10] K. Jänich: Lineare Algebra. Berlin etc.: Springer 1991 (4. verb. Auflage).

[11] K. Klingenberg: Lineare Algebra und Geometrie. Heidelberg, etc.: Springer 1984.

[12] M. Koecher: Lineare Algebra und analytische Geometrie. Grundwissen Mathematik 2. Berlin etc.: Springer 1985 (2. Auflage).

[13] A. I. Kostrikin, Yu. I. Manin: Linear algebra and geometry. New York etc.: Gordon and Breach 1989.

[14] H.–J. Kowalsky: Lineare Algebra. Berlin etc.: Walter de Gruyter 1972 (6., verb. Auflage).

[15] S. Lang: Linear Algebra. Reading, Mass.: Addison–Wesley 1966.

[16] R. Lingenberg: Lineare Algebra. Mannheim: BI 1969.

[17] F. Lorenz: Lineare Algebra I, II. Mannheim–Wien–Zürich: BI Wissenschafts–Verlag 1988, 1989 (2., überarbeitete Auflage).

[18] E. Oeljeklaus, R. Remmert: Lineare Algebra I. Berlin etc.: Springer 1974.

[19] G. Pickert: Analytische Geometrie. Leipzig: Akademische Verlagsgesellschaft 1955.

[20] G. Scheja, U. Storch: Lehrbuch der Algebra. Teil 1, 2, 3. Stuttgart: Teubner 1980, 1988, 1991.

[21] E. Sperner: Einführung in die Analytische Geometrie (1. Teil). Göttingen: Vandenhoeck & Ruprecht 1959.

[22] U. Stammbach: Lineare Algebra. Stuttgart: Teubner 1980.

[23] G. Strang: Linear algebra and its applications. New York etc.: Academic Press 1976.

(Und in diesen Werken aufgeführte weitere Literatur).

Einleitung

Mit *Lineare Algebra und Analytische Geometrie* bezeichnet man diejenige mathematische Disziplin, deren Zweck in der Grundlegung des abstrakten Begriffes der Linearität von Abbildungen zwischen geeigneten mathematischen Objekten und seiner Verwendung und Veranschaulichung bei Realisierungen in Geometrie und Analysis besteht. Es handelt sich also bei diesem Gebiet um ein Wechselspiel zwischen algebraischen und geometrischen Begriffsbildungen, wobei (heutzutage) die Algebra im Vordergrund steht, da diese den universellen Rahmen bildet, in den sich die verschiedensten Konkretisierungen einfügen lassen.

Ich möchte dieses Wechselspiel erläutern an dem Studium von *zwei linearen Gleichungen in zwei Unbekannten*:

$$(*) \qquad \begin{cases} ax + by = r \\ cx + dy = s\,. \end{cases}$$

a, b, c, d, r, s sind dabei fest vorgegebene Elemente aus einer gewissen Menge von „Zahlen", die ich mit $\mathbb{K}$ bezeichnen möchte und deren Eigenschaften später noch festgelegt werden sollen. Eine *Lösung* von $(*)$ (in $\mathbb{K}$) ist ein (geordnetes) Paar $(\overline{x}, \overline{y}) \in \mathbb{K}^2$, so daß nach Einsetzen von $\overline{x}$ statt x und $\overline{y}$ statt y in $(*)$ beide Gleichungen wahr (erfüllt) sind.

Aufgabe. Finde *alle* Lösungen des Systems $(*)$ und beschreibe diese konzeptionell (strukturell).

Mindestvoraussetzung an den Zahlbereich $\mathbb{K}$ ist natürlich, daß es zwei Verknüpfungen $+$ und $\cdot$ auf $\mathbb{K}$ gibt, d. h. Abbildungen

$$\begin{array}{ccc|ccc} \mathbb{K}\times\mathbb{K} & \longrightarrow & \mathbb{K} & \mathbb{K}\times\mathbb{K} & \longrightarrow & \mathbb{K} \\ (x, y) & \longmapsto & x+y & (x, y) & \longmapsto & x\cdot y \quad \text{oder} \quad xy\,, \end{array}$$

die gewissen Bedingungen genügen müssen. So vereinbart man (i. a. stillschweigend), daß die Multiplikation $\cdot$ „stärker bindet" als die Addition $+$, d. h. daß der Ausdruck $ax + by$ genauer als

$$((a \cdot x) + (b \cdot y))$$

zu interpretieren ist.

Beispiele.

$$\begin{array}{lll} \mathbb{N} & = \{0, 1, 2, \ldots\}\,, & \text{die Menge der } \textit{natürlichen} \text{ Zahlen} \\ & & \text{mit der üblichen Addition } + \text{ und} \\ & & \text{der üblichen Multiplikation} \cdot, \\ \mathbb{Z} & = \{0, 1, -1, 2, -2, \ldots\}\,, & \text{die Menge der } \textit{ganzen} \text{ Zahlen}\,, \\ \mathbb{Q} & = \left\{\dfrac{a}{b} \,:\, a, b \in \mathbb{Z}\,,\ b \neq 0\right\}, & \text{die Menge der } \textit{rationalen} \text{ Zahlen}\,, \\ \mathbb{R}\,, & & \text{die Menge der } \textit{reellen} \text{ Zahlen}\,, \\ \mathbb{C}\,, & & \text{die Menge der } \textit{komplexen} \text{ Zahlen}\,. \end{array}$$

Wie löst man nun das System $(*)$? Man nimmt an, daß $(*)$ überhaupt eine Lösung $(\overline{x}, \overline{y})$ besitzt, d. h. daß die Gleichungen

$$a\overline{x} + b\overline{y} = r\,, \quad c\overline{x} + d\overline{y} = s$$

erfüllt sind, und versucht dann, hieraus $\overline{x}$ und $\overline{y}$ durch *Elimination* zu bestimmen. Dazu multipliziert man die erste Gleichung mit d, die zweite mit b:

$$da\overline{x} + db\overline{y} = dr\,, \quad bc\overline{x} + bd\overline{y} = bs\,.$$

Hieraus läßt sich $\overline{y}$ eliminieren, wenn $bd = db$ gilt und man in $\mathbb{K}$ uneingeschränkt *subtrahieren* kann (was z. B. in $\mathbb{N}$ nicht der Fall ist). Wir erhalten dann

$$(ad - bc)\,\overline{x} = dr - bs \,,$$

wobei wir noch $ad = da$ und ein *Distributivgesetz* vorausgesetzt haben. - Es sind nun zwei Fälle zu unterscheiden:

A $\quad \Delta := a\,d - b\,c = 0\,.$

Ist hierbei auch noch $dr - bs = 0$, so folgt für $\overline{x}$ nichts. Ist dagegen $dr - bs \neq 0$, so kann es wegen $\Delta\overline{x} = 0$ (warum ?) kein $\overline{x}$ mit $\Delta\overline{x} = dr - bs$ geben.

B $\quad \Delta = a\,d - b\,c \neq 0\,.$

In diesem Fall ist notwendig

$$\overline{x} = \Delta^{-1}\,(dr - bs)\,,$$

wenn $\Delta^{-1} \in \mathbb{K}$ ein multiplikatives Inverses zu Δ bezeichnet: $\Delta^{-1}\Delta = 1$. Genauso ergibt sich

$$\overline{y} = \Delta^{-1}\,(as - cr)\,.$$

In $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ gibt es zu jedem $\Delta \neq 0$ ein solches Inverses. Also gibt es für diese Zahlbereiche *höchstens eine* Lösung. Umgekehrt zeigt man aber leicht, daß durch die obigen Gleichungen tatsächlich auch eine Lösung gegeben wird. Z. B. ergibt sich sofort

$$a\,\Delta^{-1}(dr - bs) + b\,\Delta^{-1}(as - cr) = \Delta^{-1}(adr - bcr) = \Delta^{-1}\,\Delta\,r = 1 \cdot r = r\,,$$

und ebenso für die zweite Gleichung.

Die Analyse des obigen „Beweises“ zeigt, daß wir nur von gewissen Axiomen Gebrauch machen (denen eines *Körpers*, siehe das nächste Kapitel), ohne auf die Bedeutung der Zahlen zurückgreifen zu müssen.

Erfüllt $\mathbb{K}$ *„hinreichend viele“ Axiome (die eines Körpers) und ist* $\Delta = ad - bc \neq 0$*, so besitzt* $(*)$ *genau eine Lösung in* $\mathbb{K}$*, die zudem durch obige Gleichungen gegeben ist. Ist* $\Delta = 0$*, so braucht* $(*)$ *keine Lösungen zu besitzen, kann aber auch mehr als eine Lösung haben.*

Die Analyse zeigt ferner, daß die Zahlbereiche $\mathbb{N}$ und $\mathbb{Z}$ nicht alle diese Axiome erfüllen. Wollen wir den obigen Sachverhalt ohne Einschränkung erhalten, so müssen wir diese Zahlbereiche, wenn möglich, in einen größeren Körper einbetten und (alle) Lösungen in diesem Körper studieren. Der *kleinste* Körper, der $\mathbb{N}$ bzw. $\mathbb{Z}$ enthält, ist der Körper $\mathbb{Q}$ der rationalen Zahlen.

Wir wollen noch die Bedingung $\Delta = 0$ bzw. $\neq 0$ geometrisch deuten, um die oben angesprochene Dychotomie von Algebra und Geometrie zu beleuchten. Dazu nehmen wir $\mathbb{K} = \mathbb{R}$ an.

$$\Delta = 0 \iff ad - bc = 0 \iff ad = bc \iff \begin{cases} \exists\,\lambda : c = \lambda\,a\,,\ d = \lambda\,b \\ \text{oder} \\ \exists\,\mu : a = \mu\,c\,,\ b = \mu\,d\,. \end{cases}$$

(Wähle $\lambda = c\,a^{-1}$, wenn $a \neq 0$, oder $\lambda = d\,b^{-1}$, wenn $b \neq 0$, oder $\lambda = 0$, wenn $c = d = 0$, etc.). Im ersten Fall lautet das Gleichungssystem:

$$(**) \qquad a\,x + b\,y = r\,, \quad \lambda\,a\,x + \lambda\,b\,y = s\,.$$

Ist $\lambda = 0$, so hat dieses System keine Lösung, wenn $s \neq 0$; sonst ist die 2. Gleichung überflüssig und die 1. Gleichung hat eventuell keine Lösung, oder alle Paare $(\overline{x}, \overline{y}) \in \mathbb{R}^2$ als Lösung (falls $a = b = r = 0$), oder eine *eindimensionale* Lösungsmenge:

$$y = b^{-1}\,(r - a\,x)\,, \quad b \neq 0\,, \quad x \in \mathbb{R} \text{ beliebig (oder umgekehrt)}\,.$$

Ist $\lambda \neq 0$, so lautet das Gleichungssystem

$$a\,x + b\,y = r\,, \quad a\,x + b\,y = \lambda^{-1}\,s\,.$$

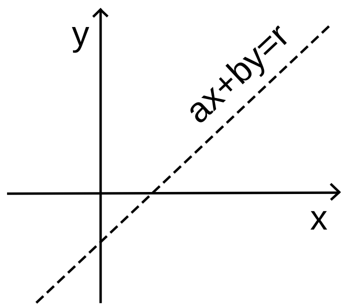
Dies hat keine Lösung, wenn $r \neq \lambda^{-1}\,s$. Sonst kann man wie oben weiterschließen: Ist $\Delta = 0$, so hat $(*)$ keine Lösung oder *viele* Lösungen, oder anders ausgedrückt:

$(*)$ *hat genau eine Lösung* $\Longleftrightarrow \Delta \neq 0$.

Der geometrische Grund für diesen Sachverhalt ist unschwer einzusehen. Die Gleichung

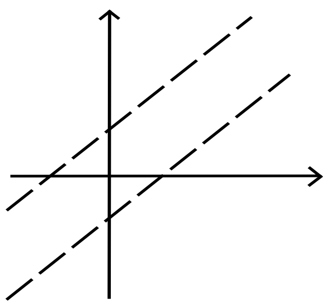
$$a\,x + b\,y = r$$

beschreibt eine Gerade in $\mathbb{R}^2$ (wenn a, b nicht beide gleich 0 sind).



Figur E.1

Zwei Geraden $ax + by = r$, $cx + dy = s$ haben die *gleiche* Steigung (und sind dann parallel oder identisch), genau dann wenn $(a,\, b)$ und $(c,\, d)$ linear abhängig über $\mathbb{R}$ sind, d. h. wenn $(a,\, b) = (\mu c,\, \mu d)$ bzw. $(c,\, d) = (\lambda a,\, \lambda b)$ gilt, und dies ist genau dann der Fall, wenn $\Delta = 0$ ist.



Figur E.2

1 Körper

Die Überlegungen der Einleitung legen die folgenden Axiome nahe. Es sei $\mathbb{K}$ eine Menge zusammen mit zwei Abbildungen $+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ und $\cdot : \mathbb{K} \times \mathbb{K}$, die wir *Addition* bzw. *Multiplikation* nennen. Wir schreiben für diesen Sachverhalt auch kurz $(\mathbb{K}, +, \cdot)$.

Definition. $(\mathbb{K}, +, \cdot)$ heißt ein *Körper*, falls die folgenden Axiome erfüllt sind:

1. Für die *Addition* $+$ gelten die folgenden Regeln:

Assoziativität :	$(a + b) + c = a + (b + c)$	für alle $a, b, c \in \mathbb{K}$.
Kommutativität :	$a + b = b + a$	für alle $a, b \in \mathbb{K}$.
Existenz eines neutralen Elements 0:	$a + 0 = a$	für alle $a \in \mathbb{K}$.
Existenz des Inversen : zu jedem $a \in \mathbb{K}$ existiert ein Element $-a \in \mathbb{K}$ mit	$a + (-a) = 0$.	

   Man sagt dann auch: $(\mathbb{K}, +)$ bildet eine *kommutative* oder *abelsche Gruppe* mit *neutralem* Element 0.

2. Für die *Multiplikation* $\cdot$ gilt:

   $(\mathbb{K}^*, \cdot) = (\mathbb{K} \setminus \{0\}, \cdot)$ bildet eine kommutative Gruppe (wobei das neutrale Element mit 1 und das multiplkative Inverse zu einem Element $a \neq 0$ mit a^{-1} bezeichnet wird).

3. Es gilt ferner das *Distributivgesetz*

$$a\,(b + c) = a\,b + a\,c\,.$$

Beispiele. 1. Die Zahlbereiche $\mathbb{N}$, $\mathbb{Z}$ besitzen eine Addition und eine Multiplikation, sind aber keine Körper (Begründung?). Dagegen erfüllen $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ alle oben aufgelisteten Axiome[2].

2. $\mathbb{F}_2 = \{0, 1\}$, $0 \neq 1$, ist die kleinstmögliche Menge, die eine Körperstruktur besitzt. Aufgrund der Körperaxiome ist die Multiplikationstabelle eindeutig festgelegt (siehe Satz 1). Bei der Additionstabelle ist a priori der Wert von $1 + 1$ noch frei. Es folgt jedoch aus $1 + 1 = 1$ sofort der Widerspruch $1 = 1 + 0 = 1 + (1 + (-1)) = (1 + 1) + (-1) = 1 + (-1) = 0$, so daß auch die Additionstabelle festgelegt ist:

$+$	0	1
0	0	1
1	1	0

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

Man überzeugt sich leicht davon, daß durch diese Tabellen tatsächlich eine Körperstruktur definiert wird.

Allgemeiner kann man zeigen: Ist $\mathbb{F}_p = \{0, 1, 2, \ldots, p - 1\}$, p eine *Primzahl*, und rechnet man in $\mathbb{F}_p \subset \mathbb{Z}$ bezüglich Addition und Multiplikation *modulo* p, d. h. bis auf Reste bzgl. p, so ist $\mathbb{F}_p$ ein Körper. (Siehe auch Aufgabe 6).

Es kommt für viele Betrachtungen nicht darauf an, um welche Zahlen es sich tatsächlich handelt, sondern nur um ihre Beziehungen zueinander. Wir werden daher i. f. meistens nur mit den obigen Axiomen rechnen und alles andere ableiten (müssen). Von Fall zu Fall machen wir noch Zusatzannahmen.

[2]Zur Konstruktion von $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ siehe mein Manuskript *Analysis I.*

Satz 1.1 $\mathbb{K}$ *sei ein Körper. Dann sind die neutralen Elemente* 0 *und* 1 *eindeutig bestimmt, und es gilt*
$$0 \cdot a = 0 \quad \textit{für alle} \quad a \in \mathbb{K}\,.$$

Beweis. $0\,,\ 0'$ seien „zwei" Nullen. Dann folgt $0' = 0' + 0 = 0 + 0' = 0$. Entsprechend ergibt sich $1' = 1' \cdot 1 = 1 \cdot 1' = 1$. Zum Beweis der letzten Aussage setzen wir $b := 0 \cdot a$. Dann folgt $b + b = 0 \cdot a + 0 \cdot a = (0 + 0) \cdot a = 0 \cdot a = b$, also
$$b = b + 0 = b + (b + (-b)) = (b + b) + (-b) = b + (-b) = 0 \ \text{, d. h. } 0 \cdot a = 0\,. \qquad \square$$

Bemerkung. Man schreibt $a - b$ für $a + (-b)$ und $\dfrac{a}{b}$ für $a\,b^{-1}$ etc.

Auch die Inversen sind eindeutig bestimmt. Allgemein gilt:

Satz 1.2 *In einem Körper* $\mathbb{K}$ *hat jede Gleichung*

$$(+) \qquad a + x = b$$

genau eine Lösung, nämlich $x = b - a$. *Die entsprechende Aussage ist richtig für die Gleichung*

$$(\times) \qquad a\,x = b\,, \quad \textit{wenn} \quad a \neq 0\,.$$

Beweis. Selbstverständlich hat $(+)$ die Lösung $b - a$: Es ist $a + (b - a) = a + (b + (-a)) = a + ((-a) + b) = (a + (-a)) + b = 0 + b = b$. Ist umgekehrt x eine beliebige Lösung, so folgt notwendig $b - a = b + (-a) = (a + x) + (-a) = (x + a) + (-a) = x + (a + (-a)) = x + 0 = x$. Für die Multiplikation schließt man ganz entsprechend. $\square$

Satz 1.3 *Es gilt in einem Körper* $a\,b = 0$ *genau dann, wenn* $a = 0$ *oder* $b = 0$.

Mit anderen Worten : *Körper sind* nullteilerfrei.

Beweis. „$\Longleftarrow$" Ist z. B. $a = 0$, so folgt $a\,b = 0$ nach Satz 1.

„$\Longrightarrow$" Umgekehrt ergibt sich aus $a\,b = 0\,,\ a \neq 0$, sofort
$$b = 1 \cdot b = (a^{-1}a)\,b = a^{-1}(a\,b) = a^{-1}0 = 0\,. \qquad \square$$

Satz 1.4 *In jedem Körper gelten die folgenden Regeln* :
$$a\,(-b) = (-a)\,b = -(a\,b)\,, \quad (-a)\,(-b) = a\,b\,.$$

Beweis. Z. B. erhält man unmittelbar aus den Definitionen
$$a\,b + a\,(-b) = a\,(b + (-b)) = a\,0 = 0\,,$$
also $a\,(-b) = -(a\,b)$. Aus $b + (-b) = 0$ folgt ferner $-(-b) = b$, und damit ergibt sich die letzte Identität aus der ersten:
$$a\,b = a\,(-(-b)) = (-a)\,(-b)\,. \qquad \square$$

Seien jetzt endlich viele Elemente $a_1, \ldots, a_n \in \mathbb{K}$ vorgegeben. Man definiert induktiv
$$\sum_{j=1}^{n} a_j := a_1 + \cdots + a_n := \sum_{j=1}^{n-1} a_j + a_n\,.$$

Also
$$a_1 + a_2 + a_3 = (a_1 + a_2) + a_3\,, \quad a_1 + a_2 + a_3 + a_4 = ((a_1 + a_2) + a_3) + a_4 \quad \text{etc.}$$

Für $n \geq 3$ hängt diese Definition nicht von der Beklammerung ab (zum Beweis siehe Satz 4.3 in meinem Manuskript *Analysis I*):

Satz 1.5 *Endliche Summen in $\mathbb{K}$ können beliebig geklammert werden. Das Gleiche gilt für endliche Produkte.*

Ist speziell $a := a_1 = a_2 = \cdots = a_n$, so setzt man

$$\sum_{j=1}^{n} a = \underbrace{a + \cdots + a}_{n-\text{mal}} =: n\,a \ ,$$
$$\prod_{j=1}^{n} a = \underbrace{a \cdot \ldots \cdot a}_{n-\text{mal}} =: a^n \ .$$

Man beachte aber, daß die Abbildung

$$\begin{cases} \mathbb{N} & \longrightarrow & \mathbb{K} \\ n & \longmapsto & n \cdot 1 \end{cases}$$

im allgemeinen *nicht injektiv* ist, d. h. es kann durchaus $n \cdot 1 = 0$ in $\mathbb{K}$ gelten, ohne daß $n = 0$ ist. Tritt diese Pathologie nicht auf wie z. B. für die Körper $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, so nennt man $\mathbb{K}$ *von der Charakteristik* 0, in Zeichen: $\operatorname{char} \mathbb{K} = 0$. In diesem Fall ist der Körper $\mathbb{Q}$ der rationalen Zahlen in $\mathbb{K}$ als *Unterkörper* enthalten. Im anderen Fall gibt es eine *kleinste* positive Zahl $p \in \mathbb{N}$ mit $p \cdot 1 = 0$ in $\mathbb{K}$, und man schreibt dann $\operatorname{char} \mathbb{K} = p$. Der kleinste Körper mit zwei Elementen hat die Charakteristik 2; allgemeiner gilt

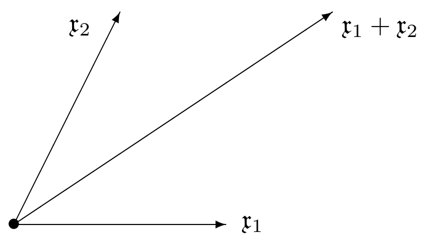
$$\operatorname{char} \mathbb{F}_p = p \ .$$

Es ist übrigens leicht einzusehen, daß die Charakteristik eines Körpers nur Null oder eine *Primzahl* sein kann. - Wir sollten also stets der folgenden Warnung eingedenk sein:

Warnung. Nicht jede Eigenschaft der uns vertrauten Körper überträgt sich auf einen beliebig vorgegebenen Körper $\mathbb{K}$, insbesondere dann nicht, wenn $\operatorname{char} \mathbb{K} = p > 0$.

2 Vektorräume

Der Begriff des *Vektors* stammt ursprünglich aus der Physik (als „gerichtete Größe“). Als Prototyp für algebraische Operationen bei solchen Vektoren gilt die Addition in Form des „Parallelogramms der Kräfte“:



Figur 2.1

Solche Kräftevektoren können nicht nur addiert, sondern auch mit reellen Skalaren multipliziert werden. In obigem Beispiel haben wir die Modellierung von Vektoren in der Ebene als geordnete Paare $(x, y) \in \mathbb{R}^2$ mit Addition

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$

und Sakalaren–Multiplikation

$$r \cdot (x, y) = (rx, ry) .$$

Wir verallgemeinern sogleich.

Definition. Sei $\mathbb{K}$ ein Körper und V eine nichtleere Menge. V heißt ein $\mathbb{K}$–*Vektorraum* bzgl. der Addition

$$+ : V \times V \longrightarrow V , \quad (v_1, v_2) \longmapsto v_1 + v_2 ,$$

und der „Skalarenmultiplikation“

$$\cdot : \mathbb{K} \times V \longrightarrow V , \quad (\lambda, v) \longmapsto \lambda v ,$$

falls die folgenden Axiome erfüllt sind:

1. $(V, +)$ ist eine abelsche Gruppe, d. h.
   a) $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$.
   b) $v_1 + v_2 = v_2 + v_1$.
   c) Es gibt (genau) ein neutrales Element (ebenfalls mit 0 bezeichnet, obwohl es nichts mit der Null $0 \in \mathbb{K}$ zu tun hat) mit

      $v + 0 = v$ für alle $v \in V$.
   d) Für alle v gibt es (genau) ein w mit $v + w = 0$. (Man schreibt dann $w = -v$).
2. Für die Skalaren–Multiplikation gelten die folgenden Regeln:
   a) $(\lambda + \mu) v = \lambda v + \mu v$.
   b) $\lambda (v + w) = \lambda v + \lambda w$.
   c) $(\lambda \mu) v = \lambda (\mu v)$.
   d) $1 \cdot v = v$.

Bemerkung. In jedem Vektorraum gelten die folgenden Rechenregeln:

1. $0 \cdot v = 0$ für alle $v \in V$. (Dabei steht links die Null in $\mathbb{K}$, rechts der Nullvektor in V).
2. $\lambda \cdot 0 = 0$ für alle $\lambda \in \mathbb{K}$.
3. $\lambda\, v = 0$ genau dann, wenn $\lambda = 0$ oder $v = 0$.
4. $-v = (-1)\, v$.

Beispiele. 1. X sei eine beliebige nichtleere Menge, $\mathbb{K}$ sei ein Körper. Man setzt

$$\mathrm{Abb}\,(X, \mathbb{K}) = \{ f : f \text{ Abbildung } X \longrightarrow \mathbb{K} \}$$

und definiert für $f, g : X \longrightarrow \mathbb{K}$, $\lambda \in \mathbb{K}$, Abbildungen $f + g$, $\lambda f : X \to \mathbb{K}$ durch

$$(f + g)(x) := f(x) + g(x)\,, \quad x \in X$$

und

$$(\lambda f)(x) = \lambda \cdot f(x)\,, \quad x \in X\,.$$

Dann sind $f + g$, $\lambda f \in \mathrm{Abb}\,(X, \mathbb{K})$, und es ist trivial nachzurechnen, daß diese Operationen eine $\mathbb{K}$–Vektorraumstruktur auf $\mathrm{Abb}\,(X, \mathbb{K})$ ergeben. Das Negative zu f ist z. B. gegeben durch $(-f)(x) = -f(x)$.

2. Wähle speziell im 1. Beispiel $X = \{1, \ldots, n\}$. Jede Abbildung $f : X \longrightarrow \mathbb{K}$ ist eindeutig bestimmt durch ihre Werte $f(j) =: x_j \in \mathbb{K}$, also durch das n–Tupel $(x_1, \ldots, x_n)$. M. a. W.:

$$\mathrm{Abb}\,(\{1, \ldots, n\}, \mathbb{K}) \cong \mathbb{K}^n := \{(x_1, \ldots, x_n) : x_j \in \mathbb{K}\,,\ j = 1, \ldots, n\}$$

mit der auf der rechten Seite induzierten Addition und Skalarenmultiplikation

$$\begin{aligned} (x_1, \ldots, x_n) + (y_1, \ldots, y_n) &= (x_1 + y_1, \ldots, x_n + y_n)\,, \\ \lambda\,(x_1, \ldots, x_n) &= (\lambda\, x_1, \ldots, \lambda\, x_n)\,. \end{aligned}$$

Wir werden stets die Menge $\mathbb{K}^n$ mit dieser Vektorraum–Struktur versehen.

3. $\mathcal{F}_{\mathbb{K}} = \{\text{unendliche Folgen } (a_j)_{j \in \mathbb{N}} = (a_0, a_1, a_2, \ldots) \text{ mit } a_j \in \mathbb{K}\} = \mathrm{Abb}\,(\mathbb{N}, \mathbb{K})$. Addition und Multiplikation lassen sich auf diesem „Folgenraum“ ähnlich (nämlich „komponentenweise“) wie im Falle $\mathbb{K}^n$ einführen.

4. Die Menge $\mathbb{K}[t] \subset \mathcal{F}_{\mathbb{K}}$ sei definiert durch

$$\mathbb{K}[t] := \{(a_j)_{j \in \mathbb{N}} : \text{ fast alle (d. h. alle bis auf endlich viele) } a_j = 0\}\,.$$

Man überzeugt sich leicht, daß die Addition und Skalarenmultiplikation in $\mathcal{F}_{\mathbb{K}}$ eine $\mathbb{K}$–Vektorraumstruktur auf $\mathbb{K}[t]$ induziert. $\mathbb{K}[t]$ heißt der *Vektorraum der Polynome mit Koeffizienten in* $\mathbb{K}$.

Bemerkung. Man kann sogar eine vernünftige Multiplikation auf $\mathbb{K}[t]$ erklären:

$$\begin{aligned} \mathbb{K}[t] \times \mathbb{K}[t] &\longrightarrow \mathbb{K}[t] \\ ((a_j), (b_k)) &\longmapsto (c_\ell)\,, \end{aligned}$$

wobei $c_\ell = \sum_{j+k=\ell} a_j b_k$. Diese macht $\mathbb{K}[t]$ zu einer sogenannten $\mathbb{K}$–*Algebra.* Die *Unbestimmte* t kann nun einfach mit dem Element $(0, 1, 0, 0, \ldots) \in \mathbb{K}[t]$ identifiziert werden, da bzgl. dieser Verknüpfung jede Folge $(a_0, a_1, a_2, \ldots, a_n, 0, 0, \ldots)$ als das *Polynom*

$$a_0 + a_1\, t + a_2\, t^2 + \cdots + a_n\, t^n$$

interpretiert werden kann.

Einem solchen Polynom $P(t) := a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$ können wir nun die *polynomiale Funktion*

$$\mathbb{K} \ni x \longmapsto P(x) := a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \in \mathbb{K}$$

zuordnen (indem wir also die *Unbestimmte* t durch die *Werte* $x \in \mathbb{K}$ ersetzen). Diese Zuordnung induziert eine Abbildung

$$\mathbb{K}[t] \longrightarrow \mathrm{Abb}(\mathbb{K}, \mathbb{K}) ,$$

die aber i. a. *nicht injektiv* ist. So sind die Polynome $P_1(t) = 1$ und $P_2(t) = t^2 + t + 1$ (über jedem Körper) per definitionem verschieden, liefern aber z. B. über $\mathbb{K} = \mathbb{F}_2 = \{0, 1\}$ *dieselbe* polynomiale Funktion. Diese Problematik besteht z. B. dann nicht, wenn der Körper $\mathbb{K}$ unendlich viele Elemente besitzt. In diesem Fall kann man $\mathbb{K}[t]$ mit seinem Bild in $\mathrm{Abb}(\mathbb{K}, \mathbb{K})$ identifizieren.

Das für uns in dieser Vorlesung wohl wichtigste Beispiel eines Vektorraums ist der *Lösungsraum* eines *homogenen Gleichungssystems*:

$$(*) \qquad \sum_{k=1}^{n} a_{jk}\, x_k = 0 , \quad j = 1, \ldots, m \quad (m \text{ Gleichungen in } n \text{ Unbekannten}) .$$

$L \subset \mathbb{K}^n$ sei die Menge aller Lösungen, also aller n–Tupel $(x_1, \ldots, x_n) \in \mathbb{K}^n$, die $(*)$ erfüllen.

Die wichtigste strukturelle Eigenschaft von L besteht darin, daß die Vektorraum–Verknüpfungen auf $\mathbb{K}^n$ mit L in dem folgenden Sinne „verträglich“ sind: Seien $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ Lösungen, und $\lambda \in \mathbb{K}$ sei beliebig. Dann gilt für alle $j = 1, \ldots, m$:

$$\begin{aligned} \sum_{k=1}^{n} a_{jk}\,(x_k + y_k) &= \sum_{k=1}^{n} a_{jk}\, x_k + \sum_{k=1}^{n} a_{jk}\, y_k = 0 + 0 = 0 , \\ \sum_{k=1}^{n} a_{jk}\,(\lambda\, x_k) &= \sum_{k=1}^{n} \lambda\,(a_{jk}\, x_k) = \lambda \sum_{k=1}^{n} a_{jk}\, x_k = \lambda\, 0 = 0 ; \end{aligned}$$

also

$$x \in L , \quad y \in L , \quad \lambda \in \mathbb{K} \quad \Longrightarrow \quad x + y \in L , \quad \lambda\, x \in L .$$

Diese Bedingung reicht aus, um die *Vektorraum*–Eigenschaft von L zu deduzieren, wie wir anschließend allgemein zeigen werden. Dies wurde auch schon bei den obigen Beispielen implizit verwendet. Wir haben also die

Folgerung 2.1 *Die Lösungen eines homogenen linearen Gleichungssystems* $(*)$ *bilden einen* $\mathbb{K}$*–Vektorraum.*

Definition. Sei V ein $\mathbb{K}$–Vektorraum und $W \subset V$ eine Teilmenge mit

1. $W \neq \emptyset$,
2. $w\,,\, w' \in W \Longrightarrow w + w' \in W$,
3. $\lambda \in \mathbb{K}\,,\, w \in W \Longrightarrow \lambda\, w \in W$.

Dann heißt W ein $\mathbb{K}$–*Untervektorraum* von V .

Bemerkung. $(W, +, \cdot)$ ist dann tatsächlich ein $\mathbb{K}$–Vektorraum an sich (mit demselben Nullvektor 0 wie in V). Also muß notwendig $0 \in W$ gelten. Man weist i. a. diese Eigenschaft anstelle von 1. nach.

Beweis (der Bemerkung). Nach Voraussetzung bilden die Addition und die Skalarenmultiplikation in V die Mengen $W \times W$ bzw. $\mathbb{K} \times W$ nach W ab. Wegen des 1. und 3. Axioms und Teil 1 der Bemerkungen im Anschluß an die Definition eines Vektorraums ist das Nullelement $0 \in V$ in W enthalten: Wähle irgendein Element $w \in W$; dann folgt $0 = 0 \cdot w \in W$. Also existiert ein neutrales Element bzgl. der Addition in W. Wegen 3. und der früheren Bemerkung 4 folgt auch: $w \in W \Longrightarrow -w = (-1)\,w \in W$. Damit ist $(W, +)$ eine abelsche Gruppe. Die restlichen Vektorraum–Axiome sind trivialerweise erfüllt. $\square$

Beispiel. Es sei $I \subset \mathbb{R}$ ein Intervall. Man beweist in der *Analysis*–Vorlesung, daß die Teilmengen

$$\mathrm{Abb}\,(I, \mathbb{R}) \quad \supset \quad \underset{\substack{\text{stetige}\\ \text{Funktionen}}}{\mathcal{C}^0(I, \mathbb{R})} \quad \supset \quad \underset{\substack{\text{differenzierbare}\\ \text{Funktionen}}}{\mathrm{Diff}\,(I, \mathbb{R})} \quad \supset \quad \underset{\substack{\text{stetig}\\ \text{differenzierbare}\\ \text{Funktionen}}}{\mathcal{C}^1(I, \mathbb{R})}$$

die obigen Bedingungen erfüllen. Also besitzen diese Räume kanonische Vektorraumstrukturen.

Bemerkung. Sind $W_j \subset V$ Untervektorräume, $j \in J$, J eine beliebige Indexmenge, so ist der Durchschnitt $\bigcap_{j \in J} W_j$ wieder ein Untervektorraum.

Warnung. Dagegen ist die Vereinigung $W_1 \cup W_2$ i. a. <u>kein</u> Untervektorraum (es sei denn, man hat $W_1 \subset W_2$ oder $W_2 \subset W_1$). - Aber es gilt der folgende Satz.

Satz 2.2 *V sei ein $\mathbb{K}$–Vektorraum, $M \subset V$ eine nichtleere Teilmenge. Dann gibt es einen (bzgl. Inklusion)* kleinsten *Untervektorraum $< M >$ von V, der M enthält*: $M \subset < M >$.

Beweis. Setze $\mathfrak{M} = \{W : W \subset V$ Untervektorraum mit $M \subset W\}$. Es ist $\mathfrak{M} \neq \emptyset$, da $V \in \mathfrak{M}$. Definiere dann

$$< M > := \bigcap_{W \in \mathfrak{M}} W\ .$$

$< M >$ ist ein Untervektorraum, der M umfaßt, und aus $W \supset M$, W Untervektorraum, folgt $< M > \subset W$. $\square$

Bemerkung und Definition. Sind W_1 und W_2 zwei Untervektorräume von V, so gilt

$$< W_1 \cup W_2 > = W_1 + W_2\ ,$$

wobei

$$W_1 + W_2 := \{v \in V : \exists\, w_1 \in W_1\,,\ \exists\, w_2 \in W_2\,,\ \text{ s. d. } v = w_1 + w_2\}\ .$$

Der Untervektorraum $W_1 + W_2$ heißt die *Summe* von W_1 und W_2.

Beweis der Bemerkung. Aus $< W_1 \cup W_2 > \supset W_1$, W_2 folgt, daß mit $w_1 \in W_1$, $w_2 \in W_2$ auch $w_1 + w_2$ in $< W_1 \cup W_2 >$ liegt. Also umfaßt $< W_1 \cup W_2 >$ die rechte Seite $W_1 + W_2$. Offensichtlich ist $W_1 + W_2$ ein Untervektorraum von V, der W_1 und W_2 umfaßt. Also gilt auch die umgekehrte Inklusion. $\square$

Wir wollen die letzte Bemerkung noch wesentlich verallgemeinern.

Definition. Eine *Familie* von Vektoren ist eine Abbildung $\varphi : I \to V$ (I eine nicht notwendig endliche Indexmenge). Man schreibt in diesem Fall auch $(v_\iota)_{\iota \in I}$, wobei $\varphi(\iota) = v_\iota \in V$. Man beachte, daß eine Familie nicht gleichgesetzt werden darf mit der *Menge* $\{v_\iota : \iota \in I\}$. Andererseits kann jede Teilmenge $M \subset V$ als eine durch M selbst indizierte Familie von Vektoren aufgefaßt werden.

Beispiel. Eine Folge $\varphi : \mathbb{N} \to V$ ist nicht das gleiche wie ihre Bildmenge $\varphi(\mathbb{N})$; betrachte z. B. $(1, 2, 0, 1, 2, 0, 1, 2, 0, \ldots)$ als Folge in $\mathbb{R}$.

Definition. Es sei $(v_\iota)_{\iota\in I}$ eine Familie von Vektoren in V. Dann heißt der kleinste Untervektorraum von V, der alle Vektoren v_ι, $\iota \in I$, enthält, der *Spann* oder das *lineare Erzeugnis* der Familie; in Zeichen:

$$\operatorname{span}(v_\iota)_{\iota\in I} := < \{v_\iota : \iota \in I\} > \quad .$$

Der Beweis des folgenden Lemmas ist eine leichte Übungsaufgabe (man kopiere den Beweis für die Identität $< W_1 \cup W_2 > = W_1 + W_2$). Manchmal wird die angegebene Charakterisierung als Definition verwendet.

Lemma 2.3 *Für jede Familie* $(v_\iota)_{\iota\in I}$ *gilt :*

$$\operatorname{span}(v_\iota)_{\iota\in I} = \{v = \sum_{\iota\in I_0} \lambda_\iota v_\iota : \ I_0 \text{ eine (von } v \text{ abhängige) } \textit{endliche} \text{ Teilmenge von } I\} \; .$$

Bemerkung. Aufgrund der obigen Beschreibung ist auch die folgende Bezeichnung üblich:

$$\operatorname{span}(v_\iota)_{\iota\in I} = \sum_{\iota\in I} \mathbb{K}\, v_\iota \; .$$

Die vielleicht wichtigste Definition der *Linearen Algebra* ist nun die folgende.

Definition. Eine Familie $(v_\iota)_{\iota\in I}$ von Vektoren heißt *linear unabhängig*, falls für alle *endlichen* Teilmengen $I_0 \subset I$ gilt:

$$\sum_{\iota\in I_0} \lambda_\iota v_\iota = 0 \; , \quad \lambda_\iota \in \mathbb{K} \implies \lambda_\iota = 0 \text{ für alle } \iota \in I_0 \; .$$

(M. a. W.: „Der Nullvektor läßt sich nur *trivial* mit Hilfe der $(v_\iota)_{\iota\in I}$ darstellen“).

Bemerkung. $(v_j)_{j=1,\ldots,n}$ sei eine *endliche* Familie. Dann ist $(v_j)_{j=1,\ldots,n}$ genau dann linear unabhängig, falls

$$\sum_{j}^{n} \lambda_j v_j = 0 \implies \lambda_j = 0 \text{ für alle } j = 1, \ldots, n \; .$$

Definition. Die Familie $(v_\iota)_{\iota\in I}$ heißt *linear abhängig*, wenn sie nicht linear unabhängig ist, d. h. wenn es eine endliche Teilmenge $I_0 \subset I$ und Koeffizienten λ_ι, $\iota \in I_0$, die nicht alle gleich Null sind, gibt, s. d.

$$\sum_{\iota\in I_0} \lambda_\iota v_\iota = 0 \; .$$

Leichte Folgerungen aus diesen Definitionen liegen auf der Hand. So ist natürlich jede *Teilfamilie* einer unabhängigen Familie wieder unabhängig, und jede *Oberfamilie* einer abhängigen Familie ebenfalls abhängig. - Zur Erläuterung werde ich eine weitere Charakterisierung geben.

Satz 2.4 *Die Familie* $(v_\iota)_{\iota\in I}$ *ist genau dann linear unabhängig, falls*

$$\mathit{span}(v_\iota)_{\iota\in I'} \subsetneqq \mathit{span}(v_\iota)_{\iota\in I} \quad \textit{für alle echten Teilmengen} \quad I' \subsetneqq I \; ,$$

(und damit genau dann linear abhängig, wenn es eine Teilmenge $I' \subsetneqq I$ *gibt, s. d.* $\mathit{span}(v_\iota)_{\iota\in I} = \mathit{span}(v_\iota)_{\iota\in I'}$ *).*

Beweis. Wir zeigen die äquivalente Aussage in Klammern (natürlich ist stets $\operatorname{span}(v_\iota)_{\iota\in I'} \subset \operatorname{span}(v_\iota)_{\iota\in I}$ für $I' \subset I$).

1. Es sei $\operatorname{span}(v_\iota)_{\iota\in I'} = \operatorname{span}(v_\iota)_{\iota\in I}$, $I' \subsetneqq I$, und sei $\iota_0 \in I \setminus I'$. Wegen der vorausgesetzten Gleichheit folgt

$$v_{\iota_0} = \sum_{\iota\in J} \lambda_\iota v_\iota \quad \text{mit} \quad J \subset I' \quad \text{endlich} \; .$$

Es sei weiter $I_0 = J \cup \{\iota_0\}$. Dann ergibt sich

$$1 \cdot v_{\iota_0} - \sum_{\iota \in J} \lambda_\iota \, v_\iota = 0 \, ,$$

d. h. die Familie $(v_\iota)_{\iota \in I_0}$ ist linear abhängig und damit auch $(v_\iota)_{\iota \in I}$.

2. Sei umgekehrt $(v_\iota)_{\iota \in I}$ linear abhängig. Dann existiert eine endliche Teilmenge $J_0 \subset I$, s. d.

$$\sum_{\iota \in J_0} c_\iota \, v_\iota = 0 \, , \quad \text{nicht alle } c_\iota = 0 \, .$$

Sei $c_{\iota_0} \neq 0$. Dann folgt $v_{\iota_0} = \sum_{\iota \in J_0'} b_\iota \, v_\iota$, $J_0' = J_0 \setminus \{\iota_0\}$ mit $b_\iota = -c_{\iota_0}^{-1} \, c_\iota$.

Sei nun $v \in \mathrm{span}\,(v_\iota)_{\iota \in I}$, also $v = \sum_{\iota \in I_0} a_\iota v_\iota$, $I_0 \subset I$ endlich. Wir unterscheiden zwei Fälle: Ist $\iota_0 \notin I_0$, so folgt $I_0 \subset I \setminus \{\iota_0\}$. Ist $\iota_0 \in I_0$, so ergibt sich

$$\begin{aligned} v &= a_{\iota_0} \, v_{\iota_0} + \sum_{\iota \in I_0'} a_\iota \, v_\iota \, , \qquad I_0' = I_0 \setminus \{\iota_0\} \\ &= a_{\iota_0} \sum_{\iota \in J_0'} b_\iota \, v_\iota + \sum_{\iota \in I_0'} a_\iota \, v_\iota = \sum_{\iota \in K_0'} c_\iota \, v_\iota \, , \end{aligned}$$

mit $K_0' = I_0' \cup J_0' \subset I \setminus \{\iota_0\}$. Also ist in beiden Fällen

$$\mathrm{span}\,(v_\iota)_{\iota \in I} \subset \mathrm{span}\,(v_\iota)_{\iota \in I'} \quad \text{mit} \quad I' = I \setminus \{\iota_0\} \, .$$

□

Satz 2.5 *Die Familie $(v_\iota)_{\iota \in I}$ ist genau dann linear unabhängig, wenn sich jeder Vektor $v \in$ span$(v_\iota)_{\iota \in I}$ auf* nur eine *Weise aus Vektoren in $\{v_\iota : \iota \in I\}$ linear kombinieren läßt.*

Folgerung 2.6 *Enthält die Familie $(v_\iota)_{\iota \in I}$ zwei gleiche Vektoren oder den Nullvektor, so ist sie linear abhängig.*

Beweis von Satz 5. „$\Longrightarrow$" Aus

$$v = \sum_{\iota \in I_1} \lambda_\iota \, v_\iota = \sum_{\iota \in I_2} \mu_\iota \, v_\iota$$

folgt (wenn man $\lambda_\iota = 0$ setzt für $\iota \in I_2 \setminus I_1$ und vice versa für μ_ι): $\sum_{\iota \in I_1 \cup I_2} (\lambda_\iota - \mu_\iota) \, v_\iota = 0$ und damit $\lambda_\iota = \mu_\iota$ für alle $\iota \in I := I_1 \cup I_2$. Also hat man Eindeutigkeit.

„$\Longleftarrow$" Wir zeigen stattdessen die „Transposition" $\neg \implies \neg$. Sei also die Familie $(v_\iota)_{\iota \in I}$ linear abhängig. Nach dem vorhergehenden Satz existiert dann mindestens ein v mit verschiedenen Darstellungen. □

Wir haben also gesehen, daß linear unabhängige Familien $(v_\iota)_{\iota \in I}$ ihren Spann $\mathrm{span}\,(v_\iota)_{\iota \in I}$ auf minimale Weise erzeugen, und dies führt dazu, daß sich jeder Vektor im Erzeugnis nur auf eine Weise als Linearkombination der Erzeuger darstellen läßt. Dies ist die Eigenschaft, die wir von sogenannten *Basen* erwarten.

Definition. i) Die Familie $(v_\iota)_{\iota \in I}$, $v_\iota \in V$, heißt ein *Erzeugendensystem* von V, falls $V = \mathrm{span}\,(v_\iota)_{\iota \in I}$, d. h. mit anderen Worten, wenn es zu jedem Vektor $v \in V$ eine endliche Teilmenge $I_0 \subset I$ und Koeffizienten $\lambda_\iota \in \mathbb{K}$, $\iota \in I_0$, gibt, s. d.

$$v = \sum_{\iota \in I_0} \lambda_\iota \, v_\iota \, .$$

ii) $(v_\iota)_{\iota \in I}$ heißt eine *Basis* von V, falls $(v_\iota)_{\iota \in I}$ ein linear unabhängiges Erzeugendensystem von V ist.

Beispiele. 1. Für $\mathbb{K}^n$ hat man die *Standardbasis* bestehend aus den *Einheitsvektoren*

$$e_j \,=\, (\delta_{ij})_{i=1,\ldots,n}\ , \quad \delta_{ij} \text{ das } \textit{Kroneckersymbol}\ ,$$

also ausgeschrieben

$$e_1 \,=\, (1,\,0,\,0,\ldots,0)\,,\ e_2 \,=\, (0,\,1,\,0,\ldots,0)\,,\ \ldots\,,\ e_n \,=\, (0,\,0,\ldots,1)\ .$$

2. $\mathbb{K}[\,t\,]$ sei der Vektorraum der Polynome über $\mathbb{K}$. Dieser hat offensichtlich die abzählbar unendliche Basis $t^0 \,=\, 1\,,\ t\,,\ t^2,\ldots$.

3. Der triviale Vektorraum $\{0\}$ bekommt die Basis $\emptyset$ zugesprochen!

Die folgende Charakterisierung ist mehr oder minder nur eine Zusammenfassung der bisherigen Ergebnisse. Wegen ihrer entscheidenden Bedeutung für alles Folgende werden wir die einzelnen Argumente, wenn nötig, noch einmal wiederholen.

Satz 2.7 *Es sei $V \neq \{0\}$ ein $\mathbb{K}$–Vektorraum und $(v_\iota)_{\iota\in I}$ eine Familie von Vektoren $v_\iota \in V$. Dann sind die folgenden Aussagen äquivalent* :

i) $(v_\iota)_{\iota\in I}$ *ist eine Basis von* V ;

ii) $(v_\iota)_{\iota\in I}$ *ist ein* unverkürzbares Erzeugendensystem *von* V *(d. h. ein Erzeugendensystem, s. d. $(v_\iota)_{\iota\in I'}$ den Vektorraum V nicht mehr erzeugt, wenn $I' \subsetneqq I$)* ;

iii) $(v_\iota)_{\iota\in I}$ *ist eine* unverlängerbare Familie von linear unabhängigen Vektoren *in* V *(d. h. jede andere Familie $(v_\iota)_{\iota\in I'}$ mit $I \subsetneqq I'$ ist linear abhängig)* ;

iv) *Jeder Vektor $v \in V$ schreibt sich eindeutig als Linearkombination der Vektoren v_ι , $\iota \in I$* .

Beweis durch „Rundlauf“:

$$\begin{array}{ccc} \text{i)} & \Longrightarrow & \text{ii)} \\ \Uparrow & & \Downarrow \\ \text{iv)} & \Longleftarrow & \text{iii)} \end{array}$$

i) $\Longrightarrow$ ii) ist gerade die Aussage von Satz 4.

ii) $\Longrightarrow$ iii). Wieder wegen Satz 4 liefert die Unverkürzbarkeit, daß das Erzeugendensystem $(v_\iota)_{\iota\in I}$ eine linear unabhängige Familie ist. Diese läßt sich nicht verlängern: Wäre nämlich $(v_\iota)_{\iota\in J}$, $I \subsetneqq J$, eine unabhängige Oberfamilie, so würde wieder mit Satz 4 der Widerspruch $V \,=\, \operatorname{span}(v_\iota)_{\iota\in I} \subsetneqq \operatorname{span}(v_j)_{j\in J} \subset V$ folgen.

iii) $\Longrightarrow$ iv). Nur zu zeigen ist, daß $(v_\iota)_{\iota\in I}$ ein Erzeugendensystem von V bildet (denn die Eindeutigkeitsaussage folgt aus Satz 5). Angenommen, dies wäre nicht der Fall. Dann gäbe es einen Vektor v, der sich nicht darstellen läßt als Linearkombination der v_ι. Dann folgt aber aus einer Gleichung

$$\lambda\, v \,+\, \lambda_1\, v_{\iota_1} \,+\cdots+\, \lambda_r\, v_{\iota_r} \,=\, 0\ , \quad \iota_1,\ldots,\iota_r \in I\ ,$$

sofort $\lambda \,=\, 0$ (denn sonst ließe sich v doch darstellen) und wegen der Unabhängigkeit der v_ι auch $\lambda_1 \,=\cdots=\, \lambda_r \,=\, 0$. Also könnte man das gegebene System durch Hinzunahme von v doch zu einem linear unabhängigen System vergrößern.

vi) $\Longrightarrow$ i). Die $(v_\iota)_{\iota\in I}$ erzeugen nach Voraussetzung den Vektorraum V und sind linear unabhängig, da insbesondere der Nullvektor 0 nur auf *eine* Weise, nämlich trivial, dargestellt werden kann. $\square$

Wir wollen uns noch mit der Frage beschäftigen, ob jeder Vektorraum eine Basis besitzt. Diese Frage ist absolut nicht trivial und führt zu tiefliegenden Problemen der Mengentheorie. Man kann sich dies klar machen an dem Beispiel der reellen Zahlen, aufgefaßt als Vektorraum über dem Unterkörper

$\mathbb{Q}$: Da die Menge der rationalen Zahlen abzählbar ist, die Menge der reellen Zahlen jedoch nicht, schließt man unmittelbar, daß eine Basis in diesem Fall notwendig *überabzählbar* sein muß! In der Tat existieren stets Basen (die man HAMELsche Basen nennt); der Nachweis ihrer Existenz benötigt aber nichtkonstruktive Methoden der Mengentheorie wie das sogenannte ZORNsche *Lemma*, das mit dem *Auswahlaxiom* äquivalent ist. Das Hauptergebnis in diesem Zusammenhang ist der folgende

Satz 2.8 (Basisauswahlsatz) *Aus jedem Erzeugendensystem eines Vektorraums läßt sich eine Basis auswählen.*

Da jeder Vektorraum selbstverständlich Erzeugendensysteme besitzt (z. B. sich selbst als indizierte Menge), so folgt hieraus unmittelbar

Satz 2.9 *Jeder* $\mathbb{K}$*–Vektorraum besitzt eine Basis* $(v_\iota)_{\iota\in I}$.

Man kann sogar zeigen, daß je zwei Basen $(v_\iota)_{\iota\in I}$, $(w_\kappa)_{\kappa\in K}$ gleiche Kardinalität besitzen, d. h. gleichmächtig sind: Es gibt eine Bijektion $I \xrightarrow{\sim} K$. Hierzu benötigt man das Bernsteinsche Lemma:

$$\operatorname{card} M \le \operatorname{card} N \ \text{ und } \ \operatorname{card} N \le \operatorname{card} M \quad \Longrightarrow \quad \operatorname{card} M = \operatorname{card} N$$

und den sogenannten

Satz 2.10 (Basisergänzungssatz) *Jede linear unabhängige Familie läßt sich durch Hinzunahme von Vektoren aus einem beliebig vorgegebenen Erzeugendensystem zu einer Basis ergänzen.*

Dieser Satz ist übrigens nicht stärker als der Auswahlsatz, wie wir später mit Hilfe der Theorie linearer Abbildungen zeigen werden.

Um den mengentheoretischen Schwierigkeiten aus dem Weg zu gehen, beschränken wir uns in Bezug auf den Auswahlsatz auf den „endlich–dimensionalen" Fall. Im ∞–dimensionalen sind ohnehin andere (topologische, analytische) Methoden interessanter, bei denen endliche Linearkombinationen durch konvergente Reihen ersetzt werden müssen (Banachräume, Hilberträume, etc.).

Definition. Ein $\mathbb{K}$–Vektorraum heißt *endlich erzeugt*, wenn $V = \operatorname{span}(v_\iota)_{\iota\in I}$ mit einer *endlichen* Indexmenge I .

Satz 2.11 (Basisauswahlsatz, endlich erzeugter Fall) *Ist* V *ein endlich erzeugter* $\mathbb{K}$*–Vektorraum, so läßt sich aus jedem (nicht notwendig endlichen) Erzeugendensystem von* V *eine endliche Basis auswählen. Insbesondere sind alle Basen endlich.*

Beweis . Es sei $v_1, \ldots, v_m$ ein endliches Erzeugendensystem. Ist dieses nicht verkürzbar, so handelt es sich schon um eine Basis. Im anderen Fall lasse man überflüssige Elemente weg und gelangt dadurch nach endlich vielen Schritten zu einer endlichen Basis.

Ist ferner $(w_\kappa)_{\kappa\in K}$ ein beliebiges Erzeugendensystem (z. B. eine Basis), so reicht zu zeigen, daß dieses zu einem *endlichen Erzeugendensystem* verkürzt werden kann. Dazu schreiben wir mit dem endlichen System $v_1, \ldots, v_m$:

$$v_j = \sum_{\kappa\in K_j} \lambda_\kappa^{(j)} w_\kappa \ , \quad j = 1, \ldots, m \ ,$$

wobei die K_j endliche Teilmengen von K sind. Da die v_j den Vektorraum V erzeugen, ist klar, daß er auch von den endlich vielen Vektoren w_κ , $\kappa \in \bigcup_{j=1}^m K_j$, erzeugt wird. □

Warnung. Dieses Argument liefert keinen Beweis für die Existenz einer Basis, wenn V *nicht* endlich erzeugt ist.

Wir kommen auf den *Basisergänzungssatz* und den Beweis der Gleichmächtigkeit aller Basen im endlich erzeugten Fall später zurück. Im nicht endlich erzeugten Fall werden wir an unbewiesenen Aussagen nur die Existenz von Basen benutzen.

3 Lineare Abbildungen

Wir betrachten zunächst noch einmal lineare Gleichungssysteme in voller Allgemeinheit:

$$(*) \qquad \sum_{k=1}^{n} a_{jk}\, x_k \,=\, b_j \;, \quad j \,=\, 1, \ldots, m \;.$$

Übereinander ausgeschrieben lautet ein solches System

$$(*) \qquad \left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \;. \end{array} \right.$$

Gegeben sind uns also die Elemente $a_{jk} \in \mathbb{K}$, $j \,=\, 1, \ldots, m$, $k \,=\, 1, \ldots, n$, und $b_j \in \mathbb{K}$, $j \,=\, 1, \ldots, m$, die wir aufgrund ihrer Herkunft von dem obigen System in der Form

$$A \,=\, \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad , \qquad b \,=\, \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

zusammenfassen. Wir nennen ein Schema A von dem obigen Typ eine *Matrix*, und zwar genauer eine *Matrix mit* m *Zeilen und* n *Spalten (oder kurz: eine* $m \times n$*–Matrix) mit Einträgen* a_{jk} *in dem Grundkörper* $\mathbb{K}$. Wir verwenden auch abgekürzte Symbole wie

$$A \,=\, (a_{jk})_{\substack{j=1,\ldots,m \\ k=1,\ldots,n}} \quad \text{oder} \quad A \,=\, (a_{jk})_{j=1,\ldots,m,\, k=1,\ldots,n}$$

oder sogar $A \,=\, (a_{jk})$, wenn aus dem Kontext heraus klar ist, welche Zahlen der *Zeilenindex* j und welche der *Spaltenindex* k durchläuft. Formal ist A aufzufassen als eine Abbildung $\{1, \ldots, m\} \times \{1, \ldots, n\} \longrightarrow \mathbb{K}$, deren Werte $A\,((j,\, k))$ man halt als a_{jk} schreibt (besser und bei konkreten Indizes zur Vermeidung von Mißverständnissen angebracht ist $a_{j,k}$, vgl. z. B. $a_{12,3}$, $a_{1,23}$ und a_{123}) und in der obigen Matrixform anordnet. In dieser Interpretation wird die Menge der $m \times n$–Matrizen, die wir im folgenden stets mit $M\,(m \times n,\, \mathbb{K})$ bezeichnen werden:

$$M\,(m \times n,\, \mathbb{K}) \,=\, \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \;:\; a_{jk} \in \mathbb{K} \right\} \,,$$

zu einem $\mathbb{K}$–Vektorraum, in dem die Vektorraumoperationen für

$$A \,=\, (a_{jk})_{\substack{j=1,\ldots,m \\ k=1,\ldots,n}} \;, \quad A' \,=\, (a'_{jk})_{\substack{j=1,\ldots,m \\ k=1,\ldots,n}}$$

komponentenweise durch

$$A \,+\, A' \,=\, (a_{jk} \,+\, a'_{jk}) \;, \quad \lambda\, A \,=\, (\lambda\, a_{jk})$$

definiert sind.

Die rechte Seite b ist in diesem Sinne eine $m \times 1$–Matrix; wir nennen b auch einen *Spaltenvektor*; bei der Interpretation von $\mathrm{Abb}\,(\{1, \ldots, n\},\, \mathbb{K})$ als Menge $\mathbb{K}^n$ der *Zeilenvektoren* mit Werten in $\mathbb{K}$ hat uns natürlich nur die Bequemlichkeit der Schreibweise geleitet. Es kann uns niemand daran hindern, den Vektorraum $\mathbb{K}^n$ auch als Raum der Spaltenvektoren aufzufassen, was wir im folgenden stets tun werden, sofern nicht ausdrücklich das Gegenteil gesagt wird. Schreibt man nun entsprechend auch noch für den *Vektor der Unbekannten* formal einen Spaltenvektor

$$x \,=\, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

und definiert das Produkt Ax durch

$$Ax := A \cdot x := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}\,x_1 & + & \cdots & + & a_{1n}\,x_n \\ & & \vdots & & \\ a_{m1}\,x_1 & + & \cdots & + & a_{mn}\,x_n \end{pmatrix},$$

so wird $(*)$ einfach zu der Gleichung

$$A\,x = b\ .$$

(Daß die obige Definition der Matrizenmultiplikation geradezu zwangsläufig ist, werden wir später noch in diesem Kapitel beweisen).

Die letztgenannte Schreibweise von linearen Gleichungssystemen macht den *Abbildungscharakter* der Aufgabenstellung wesentlich deutlicher: Bei gegebener Matrix A in $M\,(m \times n,\, \mathbb{K})$ wird durch

$$\mathbb{K}^n \ni x \longmapsto A\,x \in \mathbb{K}^m$$

eine Abbildung von $\mathbb{K}^n$ nach $\mathbb{K}^m$ definiert, die wir stets mit

$$F_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$$

bezeichnen. Der *Lösungsraum* $L_{A,b}$ von $(*)$ ist dann schlicht die Gesamtheit aller $x \in \mathbb{K}^n$ mit $F_A(x) = b$; oder mengentheoretisch ausgedrückt:

$$L_{A,b} = F_A^{-1}\,(b)\ .$$

Wenn das Gleichungssystem $(*)$ homogen ist, d. h. wenn $b = 0$, so haben wir im vorigen Kapitel schon bemerkt, daß der Lösungsraum $L_A := L_{A,0} \subset \mathbb{K}^n$ einen Untervektorraum bildet. Der Grund hierfür besteht in der *Linearität* der Abbildung F_A:

$$\begin{cases} F_A(x + x') = F_A(x) + F_A(x') & \text{für alle} \quad x\,,\ x' \in \mathbb{K}^n\ , \\ F_A(\lambda\,x) = \lambda\,F_A\,(x) & \text{für alle} \quad x \in \mathbb{K}^n\,,\ \lambda \in \mathbb{K}\ . \end{cases}$$

Diese Eigenschaft erheben wir nun zu einer allgemeinen Definition:

Definition. Eine Abbildung $F : V \to W$ zwischen $\mathbb{K}$–Vektorräumen V, W heißt ($\mathbb{K}$–)*linear*, wenn

$$\begin{aligned} F\,(v + v') &= F\,(v) + F\,(v') && \text{für alle} \quad v\,,\ v' \in V\ , \\ F\,(\lambda\,v) &= \lambda\,F\,(v) && \text{für alle} \quad v \in V\,,\ \lambda \in \mathbb{K}\ . \end{aligned}$$

Beispiele. 1. $I \subset \mathbb{R}$ sei ein Intervall, $x_0 \in I$ ein fest gewählter Punkt. Dann ist die Abbildung

$$\begin{cases} \mathrm{Abb}\,(I,\,\mathbb{R}) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & f\,(x_0) \end{cases}$$

linear.

2. Wie oben sei $I \subset \mathbb{R}$ ein Intervall, x_0 sei fest in I, und $\mathrm{Diff}\,(I,\,\mathbb{R})$ sei der Vektorraum der *differenzierbaren* Funktionen auf I. Dann sind die Abbildungen

$$\begin{cases} \mathrm{Diff}\,(I,\,\mathbb{R}) & \longrightarrow & \mathrm{Abb}\,(I,\,\mathbb{R}) \\ f & \longmapsto & f' \end{cases} \quad \text{und} \quad \begin{cases} \mathrm{Diff}\,(I,\,\mathbb{R}) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & f'(x_0) \end{cases}$$

linear.

3. Die Abbildung
$$\lim : \mathrm{Konv}\,(\mathbb{N},\, \mathbb{R}) \longrightarrow \mathbb{R}$$
ist linear, wobei $\mathrm{Konv}\,(\mathbb{N},\, \mathbb{R})$ den Vektorraum der *konvergenten* Folgen $(x_j)_{j\in\mathbb{N}}$ in $\mathbb{R}$ bezeichnet:
$$\mathrm{Konv}\,(\mathbb{N},\, \mathbb{R}) \ni (x_0,\, x_1, \ldots) \longmapsto \lim_{j\to\infty} x_j \in \mathbb{R}\,.$$

4. Es bezeichne $\mathrm{Int}\,(I,\, \mathbb{R})$ den Vektorraum der (Riemann– oder Lebesgue–) *integrierbaren* Funktionen auf dem Intervall $I \subset \mathbb{R}$. Dann ist die Abbildung
$$\begin{cases} \mathrm{Int}\,(I,\, \mathbb{R}) \longrightarrow & \mathbb{R} \\ \quad f \qquad\longmapsto & \displaystyle\int_I f\,(x)\,dx \end{cases}$$
linear.

Daß lineare Abbildungen automatisch weitere charakteristische Eigenschaften der betroffenen Vektorräume respektieren, dürfte nicht weiter verwundern. Es gilt zum Beispiel:

Lemma 3.1 *Es sei* $F : V \to W$ *eine lineare Abbildung. Dann gelten die folgenden Aussagen.*

1. $F\,(0) = 0$.
2. $F\,(-v) = -F\,(v)$.
3. *Ist* V' *ein Untervektorraum von* V *, dann ist* $F\,(V') \subset W$ *ein Untervektorraum von* W .
4. *Ist* W' *ein Untervektorraum von* W *, so ist auch* $F^{-1}(W') \subset V$ *ein Untervektorraum von* V .

Beweis. 1. $F\,(0) = F\,(0 \cdot 0) = 0 \cdot F\,(0) = 0$.

2. $F\,(v) + F\,(-v) = F\,(v + (-v)) = F\,(0) = 0$.

3. w', $w'' \in F\,(V') \Longrightarrow w' = F\,(v')$, $w'' = F\,(v'')$, v', $v'' \in V' \Longrightarrow v' + v'' \in V'$ und $w' + w'' = F\,(v') + F\,(v'') = F\,(v' + v'') \in F\,(V')$. Entsprechend ist $\lambda\, w' \in F\,(V')$, $0 \in F\,(V')$.

4. beweist man genauso wie 3. □

Folgerung und Definition. a) Wählt man speziell $W' = \{0\}$ als Untervektorraum von W, so folgt mit 4. aus dem obigen Lemma, daß die Menge
$$\ker F := F^{-1}(0) = \{v \in V : F\,(v) = 0\}$$
ein Untervektorraum von V ist. Man nennt ihn den *Kern* von F .

b) Wählt man in 3. den Untervektorraum $V' = V$, so nennt man den Untervektorraum
$$\mathrm{im}\, F := F\,(V) = \{w \in W : \text{es gibt ein } v \in V \text{ mit } F\,(v) = w\}$$
das *Bild* von F .

Bemerkung. Ist insbesondere F_A die lineare Abbildung $\mathbb{K}^n \to \mathbb{K}^m$ zur Matrix A, so ist $\ker F_A$ gerade der Lösungsraum L_A des homogenen Systems $A\,x = 0$.

Die Begriffe Kern und Bild von linearen Abbildungen sind mit Injektivität und Surjektivität durch die beiden folgenden Sätze verbunden.

Satz 3.2 *$F : V \to W$ sei linear. Dann sind äquivalent :*

i) *F ist injektiv.*

ii) $\ker F = \{0\}$.

iii) *Ist $(v_\iota)_{\iota \in I}$ eine beliebige linear unabhängige Familie in V , so ist auch die Familie $(F(v_\iota))_{\iota \in I}$ linear unabhängig in W .*

iv) *Ist $(v_1, \ldots, v_r)$ eine beliebige endliche linear unabhängige Familie in V , so ist auch die Familie $(F(v_1), \ldots, F(v_r))$ linear unabhängig in W .*

Beweis. i) $\Longrightarrow$ ii). Aus $F(v) = 0 = F(0)$ und der Injektivität von F folgt $v = 0$.

ii) $\Longrightarrow$ iii). Die Relation

$$0 = \sum_{\iota \in I_0} \lambda_\iota F(v_\iota) = F\Big(\sum_{\iota \in I_0} \lambda_\iota v_\iota\Big) , \quad I_0 \subset I \quad \text{endlich} ,$$

impliziert $\sum \lambda_\iota v_\iota \in \ker F = \{0\}$, also $\lambda_\iota = 0$ für alle $\iota \in I_0$.

iii) $\Longrightarrow$ iv) ist trivial.

iv) $\Longrightarrow$ i). $F(v_1) = F(v_2) \Longrightarrow F(v) = 0$, wobei $v = v_2 - v_1$. Also ist $(F(v))$ eine linear abhängige Familie in W , sodaß nach Voraussetzung auch das System (v) linear abhängig in V sein muß, was aber nichts anderes als $v_2 - v_1 = v = 0$, d. h. $v_1 = v_2$ bedeutet. □

Man beachte, daß die Umkehrung von iii) und iv) wegen Lemma 1 für *beliebige* lineare Abbildungen richtig ist. Man kann daher in beiden Aussagen die Implikation durch eine Äquivalenz ersetzen.

Satz 3.3 *$F : V \to W$ sei linear. Dann sind äquivalent :*

i) *F ist surjektiv.*

ii) $\operatorname{im} F = W$.

iii) *Für jedes Erzeugendensystem $(v_\iota)_{\iota \in I}$ von V ist $(F(v_\iota))_{\iota \in I}$ ein Erzeugendensystem von W .*

iv) *Es existiert ein Erzeugendensystem $(v_\iota)_{\iota \in I}$ von V , s. d. $(F(v_\iota))_{\iota \in I}$ ein Erzeugendensystem von W ist.*

Beweis. i) $\Longleftrightarrow$ ii) und iii) $\Longrightarrow$ iv) sind trivial aufgrund der Definition bzw. der Tatsache, daß jeder Vektorraum Erzeugendensysteme besitzt.

ii) $\Longrightarrow$ iii). Es sei $w \in W = \operatorname{im} F$, also $w = F(v)$ und $v = \sum_{\iota \in I_0} \lambda_\iota v_\iota$. Dann folgt

$$w = F(v) = F\Big(\sum_{\iota \in I_0} \lambda_\iota v_\iota\Big) = \sum_{\iota \in I_0} \lambda_\iota F(v_\iota) .$$

iv) $\Longrightarrow$ ii) Für beliebiges $w \in W$ gibt es eine Darstellung

$$w = \sum_{\iota \in I_0} \lambda_\iota F(v_\iota) = F\Big(\sum_{\iota \in I_0} \lambda_\iota v_\iota\Big) \in \operatorname{im} F . \qquad □$$

Insbesondere ergibt sich aus dem vorstehenden Satz:

Folgerung 3.4 *Ist V endlich erzeugt, so auch das Bild* $\operatorname{im} F$ *unter einer linearen Abbildung $F : V \to W$.*

Daß auch Kerne in diesem Fall oder ganz allgemein Untervektorräume von endlich erzeugten Vektorräumen wieder endlich erzeugt sind, werden wir im nächsten Kapitel beweisen, auch wenn dies an dieser Stelle schon ohne Schwierigkeiten möglich wäre.

Definition. Lineare Abbildungen werden auch als *Homomorphismen* (genauer als *Vektorraum–Homomorphismen*) bezeichnet. Für *injektive* Homomorphismen ist der Begriff *Monomorphismus*, für *surjektive* der Begriff *Epimorphismus* gebräuchlich. Schließlich spricht man von einem *Isomorphismus*, falls die gegebene lineare Abbildung *bijektiv* ist. Man nennt zwei Vektorräume *isomorph*, wenn es einen Isomorphismus $F : V \to W$ gibt.

In diesem Zusammenhang ist der folgende Satz von Bedeutung. Er beinhaltet insbesondere, daß Isomorphie nicht von der Reihenfolge der Vektorräume V, W abhängig ist.

Satz 3.5 *$F : V \to W$ sei linear und bijektiv, also ein Isomorphismus. Dann ist auch die Umkehrabbildung F^{-1} linear (und bijektiv).*

Bemerkung. Erst diese Aussage rechtfertigt den Begriff des *Isomorphismus*, der stets beinhaltet, daß mit einer bijektiven Abbildung auch ihr Inverses die gleichen strukturellen Merkmale aufweist.

Beweis. a) $w_1 = F(v_1)\,,\ w_2 = F(v_2) \Longrightarrow F(v_1 + v_2) = w_1 + w_2 \Longrightarrow F^{-1}(w_1 + w_2) = v_1 + v_2 = F^{-1}(w_1) + F^{-1}(w_2)\,.$

b) $w = F(v) \Longrightarrow F(\lambda v) = \lambda F(v) = \lambda w \Longrightarrow F^{-1}(\lambda w) = \lambda v = \lambda F^{-1}(w)\,.$ □

Selbstverständlich haben die Sätze 2 und 3 auch ein Analogon für den bijektiven Fall.

Satz 3.6 *$F : V \to W$ sei linear. Dann sind äquivalent :*

i) *F ist ein Isomorphismus.*

ii) *Für jede Basis $(v_\iota)_{\iota\in I}$ von V ist $(F(v_\iota))_{\iota\in I}$ eine Basis von W.*

iii) *Es existiert eine Basis $(v_\iota)_{\iota\in I}$ von V, s. d. $(F(v_\iota))_{\iota\in I}$ eine Basis von W bildet.*

Beweis. i) $\Longrightarrow$ ii) folgt unmittelbar aus den Sätzen 2 und 3; ii) $\Longrightarrow$ iii) ist trivial (vorausgesetzt, wir hätten allgemein die Existenz von Basen in V bewiesen). Bei iii) $\Longrightarrow$ i) ist ebenfalls nach Satz 3 die Surjektivität von F klar. Ist schließlich $v = \sum_{\iota\in I_0} \lambda_\iota v_\iota \in \ker F$, so impliziert

$$\sum_{\iota\in I_0} \lambda_\iota F(v_\iota) = F\Big(\sum_{\iota\in I_0} \lambda_\iota v_\iota\Big) = F(v) = 0$$

sofort $\lambda_\iota = 0$ für alle $\iota \in I_0$, d. h. $v = 0$. □

Wichtig ist noch der folgende Satz, dessen Beweis wir erst später durchführen können, da uns der *Dimensionsbegriff* noch fehlt.

Satz 3.7 *Es seien V und W endlich erzeugte $\mathbb{K}$–Vektorräume der* gleichen *Dimension, und $F : V \to W$ sei ein Homomorphismus. Dann sind die folgenden Aussagen äquivalent :*

i) *F ist injektiv, also ein Monomorphismus.*

ii) *F ist surjektiv, also ein Epimorphismus.*

iii) *F ist bijektiv, also ein Isomorphismus.*

Definition. Die Menge aller Homomorphismen $F : V \longrightarrow W$ bezeichnet man oft mit dem Symbol

$$\mathrm{Hom}_{\mathbb{K}}(V, W) \quad \text{oder} \quad \mathrm{Hom}\,(V, W)\,.$$

Ist $V = W$, so setzt man auch

$$\mathrm{End}_{\mathbb{K}}(V) = \mathrm{Hom}_{\mathbb{K}}(V, V)$$

und läßt auch hier gegebenenfalls den Index $\mathbb{K}$ weg, wenn aus dem Zusammenhang der Grundkörper eindeutig hervorgeht. Man nennt die linearen Selbstabbildungen $F : V \to V$ auch *Endomorphismen*. Schließlich bezeichnet man die bijektiven Endomorphismen auch als *Automorphismen* (oder auch weiter als *Isomorphismen*) und verwendet für ihre Gesamtheit das Symbol

$$\mathrm{Aut}_{\mathbb{K}}(V)\,.$$

Es ist klar, daß man je zwei Homomorphismen $F, G : V \longrightarrow W$ addieren und F mit Skalaren $\lambda \in \mathbb{K}$ multiplizieren kann. Eine leichte Überlegung zeigt einem dann die Richtigkeit der folgenden Aussage:

Satz 3.8 *Die Mengen*

$$\mathrm{Hom}_{\mathbb{K}}(V, W) \quad \text{und} \quad \mathrm{End}_{\mathbb{K}}(V)$$

tragen kanonische $\mathbb{K}$*–Vektorraumstrukturen.*

Haben wir noch einen weiteren $\mathbb{K}$–Vektorraum U, so können wir jeden Homomorphismus $F : V \to W$ mit Homomorphismen $G : U \to V$ verknüpfen und erhalten neue Abbildungen

$$F \circ G : U \longrightarrow W\,,$$

die unmittelbar als linear erkannt werden können. Wir dürfen diesen Prozeß des Verknüpfens selbst wieder als eine kanonische Abbildung auffassen

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbb{K}}(V, W) \times \mathrm{Hom}_{\mathbb{K}}(U, V) & \longrightarrow & \mathrm{Hom}_{\mathbb{K}}(U, W) \\ (F \quad , \quad G) & \longmapsto & F \circ G \end{array} \quad ,$$

wobei die unglückliche Reihenfolge der Homomorphismenräume in dem Produkt der linken Seite der historisch gewachsenen Schreibweise für die Komposition von Abbildungen entspringt, die leider oft nachteilig ist. Klar ist, daß mit F und G auch $F \circ G$ ein Isomorphismus ist. Dies impliziert insbesondere, daß Isomorphie eine Äquivalenzrelation induziert (siehe Kapitel 5).

Wir wollen uns noch mit der Frage beschäftigen, wodurch lineare Abbildungen bestimmt sind. Grundlegend hierfür ist

Satz 3.9 *Es sei* $(v_\iota)_{\iota \in I}$ *eine Basis von* V *, und* $(\widetilde{w}_\iota)_{\iota \in I}$ *sei eine beliebige (gleichindizierte) Familie in* W *. Dann gibt es genau eine lineare Abbildung*

$$F : V \longrightarrow W \quad \textit{mit} \quad F\,(v_\iota) = \widetilde{w}_\iota \quad \textit{für alle} \quad \iota \in I\,.$$

Beweis. Es kann nur höchstens ein solches $F \in \mathrm{Hom}\,(V, W)$ geben. Denn aus $V \ni v = \sum_{\iota \in I_0} \lambda_\iota\, v_\iota$, I_0 eine endliche Teilmenge von I, folgt automatisch

$$F\,(v) = F\,(\sum_{\iota \in I_0} \lambda_\iota\, v_\iota) = \sum_{\iota \in I_0} \lambda_\iota\, F\,(v_\iota) = \sum_{\iota \in I_0} \lambda_\iota\, \widetilde{w}_\iota\,.$$

Wegen der Eindeutigkeit der Darstellung von v wird durch diese Gleichung aber auch in eindeutiger Weise eine Abbildung von V nach W erklärt, von der leicht die Linearität nachzuweisen ist. $\square$

Als erste Anwendung dieses Satzes zeigen wir

Satz 3.10 *Es gibt eine kanonische Vektorraum–Isomorphie*

$$M\,(m \times n,\, \mathbb{K}) \longrightarrow \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n,\, \mathbb{K}^m)\ .$$

Beweis. Wir haben schon jeder Matrix $A = (a_{jk}) \in M\,(m \times n,\, \mathbb{K})$ die lineare Abbildung $F_A : \mathbb{K}^n \to \mathbb{K}^m$ mit

$$F_A(x) = A \cdot x$$

zugeordnet. Diese Zuordnung liefert eine Abbildung der gesuchten Art, die zudem leicht als $\mathbb{K}$–linear zu erkennen ist, da offensichtlich $(A + B)\,x = A\,x + B\,x$ und $(\lambda\,A)\,x = \lambda\,(A\,x)$ gilt. Ist $(e_1, \ldots, e_n)$ die Standardbasis von $\mathbb{K}^n$, $(f_1, \ldots, f_m)$ die von $\mathbb{K}^m$, so berechnet man sofort

$$F_A(e_k) = A \cdot e_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix} = \sum_{j=1}^{m} a_{jk}\, f_j\ , \quad k = 1, \ldots, n\ ,$$

d. h. der k–te Spaltenvektor von A ist gerade das Bild des k–ten Einheitsvektors unter F in $\mathbb{K}^m$. Ist nun umgekehrt $F \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n,\, \mathbb{K}^m)$ gegeben, so kann man F die Matrix $A = (a_{jk})$ mit

$$F\,(e_k) = \sum_{j=1}^{m} a_{jk}\, f_j\ , \quad k = 1, \ldots, n\,,$$

zuordnen. Man sieht unmittelbar, daß $F = F_A$ und daß die einer Abbildung F_A zugeordnete Matrix wieder A ist, d. h. daß die hierdurch gegebene (lineare) Abbildung

$$\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n\,,\ \mathbb{K}^m) \longrightarrow M\,(m \times n,\, \mathbb{K})$$

zu der obigen Abbildung invers ist. □

Bemerkung. Aufgrund dieses Satzes werden wir i. f. die Matrix A stets mit dem Homomorphismus $F_A : \mathbb{K}^n \to \mathbb{K}^m$ identifizieren, also z. B. einfach $A : \mathbb{K}^n \to \mathbb{K}^m$ schreiben.

Als zweite Anwendung notieren wir:

Satz 3.11 *Das System* $(v_1, \ldots, v_n)$ *ist genau dann eine Basis von* V*, wenn es einen Isomorphismus* $\Phi : \mathbb{K}^n \to V$ *mit* $\Phi\,(e_k) = v_k$*,* $k = 1, \ldots, n$*, gibt, wobei* $(e_1, \ldots, e_n)$ *die Standardbasis von* $\mathbb{K}^n$ *bezeichnet.*

Beweis. Eine Richtung ist nach Satz 6 klar. Sei umgekehrt $(v_1, \ldots, v_n)$ eine Basis von V. Dann gibt es nach Satz 9 einen Homomorphismus $\Phi : \mathbb{K}^n \to V$ mit $\Phi\,(e_k) = v_k$, der wiederum nach Satz 6 ein Isomorphismus ist. □

Bemerkung. Unter dem Isomorphismus $\Phi : \mathbb{K}^n \to V$ geht der Spaltenvektor

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{K}^n$$

gerade in den Vektor $v = \sum_{k=1}^{n} \lambda_k\, v_k$ über. Wir nennen ihn den *Koordinatenvektor* (oder die *Koordinaten*) von v *relativ* zu der Basis $\mathcal{B} = (v_1, \ldots, v_n)$. Wir werden i. f. genauer $\Phi_{\mathcal{B}}$ statt Φ schreiben, wenn dies notwendig erscheint.

Wenn man beide Anwendungen kombiniert, so erhält man eine zu dem ersten Satz analoge Aussage für endlich erzeugte Vektorräume $V,\, W$ *nach Wahl fester Basen.* Zuvor formulieren wir ein einfaches Lemma, dessen Beweis wir dem Leser überlassen.

Lemma 3.12 *Es seien* $\Phi : V' \to V$, $\Psi : W' \to W$ *Isomorphismen. Dann wird durch*

$$\mathrm{Hom}\,(V',\, W') \ni F' \longmapsto \Psi \circ F' \circ \Phi^{-1} \in \mathrm{Hom}\,(V,\, W)$$

ein Isomorphismus

$$\mathrm{Hom}\,(V',\, W') \longrightarrow \mathrm{Hom}\,(V,\, W)$$

induziert.

Bemerkung. Man mache sich klar, daß die Situation des Lemmas sich augenfällig in der *Kommutativität* des Diagramms

$$\begin{tikzcd} V' \arrow[r, "F'"] \arrow[d, "\Phi"'] & W' \arrow[d, "\Psi"] \\ V \arrow[r, "\Psi \circ F' \circ \Phi^{-1}"'] & W \end{tikzcd}$$

ausdrückt. (Allgemein heißt ein Diagramm von Vektorräumen und Homomorphismen zwischen diesen kommutativ, wenn für je zwei Vektorräume in diesem Diagramm alle sie verbindenden Kompositionen von Homomorphismen – unabhängig von den gewählten Wegen – gleich sind).

Es seien nun V und W endlich erzeugte $\mathbb{K}$–Vektorräume mit Basen $\mathcal{B} = (v_1, \ldots, v_n)$ von V und $\mathcal{C} = (w_1, \ldots, w_m)$ für W. Wir wenden dann das obige Lemma auf die Isomorphismen

$$\Phi = \Phi_{\mathcal{B}} : V' = \mathbb{K}^n \longrightarrow V \quad \text{und} \quad \Psi = \Phi_{\mathcal{C}} : W' = \mathbb{K}^m \longrightarrow W$$

an und erhalten eine Komposition von Isomorphismen

$$M\,(m \times n,\, \mathbb{K}) \xrightarrow{\sim} \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n,\, \mathbb{K}^m) \xrightarrow{\sim} \mathrm{Hom}_{\mathbb{K}}(V,\, W)\ ,$$

also wieder einen Isomorphismus, den wir mit $L_{\mathcal{C}}^{\mathcal{B}}$ bezeichnen. Man beachte, daß der Isomorphismus

$$M\,(m \times n,\, \mathbb{K}) \xrightarrow{\sim} \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n,\, \mathbb{K}^m)$$

hiernach mit $L_{\mathcal{K}'}^{\mathcal{K}}$ zu bezeichnen ist, wenn $\mathcal{K}$ bzw. $\mathcal{K}'$ die kanonischen Basen von $\mathbb{K}^n$ bzw. $\mathbb{K}^m$ bezeichnen. – Wir haben insgesamt gezeigt:

Satz 3.13 *Zu vorgegebenen (endlichen) Basen* $\mathcal{B} = (v_1, \ldots, v_n)$ *von* V *und* $\mathcal{C} = (w_1, \ldots, w_m)$ *von* W *gibt es eine Isomorphie*

$$L_{\mathcal{C}}^{\mathcal{B}} : M\,(m \times n,\, \mathbb{K}) \longrightarrow \mathrm{Hom}_{\mathbb{K}}(V,\, W)\ .$$

Hierdurch wird einer Matrix $A = (a_{jk})_{j,k} \in M\,(m \times n,\, \mathbb{K})$ *gerade die durch*

$$F\,(v_k) = \sum_{j=1}^{m} a_{jk}\, w_j\,, \quad k = 1, \ldots, n\,,$$

bestimmte lineare Abbildung $F : V \to W$ *zugeordnet.*

Bemerkung und *Definition*. Die Umkehrabbildung wird durch die gleichen Bedingungsgleichungen bestimmt. Wir bezeichnen sie stets mit

$$M_{\mathcal{C}}^{\mathcal{B}} : \mathrm{Hom}_{\mathbb{K}}(V,\, W) \longrightarrow M\,(m \times n,\, \mathbb{K})\ .$$

Ist nun ein weiterer Vektorraum U mit Basis $\mathcal{A} = (u_1, \ldots, u_p)$ gegeben, so haben wir ein Diagramm

$$\begin{tikzcd} M(m \times n, \mathbb{K}) \times M(n \times p, \mathbb{K}) \arrow[r, dashed] \arrow[d, "\cong"'] & M(m \times p, \mathbb{K}) \arrow[d, "\cong"] \\ \mathrm{Hom}_{\mathbb{K}}(V, W) \times \mathrm{Hom}_{\mathbb{K}}(U, V) \arrow[r] & \mathrm{Hom}_{\mathbb{K}}(U, W) \end{tikzcd}$$

wobei die untere Zeile durch $(F, G) \longmapsto F \circ G$ gegeben wird. Entspricht dem Paar (F, G) in der oberen Zeile ein Paar von Matrizen (A, B), so definiert dieses Diagramm ein kanonisches Produkt

$$\begin{cases} M\,(m \times n, \mathbb{K}) \times M\,(n \times p, \mathbb{K}) & \longrightarrow \quad M\,(m \times p, \mathbb{K}) \\ \qquad (A, B) & \longmapsto \quad A \cdot B \end{cases} ,$$

wobei $C = A \cdot B$ die $F \circ G$ zugeordnete Matrix bezeichnet. Wir müssen

$$C = (c_{j\ell})_{\substack{j=1,\ldots,m \\ \ell=1,\ldots,p}}$$

noch ausrechnen. Nach Definition ist

$$F\,(v_k) = \sum_{j=1}^{m} a_{jk}\, w_j\ , \qquad k = 1, \ldots, n\,,$$

$$G\,(u_\ell) = \sum_{k=1}^{n} b_{k\ell}\, v_k\ , \qquad \ell = 1, \ldots, p\,,$$

$$\begin{aligned} F \circ G\,(u_\ell) &= F\,\Big(\sum_{k=1}^{n} b_{k\ell}\, v_k \Big) = \sum_{k=1}^{n} b_{k\ell}\, F\,(v_k) = \sum_{k=1}^{n} b_{k\ell} \sum_{j=1}^{m} a_{jk}\, w_j \\ &= \sum_{j=1}^{m} \Big(\sum_{k=1}^{n} a_{jk}\, b_{k\ell} \Big)\, w_j\ , \qquad \ell = 1, \ldots, p\,, \end{aligned}$$

also

$$(*) \qquad c_{j\ell} = \sum_{k=1}^{n} a_{jk}\, b_{k\ell}\ , \quad j = 1, \ldots, m\ , \quad \ell = 1, \ldots, p\,.$$

Satz 3.14 *Unter den obigen Voraussetzungen gilt die* Kompositionsregel

$$M_{\mathcal{C}}^{\mathcal{A}}\,(F \circ G) = M_{\mathcal{C}}^{\mathcal{B}}\,(F) \cdot M_{\mathcal{B}}^{\mathcal{A}}\,(G)\ ,$$

wobei auf der rechten Seite das Matrizenprodukt gemäß der Formel $(*)$ *zu bilden ist.*

Bemerkung. Nach Satz 11 ist jeder endlich erzeugte $\mathbb{K}$–Vektorraum zu einem der Standardräume $\mathbb{K}^n$ isomorph. (Im nächsten Kapitel werden wir sehen, daß die Zahl n eindeutig bestimmt ist; m. a. W.: $\mathbb{K}^n$ und $\mathbb{K}^m$ sind nur dann isomorph, wenn $m = n$). Die endlich erzeugten Vektorräume werden somit bis auf Isomorphie *durch die natürlichen Zahlen klassifiziert*. Dennoch reicht es nicht, nur die Räume $\mathbb{K}^n$ mit ihren Standardbasen zu untersuchen, da damit schon Unterräume wie z. B. Lösungsräume von homogenen Gleichungssystemen nicht adäquat zu behandeln wären.

Zum Schluß dieses Kapitels wollen wir noch kurz die Frage streifen, was anstelle von $\mathbb{K}^n$ in Satz 11 im unendlich dimensionalen Fall zu treten hat. Es ist naheliegend, den Raum *aller* Abbildungen $\mathbb{K}^I := \mathrm{Abb}\,(I, \mathbb{K})$ zu betrachten, wenn I die Mächtigkeit einer Basis $(v_\iota)_{\iota \in I}$ von V bezeichnet. Ist I jedoch nicht endlich, so ist dieser Vektorraum viel zu groß (dies mache man sich am Beispiel $V = \mathbb{K}\,[\,x\,]$ klar!). Man hat aber auf jeden Fall einen Monomorphismus

$$\begin{cases} V & \longrightarrow \quad \mathbb{K}^I \\ v & \longmapsto \quad \alpha_v\ , \end{cases}$$

wobei $\alpha_v : I \to \mathbb{K}$ die durch
$$\alpha_v(\iota) \,=\, \lambda_\iota\,, \quad v \,=\, \sum_\iota \lambda_\iota v_\iota$$
definierte Abbildung ist. Das Bild von V läßt sich leicht bestimmen. Da nur *endliche* Linearkombinationen erlaubt sind, ergibt sich sofort

Satz 3.15 *Es sei V ein $\mathbb{K}$–Vektorraum mit Basis $(v_\iota)_{\iota\in I}$. Dann ist V isomorph zu dem $\mathbb{K}$–Vektorraum*
$$\mathbb{K}^{(I)} \,:=\, \{\alpha \in \mathbb{K}^I : \alpha\,(\iota) \,=\, 0 \text{ für fast alle } \iota \in I\}\,.$$

4 Direkte Summen und Dimensionsformeln

V sei ein fester Vektorraum, U, U' seien Untervektorräume von V, und es gelte $V = U + U'$, d. h. für alle v existiere eine Zerlegung $v = u + u'$ mit $u \in U$, $u' \in U'$. Es gibt ein einfaches Kriterium dafür, daß diese Zerlegung für alle $v \in V$ *eindeutig* ist.

Satz 4.1 *Äquivalent sind (für $V = U + U'$) :*

i) *Die Darstellung $v = u + u'$, $u \in U$, $u' \in U'$ ist eindeutig für alle $v \in V$* ;

ii) $U \cap U' = \{0\}$.

Definition. Man nennt unter der Voraussetzung des vorigen Satzes die Summe $V = U + U'$ *direkt* und schreibt

$$V = U \oplus U' .$$

Beweis von Satz 1. i) $\Longrightarrow$ ii). Sei $u \in U \cap U'$. Die Eindeutigkeit der Zerlegung $u = 0 + u = u + 0$ impliziert dann $u = 0$.

ii) $\Longrightarrow$ i). $v = u + u' = u_1 + u_1' \Longrightarrow u - u_1 = u_1' - u' \in U \cap U' = \{0\} \Longrightarrow u_1 = u ,\ u_1' = u'$. $\square$

Bemerkung. Am Ende dieses Kapitels werden wir beweisen: *Ist V endlich dimensional, $V = U + U'$, so gilt $V = U \oplus U'$ genau dann, wenn* $\dim V = \dim U + \dim U'$.

Wir geben noch eine weitere wichtige

Definition. Es sei V ein $\mathbb{K}$–Vektorraum und $U \subset V$ ein Untervektorraum. Ein Untervektorraum $U' \subset V$ heißt ein *direkter Summand* oder ein *direktes Komplement* von U (in V), wenn

$$V = U \oplus U' .$$

Bemerkung. Solche direkten Summanden existieren immer, wie wir gleich mit Hilfe des *Basisergänzungssatzes* zeigen werden, obwohl diese Aussage viel leichter auch aus dem *Rangsatz* für Homomorphismen deduziert werden kann und wird.

Satz 4.2 *Jeder Untervektorraum U in V besitzt ein Komplement.*

Beweis. Sei $(u_\iota)_{\iota \in I}$ eine Basis von U. Der Basisergänzungssatz liefert dann die Existenz eines Systems $(u_\iota)_{\iota \in I'}$, s. d. $(u_\iota)_{\iota \in J}$, $J = I \cup I'$, eine Basis von V darstellt. Es sei dann U' erzeugt von u_ι, $\iota \in I'$. Damit ergibt sich sofort

$$V = U \oplus U' . \qquad \square$$

Man beachte, daß dieses Argument nicht einmal für endlich dimensionales V ohne den allgemeinen Basissatz auskommt, da wir noch nicht wissen, daß mit V auch U endlich erzeugt ist. Wir geben daher noch einen weiteren direkten Beweis von Satz 2 in diesem Spezialfall: Es sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine endliche Basis von V. Dann schreibt sich jedes Element $v \in V$ in der Form

$$(*) \qquad v = u + \sum_{j=1}^{n} \lambda_j v_j , \quad \lambda_j \in \mathbb{K} ,\ u \in U ,$$

(man braucht nur $u = 0$ zu setzen). Wir nehmen solange Vektoren aus $\mathcal{B}$ fort, bis $(*)$ gerade noch entsprechend erfüllt ist. Ohne Einschränkung können wir annehmen, daß das verbleibende System aus den r ersten Vektoren $v_1, \ldots, v_r$ besteht. Ist hierbei $r = 0$, so ist $U = V$ und nichts weiter zu zeigen. Im anderen Fall sei $U' = \text{span}(v_j)_{j=1,\ldots,r}$, so daß nach Konstruktion $V = U + U'$ gilt. Ist nun $u_0 \in U \cap U'$, also

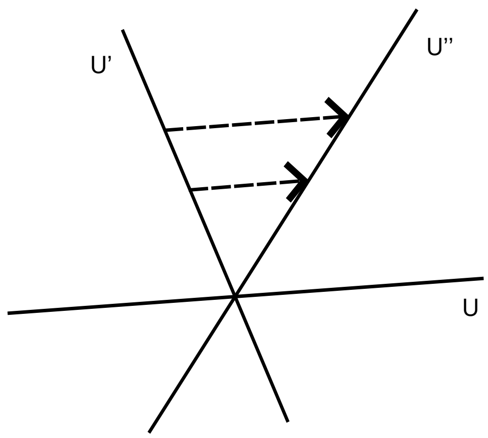
$$u_0 = \sum_{j=1}^{r} \mu_j v_j \in U ,$$

so kann kein μ_j ungleich Null sein. Denn ist z. B. $\mu_r \neq 0$, so ergibt sich für jedes $v \in V$ eine Darstellung

$$v = (u + \lambda_r \mu_r^{-1} u_0) + \sum_{j=1}^{r-1} (\lambda_j - \lambda_r \mu_j \mu_r^{-1})\, v_j$$

im Gegensatz zur Minimalität von r . Somit ist notwendig $u_0 = 0$ und daher U' ein Komplement zu U . □

Direkte Summanden sind nicht eindeutig bestimmt, aber (jedenfalls nach der folgenden Zeichnung) offensichtlich isomorph!



Figur 4.1

Diese Idee kann man ohne weiteres präzisieren.

Satz 4.3 *Direkte Summanden zu festem Unterverktorraum $U \subset V$ sind isomorph. D. h. :*

$$V = U \oplus U' = U \oplus U'' \Longrightarrow U' \cong U'' .$$

Die Isomorphie entsteht durch Projektion entlang des Unterraumes U .

Beweis. Der Beweis ist denkbar einfach (man mache sich die Situation an der obigen Zeichnung klar): Ist $u' \in U' \subset V$, so gibt es eine eindeutige Zerlegung

$$u' = u + u'' , \quad u \in U , \; u'' \in U'' .$$

Wir setzen $\pi\,(u') = u''$ und erhalten so eine Abbildung

$$\pi : U' \longrightarrow U'' ,$$

die wegen der Eindeutigkeit der Zerlegung $u' = u + u''$ linear ist (z. B. ist $\lambda\, u' = \lambda\, u + \lambda\, u''$ die eindeutige Zerlegung von $\lambda\, u'$ und damit $\pi\,(\lambda\, u') = \lambda\, u'' = \lambda\, \pi\,(u')$) . Vertauscht man die Rollen von U' und U'' , so folgt sofort die Existenz einer (linearen) Umkehrabbildung von π . □

Wir wenden diese Überlegungen nun zunächst auf *endlich erzeugte Vektorräume* an. Um frühere Bemerkungen zu konkretisieren, sagen wir, daß der Vektorraum V von der Dimension I ist, wenn *jede* Basis $(v_\iota)_{\iota \in J}$ von V die Kardinalität der Menge I besitzt, d. h. wenn es stets eine Bijektion $J \xrightarrow{\sim} I$ gibt. Tatsächlich besitzen in diesem Sinne, wie wir früher schon behauptet haben, *alle* Vektorräume eine Dimension.

Wir wollen uns, wie oben schon gesagt, auf endlich erzeugte Vektorräume V beschränken. Statt von der Dimension $I = \{1, \ldots, n\}$ zu sprechen, sagen wir natürlich, die Dimension von V sei gleich n; in Zeichen:

$$\dim_{\mathbb{K}} V = n .$$

M. a. W.: $\dim_{\mathbb{K}} V = n$ genau dann, wenn jede Basis von V endlich ist und alle Basen die gleiche Kardinalität n besitzen.

Satz 4.4 *Jeder endlich erzeugte* $\mathbb{K}$*–Vektorraum* V *besitzt eine (endliche) Dimension.*

Wir benötigen zum Beweis den *Austauschsatz* in einer sehr schwachen Form (*Austauschlemma*).

Lemma 4.5 *Es sei* $(v_\iota)_{\iota \in I}$ *eine Basis von* V *und* $v_0 \neq 0$ *ein beliebiger Vektor. Dann kann man einen der Vektoren* v_ι *durch* v_0 *ersetzen, ohne die Basiseigenschaft zu verlieren.*

Beweis. Es gibt eine nichtleere endliche Teilmenge $I_0 \subset I$ mit

$$v_0 = \sum_{\iota \in I_0} \lambda_\iota v_\iota , \quad \lambda_\iota \neq 0 , \quad \iota \in I_0 .$$

Jedes solche ι erfüllt die Behauptung des Satzes, wie man sich leicht überzeugt. □

Beweis von Satz 4. Wir führen Induktion nach der Minimalzahl n von Erzeugenden für V (da V endlich erzeugt ist, gibt es eine kleinste solche Zahl).

Ist $n = 0$, so muß $V = \{0\}$ sein, und es gilt tatsächlich $\dim V = 0$.

Sei nun $n \geq 1$ und $v_1, \ldots, v_n$ ein minimales Erzeugendensystem von V, also eine Basis, und $(w_\kappa)_{\kappa \in K}$ sei eine beliebige Basis. Nach dem Austauschlemma können wir ohne Einschränkung voraussetzen, daß $v_n = w_{\kappa_0}$, $\kappa_0 \in K$. Dann gilt:

$$V = V_1 \oplus V_2 = V_1 \oplus W_2$$

mit $V_1 = \operatorname{span}(v_n) = \operatorname{span}(w_{\kappa_0})$, $V_2 = \operatorname{span}(v_1, \ldots, v_{n-1})$, $W_2 = \operatorname{span}(w_\kappa)_{\kappa \in K \setminus \{\kappa_0\}}$. Nach Satz 3 gilt nun $V_2 \cong W_2$, so daß nach Induktionsvoraussetzung

$$\dim W_2 = \dim V_2 = n - 1$$

und damit $\operatorname{card} K = \operatorname{card}(K \setminus \{\kappa_0\}) + 1 = (n - 1) + 1 = n$ ist. □

Wir kommen nun zu der zentralen *Dimensionsformel*, von der alle später noch abzuleitenden abstammen.

Satz 4.6 *Es seien* U_1, U_2 *Untervektorräume des endlich erzeugten Vektorraumes* V, *und die Summe* $V = U_1 + U_2$ *sei direkt. Dann sind die Vektorräume* U_1, U_2 *endlich erzeugt, und es gilt*

$$\dim V = \dim U_1 + \dim U_2 .$$

Beweis. Es gibt kanonische Projektionen $V = U_1 \oplus U_2 \to U_j$, $j = 1, 2$. Hieraus folgt die endliche Erzeugtheit von U_1 und U_2. Man wählt dann (endliche) Basen von U_1 und U_2 und zeigt ohne Schwierigkeit, daß deren Vereinigung eine Basis von $V = U_1 \oplus U_2$ bildet. □

Da nach Satz 2 Untervektorräume U von endlich erzeugten Vektorräumen V direkte Summanden besitzen, folgt hieraus unmittelbar

Satz 4.7 *Es sei U ein Untervektorraum des endlich erzeugten Vektorraums V. Dann gilt*

$$\dim U \leq \dim V ,$$

und Gleichheit gilt dann und nur dann, wenn $U = V$.

Weiter ergibt sich damit sofort der außerordentlich nützliche

Satz 4.8 *Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}$. Dann gilt:*

i) *Je m Vektoren $v_1, \ldots, v_m$ sind linear abhängig, wenn $m > n$.*

ii) *m Vektoren $v_1, \ldots, v_m$ erzeugen niemals V, wenn $m < n$.*

iii) *n Vektoren bilden schon dann eine Basis von V, wenn sie V erzeugen oder linear unabhängig sind.*

Diese Resultate werden besonders fruchtbar im Zusammenhang mit linearen Abbildungen. Hier ist die Hauptaussage der folgende Satz, den wir weiter unten beweisen.

Satz 4.9 *Es sei $F : V \to W$ ein Homomorphismus von Vektorräumen. Dann gibt es direkte Summanden $V' \subset V$ von* $\ker F$:

$$V = \ker F \oplus V' ,$$

und jeder solche direkte Summand V' ist kanonisch isomorph zu $\operatorname{im} F$.

Folgerungen aus diesem Satz sind Legion. Z. B. ergibt sich sofort die *Rangformel*.

Definition. Der *Rang* einer linearen Abbildung $F : V \to W$ ist die Dimension des Bildes $\operatorname{im} F$; in Zeichen:

$$\operatorname{rang} F = \dim \operatorname{im} F .$$

Satz 4.10 (Rangformel) *$F : V \longrightarrow W$ sei eine lineare Abbildung, V sei endlich dimensional. Dann gilt*

$$\dim V = \operatorname{rang} F + \dim \ker F .$$

Wir wenden die Rangformel sogleich auf lineare Gleichungssysteme an.

Definition. Es sei $A \in M(m \times n, \mathbb{K})$ eine Matrix. Dann ist der *Rang* von A der Rang der A zugeordneten linearen Abbildung $F_A : \mathbb{K}^n \to \mathbb{K}^m$. Anders ausgedrückt: Der Rang $\operatorname{rang} A$ von A ist gleich der Maximalzahl linear unabhängiger Spalten der Matrix A. Wir nennen diese Zahl daher auch manchmal genauer den *Spaltenrang* von A. Entsprechend gibt es auch einen *Zeilenrang*, von dem wir später nachweisen werden, daß er mit dem Spaltenrang überraschenderweise übereinstimmt.

Mit Hilfe dieser Begriffsbildung können wir feststellen:

Satz 4.11 *Für den Lösungsraum L_A eines homogenen linearen Gleichungssystems $Ax = 0$, $A \in M(m \times n, \mathbb{K})$, gilt*

$$\dim L_A = n - \operatorname{rang} A .$$

Eine weitere Folgerung ist der Beweis eines früher schon formulierten Satzes.

Satz 4.12 *Es seien V und W endlich erzeugt und gleichdimensional* : $\dim V = \dim W = n < \infty$, *und es sei $F : V \to W$ eine lineare Abbildung. Dann sind äquivalent* :

i) *F ist injektiv.*

ii) *F ist surjektiv.*

iii) *F ist ein Isomorphismus.*

Beweis. Es gilt nach Voraussetzung $\dim \ker F + \dim \operatorname{im} F = \dim V = \dim W$. Also ist F injektiv $\Longleftrightarrow \dim \ker F = 0 \Longleftrightarrow \dim \operatorname{im} F = \dim W \Longleftrightarrow \operatorname{im} F = W \Longleftrightarrow F$ ist surjektiv. $\square$

Wir beweisen nun Satz 9. Dazu nehmen wir zunächst die Existenz eines Komplementes $V = \ker F \oplus V'$ an und setzen $\varphi = F|\,V' : V' \longrightarrow \operatorname{im} F$. Wir behaupten: φ ist ein Isomorphismus.

a) Es sei $v' \in \ker \varphi$. Dann ist $v' \in \ker F \cap V' = \{0\}$, also $v' = 0$, d. h. φ ist injektiv.

b) $w \in \operatorname{im} F \Longrightarrow w = F(v)\,,\ v \in V = \ker F \oplus V' \Longrightarrow v = v_0 + v' \Longrightarrow w = F(v) = F(v_0 + v') = F(v_0) + F(v') = 0 + \varphi(v') = \varphi(v')$. Also ist φ surjektiv.

Für die Existenz von Komplementen benutzen wir nun im allgemeinen Fall den *Basissatz*. Die mit F „übereinstimmende“ lineare Abbildung f in dem folgenden Diagramm

$$\begin{tikzcd} V \arrow[rr, "F"] \arrow[dr, "f"'] & & W \\ & \operatorname{im} F \arrow[ur] & \end{tikzcd}$$

ist surjektiv. Wählt man irgendeine Basis $(w_\lambda)_{\lambda \in \Lambda}$ von $\operatorname{im} F$ und Urbilder $v_\lambda \in V$ von w_λ unter f, so hat man eine lineare Abbildung $g : \operatorname{im} F \to V$ mit $g(w_\lambda) = v_\lambda$ und folglich $f \circ g = \mathrm{id}_{\operatorname{im} F}$. Hieraus ergibt sich unmittelbar die Injektivität von g und die Isomorphie $\operatorname{im} F \xrightarrow{\sim} g(\operatorname{im} F) =: V'$:

$$\begin{tikzcd} \operatorname{im} g = V' \arrow[rr] & & V \\ & \operatorname{im} F \arrow[ul, "\cong"] \arrow[ur, "g"'] & \end{tikzcd}$$

Es bleibt zu zeigen:

(+) $$V \cong \ker F \oplus V' \,.$$

Ist also $v \in V$, so ist $v' := g \circ f(v) \in V'$ und

$$F(v - v') = f(v) - f(v') = f(v) - f(v) = 0\,,$$

d. h.

$$v = (v - v') + v'$$

mit $(v - v') \in \ker F$ und $v' \in V'$. Ist aber $v \in \ker F \cap V'$, so gilt $v = g(w)$ mit einem Element $w \in \operatorname{im} F$ und damit wegen $v \in \ker F$:

$$0 = f(v) = f(g(w)) = \mathrm{id}(w) = w \quad \text{und} \quad v = g(w) = g(0) = 0\,. \qquad \square$$

Bemerkung. Man beachte, daß dieser Beweis für endlich dimensionales V ohne Satz 2 auskommt. Tatsächlich kann man Satz 2 aus der Rangformel (bzw. ihrem Vorläufer Satz 9) folgern, wenn man zeigen kann, daß jeder Unterraum Kern einer linearen Abbildung ist. Dazu kann man das Konzept des *Quotientenraumes*, das wir im folgenden Kapitel entwickeln wollen, heranziehen. - Als eines der dortigen Ergebnisse formulieren wir schon hier:

Satz 4.13 *Jeder Unterverktorraum $U \subset V$ ist Kern eines Epimorphismus $\varepsilon : V \longrightarrow W$.*

Wir geben hiermit einen weiteren *Beweis* von Satz 7. Nach der Rangformel für den Epimorphismus $\varepsilon : V \to W$ mit $\ker \varepsilon = U$ ergibt sich

$$\dim V = \dim U + \dim W ,$$

und $\dim W$ ist nur dann gleich Null, wenn $W = \{ 0 \}$, was offenbar zu $V = U$ äquivalent ist. □

Genauso können wir die Existenz von Komplementen folgern. Wir zeigen ferner noch, wie mit derselben Methode aus dem Basisauswahlsatz der Austauschsatz folgt: Es seien $(u_\iota)_{\iota \in I}$ ein linear unabhängiges und $(v_\kappa)_{\kappa \in K}$ ein beliebiges Erzeugendensystem von V. Dann ist nach früheren Überlegungen $(u_\iota)_{\iota \in I}$ eine Basis des Erzeugnisses

$$U = \operatorname{span} (u_\iota)_{\iota \in I} ,$$

und die $\varepsilon(v_\kappa)$ bilden ein Erzeugendensystem von W unter der surjektiven Abbildung $\varepsilon : V \longrightarrow W$. Wir wählen nun aus $(\varepsilon(v_\kappa))_{\kappa \in K}$ eine Basis $(\varepsilon(v_\kappa))_{\kappa \in K'}$ aus. Aus dem Beweis der Existenz von Komplementen zu Kernen folgt dann

$$V = U \oplus \operatorname{span} (v_\kappa)_{\kappa \in K'} ,$$

und die $(v_\kappa)_{\kappa \in K'}$ bilden eine Basis des zweiten Anteils wegen $\operatorname{span} (\varepsilon(v_\kappa))_{\kappa \in K'} = W$.

Die *Dimensionsformel für nichtdirekte Summen* erhält man wie folgt: Es seien zunächst V_1 und V_2 zwei beliebige Vektorräume. Man bildet dann das *direkte Produkt*

$$V_1 \times V_2 = \{(v_1, v_2) : v_1 \in V_1, v_2 \in V_2\}$$

mit der komponentenweisen Addition und skalaren Multiplikation. Es ist sofort einzusehen, daß die Abbildung

$$V_1 \ni v_1 \longmapsto (v_1, 0) \in V_1 \times V_2$$

einen Isomorphismus von V_1 auf den Untervektorraum

$$V_1' = \{(v_1, 0) \in V_1 \times V_2 : v_1 \in V_1\}$$

von $V_1 \times V_2$ stiftet und entsprechend für V_2 und

$$V_2' = \{(0, v_2) \in V_1 \times V_2 : v_2 \in V_2\} .$$

Außerdem ist trivialerweise

$$V_1 \times V_2 = V_1' \oplus V_2'$$

und damit für endlich dimensionale V_1 und V_2:

$$\dim V_1 \times V_2 = \dim V_1' + \dim V_2' = \dim V_1 + \dim V_2 .$$

(Man bezeichnet aus diesem Grunde $V_1 \times V_2$ auch manchmal als *(äußere) direkte Summe* $V_1 \oplus V_2$ von V_1 und V_2; Vorsicht mit dieser Schreibweise ist nur geboten, wenn V_1 und V_2 Untervektorräume eines festen Vektorraumes sind und ihre Summe in diesem Vektorraum nicht direkt ist).

Sind nun U und U' Untervektorräume eines festen $\mathbb{K}$–Vektorraumes V, so hat man eine (offensichtlich lineare) Abbildung

$$\sigma : \begin{cases} U \times U' & \longrightarrow \quad V \\ (u, u') & \longmapsto \quad u + u' . \end{cases}$$

Was ist das Bild $\operatorname{im} \sigma$ von σ? Nun, direkt nach Definition ist

$$\operatorname{im} \sigma = \{v \in V : \exists u \in U , \exists u' \in U' \text{ mit } v = u + u' \} ,$$

d. h.

$$\operatorname{im} \sigma = U + U' \, .$$

Was ist der Kern von σ? Offensichtlich gilt $(u, u') \in \ker \sigma \iff 0 = \sigma(u, u') = u + u' \iff u' = -u \in U \cap U'$. M. a. W.:

$$\ker \sigma = \{(u, -u) \in U \times U' : u \in U \cap U'\} \, ,$$

und die Abbildung

$$\begin{cases} U \cap U' & \longrightarrow & \ker \sigma \\ u & \longmapsto & (u, -u) \end{cases}$$

ist ein Isomorphismus. Infolgedessen ist bei endlich dimensionalem V:

$$\begin{aligned} \dim U + \dim U' &= \dim U \times U' = \dim \operatorname{im} \sigma + \dim \ker \sigma \\ &= \dim (U + U') + \dim (U \cap U') \, . \end{aligned}$$

Wir haben bewiesen:

Satz 4.14 *V sei endlich dimensional, U, $U' \subset V$ seien Untervektorräume. Dann gilt*

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

Folgerung 4.15 *Der Vektorraum V sei endlich dimensional, und es gelte $V = U + U'$. Genau dann ist die Summe direkt, wenn $\dim V = \dim U + \dim U'$.*

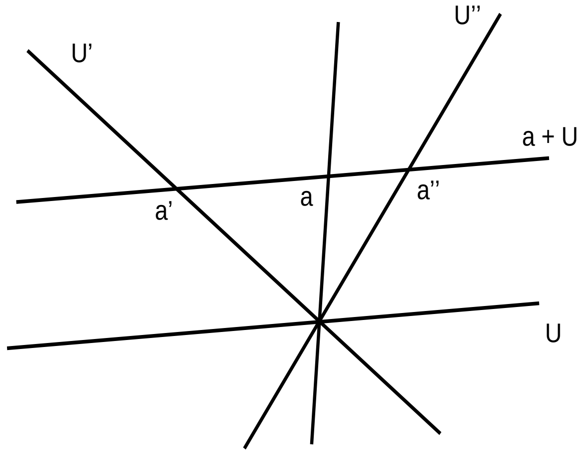
Denn: $V = U + U' \iff \dim V = \dim (U + U')$. Also ist

$$\begin{aligned} \dim V &= \dim U + \dim U' = \dim V + \dim (U \cap U') \\ \iff \quad & \dim (U \cap U') = 0 \iff U \cap U' = \{0\} \iff V = U \oplus U' \, . \end{aligned}$$

Die Umkehrung ist schon bekannt. $\square$

5 Quotientenvektorräume

U sei ein $\mathbb{K}$–Untervektorraum von V. Wir wissen: Komplemente zu U in V sind nicht eindeutig bestimmt, aber isomorph. Die Isomorphie bekommt man durch „Projektion“ parallel zu U. Anders ausgedrückt: „Alle Punkte auf dem affinen Raum parallel zu U sind gleich gut.“



Figur 5.1

Also betrachten wir eine *neue* Menge bestehend aus den *affinen* Mengen

$$[a]_U := [a] := a + U := \{v \in V : \exists u \in U \text{ mit } v = a + u\}.$$

Wir bezeichnen die Menge $\{[a]_U : a \in V\}$ mit V/U und nennen sie die *Menge der Restklassen von* V *modulo* U (oder kurz V *modulo* U). Wir sprechen manchmal auch von dem *Quotienten von* V *nach dem Unterraum* U.

Bemerkungen. 1. Wir wissen von früher: $[a_1] = [a_2] \iff a_1 - a_2 \in U$.

2. Man hat eine kanonische surjektive Abbildung

$$\varepsilon : \begin{cases} V \longrightarrow V/U \\ a \longmapsto [a] \end{cases}.$$

Satz 5.1 *Es gibt genau eine* $\mathbb{K}$*–Vektorraumstruktur auf* V/U*, so daß* ε *linear wird. Der Kern von* ε *ist gleich* U*.*

Beweis. Um die Vektorraumstruktur auf V/U zu finden, setzen wir einfach ε als linear voraus. Dann muß notwendig

$$[a + b] = \varepsilon(a + b) = \varepsilon(a) + \varepsilon(b) = [a] + [b]$$

und

$$[\lambda a] = \varepsilon(\lambda a) = \lambda \varepsilon(a) = \lambda [a]$$

gelten. Man definiert also die Operationen $+$ und $\cdot$ auf diese Weise und stellt denkbar einfach fest, daß sie tatsächlich unabhängig von der Auswahl der speziellen Repräsentanten der Restklassen sind:

$[a_1] = [a_2],\ [b_1] = [b_2] \Longrightarrow a_1 - a_2,\ b_1 - b_2 \in U \Longrightarrow (a_1 + b_1) - (a_2 + b_2) = (a_1 - a_2) + (b_1 - b_2) \in U \Longrightarrow [a_1 + b_1] = [a_2 + b_2]$.

Entsprechend folgt $[\lambda\, a_1] \,=\, [\lambda\, a_2]$.

Der Rest ist leicht nachzurechnen: $V/\,U$ ist ein $\mathbb{K}$–Vektorraum, $\varepsilon : V \to V/\,U$ ist ein Epimorphismus, und es gilt $\ker \varepsilon \,=\, U$. □

Insbesondere haben wir das folgende Resultat, das wir im vorigen Kapitel schon als Satz verwendet haben.

Folgerung 5.2 *Jeder Untervektorraum* $U \subset V$ *ist Kern eines Epimorphismus* $\varepsilon : V \to W$ *.*

Warnung. $V/\,U$ ist *nicht kanonisch* in V enthalten! Also gibt es kein *kanonisches* Komplement. Aber nach Wahl einer Basis von V/U findet man leicht ein Rechtsinverses ι zu ε :

$$\iota : V/\,U \hookrightarrow V \;, \quad \varepsilon \circ \iota \,=\, \mathrm{id} \;.$$

ι ist automatisch injektiv, und es gilt $V \cong U \oplus \mathrm{im}\ \iota$ (und jedes Komplement zu U entsteht auf diese Weise).

Die Konstruktion von $V/\,U$ ist ein Spezialfall der Bildung von Klassen bzgl. einer *Äquivalenzrelation*, die wir noch kurz erläutern wollen.

Definition. 1. Sei X eine Menge. Eine *(binäre) Relation* ist eine Teilmenge $R \subset X \times X$. Man schreibt

$$x_1\, R\, x_2 \quad (x_1 \text{ steht in der Relation } R \text{ zu } x_2) \;:\Longleftrightarrow\; (x_1,\, x_2) \in R \;.$$

2. Eine Relation $R \subset X \times X$ heißt eine *Äquivalenzrelation*, wenn die folgenden Axiome erfüllt sind:

i) $\forall\, x :\; x\, R\, x$ (Reflexivität)

ii) $\forall\, x_1,\, x_2 :\; x_1\, R\, x_2 \Longrightarrow x_2\, R\, x_1$ (Symmetrie)

iii) $\forall\, x_1,\, x_2,\, x_3 :\; x_1\, R\, x_2,\; x_2\, R\, x_3 \Longrightarrow x_1\, R\, x_3$ (Transitivität).

Man schreibt dann i. a. $x_1 \sim_R x_2$ statt $x_1\, R\, x_2$ oder kürzer $x_1 \,\sim\, x_2$.

Beispiele. 1. $R \,=\, X \times X$ ist eine Äquivalenzrelation; hierbei sind alle Elemente paarweise äquivalent.

2. $R \,=\, \Delta \,=\, \{(x_1,\, x_2) \in X \times X \,:\, x_1 \,=\, x_2\}$ ist eine Äquivalenzrelation.

3. $X \,=\, \mathbb{R}\,,\; R \,=\, \{(x_1,\, x_2) \,:\, x_1 \,\le\, x_2\}$. Hier ist ii) verletzt, es liegt also <u>keine</u> Äquivalenzrelation vor.

4. $U \subset V$ sei ein Untervektorraum, $a \sim b :\Longleftrightarrow a \,-\, b \in U$ ist eine Äquivalenzrelation:

$a \,-\, a \,=\, 0 \in U \Longrightarrow a \sim a$.

$a \sim b \Longrightarrow a \,-\, b \in U \Longrightarrow b \,-\, a \,=\, -(a \,-\, b) \in U \Longrightarrow b \sim a$.

$a \sim b\,,\; b \sim c \Longrightarrow a \,-\, b\,,\; b \,-\, c \in U \Longrightarrow a \,-\, c \,=\, (a \,-\, b) \,+\, (b \,-\, c) \in U \Longrightarrow a \sim c$.

Man schreibt in diesem Fall statt $a \sim b$ auch $a \equiv b \bmod U$.

5. Es sei $\mathbb{Z} \ni p\,,\; p \,\ge\, 1$. Man setzt dann für Paare ganzer Zahlen $x_1 \,\equiv\, x_2 \mod p$, falls $p\,|_{\mathbb{Z}}\,(x_1 \,-\, x_2)$ (p teilt $x_1 \,-\, x_2$ in $\mathbb{Z}$). Dies ist eine Äquivalenzrelation. Denn $p\,|\,0 \,=\, x \,-\, x \Longrightarrow x \,\equiv\, x \mod p$, etc.

Definition. Sei X eine Menge mit der Äquivalenzrelation $\sim$. Eine *Äquivalenzklasse* $A \subset X$ ist eine nichtleere Teilmenge $A \subset X$ mit:

i) $x_1,\, x_2 \in A \Longrightarrow x_1 \sim x_2$.

ii) $x_1 \in A,\, x_2 \sim x_1 \Longrightarrow x_2 \in A$.

Satz 5.3 1. *Jede Äquivalenzklasse ist von der Gestalt*

$$A = [a] := \{x \in X : x \sim a\}\,, \quad a \in X\,.$$

2. *Zwei Klassen* $[a_1], [a_2]$ *sind entweder disjunkt oder gleich. (Genauer gilt :* $[a_1] \cap [a_2] \neq \emptyset \iff a_1 \sim a_2 \iff [a_1] = [a_2]$ *).*

Bemerkung. Es existiert also die *Menge* der Äquivalenzklassen

$$X/\sim := \{[a] : a \in X\} \subset \mathfrak{P}(X)\,,$$

und man hat eine kanonische surjektive Abbildung

$$X \xrightarrow{\varepsilon} X/\sim\,.$$

Beweis von Satz 3. 1. Ist $A \neq \emptyset$, so existiert ein Element $a \in A$, und es gilt $x \sim a \overset{\text{ii)}}{\Longrightarrow} x \in A \Longrightarrow [a] \subset A$. Ist umgekehrt $a' \in A$, so gilt $a' \sim a \Longrightarrow a' \in [a]$, also $A \subset [a]$.

2. $[a_1] \cap [a_2] \neq \emptyset \Longrightarrow \exists\, a$ im Durchschnitt $\Longrightarrow a_1 \sim a$ und $a \sim a_2 \Longrightarrow a_1 \sim a_2 \Longrightarrow [a_1] = [a_2]$ (und daraus folgt wiederum $[a_1] \cap [a_2] \neq \emptyset$). □

Beispiele (Numerierung wie oben).

1. $X/R = \{X\}$.

2. $X/\Delta = X$.

4. $V/\sim = V/U$.

5. $\mathbb{Z}/\equiv_p =: \mathbb{Z}/p\mathbb{Z}$. (Diese Menge erbt von $\mathbb{Z}$ sogar eine *Ringstruktur*, ist aber nur dann ein Körper, wenn p eine *Primzahl* ist).

6 Lineare Gleichungssysteme und affine Räume

Wir erläutern zuerst die *Strukturtheorie* für den Lösungsraum des Systems

$$(*) \qquad A\,x \,=\, b\,, \quad A \in M\,(m \times n, \mathbb{K})\,, \quad b \in \mathbb{K}^m\,.$$

Ordnen wir der Matrix A wie früher die lineare Abbildung $F \,=\, F_A$ zu:

$$F : \begin{cases} V \,=\, \mathbb{K}^n & \longrightarrow \quad \mathbb{K}^m \,=\, W \\ \qquad\; x & \longmapsto \quad A\,x\,, \end{cases}$$

so können wir in dieser Formulierung den Lösungsraum von $(*)$ beschreiben durch

$$L_{A,b} \,=\, \{x \in \mathbb{K}^n : A\,x \,=\, b\} \,=\, F^{-1}(b)\,.$$

Es erscheint in diesem Zusammenhang sinnvoll, noch einmal die Definition einer affinen Menge aus dem vorigen Kapitel zu wiederholen.

Definition. Eine Teilmenge $X \subset V$ eines Vektorraumes V heißt *affin*, wenn entweder $X \,=\, \emptyset$ oder

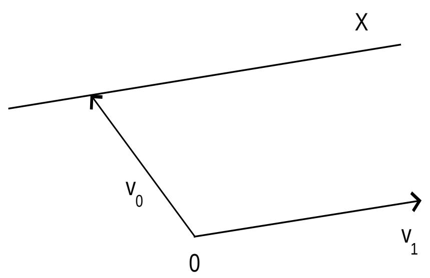
$$X \,=\, v_0 \,+\, U := \{v \,=\, v_0 \,+\, u : u \in U\}$$

für einen Vektor $v_0 \in V$ und einen Untervektorraum $U \subset V$. (Es handelt sich also bei einer affinen Teilmenge sozusagen um einen „verschobenen Untervektorraum“).

Beispiel. Der affine Raum

$$X \,=\, v_0 \,+\, U\,, \quad U \,=\, \operatorname{span}(v_1)\,, \quad v_1 \,\neq\, 0\,,$$

beschreibt nichts anderes als die *Gerade durch* v_0 *mit Richtungsvektor* v_1.



Figur 6.1

Der Untervektorraum U zu einem affinen Raum $X \,=\, v_0 \,+\, U$ ist durch X eindeutig bestimmt; man setzt daher z. B.

$$\dim X := \dim U\,.$$

v_0 ist dagegen nicht durch X bestimmt. Es gilt genauer:

Satz 6.1 $X \,=\, v_0 \,+\, U \,=\, v_0' \,+\, U' \iff U \,=\, U'\,,\; v_0' \,-\, v_0 \in U\,.$

Beweis. „$\Longleftarrow$“ ist trivial. „$\Longrightarrow$“: $v_0 \,+\, U \,=\, v_0' \,+\, U' \Longrightarrow v_0' \,=\, v_0' \,+\, 0 \in v_0' \,+\, U' \,=\, v_0 \,+\, U \Longrightarrow v_0' \,-\, v_0 \in U$. Ist nun $u' \in U'$ beliebig, so folgt $v_0' \,+\, u' \,=\, v_0 \,+\, u$, also $u' \,=\, u \,-\, (v_0' \,-\, v_0) \in U$, d. h. $U' \subset U$. Die umgekehrte Inklusion folgt genauso. $\square$

Satz 6.2 *$F : V \longrightarrow W$ sei linear, $b \in W$. Dann ist $F^{-1}(b)$ (leer oder) affin. Genauer: Ist $v_0 \in F^{-1}(b)$, so gilt*

$$(*) \qquad F^{-1}(b) = v_0 + \ker F .$$

Insbesondere ist dann $\dim F^{-1}(b) = \dim \ker F = \dim V - \operatorname{rang} F$, *falls* $\dim V < \infty$.

Beweis. Nur $(*)$ ist zu zeigen.

a) $v \in F^{-1}(b) \Longrightarrow F(v) = b \Longrightarrow F(v - v_0) = F(v) - F(v_0) = b - b = 0 \Longrightarrow v - v_0 \in \ker F \Longrightarrow v \in v_0 + \ker F$. Also ist $F^{-1}(b) \subset v_0 + \ker F$.

b) $v \in v_0 + \ker F \Longrightarrow v = v_0 + u,\ u \in \ker F \Longrightarrow F(v) = F(v_0 + u) = F(v_0) + F(u) = b + 0 = b \Longrightarrow v \in F^{-1}(b)$; also gilt $v_0 + \ker F \subset F^{-1}(b)$. □

Die Anwendung auf lineare Gleichungssysteme ist jetzt fast automatisch:

i) $L_{A,0}$ sei der Lösungsraum des homogenen Systems $Ax = 0$. Dann ist $L_{A,0}$ ein Untervektorraum von $\mathbb{K}^n$ der Dimension
$$\dim L_{A,0} = n - \operatorname{rang} A .$$

ii) $L_{A,b}$, $b \neq 0$, der Lösungsraum des inhomogenen Systems $Ax = b$, ist entweder leer oder affin von der Gestalt

$$L_{A,b} = \underset{\text{spezielle Lösung}}{x_0} + \underset{\text{Lösungsraum des homogenen Systems}}{L_{A,0}} .$$

Zur vollständigen Klärung des Sachverhalts braucht man noch ein Kriterium dafür, wann die affine Menge $L_{A,b}$ nicht leer ist. Wir setzen

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad A' = (A, b) = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}.$$

A' heißt die *erweiterte Koeffizientenmatrix* des Systems. Im folgenden bezeichnen wir mit $a^1, \ldots, a^n, b \in \mathbb{K}^m$ stets die *Spaltenvektoren* der erweiterten Matrix, und mit $a_1, \ldots, a_m \in \mathbb{K}^n$ die *Zeilenvektoren* von A.

Satz 6.3 *Der Lösungsraum $L_{A,b}$ ist genau dann nicht leer, wenn*

$$\operatorname{rang} A = \operatorname{rang}(A, b) .$$

Beweis. a) $L_{A,b} \neq \emptyset \Longrightarrow \exists x \in \mathbb{K}^n$ mit $F(x) = Ax = b \Longrightarrow b \in \operatorname{im} F = \operatorname{span}(a^1, \ldots, a^n) \Longrightarrow \operatorname{span}(a^1, \ldots, a^n, b) = \operatorname{span}(a^1, \ldots, a^n)$. Also ist $\operatorname{rang}(A, b) = \dim \operatorname{span}(a^1, \ldots, a^n, b) = \dim \operatorname{span}(a^1, \ldots, a^n) = \operatorname{rang} A$.

b) Umgekehrt gilt stets $\operatorname{span}(a^1, \ldots, a^n) \subset \operatorname{span}(a^1, \ldots, a^n, b)$. Ist nun $\operatorname{rang} A = \operatorname{rang}(A, b)$, so sind die Dimensionen der beiden Räume gleich, so daß also

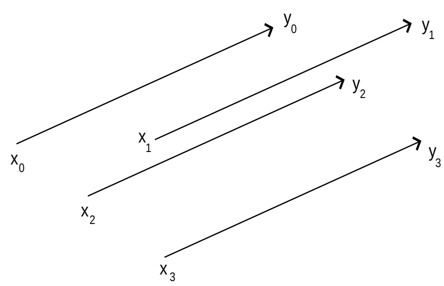
$$\operatorname{span}(a^1, \ldots, a^n) = \operatorname{span}(a^1, \ldots, a^n, b) \ni b .$$

Folglich gilt $b \in \operatorname{span}(a^1, \ldots, a^n)$, d. h. $\exists x = {}^t(x_1, \ldots, x_n)$, s. d. $b = x_1 a^1 + \cdots + x_n a^n$. Dies bedeutet $Ax = b$, d. h. $L_{A,b} \neq \emptyset$. □

In der *Geometrie* gibt es keinen Ursprung (ausgezeichneten Punkt) und kein ausgezeichnetes Koordinatensystem. Das geometrische set–up muß hier also mehr beinhalten als nur die Vorgabe eines Vektorraums.

Nach Felix Klein besteht eine Geometrie aus einer Menge zusammen mit einer auf ihr operierenden Gruppe von bijektiven Abbildungen. Geometrie in diesem Entwurf ist das Studium der Invarianten unter dieser Operation.

Für die n–dimensionale *affine Geometrie* ist die unterliegende Menge der $\mathbb{R}^n$, schlicht als *Menge* aller n–Tupel $x = (x_1, \ldots, x_n)$. Wir illustrieren im folgenden alle Begriffe am Spezialfall der Ebene, also für $n = 2$. Die entscheidenden Abbildungen sind die *Translationen*:



Figur 6.2

Offensichtlich kann man Translationen hintereinanderschalten, „rückgängig machen", und es gibt eine *neutrale* Translation. Also bildet die Menge

$$\mathrm{Trans}\,(\mathbb{R}^n)$$

eine Gruppe, die zudem abelsch ist (dies mache man sich an einem Parallelogramm klar). Wir können dies alles analytisch beschreiben: Durch

$$\mathbb{R}^n \ni a \longmapsto T_a\,, \quad \text{wobei} \quad T_a(x) = x + a$$

(jetzt ist $\mathbb{R}^n$ als abelsche Gruppe mit der üblichen Addition versehen) wird eine Abbildung $\mathbb{R}^n \to \mathrm{Trans}\,(\mathbb{R}^n)$ erklärt, für die gilt:

$$T_a = \mathrm{id} \Longleftrightarrow a = 0\,, \quad T_{a+b} = T_a\,T_b = T_b\,T_a\,.$$

Also ist

$$\mathbb{R}^n \xrightarrow{\sim} \mathrm{Trans}\,(\mathbb{R}^n)$$

ein bijektiver Gruppenhomomorphismus, also ein Gruppenisomorphismus. Nun ist natürlich auch $T_{na} = n\,T_a$ (wenn wir besser $T_a + T_b$ anstelle von $T_a\,T_b$ schreiben), und für jedes T gibt es ein T' mit $n\,T' = T$ bei vorgegebenem $n \in \mathbb{N}^*$, nämlich

$$T' = T_{a/n}\,, \quad \text{wenn} \quad T = T_a\,.$$

Damit wird, wie man leicht nachrechnet, $\mathrm{Trans}\,(\mathbb{R}^n)$ zu einem $\mathbb{Q}$–Vektorraum, und durch einen (zumindest „ideellen") Grenzübergang wird

$$\mathrm{Trans}\,(\mathbb{R}^n)$$

sogar mit der Struktur eines $\mathbb{R}$–Vektorraumes versehen, der isomorph zu $\mathbb{R}^n$ ist.

Ist allgemein X eine Menge, G eine Gruppe, so ist eine *Operation* von G auf X ein Gruppen–Homomorphismus

$$\begin{aligned} \tau : G \longrightarrow S\,(X) &= \textit{symmetrische Gruppe von } X \\ &= \text{Menge aller Bijektionen von } X \text{ mit } \circ \text{ als Verknüpfung}\,, \end{aligned}$$

d. h. $\tau_g := \tau(g) \in S(X)$, $g \in G$, und

$$\tau_e = \mathrm{id}\,, \quad \tau_{gg'} = \tau_g \, \tau_{g'}\,, \quad \text{etc.}$$

Die Operation τ heißt *eigentlich*, wenn $\ker \tau = \{e\}$, d. h. $\tau_g = \mathrm{id} \Longleftrightarrow g = e$, und *transitiv*, falls es für alle $x_1, x_2 \in X$ genau ein $g \in G$ gibt mit $\tau_g(x_1) = x_2$. Transitivität impliziert natürlich Eigentlichkeit.

In diesem Sinne operiert die Gruppe $\mathrm{Trans}(\mathbb{R}^n)$ (eigentlich und) transitiv auf $\mathbb{R}^n$.

Definition. Es sei $\mathbb{K}$ ein beliebiger Körper. Ein *affiner Raum* über $\mathbb{K}$ ist ein Tripel

$$(X, T(X), \tau)\,,$$

wobei $X \neq \emptyset$, $T(X)$ einen (endlich dimensionalen) $\mathbb{K}$–Vektorraum und $\tau : (T(X), +) \longrightarrow S(X)$ eine transitive Operation bezeichnet.

Beispiel. $(\mathbb{K}^n, \mathrm{Trans}(\mathbb{K}^n), \tau)$ heißt *der affine Raum* $\mathbb{A}^n$ (oder besser: $\mathbb{A}^n(\mathbb{K})$).

Dies ist aber auch schon das Standardbeispiel. Sei nämlich $(X, T(X), \tau)$ beliebig vorgegeben, also $T(X) \cong \mathbb{K}^n$. Wähle dann x_0 aus X fest und betrachte einen beliebigen Punkt $x \in X$. Dann gibt es genau ein Gruppenelement $g \in T(X)$ mit $\tau_g(x_0) = x$. Man schreibt auch

$$g = \overrightarrow{x_0 x}\,.$$

Aus dieser Definition folgt unmittelbar

$$\overrightarrow{x_0 x} + \overrightarrow{x x_1} = \overrightarrow{x_0 x_1}\,, \quad \overrightarrow{x x_0} + \overrightarrow{x_0 x} = \mathrm{id}\,.$$

Also hat man eine offensichtlich injektive und surjektive Abbildung

$$X \ni x \longmapsto \overrightarrow{x_0 x} \in T(X) \cong \mathbb{K}^n\,,$$

mit deren Hilfe man X und $T(X)$ identifizieren kann. - Als Fazit können wir festhalten:

Ein affiner Raum entsteht aus einem Vektorraum, indem man die Auszeichnung eines festen Punktes als Ursprung aufhebt. Umgekehrt erhält man einen Vektorraum, wenn man in einem affinen Raum einen Punkt als Ursprung auszeichnet.

Als nächstes müssen wir die zulässigen Abbildungen zwischen solchen affinen Räumen festlegen.

Definition. $(X, T(X), \tau)$ und $(Y, T(Y), \sigma)$ seien zwei affine Räume über $\mathbb{K}$. Eine Abbildung $f : X \to Y$ heißt *affin*, falls es eine $\mathbb{K}$–lineare Abbildung

$$T(f) : T(X) \longrightarrow T(Y)$$

gibt, s. d. für alle x_0, $x \in X$ mit $y_0 = f(x_0)$, $y = f(x)$ gilt:

$$\overrightarrow{y_0 y} = T(f)\left(\overrightarrow{x_0 x}\right)\,.$$

Satz 6.4 *f sei affin. Dann ist f injektiv, surjektiv oder bijektiv, falls $T(f)$ die „entsprechende“ Eigenschaften besitzt. Ist f bijektiv, so ist auch f^{-1} affin mit $T(f^{-1}) = T(f)^{-1}$. Solche bijektive affine Abbildungen heißen auch* Affinitäten.

Satz 6.5 *Die Translationen sind genau die Affinitäten $X \longrightarrow X$ mit $T(f) = \mathrm{id}_{T(X)}$.*

Bemerkung. Für die Definition einer affinen Abbildungen reicht tatsächlich die Auswahl eines *einzigen* Punktes x_0; setzt man nämlich wie oben y, y_0 und $y_1 = f(x_1)$ für einen beliebigen weiteren Punkt x_1, so ergibt sich

$$\begin{aligned}\overrightarrow{y_1 y} = \overrightarrow{y_0 y} - \overrightarrow{y_0 y_1} &= T(f)(\overrightarrow{x_0 x}) - T(f)(\overrightarrow{x_0 x_1}) \\ = T(f)(\overrightarrow{x_0 x} - \overrightarrow{x_0 x_1}) &= T(f)(\overrightarrow{x_1 x}) .\end{aligned}$$

Beweis (Satz 4): $y_1 = y_2 \iff \overrightarrow{y_0 y_1} = \overrightarrow{y_0 y_2} \iff T(f)(\overrightarrow{x_0 x_1}) = T(f)(\overrightarrow{x_0 x_2}) \iff T(f)(\overrightarrow{x_1 x_2}) = 0$. Also ist f injektiv genau dann, wenn aus $T(f)(\overrightarrow{x_1 x_2}) = 0$ stets $\overrightarrow{x_1 x_2} = 0$ folgt. Dies ist aber gleichbedeutend mit der Injektivität von $T(f)$. Der Rest funktioniert genauso. □

Beweis (Satz 5). f ist eine Translation, falls $\overrightarrow{x_0 y_0} = \overrightarrow{x_1 y_1}$ für alle x_0, x_1 gilt. Dies ist äquivalent zu $\overrightarrow{y_0 y_1} = \overrightarrow{x_0 x_1}$ für alle x_0, x_1, d. h. $T(f) = \mathrm{id}$. □

Definition und *Bemerkung*. $(X, T(X), \tau)$ sei ein affiner Raum. Eine Teilmenge $Y \subset X$ heißt ein *affiner Unterraum*, falls es ein Element $y_0 \in Y$ gibt, s. d.

$$T(Y) = \{ \overrightarrow{y_0 y} \in T(X) : y \in Y \}$$

ein Untervektorraum von $T(X)$ ist.

Man zeigt dann leicht: Ist $Y \neq \emptyset$, so ist $T(Y)$ durch Y eindeutig bestimmt, und

$$(Y, T(Y), \tau \,|\, T(Y))$$

ist ein affiner Raum.

Definition. $Y \subset X$ sei ein affiner Unterraum. Dann setzt man

$$\dim Y = \begin{cases} -\infty & , \quad Y = \emptyset , \\ \dim T(Y) & , \quad Y \neq \emptyset . \end{cases}$$

Speziell heißt ein affiner Raum Y mit $\dim Y = 1$ eine *Gerade*, mit $\dim Y = 2$ eine *Ebene*, mit $\dim Y = \dim X - 1$ eine *Hyperebene*.

Durchschnitte $\bigcap_{\iota \in I} Y_\iota$ von affinen Unterräumen $Y_\iota \subset X$ sind wieder affin, und es gilt

$$T(\bigcap_{\iota \in I} Y_\iota) = \bigcap_{\iota \in I} T(Y_\iota) .$$

Für Vereinigungen von affinen Unterräumen gilt Entsprechendes offensichtlich nicht. Man definiert deshalb den *Verbindungsraum*

$$\bigvee\nolimits_{i \in I} Y_\iota := \bigcap_{\substack{Y \text{affin} \\ Y_\iota \subset Y}} Y .$$

Z. B. gilt $\{x_1\} \vee \{x_2\} = \{(\lambda\, \overrightarrow{x_1 x_2})(x_1) : \lambda \in \mathbb{K}\}$; dies ist natürlich die Gerade durch x_1 und x_2 falls $x_1 \neq x_2$. – Der folgende Satz ist nicht schwer zu beweisen.

Satz 6.6 *Ist* $\operatorname{char} \mathbb{K} \neq 2$, X *ein affiner Raum und* $Y \subset X$, *dann gilt* : Y *ist ein affiner Teilraum genau dann, wenn für alle* $y_1, y_2 \in Y$ *der Verbindungsraum* $\{y_1\} \vee \{y_2\}$ *in* Y *enthalten ist.*

Für die Dimension von Verbindungsräumen gelten unterschiedliche Regeln bei geeigneter Fallunterscheidung:

a) $Y_1 \cap Y_2 \neq \emptyset$. Man zeigt dann leicht:

$$T(Y_1 \vee T_2) = T(Y_1) + T(Y_2) .$$

Nach einer früheren Formel ist

$$\begin{aligned}\dim Y_1 \vee Y_2 &= \dim T(Y_1 \vee Y_2)\\ &= \dim T(Y_1) + \dim T(Y_2) - \dim T(Y_1) \cap T(Y_2)\\ &= \dim Y_1 + \dim Y_2 - \dim T(Y_1 \cap T_2)\,,\end{aligned}$$

also

$$\dim(Y_1 \vee Y_2) = \dim Y_1 + \dim Y_2 - \dim(Y_1 \cap Y_2)\,.$$

b) $Y_1 \cap Y_2 = \emptyset$. Wähle in diesem Fall $y_j \in Y_j$ beliebig und schreibe für die Verbindungsgerade $Y = y_1 \vee y_2$. Man überzeugt sich ebenso leicht wie oben davon, daß

$$T(Y_1 \vee Y_2) = (T(Y_1) + T(Y_2)) \oplus T(Y)\,,$$

so daß in dieser Situation

$$\dim(Y_1 \vee Y_2) = \dim Y_1 + \dim Y_2 - \dim(T(Y_1) \cap T(Y_2)) + 1\,.$$

Beispiele. a) Für (verschiedene) parallele Geraden im $\mathbb{R}^3$ ist nach Definition $T(Y_1) = T(Y_2)$. Also ergibt sich $\dim T(Y_1 \vee Y_2) = 1 + 1 - 1 + 1 = 2$. Mit anderen Worten:

Der Verbindungsraum von zwei parallelen, nicht übereinstimmenden Geraden ist eine Ebene.

b) *Windschiefe Geraden* in $\mathbb{R}^3$ sind charakterisiert durch $T(Y_1) \neq T(Y_2)$ und $Y_1 \cap Y_2 = \emptyset$. Daraus folgt $T(Y_1) \cap T(Y_2) = \{0\}$, da die beiden betroffenen Vektorräume eindimensional und verschieden sind, und $\dim(Y_1 \vee Y_2) = 1 + 1 - 1 + 1 = 3$.

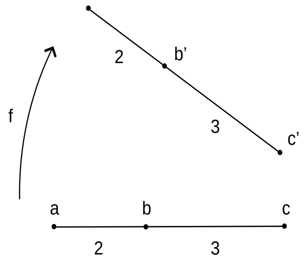
Wir stellen uns nun die Frage: *Was bleibt invariant unter Affinitäten?* Natürlich gehen z. B. k–dimensionale affine Unterräume in ebensolche über, insbesondere Geraden in Geraden. Abbildungen mit der letzten Eigenschaft nennt man auch *Kollineationen* (äquivalent ist: sind x_1, x_2, x_3 *kollinear*, d. h. liegen diese drei Punkte auf einer Geraden, so auch $f(x_1), f(x_2), f(x_3)$).

Der folgende Satz wird manchmal auch der *Hauptsatz der affinen Geometrie* genannt. Zu seinem Beweis siehe z. B. das Buch von Fischer über *Analytische Geometrie*.

Satz 6.7 $\mathbb{K}$ *habe mindestens drei Elemente,* $X/\mathbb{K}$ *sei affin,* $\dim X \geq 2$. *Dann gilt:* $f : X \to X$ *ist kollinear genau dann, wenn* f *eine* Semiaffinität *ist (d. h. wenn* $f : X \to X$ *bijektiv und* $T(f) : T(X) \to T(X)$ *additiv ist und* $T(f)(\lambda v) = \alpha(\lambda)\, T(f)(v)$ *für einen (festen) Automorphismus* α *von* $\mathbb{K}$ *gilt).*

Bemerkung. Man kann zeigen, daß $\mathbb{R}$ nur einen Automorphismus besitzt, so daß also in dem obigen Satz f automatisch eine *Affinität* sein muß.

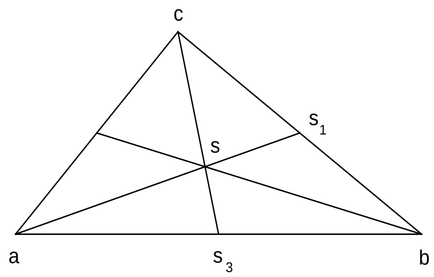
Des Weiteren bleiben bei Affinitäten sogenannte *Teilungsverhältnisse* (im Falle $\mathbb{K} = \mathbb{R}$) erhalten (siehe Aufgabe 33).



Figur 6.3

Als typisches Beispiel einer Anwendung dieses Begriffes notieren wir

Satz 6.8 *Die Seitenhalbierenden in einem (nicht ausgearteten) Dreieck schneiden sich in einem Punkt und teilen sich gegenseitig im Verhältnis* $2:1$.



Figur 6.4

Beweis. Ohne Einschränkung sei $a = 0$, $\overrightarrow{ab} = e_1$, $\overrightarrow{ac} = e_2$. Dann gilt offensichtlich nach der Zeichnung

$$\overrightarrow{as_1} = \overrightarrow{ab} + \frac{1}{2}\,\overrightarrow{bc} = \overrightarrow{ab} + \frac{1}{2}\,(\overrightarrow{ac} - \overrightarrow{ab}) = \frac{1}{2}\,(\overrightarrow{ab} + \overrightarrow{ac}) = \frac{1}{2}\,(e_1 + e_2)$$

und

$$\overrightarrow{cs_3} = -\,\overrightarrow{ac} + \frac{1}{2}\,\overrightarrow{ab} = -e_2 + \frac{1}{2}\,e_1\,.$$

Damit ist die Strecke $\overrightarrow{as}$ ein Vielfaches von $(1/2)\,(e_1 + e_2)$:

$$\overrightarrow{as} = \frac{\sigma}{2}\,(e_1 + e_2)\,,$$

$\sigma \in [\,0,\,1\,]$, und entsprechend ist

$$\overrightarrow{as} = e_2 + \tau\,\left(\frac{1}{2}\,e_1 - e_2\right)$$

mit $\tau \in [\,0,\,1\,]$. Gleichheit gilt damit genau dann, wenn $\sigma/2 = \tau/2$ und $\sigma/2 = 1 - \tau$, was zu $\sigma = \tau = 2/3$ äquivalent ist. □

Wir können in der affinen Geometrie zwar über *Seitenhalbierende* und *Teilungsverhältnisse* reden, nicht aber über *Mittelsenkrechte*, *Umkreis*, *Winkelhalbierende*, etc. Dazu benötigen wir mehr als nur den Vektorraumbegriff, nämlich den des *Skalarproduktraumes*. Wir kommen darauf im Kapitel 10 zurück.

7 Basiswechsel und Matrizen

Wir haben in Kapitel 3 jedem Homomorphismus $F : V \to W$ von $\mathbb{K}$–Vektorräumen relativ zu gegebenen Basen $\mathcal{B}$ und $\mathcal{C}$ eine Matrix $M_{\mathcal{C}}^{\mathcal{B}}(F)$ zugeordnet. Es stellt sich die Frage, wie sich diese Matrix verändert, wenn man die vorgegebenen Basen wechselt. Es wird sich herausstellen, daß solche Basiswechsel durch spezielle quadratische Matrizen beschrieben werden und das Transformationsgesetz durch geeignete Multiplikation aller betroffenen Matrizen zu formulieren ist. Es wird ferner darum gehen, die Basiswechselmatrizen auf verschiedene Weisen zu charakterisieren. Außerdem werden wir in einem einfachen Fall das Problem der Vereinfachung der Matrizen durch Wahl geschickter Basen lösen. Dieses *Normalformenproblem* wird uns bis zum Ende der Vorlesung begleiten.

Es sei uns also in dem Vektorraum V (der Dimension $n < \infty$) neben der Basis $\mathcal{B} = (v_1, \ldots, v_n)$ noch eine weitere $\mathcal{A} = (w_1, \ldots, w_n)$ gegeben. Als *Basiswechselmatrix* von $\mathcal{A}$ nach $\mathcal{B}$ definieren wir dann die quadratische Matrix

$$S = S_{\mathcal{B}}^{\mathcal{A}} := M_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id}) \in M(n \times n, \mathbb{K}) .$$

Sie wird aufgrund der Definition der Abbildung $M_{\mathcal{B}}^{\mathcal{A}}$ charakterisiert durch die Kommutativität des Diagramms

$$\begin{tikzcd}
\mathbb{K}^n \arrow[r, "M_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id})"] \arrow[d, "\Phi_{\mathcal{A}}"'] & \mathbb{K}^n \arrow[d, "\Phi_{\mathcal{B}}"] \\
V \arrow[r, "\mathrm{id}"'] & V
\end{tikzcd}$$

Oder anders gewendet: die Matrix $M_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id}) = (a_{jk})$ ist eindeutig definiert durch die Darstellung (man beachte die Reihenfolge!)

$$w_k = \mathrm{id}(w_k) = \sum_{j=1}^{n} a_{jk}\, v_j .$$

Die Frage nach dem Transformationsverhalten beantwortet sich ganz elementar ohne jegliche Rechnung mit Satz 3. 14. Nach der Kompositionsregel ist nämlich

$$M_{\mathcal{C}}^{\mathcal{A}}(F) = M_{\mathcal{C}}^{\mathcal{A}}(F \circ \mathrm{id}) = M_{\mathcal{C}}^{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id}) .$$

Mit anderen Worten:

Der Übergang von $M_{\mathcal{C}}^{\mathcal{B}}(F)$ nach $M_{\mathcal{C}}^{\mathcal{A}}(F)$ geschieht durch Multiplikation von rechts *mit der Basiswechselmatrix $S_{\mathcal{B}}^{\mathcal{A}}$.*

Es bezeichne nun noch $S_{\mathcal{D}}^{\mathcal{C}}$ die Matrix, die den Wechsel von der Basis $\mathcal{C}$ nach $\mathcal{D}$ in dem Vektorraum W beschreibt. Mit dem gleichen Argument wie oben erhält man dann sofort die *Transformationsformel*:

Satz 7.1 (Transformationsformel) *Für jeden Homomorphismus $F : V \to W$ gilt*

$$M_{\mathcal{D}}^{\mathcal{A}}(F) = S_{\mathcal{D}}^{\mathcal{C}} \cdot M_{\mathcal{C}}^{\mathcal{B}}(F) \cdot S_{\mathcal{B}}^{\mathcal{A}} .$$

Man sollte sich diese Formel klar machen und merken können mittels der Kommutativität des folgenden Diagramms.

$$\begin{tikzcd}
\mathbb{K}^n \arrow[rrr, "M^{\mathcal{A}}_{\mathcal{D}}(F)"] \arrow[d, "\mathrm{id}"'] & & & \mathbb{K}^m \arrow[d, "\mathrm{id}"] \\
\mathbb{K}^n \arrow[r, "M^{\mathcal{A}}_{\mathcal{B}}(\mathrm{id})"] \arrow[d, "\Phi_{\mathcal{A}}"'] & \mathbb{K}^n \arrow[r, "M^{\mathcal{B}}_{\mathcal{C}}(F)"] \arrow[d, "\Phi_{\mathcal{B}}"'] & \mathbb{K}^m \arrow[r, "M^{\mathcal{C}}_{\mathcal{D}}(\mathrm{id})"] \arrow[d, "\Phi_{\mathcal{C}}"] & \mathbb{K}^m \arrow[d, "\Phi_{\mathcal{D}}"] \\
V \arrow[r, "\mathrm{id}"'] \arrow[d, "\mathrm{id}"'] & V \arrow[r, "F"'] & W \arrow[r, "\mathrm{id}"'] & W \arrow[d, "\mathrm{id}"] \\
V \arrow[rrr, "F"'] & & & W
\end{tikzcd}$$

Bevor wir Anwendungen dieser grundlegenden Formel bringen, wollen wir erst die quadratischen Matrizen kennzeichnen, die zu Basiswechseln gehören. Wählen wir z. B. $\mathcal{B} = \mathcal{A}$, so ist offensichtlich

$$M^{\mathcal{A}}_{\mathcal{A}}(\mathrm{id}) = (\delta_{jk}) = E_n \, ,$$

die sogenannte *Einheitsmatrix* in $M(n \times n, \mathbb{K})$. Wegen der Multiplikationsformel für darstellende Matrizen von Produkten folgt dann sofort

$$M^{\mathcal{A}}_{\mathcal{B}}(\mathrm{id})\, M^{\mathcal{B}}_{\mathcal{A}}(\mathrm{id}) = M^{\mathcal{B}}_{\mathcal{B}}(\mathrm{id}) = E_n \, ,$$
$$M^{\mathcal{B}}_{\mathcal{A}}(\mathrm{id})\, M^{\mathcal{A}}_{\mathcal{B}}(\mathrm{id}) = M^{\mathcal{A}}_{\mathcal{A}}(\mathrm{id}) = E_n \, .$$

Dies ist gerade die Definition einer *invertierbaren Matrix*.

Definition. Eine Matrix $A \in M(n \times n, \mathbb{K})$ heißt *invertierbar*, falls es eine weitere quadratische Matrix $A' \in M(n \times n, \mathbb{K})$ gibt, s. d.

$$A \cdot A' = A' \cdot A = E_n \, .$$

Da die Einheitsmatrix E_n per definitionem das neutrale Element bezüglich der (assoziativen) Multiplikation in $M(n \times n, \mathbb{K})$ ist, besagt die obige Definition gerade, daß A das Inverse A' besitzt, das dann aber bekanntlich eindeutig bestimmt ist und wie üblich mit A^{-1} bezeichnet wird. M. a. W:

Satz 7.2 *Die Menge der invertierbaren $n \times n$–Matrizen bildet bezüglich der Matrizenmultiplikation eine (i. a. nicht kommutative) Gruppe. Es gilt für solche Matrizen A, B:*

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

und

$$(A\,B)^{-1} = B^{-1}\,A^{-1} \, .$$

Die letzten beiden Aussagen sind in jeder Gruppe erfüllt. Wir werden einen Beweis nachher mit einer anderen Charakterisierung der invertierbaren Matrizen geben.

Definition. Die Gruppe

$$\mathrm{GL}(n, \mathbb{K}) := \{A \in M(n \times n, \mathbb{K}) : A \text{ ist invertierbar } \}$$

heißt die *allgemeine lineare Gruppe („general linear group“)* in n Veränderlichen mit Koeffizienten in $\mathbb{K}$.

Daß die allgemeine lineare Gruppe genau aus den Basiswechselmatrizen besteht, wird durch die folgende Überlegung bestätigt: Ist eine invertierbare Matrix $A = (a_{jk})$ mit der Inversen $A' = B = (b_{kj})$ gegeben und gibt man eine Basis $\mathcal{B}$ vor, so definiere man

$$w_k = \sum_{j=1}^n a_{jk} v_j , \quad k = 1, \ldots, n .$$

Wegen

$$v_\ell = \sum_{j=1}^n \delta_{j\ell} v_j = \sum_{j=1}^n (\sum_{k=1}^n a_{jk} b_{k\ell}) v_j = \sum_{k=1}^n b_{k\ell} (\sum_{j=1}^n a_{jk} v_j) = \sum_{k=1}^n b_{k\ell} w_k$$

bildet das n–Tupel $(w_1, \ldots, w_n)$ ein Erzeugendensystem von V der minimalen Länge $n = \dim V$, also eine Basis $\mathcal{A}$, und es gilt

$$A = M_{\mathcal{B}}^{\mathcal{A}} (\mathrm{id}) .$$

Als erstes Beispiel dafür, wie sich allgemeine Sätze über lineare Abbildungen in solche über Matrizen uminterpretieren lassen, zitieren wir

Satz 7.3 *Eine Matrix $A \in M(m \times n, \mathbb{K})$ hat genau dann den Rang r, wenn es invertierbare Matrizen $S \in \mathrm{GL}(m, \mathbb{K})$ und $T \in \mathrm{GL}(n, \mathbb{K})$ gibt, so daß die Matrix $S A T^{-1}$ die einfache Gestalt*

$$S A T^{-1} = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right)$$

besitzt. Dabei soll die Matrix auf der rechten Seite eine $m \times n$–Matrix $B = (b_{jk})$ bezeichnen mit $b_{jj} = 1$ für $j = 1, \ldots, r$ und $b_{jk} = 0$ sonst.

Beweis. Es bezeichne $F = F_A : V = \mathbb{K}^n \longrightarrow \mathbb{K}^m = W$ die durch A induzierte Abbildung, die in den *kanonischen* Basen $\mathcal{B}$ bzw. $\mathcal{C}$ gerade durch die Matrix A beschrieben wird:

$$A = M_{\mathcal{B}}^{\mathcal{C}} .$$

Für beliebige Basiswechsel von $\mathcal{A}$ nach $\mathcal{B}$ bzw. $\mathcal{C}$ nach $\mathcal{D}$ schreiben wir die Transformationsformel in der Form

$$M_{\mathcal{D}}^{\mathcal{A}} (F) = S_{\mathcal{D}}^{\mathcal{C}} \cdot A \cdot (S_{\mathcal{A}}^{\mathcal{B}})^{-1} .$$

Sind nun S und T wie in dem Satz gegeben, so wähle man gemäß der obigen Bemerkung Basen $\mathcal{A}$ resp. $\mathcal{D}$, so daß

$$T := S_{\mathcal{A}}^{\mathcal{B}} \quad \text{und} \quad S := S_{\mathcal{D}}^{\mathcal{C}} .$$

Bezüglich dieser Basen wird dann F durch die sehr einfache Matrix auf der rechten Seite beschrieben, aus der man ohne Rechnung

$$\mathrm{rang}\, A = \dim \mathrm{im}\, F = r$$

abliest.

Hat umgekehrt A den Rang r, so wähle man für $F = F_A$ Zerlegungen

$$V = V' \oplus \ker F , \quad W = \mathrm{im}\, F \oplus W'$$

und Basen der Summanden V', $\ker F$ und W'. Wegen $V' \cong \mathrm{im}\, F$ wird die erste in eine Basis von $\mathrm{im}\, F$ überführt. Faßt man diese Basen zu solchen von V bzw. W zusammen, so wird wegen $r = \dim \mathrm{im}\, F$ die Abbildung F offensichtlich bzgl. dieser durch die einfache Matrix in dem Satz dargestellt. Setzt man wieder für die entsprechenden Basiswechsel wie oben

$$T := S_{\mathcal{A}}^{\mathcal{B}} \quad \text{und} \quad S := S_{\mathcal{D}}^{\mathcal{C}} ,$$

so ergibt sich die Behauptung erneut aus der Transformationsformel. □

Wir werden im letzten Teil dieses Kapitels darauf eingehen, wie man die Matrizen S und T tatsächlich algorithmisch bestimmen und damit lineare Gleichungssysteme effektiv lösen kann.

Bevor wir dies tun, wollen wir aber die oben angekündigte weitere Charakterisierung von invertierbaren Matrizen besprechen. Es handelt sich dabei um die Auszeichnung der Automorphismen innerhalb der Endomorphismen eines endlich dimensionalen Vektorraumes V. Als $\mathbb{K}$–Vektorräume sind nach den früheren Überlegungen die Räume

$$\mathrm{End}\,(V) \quad \text{und} \quad M\,(n \times n, \mathbb{K})$$

vermöge der Abbildung $M_{\mathcal{B}}^{\mathcal{A}}$ bezgl. zweier beliebiger Basen isomorph. Nun tragen beide Räume aber auch eine multiplikative Struktur, nämlich bezüglich der Komposition von Abbildungen bzw. der Multiplikation von Matrizen. I. a. hat aber für zwei Endomorphismen F und G das Produkt

$$M_{\mathcal{B}}^{\mathcal{A}}(F) \cdot M_{\mathcal{B}}^{\mathcal{A}}(G)$$

keinerlei Bedeutung für die Hintereinanderschaltung $F \circ G$. Dies wird grundsätzlich anders, wenn man jetzt die Basen $\mathcal{A}$ und $\mathcal{B}$ als *gleich* voraussetzt, denn dann gilt nach der Kompositionsregel

$$M_{\mathcal{B}}^{\mathcal{B}}\,(F \circ G) \,=\, M_{\mathcal{B}}^{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{B}}\,(G)\,.$$

Man nennt einen $\mathbb{K}$–Vektorraum zusammen mit einer „vernünftigen“ assoziativen, nicht notwendig kommutativen Multiplikation auch eine $\mathbb{K}$–*Algebra*. Was hier vernünftig bedeutet, soll stellvertretend durch die Rechenregeln in dem folgenden Satz über die Algebra der quadratischen Matrizen angedeutet werden:

Satz 7.4 *Es seien A, B, C etc. quadratische Matrizen mit Einträgen in dem Körper $\mathbb{K}$, und ferner sei λ ein Skalar in $\mathbb{K}$. Dann gelten die folgenden Gesetze* :

i) $A\,(B + B') = A\,B + A\,B'\,, \quad (A + A')\,B = A\,B + A'\,B\,,$

ii) $A\,(\lambda\,B) = (\lambda\,A)\,B = \lambda\,(A\,B)\,,$

iii) $(A\,B)\,C = A\,(B\,C)\,.$

Eine $\mathbb{K}$–lineare Abbildung zwischen solchen $\mathbb{K}$–Algebren, die die multiplikative Struktur respektiert, nennt man einen $\mathbb{K}$–*Algebrahomomorphismus*. In diesem Sinne gilt also

Satz 7.5 *Ist $\mathcal{B}$ eine Basis eines $\mathbb{K}$–Vektorraumes der endlichen Dimension n, so ist die Abbildung*

$$M_{\mathcal{B}}^{\mathcal{B}} : \mathrm{End}\,(V) \longrightarrow M\,(n \times n, \mathbb{K})$$

ein $\mathbb{K}$–Algebraisomorphismus mit Umkehrabbildung $L_{\mathcal{B}}^{\mathcal{B}}$. Ist $\mathcal{C}$ eine weitere Basis mit Basiswechselmatrix $S_{\mathcal{C}}^{\mathcal{B}}$, so gilt die Transformationsformel in der Form

$$M_{\mathcal{C}}^{\mathcal{C}} \,=\, S_{\mathcal{C}}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{B}} \cdot (S_{\mathcal{C}}^{\mathcal{B}})^{-1}\,.$$

Das *Normalformenproblem* für Endomorphismen ist daher wesentlich schwieriger als für allgemeine Homomorphismen (die vollständig durch ihren Rang bestimmt sind), da uns jetzt nur eine *einzige* invertierbare Matrix S auf beiden Seiten gleichzeitig zur Verfügung steht. Die Frage nach einfachen Formen von $S \cdot A \cdot S^{-1}$ für gegebene Matrix A wird in späteren Kapiteln beantwortet (Jordansche Normalform, Hauptachsentransformation).

Kommen wir wieder zu der Untersuchung von $\mathrm{End}\,(V)$ bzw. $M\,(n \times n, \mathbb{K})$ zurück. Für $n = 1$ ist diese $\mathbb{K}$–Algebra natürlich isomorph zu $\mathbb{K}$ selbst, also ein Körper. Diese Eigenschaft geht aber für größere n sofort verloren, da es dann stets nichttriviale Endomorphismen F und G gibt, deren Komposition aber identisch Null ist (z. B. $F(v_1) = v_1, F(v_j) = 0$, $j = 2, \ldots, n$, und $G(v_n) = v_n, G(v_j) = 0$, $j = 1, \ldots, n-1$). Der Endomorphismenring besitzt also für diese n stets Nullteiler. Wir verändern unsere Fragestellung daher zu der folgenden: Welches sind die bzgl. der Multiplikation *invertierbaren Elemente* in dem Endomorphismenring von V? Direkt nach Definition sind dies selbstverständlich die Automorphismen, die genauso offensichtlich eine Gruppe bilden. Da $M_{\mathcal{B}}^{\mathcal{B}}$ die Multiplikation respektiert, wird infolgedessen die Automorphismengruppe unter dieser Abbildung auf eine Untergruppe von $M(n \times n, \mathbb{K})$ bijektiv abgebildet. Dies ist gerade die allgemeine lineare Gruppe, wie man sich sofort überlegt:

Satz 7.6 $M_{\mathcal{B}}^{\mathcal{B}} : \mathrm{Aut}(V) \longrightarrow \mathrm{GL}(n, \mathbb{K})$ *ist ein Gruppenisomorphismus mit Umkehrabbildung* $L_{\mathcal{B}}^{\mathcal{B}} : \mathrm{GL}(n, \mathbb{K}) \longrightarrow \mathrm{Aut}(V)$.

Den Inhalt dieses Satzes sollte man sich an dem folgenden Diagramm klarmachen:

$$\begin{tikzcd}
\mathrm{End}(V) \arrow[r, "M_{\mathcal{B}}^{\mathcal{B}}"] & M(n \times n, \mathbb{K}) \\
\cup & \cup \\
\mathrm{Aut}(V) \arrow[r, "M_{\mathcal{B}}^{\mathcal{B}}"'] & \mathrm{GL}(n, \mathbb{K})
\end{tikzcd}$$

Es sei noch einmal betont, daß es gerade in diesem Kontext absolut unerläßlich ist, nur *eine* Basis zur Beschreibung von Endomorphismen zu wählen! (Siehe auch Aufgabe 40).

Wir geben in Satz 8 noch einige weitere Charakterisierungen invertierbarer Matrizen, von denen wir den Teil vii) aber erst in Kapitel 8 abhandeln können. Bei der Formulierung benötigen wir noch die Operation des *Transponierens* einer Matrix. Diese ist definiert durch

$$M(m \times n, \mathbb{K}) \quad \ni \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

$$\longmapsto \quad {}^tA = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix} \quad \in \quad M(n \times m, \mathbb{K}) .$$

Oder anders ausgedrückt:

$${}^tA = (b_{kj})_{\substack{k=1,\ldots,n \\ j=1,\ldots,m}} , \quad b_{kj} = a_{jk} .$$

Es ist sofort klar, daß *der Spaltenrang von* A *gleich dem Zeilenrang der transponierten Matrix* ${}^tA$ *ist und umgekehrt*. Es genügt also grundsätzlich, Algorithmen zur Bestimmung von einem der Ränge anzugeben. – Man beweist ohne sonderliche Mühe den folgenden

Satz 7.7 *Für beliebige Matrizen* A, $B \in M(m \times n, \mathbb{K})$ *und Skalare* $\lambda \in \mathbb{K}$ *gelten die folgenden Regeln*:

1. ${}^t(A + B) = {}^tA + {}^tB$;
2. ${}^t(\lambda A) = \lambda {}^tA$;
3. ${}^t({}^tA) = A$.

M. a. W. : $M(m \times n, \mathbb{K}) \xrightarrow{t} M(n \times m, \mathbb{K})$ *ist ein* $\mathbb{K}$*–Vektorraumisomorphismus mit erneutem Transponieren als inverser Operation.*

Bezüglich der Multiplikation von Matrizen gilt (man beachte die Reihenfolge !)

4. ${}^t(A \cdot B) = {}^tB \cdot {}^tA$.

Beweis. Nur die 4. Aussage ist zu begründen. Setze $C = A \cdot B$, $C = (c_{j\ell})$. Dann ist nach Definition ${}^tC = ({}^tc_{\ell j})$ mit ${}^tc_{\ell j} = c_{j\ell}$. Also ist nach Definition der Matrizenmultiplikation

$$ {}^tc_{\ell j} = c_{j\ell} = \sum_{k=1}^{n} a_{jk}\, b_{k\ell} = \sum_{k=1}^{n} {}^tb_{\ell k}\, {}^ta_{kj} \ , $$

d. h. ${}^t(A \cdot B) = {}^tB \cdot {}^tA$. □

Satz 7.8 *Für* $A \in M(n \times n, \mathbb{K})$ *sind äquivalent :*

i) $A \in \mathrm{GL}(n, \mathbb{K})$.

ii) *Es existiert* A' *mit* $A' A = E_n$ *, also ein Linksinverses zu* A *.*

iii) *Es existiert ein Rechtsinverses zu* A *.*

iv) ${}^tA \in \mathrm{GL}(n, \mathbb{K})$.

v) *Der Spaltenrang von* A *ist gleich* n *.*

vi) *Der Zeilenrang von* A *ist gleich* n *.*

vii) $\det A \neq 0$.

Beweis. A ist genau dann invertierbar, wenn die zugeordnete lineare Abbildung F_A invertierbar ist. Nun folgt aus der Existenz eines Linksinversen zu A und damit zu F_A die Injektivität von F_A und entsprechend aus der Existenz eines Rechtsinversen die Surjektivität von F_A . Nach Satz 4.12 reicht dies aber für die Bijektivität von F_A aus. Schließlich impliziert v) ebenfalls die Surjektivität von F_A . Dies beweist die Äquivalenz von i), ii), iii) und v). Ist weiter $A' A = E_n$, so folgt unmittelbar ${}^tA\, {}^tA' = {}^tE_n = E_n$. Somit ist auch i) $\Longleftrightarrow$ iv) bewiesen. Der Rest folgt aus dem folgenden allgemeineren Satz (und aus Kapitel 8 bzgl. vii)). □

Satz 7.9 *Für jede Matrix* A *stimmen Spaltenrang und Zeilenrang überein.*

Beweis. Es ist $\mathrm{rang}\, A = \dim(\mathrm{im}\, A)$, A aufgefaßt als lineare Abbildung $\mathbb{K}^n \to \mathbb{K}^m$. Es sei $S \in \mathrm{GL}(m, \mathbb{K})$. Dann gilt für $x \in \mathbb{K}^n$:

$$ A\,x = 0 \Longleftrightarrow (S\,A)\,x = 0 \, . $$

Also ist $\ker A = \ker S\,A$, so daß

$$ \mathrm{rang}\, S\,A = n - \dim \ker S\,A = n - \dim \ker A = \mathrm{rang}\, A \, . $$

Alternativ kann man S als Element in $\mathrm{Aut}\, \mathbb{K}^m$ auffassen und erhält einen Isomorphismus $\mathrm{im}\, A \xrightarrow{S} S(\mathrm{im}\, A) = \mathrm{im}\, S\,A$, der ebenfalls die obige Rangbeziehung impliziert.

Ist weiter $T \in \mathrm{GL}(n, \mathbb{K})$, so ist natürlich $\mathrm{im}\, A = \mathrm{im}\,(A\,T)$, da T als Automorphismus von $\mathbb{K}^n$ interpretiert werden kann. Folglich ist $\mathrm{rang}\, A = \mathrm{rang}\, A\,T$. Also ergibt sich insgesamt:

$$ \mathrm{rang}\, A = \mathrm{rang}\, S\,A\,T \, . $$

Wählt man nun speziell S und T so, daß

$$ S\,A\,T = N := \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right) \ , $$

so folgt

$$\mathrm{rang}^t A \,=\, \mathrm{rang}^t T^t A^t S \,=\, \mathrm{rang}^t (SAT) = \mathrm{rang}^t N \,=\, \mathrm{rang} N \,=\, \mathrm{rang}\, A \;. \qquad \square$$

Bemerkung. In Kapitel 9 werden wir die Transponierte einer Matrix konzeptionell als Matrix zu der dualen Abbildung interpretieren, was uns eine weiteren Zugang zu dem obigen Satz liefert.

Wenn man den obigen Beweis genauer analysiert, so haben wir die folgende abbildungstheoretische Aussage benutzt.

Lemma 7.10 *Es seien lineare Abbildungen*

$$V' \xrightarrow{\varphi} V \xrightarrow{F} W \xrightarrow{\psi} W'$$

zwischen endlich dimensionalen Vektorräumen gegeben, wobei φ *surjektiv und* ψ *injektiv ist. Dann gilt*

$$\mathrm{rang}\, F \,=\, \mathrm{rang}\, (\psi \circ F \circ \varphi) \;.$$

Der *Beweis* kann wie oben geführt werden, wobei man sich auf die beiden Fälle $\varphi = \mathrm{id}$ bzw. $\psi = \mathrm{id}$ beschränken kann. $\square$

Wir wollen uns für den Rest dieses Kapitels mit der Frage nach *algorithmischen Verfahren* zur Bestimmung von Rängen, Lösungen etc. beschäftigen. Aus historischen Gründen und wegen der Verwendbarkeit bei der Lösung von linearen Gleichungssystemen (*Gaußalgorithmus*) werden wir den *Zeilenrang* in den Vordergrund stellen, was jedoch nach Satz 9 keine Einschränkung bedeutet. Wir schreiben im folgenden $a_j \,=\, (a_{j1}, \ldots, a_{jn})$, $j \,=\, 1, \ldots, m$, für die Zeilen einer $m \times n$–Matrix A, so daß wir also A kurz in der Form

$$A \,=\, \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$$

notieren können. Den von den Zeilen a_j erzeugten Unterraum in $\mathbb{K}^n$ nennen wir auch den *Zeilenraum* von A, in Zeichen

$$\mathrm{ZR}\,(A) \,=\, \mathrm{span}\,(a_j)_{j=1,\ldots,m} \;,$$

so daß also

$$\text{Zeilenrang von } A \,=\, \dim \mathrm{ZR}\,(A) \;.$$

Bemerkung. Man kann den Zeilenrang auch konzeptionell erklären als $\dim \mathrm{im}\, F_A^*$, wobei der Homomorphismus

$$F_A^* : \mathbb{K}^n \,\cong\, \mathrm{Hom}\,(\mathbb{K}^n, \mathbb{K}) \longrightarrow \mathrm{Hom}\,(\mathbb{K}^m, \mathbb{K}) \,\cong\, \mathbb{K}^m$$

durch Hintereinanderschaltung mit F definiert ist (siehe Kapitel 9).

Zur Bestimmung des Zeilenranges dienen nun Operationen, die man vom Lösen linearer Gleichungssysteme schon gewöhnt ist, nämlich die sogenannten *elementaren Zeilenumformungen*. Von diesen gibt es vier Typen.

I. Multiplikationen der j–ten Zeile mit $\lambda \in \mathbb{K}^* \,=\, \mathbb{K} \setminus \{0\}$:

$$A \,=\, \begin{pmatrix} a_1 \\ \vdots \\ a_{j-1} \\ a_j \\ a_{j+1} \\ \vdots \\ a_m \end{pmatrix} \quad\longmapsto\quad \begin{pmatrix} a_1 \\ \vdots \\ a_{j-1} \\ \lambda\, a_j \\ a_{j+1} \\ \vdots \\ a_m \end{pmatrix} \;=:\, A_I \;.$$

II. Addition der k–ten Zeile zur j–ten Zeile, $k \neq j$.

III. Addition des λ–fachen der k–ten Zeile zur j–ten Zeile, $k \neq j$ $(\lambda \in \mathbb{K}^*)$.

IV. Vertauschen der j–ten Zeile mit der k–ten Zeile, $k \neq j$.

III. und IV. entstehen jedoch aus I. und II. Schematisch sieht das so aus:

$$\begin{pmatrix} a_j \\ a_k \end{pmatrix} \xmapsto{I_\lambda} \begin{pmatrix} a_j \\ \lambda a_k \end{pmatrix} \xmapsto{II} \begin{pmatrix} a_j + \lambda a_k \\ \lambda a_k \end{pmatrix} \xmapsto{I_{\lambda^{-1}}} \begin{pmatrix} a_j + \lambda a_k \\ a_k \end{pmatrix}$$

$$\begin{pmatrix} a_j \\ a_k \end{pmatrix} \longmapsto \begin{pmatrix} a_j \\ -a_k \end{pmatrix} \longmapsto \begin{pmatrix} a_j \\ a_j - a_k \end{pmatrix} \longmapsto \begin{pmatrix} a_j - (a_j - a_k) \\ a_j - a_k \end{pmatrix} = \begin{pmatrix} a_k \\ a_j - a_k \end{pmatrix} \longmapsto \begin{pmatrix} a_k \\ a_j \end{pmatrix} .$$

Satz 7.11 *B entstehe aus A durch endlich viele elementare Zeilentransformationen. Dann ist* $\mathrm{ZR}\,(A) = \mathrm{ZR}\,(B)$ *(und die Zeilenränge von A und B stimmen überein).*

Beweis. Es genügt der Nachweis für die Typen I und II.

a) Ist $v \in \mathrm{ZR}\,(A)$, so gilt nach Transformation I_λ :

$$\begin{aligned} v &= \mu_1 a_1 + \cdots + \mu_j a_j + \cdots + \mu_m\, a_m \\ &= \mu_1 a_1 + \cdots + \frac{\mu_j}{\lambda}(\lambda a_j) + \cdots + \mu_m a_m \in \mathrm{ZR}\,(B) . \end{aligned}$$

Da die Umkehrung von I_λ offensichtlich die Transformation $I_{\lambda^{-1}}$ ist, ergibt sich die umgekehrte Inklusion genauso.

b) Die Gleichung

$$\begin{aligned} v &= \mu_1 a_1 + \cdots + \mu_j a_j + \cdots + \mu_m a_m \\ &= \mu_1 a_1 + \cdots + \mu_j(a_j + a_k) + \cdots + (\mu_k - \mu_j) a_k + \cdots + \mu_m a_m \end{aligned}$$

impliziert wie oben die Behauptung im Fall II. □

Bemerkung. Elementare Zeilentransformationen werden durch *Multiplikation von links* mit *elementaren Matrizen* bewerkstelligt. Es ist nicht schwer, sich diese Matrizen aufzuschreiben. So gehört natürlich zum Typ I die $m \times m$–Diagonalmatrix mit lauter Einsen in der Diagonalen mit Ausnahme des j–ten Eintrages, der gleich λ gesetzt werden muß. Nach Definition sind diese Matrizen invertierbar, und ihre Inversen sind wieder elementar. Sie gehören zu einfachen Basiswechseln im Bildraum; für Typ I z. B. ist das der Basiswechsel $(w_1, \ldots, w_j, \ldots, w_m) \longmapsto (w_1, \ldots, \lambda w_j, \ldots, w_m)$.

Wir sagen nun, eine Matrix $B = (b_{jk})_{j=1,\ldots,m,k=1,\ldots,n}$ sei von *Zeilenstufenform*, falls es eine aufsteigende Folge $1 \leq k_1 < k_2 < \cdots < k_r \leq m$ gibt, so daß gilt:

1. $b_{1k_1} \neq 0 , \ldots , b_{rk_r} \neq 0$;
2. $b_{jk} = 0 , \quad k < k_j \quad \text{oder} \quad j > r$.

Offensichtlich (man führe Induktion) ist für eine solche Matrix der Zeilenrang gleich r .

Dem bekannten Eliminationsverfahren für lineare Gleichungssysteme entspricht der folgende

Satz 7.12 *Jede Matrix läßt sich durch endlich viele elementare Zeilenumformungen vom Typ III und IV in Zeilenstufenform bringen.*

Beweis. Es sei k_1 der Index der ersten nicht verschwindenden Spalte in A . Durch eine Transformation vom Typ IV bringt man dann ein von Null verschiedenes Element dieser Spalte in die erste Zeile und kann dann durch Transformationen vom Typ III alle darunter liegenden Einträge zum Verschwinden bringen. Danach verfährt man induktiv auf dieselbe Weise. □

Bemerkung. Man kann in Satz 12 für die Zeilenstufenform zusätzlich

$$b_{1k_1} = \cdots = b_{rk_r} = 1$$

voraussetzen.

Die oben beschriebene Methode gibt auch ein effektives Verfahren zu bestimmen, ob n vorgegebene Vektoren $v_1, \ldots, v_n \in \mathbb{K}^n$ eine Basis bilden: Man schreibe die Vektoren $v_1, \ldots, v_n$ als Zeilen $a_1, \ldots, a_n$ in eine $n \times n$–Matrix A und führe so oft elementare Zeilenumformungen durch, bis Zeilenstufenform B vorliegt.

Satz 7.13 *Unter den obigen Bezeichnungen gilt : Die Vektoren $(v_1, \ldots, v_n)$ bilden genau dann eine Basis, wenn $r = n$ ist, d. h. wenn B eine invertierbare obere Dreiecksmatrix darstellt :*

$$\begin{pmatrix} b_{11} & * & \cdots & & * \\ 0 & b_{22} & * & \cdots & * \\ \vdots & & \ddots & & \vdots \\ 0 & & & b_{n-1,n-1} & * \\ 0 & & \cdots & 0 & b_{nn} \end{pmatrix} \quad , \quad b_{jj} \neq 0\,,\ j = 1, \ldots, n\,.$$

Diesem Satz können wir eine völlig andere Wendung geben:

Folgerung 7.14 *Jede Matrix $A \in \mathrm{GL}\,(n, \mathbb{K})$ läßt sich durch elementare Zeilentransformationen in die Einheitsmatrix transformieren. M. a. W. : Jede Matrix $A \in \mathrm{GL}\,(n, \mathbb{K})$ ist ein endliches Produkt von Elementarmatrizen.*

Eine andere Interpretation dieser Folgerung besteht darin zu sagen, daß *jede Basistransformation durch endlich viele elementare Basistransformationen zusammengesetzt werden kann.*

Dies führt nun zu einem effektiven Verfahren zur Berechnung der Inversen A^{-1} zu gegebener Matrix $A \in \mathrm{GL}\,(n, \mathbb{K})$: Es seien $B_1, \ldots, B_s$ Elementarmatrizen, so daß

$$B_s \cdot \ldots \cdot B_1 \cdot A = E_n\,.$$

Dann ist natürlich

$$A^{-1} = B_s \cdot \ldots \cdot B_1\,.$$

Man betrachte also das folgende Schema:

A	E_n
$B_1 A$	$B_1 E_n$
$\vdots$	$\vdots$
$B_s \cdot \ldots \cdot B_1 A$	$B_s \cdot \ldots \cdot B_1 E_n$

d. h.: man schreibe E_n rechts neben A und führe parallel zu A die gleichen Zeilentransformationen an E_n durch. Ist man links bei E_n angekommen, so steht rechts die inverse Matrix A^{-1}.

Bemerkungen: 1. Das Verfahren bricht automatisch ab, wenn A nicht invertierbar ist.

2. Man kann stattdessen auch (*nur*) Spaltentransformationen bei diesem Prozeß verwenden.

Man kann dieses Verfahren für beliebige Matrizen verallgemeinern und zu einem Algorithmus zum Lösen des linearen Gleichungssystem

$$(*) \qquad A\,x \,=\, b\,, \quad A \in M\,(m \times n,\, \mathbb{K})\,, \quad b \in \mathbb{K}^n$$

heranziehen:

Nach Satz 12 gibt es ein Produkt $S \,:=\, B_1 \cdot \ldots \cdot B_s$ von Elementarmatrizen, s. d. $B \,:=\, S\,A$ Zeilenstufenform mit Einsen als erstem nichtverschwindenden Eintrag in jeder Zeile besitzt. Es ist aber klar, daß sich dann durch *Spaltenpermutationen*, also einfache *Spaltenumformungen* $C_1, \ldots, C_t$, die Matrix B in unsere Normalform N vom Rang $r \,=\, \operatorname{rang} A$ überführen läßt:

$$S\,A\,T^{-1} \,=\, N\,, \quad T^{-1} \,=\, C_1 \cdot \ldots \cdot C_t\,.$$

Wir visualisieren diese Operationen an dem folgenden Schema:

E_m	A	b	
B_1E_m	B_1A	B_1b	
$\vdots$	$\vdots$	$\vdots$	
$S \,=\, B_s \cdots B_1$	$B \,=\, SA$	$b' \,=\, Sb$	E_n
	BC_1		E_nC_1
	$\vdots$		$\vdots$
	$N \,=\, BT^{-1}$		$T^{-1} \,=\, C_1 \cdots C_t$

Dies benutzt man jetzt wie folgt für das Lösen des Gleichungssystems $(*)$: Man schleppt bei den Zeilentransformationen von A den Vektor b mit (ohne auf die Veränderungen bei b zu achten) und setzt $b' = Sb$. Man schreibt ferner $X \,=\, (x^{(1)}, \ldots, x^{(r)},\, x^{(r+1)}, \ldots, x^{(n)})$ für die Spalten von T^{-1}, die augenscheinlich linear unabhängig über $\mathbb{K}^n$ sind. - Dann gilt:

Satz 7.15 *Es sei* $A \in M\,(m \times n,\, \mathbb{K})$, $b \in \mathbb{K}^n$, $S \in \mathrm{GL}\,(m,\, \mathbb{K})$ *und* $X \,=\, (x^{(1)}, \ldots, x^{(n)}) \in \mathrm{GL}\,(n,\, \mathbb{K})$, *so daß*

a) $S\,A$ *Zeilenstufenform vom Rang* r *hat und*

b) $N \,=\, S\,A\,X$ *Normalgestalt*

$$N \,=\, \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right) \in M\,(m \times n,\, \mathbb{K})$$

besitzt. Dann gilt :

i) *Die Vektoren* $x^{(r+1)}, \ldots, x^{(n)}$ *bilden eine Basis des Lösungsraumes des homogenen Systems* $A\,x \,=\, 0$.

ii) $A\,x \,=\, b$ *ist genau dann lösbar, wenn für* $b' \,=\, S\,b \,=\, \begin{pmatrix} b'_1 \\ \vdots \\ b'_m \end{pmatrix}$ *gilt* :

$b'_{r+1} \,=\, \cdots \,=\, b'_m \,=\, 0$. *Eine spezielle Lösung ist unter dieser Bedingung gegeben durch*

$$b'_1\,x^{(1)} \,+ \cdots +\, b'_r\,x^{(r)}\,.$$

Bemerkung. Bei der Anwendung des Satzes braucht die Matrix S nicht einmal explizit berechnet zu werden; man benötigt nur b' und X.

Beweis. i) Aus der Beziehung $(S\,A)\,X = N$ folgt durch Vergleich der letzten $n-r$ Spalten unmittelbar

$$(S\,A)\,x^{(\rho)} = 0\,, \quad \rho = r+1, \ldots, n\,.$$

Ferner ist S invertierbar, so daß wir sofort $A\,x^{(r+1)} = \cdots = A\,x^{(n)} = 0$ schließen können. Da die Vektoren $(x^{(r+1)}, \ldots, x^{(n)})$ als Spalten einer invertierbaren Matrix linear unabhängig sind und r gleich dem Rang von A ist, müssen sie den Lösungsraum des homogenen Systems aufspannen.

ii) Da S invertierbar ist, gilt $A\,x = b$ genau dann, wenn die Gleichung $S\,A\,x = S\,b = b'$ erfüllt ist. Nun ist $S\,A$ in Zeilenstufenform mit Rang $= r$. Also ist nach dem allgemeinen Lösungskriterium $r = \operatorname{rang}(S\,A,\, b')$ eine notwendige und hinreichende Bedingung für die Lösbarkeit, und dies ist wegen der besonderen Gestalt von $S\,A$ gleichbedeutend mit $b'_{r+1} = \cdots = b'_m = 0$. Umgekehrt berechnet man unter dieser Voraussetzung leicht

$$S\,A\,T^{-1}\,b' = N\,b' = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right) \left(\begin{array}{c} b'_1 \\ \vdots \\ b'_r \\ \hline 0 \\ \vdots \\ 0 \end{array}\right) = b' = S\,b\,,$$

also $A\,(T^{-1}\,b') = S^{-1}\,S\,b = b$. Damit ist

$$T^{-1}b' = X \cdot \left(\begin{array}{c} b'_1 \\ \vdots \\ b'_r \\ \hline 0 \\ \vdots \\ 0 \end{array}\right) = b'_1\,x^{(1)} + \cdots + b'_r\,x^{(r)}$$

als spezielle Lösung erkannt. □

Beispiel. Wir betrachten das folgende Gleichungssystem:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 10 \\ 14 \\ 18 \end{pmatrix}\,.$$

Durch elementare Zeilenumformungen der erweiterten Matrix $(A,\, b)$ erhält man in drei Schritten

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 10 \\ 2 & 3 & 4 & 5 & 14 \\ 3 & 4 & 5 & 6 & 18 \end{array}\right) \longmapsto \cdots \longmapsto \cdots \longmapsto \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 10 \\ 0 & 1 & 2 & 3 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)\,.$$

Hier hat SA Zeilenstufenform, und b' hat die gewünschte Gestalt. Als nächstes führen wir elementare

Spaltentransformationen simultan für das Paar $(SA\,,\ E_n)$ durch:

$$\left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ & & & & 0 & 0 & 0 & 1 \end{array}\right) \longmapsto \left(\begin{array}{cccc|rrrr} 1 & 0 & 0 & 0 & 1 & -2 & -3 & -4 \\ 0 & 1 & 2 & 3 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ & & & & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\longmapsto \quad \cdots \quad \longmapsto \left(\begin{array}{cccc|rrrr} 1 & 0 & 0 & 0 & 1 & -2 & 1 & 2 \\ 0 & 1 & 0 & 0 & 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ & & & & 0 & 0 & 0 & 1 \end{array}\right).$$

Damit bilden die beiden letzten Spaltenvektoren

$$\begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

eine Basis des homogenen Gleichungssystems $A\,x = 0$, und eine spezielle Lösung wird gegeben durch

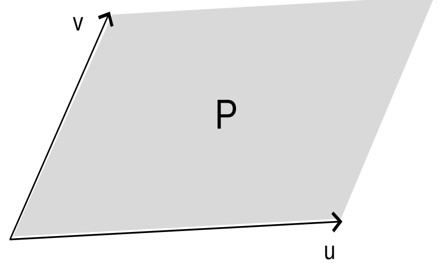
$$10 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 6 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 6 \\ 0 \\ 0 \end{pmatrix}.$$

8 Determinanten

Wir wollen die Theorie der *Determinanten* zuerst anhand von *geometrischen* Überlegungen motivieren.

Dazu betrachten wir zwei Vektoren u, v im (nicht notwendig euklidischen) Vektorraum $\mathbb{R}^2$. Diese definieren ein *Parallelogramm*

$$P = P(u, v) = \{x = ru + sv \in \mathbb{R}^2 : 0 \leq r, s \leq 1\} .$$



Figur 8.1

Wir fragen uns, welche Eigenschaften der *Flächeninhalt*

$$F(u, v) = \text{ Flächeninhalt von } P(u, v)$$

hat (oder zumindest haben sollte). Offensichtlich sind die folgenden Regeln:

a) $F(tu, v) = |t| F(u, v)$,

b) $F(u + tv, v) = F(u, v)$,

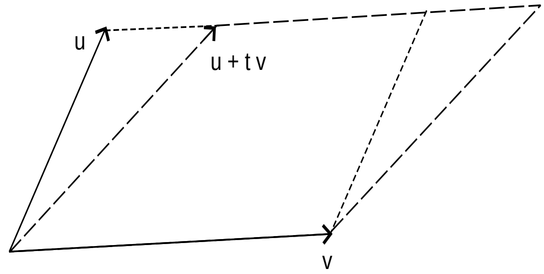
(es handelt sich hierbei um eine *Scherung*, siehe Figur 2),

c) $F(u, u) = 0$,

d) $F(u, v) = F(v, u)$,

und aus Normierungsgründen verlangt man z. B. noch

e) $F(e_1, e_2) = 1$.



Figur 8.2

Insbesondere interessiert uns die Frage: Ist F bestimmt durch die obigen Eigenschaften ? D. h. genauer: Welche Aussagen kann man machen über die Funktion

$$F(u_1 e_1 + u_2 e_2, v_1 e_1 + v_2 e_2)$$

in Abhängigkeit von den Variablen u_1, u_2, v_1, v_2? Ist also z. B. diese Funktion additiv in beiden Argumenten ? Die Antwort hierzu ist leider nein wegen a) (man braucht dort nur t negativ zu wählen). Es gilt genauer:

$$F(u + v, w) = \begin{cases} F(u, w) + F(v, w)\,, & u, v \text{ auf „derselben Seite" von } w\,, \\ |\,F(v, w) - F(u, w)\,|\,, & u, v \text{ auf „verschiedenen Seiten" von } w\,. \end{cases}$$

Es kommt also auf die *Orientierung* des Systems der betroffenen Vektoren an. Dieses Konzept werden wir später noch präzisieren (Anhang zu Kapitel 13). Im Moment genügt uns die intuitive Vorstellung eines positiv oder negativ orientierten Systems von zwei Vektoren u, v. Die Schwierigkeiten werden tatsächlich behoben, wenn man den *orientierten* Flächeninhalt einführt:

$$D(u, v) = \begin{cases} F(u, v)\,, & v \text{ „links" von } u\,, \\ -F(u, v)\,, & v \text{ „rechts" von } u\,, \\ 0\,, & u, v \text{ linear abhängig}\,. \end{cases}$$

Man zeigt damit leicht:

a) $D(tu, v) = t\,D(u, v)$.

b) $D(u + tv, v) = D(u, v)$.

c) $D(u_1 + u_2, v) = D(u_1, v) + D(u_2, v)$.

d) $D(u, u) = 0$.

e) $D(u, v) = -D(v, u)$ $\quad (\Longrightarrow D(u, v_1 + v_2) = D(u, v_1) + D(u, v_2))$.

f) $D(e_1, e_2) = 1$, falls (e_1, e_2) eine positiv orientierte Basis bezeichnet.

D ist tatsächlich, wie wir bald sehen werden, allein durch die Eigenschaften a), c), e) und f) bestimmt und damit auch $F = |\,D\,|$.

Ähnliche Überlegungen im n–dimensionalen Raum ($n \geq 3$) führen unmittelbar zu dem folgenden abstrakten Begriff.

Definition. Es sei V ein $\mathbb{K}$–Vektorraum. Eine (*Multi–*) *Linearform* vom Grad ℓ (oder kurz: eine ℓ–*Form*) auf V ist eine Abbildung

$$F : V^\ell := \underbrace{V \times \cdots \times V}_{\ell-\text{mal}} \longrightarrow \mathbb{K}\,,$$

die in jedem Argument linear ist; d. h. bei festem $j \in \{1, \ldots, \ell\}$ und festen Vektoren $u_1, \ldots, u_{j-1}$, $u_{j+1}, \ldots, u_\ell \in V$ ist die Abbildung

$$V \ni u \longmapsto F(u_1, \ldots, u_{j-1}, u, u_{j+1}, \ldots, u_\ell) \in \mathbb{K}$$

$\mathbb{K}$–linear. Wir schreiben für die Gesamtheit dieser Formen

$$\mathrm{Mult}^\ell_{\mathbb{K}}(V) = \{F : F \text{ ist eine } \ell\text{–Form auf } V\}\,.$$

Die Vorstellung von einem ℓ–dimensionalen Volumen des *Polytops*

$$P(u_1, \ldots, u_\ell) = \{v = t_1 u_1 + \cdots + t_\ell u_\ell \in \mathbb{R}^n : 0 \leq t_\lambda \leq 1\,,\ \lambda = 1, \ldots, \ell\}$$

führt zu der

Definition. Eine ℓ–Form $F : V^\ell \to \mathbb{K}$ heißt *alternierend*, falls $F(u_1, \ldots, u_\ell) = 0$, sofern mindestens zwei der Vektoren $u_1, \ldots, u_\ell$ übereinstimmen. Wir schreiben

$$\mathrm{Alt}^\ell_{\mathbb{K}}(V)$$

für die Gesamtheit der alternierenden ℓ–Formen auf V.

Leicht zu sehen ist: $\mathrm{Mult}^\ell_{\mathbb{K}}(V)$ bildet als Teilmenge von $\mathrm{Abb}\,(V^\ell, \mathbb{K})$ einen Untervektorraum, ebenso $\mathrm{Alt}^\ell_{\mathbb{K}}(V)$. Wir versehen beide Räume stets mit dieser $\mathbb{K}$–Vektorraumstruktur. Ist nun V endlich dimensional, so ist die Dimension dieser Räume interessant. Wir setzen $\dim V = n$. Klar ist (warum ?):

Satz 8.1
$$\dim_{\mathbb{K}} \mathrm{Mult}^\ell_{\mathbb{K}}(V) = n^\ell .$$

Wir verzichten auch auf einen Beweis des folgenden allgemeinen Satzes über den Vektorraum der alternierenden Formen, den man ohne größere Schwierigkeiten hier durchführen könnte. (Siehe auch die Bemerkung am Ende des Beweises von Satz 3).

Satz 8.2
$$\dim_{\mathbb{K}} \mathrm{Alt}^\ell_{\mathbb{K}}(V) = \begin{cases} \binom{n}{\ell} , & 1 \le \ell \le n , \\ \\ 0 , & \ell > n . \end{cases}$$

Bemerkung. Im Falle $\ell = 1$ ist $\mathrm{Mult}^1_{\mathbb{K}}(V) = \mathrm{Alt}^1_{\mathbb{K}}(V) = V^*$ der *Dualraum* zu V mit

$$\dim V^* = \dim V = n = \binom{n}{1}$$

(siehe Kapitel 9).

Wir interessieren uns hier speziell für den Fall $\ell = n$, wo nach dem obigen Satz gelten muß:

$$\dim_{\mathbb{K}} \mathrm{Alt}^n_{\mathbb{K}}(V) = 1 , \quad \text{wenn} \quad n = \dim V .$$

Wir werden tatsächlich die folgende präzisere Aussage beweisen:

Satz 8.3 *Es sei V ein n–dimensionaler $\mathbb{K}$–Vektorraum. Dann gibt es zu jeder geordneten Basis $\mathcal{A} = (v_1, \ldots, v_n)$ von V genau eine alternierende n–Form $D_{\mathcal{A}} : V^n \to \mathbb{K}$ mit $D_{\mathcal{A}}(v_1, \ldots, v_n) = 1$. Ist $F \in \mathrm{Alt}^n_{\mathbb{K}}(V)$ beliebig und $\lambda := F(v_1, \ldots, v_n)$, so gilt*

$$F = \lambda \cdot D_{\mathcal{A}} .$$

Zum *Beweis* dieses Satzes muß man das Verhalten von alternierenden Formen unter *Permutationen der Einträge* genauer studieren. Es sei $F \in \mathrm{Alt}^\ell_{\mathbb{K}}(V)$, $u_1, \ldots, u_\ell \in V$ seien beliebig, und es seien die Indizes j, k mit $1 \le j < k \le \ell$ fest gewählt. Dann ergibt sich:

$$\begin{aligned} 0 &= F(u_1, \ldots, \overbrace{u_j + u_k}^{j}, \ldots, \overbrace{u_j + u_k}^{k}, \ldots, u_\ell) \\ &= F(u_1, \ldots, u_j, \ldots, u_j + u_k, \ldots, u_\ell) + F(u_1, \ldots, u_k, \ldots, u_j + u_k, \ldots, u_\ell) \quad . \\ &= F(u_1, \ldots, u_j, \ldots, u_j, \ldots, u_\ell) + F(u_1, \ldots, u_j, \ldots, u_k, \ldots, u_\ell) + \\ & F(u_1, \ldots, u_k, \ldots, u_j, \ldots, u_\ell) + F(u_1, \ldots, u_k, \ldots, u_k, \ldots, u_\ell) . \end{aligned}$$

Wir haben damit gezeigt:

Satz 8.4 *Ist* $F \in \mathrm{Alt}^{\ell}_{\mathbb{K}}(V)$ *, so gilt*

$$F(u_1, \dots, u_k, \dots, u_j, \dots, u_\ell) = -F(u_1, \dots, u_j, \dots, u_k, \dots, u_\ell) .$$

M. a. W.: F vertauscht das Vorzeichen beim Vertauschen zweier Einträge.

Bemerkung. Diese Bedingung rechtfertigt den Begriff *alternierend* . Sie ist gleichbedeutend mit der obigen Definition für alternierende Formen, wenn in $\mathbb{K}$ gilt: $1 + 1 \neq 0$, d. h. wenn $\mathrm{char}\, \mathbb{K} \neq 2$ ist. Denn:

$$F(u_1, \dots, u, \dots, u, \dots, u_\ell) = -F(u_1, \dots, u, \dots, u, \dots, u_\ell)$$

impliziert $2F(u_1, \dots, u, \dots, u, \dots, u_\ell) = 0$ und damit

$$F(u_1, \dots, u, \dots, u, \dots, u_\ell) = 0 .$$

Vertauschen ist eine spezielle *Permutation*: Es bezeichne $\mathfrak{S}_\ell$ die sogenannte *symmetrische Gruppe* in ℓ Elementen:

$$\mathfrak{S}_\ell = \{\text{bijektive Abbildungen } \sigma : \{1, \dots, \ell\} \longrightarrow \{1, \dots, \ell\} \} .$$

Natürlich bildet diese Menge eine Gruppe bzgl. der Hintereinanderschaltung von Abbildungen. Man schreibt auch

$$\sigma = \begin{pmatrix} 1 & \cdots & \ell \\ \sigma(1) & \cdots & \sigma(\ell) \end{pmatrix} .$$

Man beachte, daß die *Mengen* $\{\sigma(1), \dots, \sigma(\ell)\}$ und $\{1, \dots, \ell\}$ übereinstimmen. Ist nun $F \in \mathrm{Alt}^{\ell}_{\mathbb{K}}(V)$, so setze man

$$(\sigma F)(u_1, \dots, u_\ell) := F(u_{\sigma(1)}, \dots, u_{\sigma(\ell)}) .$$

Leicht zu beweisen ist:

$$\sigma F \in \mathrm{Alt}^{\ell}_{\mathbb{K}}(V) , \quad (\sigma_2\sigma_1)F = \sigma_2(\sigma_1 F) .$$

σ heißt eine *Transposition*, falls es ein Paar (j, k) mit $1 \leq j < k \leq \ell$ und $\sigma(j) = k$, $\sigma(k) = j$, $\sigma(i) = i$, $i \neq j, k$ gibt. (Man schreibt dann lieber τ statt σ). Satz 4 impliziert:

$$\tau F = -F , \quad \tau \text{ eine Transposition} .$$

Ist $\sigma = \tau_m \cdot \ldots \cdot \tau_1$ ein Produkt von Transpositionen, so folgt nach der gerade gemachten Bemerkung

$$\sigma F = \tau_m(\cdots(\tau_1 F)) = (-1)^m F .$$

Insbesondere ergibt sich, daß der Faktor $(-1)^m$ eindeutig bestimmt ist, wenn $F \neq 0$.

Eine beliebte elementare kombinatorische Frage lautet: *Wie ordnet man einen in Unordnung geratenen 24–bändigen Brockhaus*? Ein zum Ziele führender, wenn auch evtl. nicht schneller Algorithmus besteht darin, daß man Band 1 sucht und mit dem an erster Stelle stehenden Band vertauscht, dann sucht man Band 2 und vertauscht diesen mit dem an zweiter Stelle stehenden Band u.s.w. Auf jeden Fall schließt man hieraus leicht die erste Aussage in dem folgenden

Satz 8.5 *Jede Permutation σ ist Produkt von Transpositionen: $\sigma = \tau_1 \cdot \ldots \cdot \tau_m$, und die Zahl*

$$\mathrm{sign}\, \sigma = (-1)^m \quad (\text{Signum } \textit{oder } \text{Parität } \textit{von } \sigma)$$

ist eindeutig durch σ bestimmt (nicht jedoch m). Die Abbildung

$$\mathrm{sign} : \mathfrak{S}_\ell \longrightarrow \{1, -1\}$$

ist ein Gruppenhomomorphismus (rechts bzgl. Multiplikation). Die Menge

$$\mathfrak{A}_\ell = \{\sigma \in \mathfrak{S}_\ell : \mathrm{sign}\, \sigma = 1\}$$

bildet eine Untergruppe (die sogenannte alternierende Gruppe) von $\mathfrak{S}_\ell$ der Ordnung $\frac{1}{2}\, \mathrm{ord}\, \mathfrak{S}_\ell = \frac{1}{2} \ell\,!$.

Folgerung 8.6 $\sigma F \,=\, (\operatorname{sign}\sigma)\, F\,, \quad F \in \mathrm{Alt}_{\mathbb{K}}^{\ell}(V)\,, \quad \sigma \in \mathfrak{S}_\ell\,.$

Wir können jetzt im Falle $\ell = n$ den Hauptsatz beweisen , indem wir zunächst zeigen, daß ein beliebiges Element $F \in \mathrm{Alt}_{\mathbb{K}}^{n}(V)$ durch den Wert $F\,(v_1, \ldots, v_n)$ bei vorgegebener Basis $\mathcal{A} = (v_1, \ldots, v_n)$ bestimmt ist. Es gilt

$$\begin{aligned} F\,(u_1, \ldots, u_n) \,&=\, F\,(a_{11}\, v_1 \,+ \cdots +\, a_{n1}\, v_n, \ldots, a_{1n}\, v_1 \,+ \cdots +\, a_{nn}\, v_n) \\ &=\, \sum_{j_1=1}^{n} a_{j_1 1}\, F\,(v_{j_1},\, a_{12}\, v_1 \,+ \cdots +\, a_{n2}\, v_n, \ldots) \,=\, \cdots \\ &=\, \sum_{j_1,\ldots,j_n=1}^{n} a_{j_1 1}\, a_{j_2 2} \cdot \ldots \cdot a_{j_n n}\, F\,(v_{j_1},\, v_{j_2}, \ldots, v_{j_n})\,. \end{aligned}$$

Wir unterscheiden zwei Fälle:

a) $\{j_1, \ldots, j_n\} \subsetneqq \{1, \ldots, n\} \Longrightarrow F\,(v_{j_1}, \ldots, v_{j_n}) \,=\, 0$, da mindestens zwei Einträge gleich sind.

b) $\{j_1, \ldots, j_n\} \,=\, \{1, \ldots, n\}$, d. h. $\exists\, \sigma \in \mathfrak{S}_n$ mit $j_1 \,=\, \sigma\,(1), \ldots, j_n \,=\, \sigma\,(n)$. Damit folgt

$$\begin{aligned} F\,(u_1, \ldots, u_n) \,&=\, \sum_{\sigma \in \mathfrak{S}_n} a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} F\,(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) \\ &=\, \Big(\sum_{\sigma \in \mathfrak{S}_n} (\operatorname{sign}\sigma)\, a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} \Big) \cdot F\,(v_1, \ldots, v_n)\,. \end{aligned}$$

Also ist der Hauptsatz äquivalent zu der Aussage, daß für

$$u_1 \,=\, a_{11}\, v_1 \,+ \cdots +\, a_{n1}\, v_n, \ldots, u_n \,=\, a_{1n}\, v_1 \,+ \cdots +\, a_{nn}\, v_n$$

durch

$$D_{\mathcal{A}}\,(u_1, \ldots, u_n) \,:=\, \sum_{\sigma \in \mathfrak{S}_n} (\operatorname{sign}\sigma)\, a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}$$

eine Form in $\mathrm{Alt}_{\mathbb{K}}^{n}(V)$ definiert wird mit $D_{\mathcal{A}}\,(v_1, \ldots, v_n) \,=\, 1\,.$

Dies kann man nun direkt nachrechnen, obwohl die Durchführung nicht völlig trivial ist; so benutzt der Nachweis der Alternierungs–Eigenschaft von $D_{\mathcal{A}}$ die Zerlegung $\mathfrak{S}_n \,=\, \mathfrak{A}_n \cup \tau\mathfrak{A}_n$, τ eine beliebige Transposition. $\square$

Bemerkung. Vollständig analoge Überlegungen zeigen, daß eine alternierende ℓ–Form $F \in \mathrm{Alt}_{\mathbb{K}}^{\ell}(V)$ bei beliebigem ℓ durch die Werte $F\,(v_{j_1}, \ldots, v_{j_\ell})$, $1 \le j_1 < \cdots < j_\ell \le n$, bestimmt ist, und das sind genau $\binom{n}{\ell}$ Vorgaben in Übereinstimmung mit Satz 2.

Als Konsequenz unserer bisherigen Untersuchungen notieren wir:

Satz 8.7 (und Definition) *Es gibt genau eine Funktion*

$$\det : M\,(n \times n,\, \mathbb{K}) \longrightarrow \mathbb{K}\,,$$

die

1. *multilinear ist in den Spalten,*
2. *alternierend ist in den Spalten,*
3. *normiert ist in dem Sinne, daß* $\det E_n \,=\, 1\,.$

Sie wird gegeben durch

$$(*) \qquad \det\,(a_{jk})_{1\le j,\,k\le n} \;=\; \sum_{\sigma\in\mathfrak{S}_n} (\operatorname{sign}\sigma)\, a_{\sigma(1),1}\cdot\ldots\cdot a_{\sigma(n),n}\,.$$

Für jede andere Funktion D *mit* 1. *und* 2. *gilt*

$$D \,=\, \lambda \det\,, \quad \lambda \,=\, D\,(e_1,\ldots,e_n)\,.$$

Für $A\in M\,(n\times n,\,\mathbb{K})$ *heißt* $\det A$ *die Determinante von* A.

Wir sammeln noch einige Eigenschaften der Determinante. Schreiben wir wie früher

$$A \,=\, \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \,=\, (a^1,\ldots,a^n)\,, \quad a^j \,=\, \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}\,,$$

so gelten die folgenden Regeln:

1. $\det\,(a^1,\ldots,a^j \,+\, \lambda\, b^j,\ldots,a^n) \,=\, \det\,(a^1,\ldots,a^j,\ldots,a^n) \,+\, \lambda\,\det\,(a^1,\ldots,b^j,\ldots,a^n)\,.$
2. $\det\,(a^1,\ldots,a^n) \,=\, 0$, falls $a^j \,=\, a^k$ für ein Paar $j\,\neq\,k$.
3. $\det\,(\lambda\,A) \,=\, \det\,(\lambda a^1,\ldots,\lambda a^n) \,=\, \lambda^n\,\det A\,.$
4. $\det A \,=\, 0$, falls ein $a^j \,=\, 0$.
5. $\det\,(a^{\sigma(1)},\ldots,a^{\sigma(n)}) \,=\, (\operatorname{sign}\sigma)\,\det\,(a^1,\ldots,a^n)\,.$
6. $\det\,(a^1,\ldots,a^j+\lambda\,a^k,\ldots,a^n) \,=\, \det\,(a^1,\ldots,a^n)\,, \quad j\neq k\,.$
7. Für eine obere Dreiecksmatrix
$$A \,=\, \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}\,.$$
ist $\det A \,=\, \lambda_1\cdot\ldots\cdot\lambda_n$.
8. $\det {}^tA \,=\, \det A\,.$ (Man kommt also zur gleichen Determinanten–Funktion, wenn man mit Zeilen statt mit Spalten arbeitet.)
9. $\det\,(A\,B) \,=\, \det A\cdot\det B$ für alle $A,\,B\in M\,(n\times n,\,\mathbb{K})\,.$
10. $\det A\neq 0$ genau dann, wenn A invertierbar ist. Es gilt dann $\det A^{-1} \,=\, (\det A)^{-1}\,.$

Beweis. Wir behandeln nur die nichttrivialen Punkte 7 bis 10.

ad 7: $\lambda_1 \,=\, 0 \Longrightarrow \det A \,=\, 0 = \lambda_1\cdot\ldots\cdot\lambda_n$ wegen 4. Ist dagegen $\lambda_1\neq 0$, so folgt aus der Definition der Determinante

$$\det A \,=\, \lambda_1\cdot\det \begin{pmatrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

und damit die Behauptung per Induktion. □

Diese Eigenschaft ist übrigens ein Spezialfall von

Satz 8.8 (Laplace - Entwicklung) *Seien j, k mit $1 \le j, k \le n$ fest, und A'_{jk} sei die Matrix in $M((n-1) \times (n-1), \mathbb{K})$, die man erhält, wenn man aus der Matrix A die j–te Zeile und die k–te Spalte streicht. Dann gilt (bei festem k):*

$$\det A = \sum_{j=1}^{n} (-1)^{j+k} a_{jk} \det A'_{jk} .$$

Beweis. Sammeln in $(*)$ liefert

$$\det A = \sum_{j=1}^{n} a_{jk} \sum_{\substack{\sigma \in \mathfrak{S}_n \\ \sigma(k)=j}} \cdots . \qquad \square$$

Wir führen den *Beweis* der Determinanten–Eigenschaften fort.

ad 8: Nach Definition gilt

$$\begin{aligned} \det {}^tA &= \sum_{\sigma \in \mathfrak{S}_n} (\operatorname{sign} \sigma)\, {}^ta_{\sigma(1)1} \cdot \ldots \cdot {}^ta_{\sigma(n)n} \\ &= \sum_{\sigma \in \mathfrak{S}_n} (\operatorname{sign} \sigma)\, a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathfrak{S}_n} (\operatorname{sign} \sigma)\, a_{\sigma^{-1}(1)1} \cdot \ldots \cdot a_{\sigma^{-1}(n)n} . \end{aligned}$$

Nun induziert die Abbildung $\sigma \longmapsto \sigma^{-1}$ einen Isomorphismus $\mathfrak{S}_n \xrightarrow{\sim} \mathfrak{S}_n$, so daß

$$\det {}^tA = \sum_{\sigma \in \mathfrak{S}_n} (\operatorname{sign} \sigma^{-1})\, a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} .$$

Aus $\sigma^{-1} \cdot \sigma = \mathrm{id}$ folgt aber $1 = \operatorname{sign} \mathrm{id} = \operatorname{sign} \sigma^{-1} \cdot \operatorname{sign} \sigma$, also $\operatorname{sign} \sigma^{-1} = \operatorname{sign} \sigma$ und damit die Behauptung. $\square$

ad 9: Wir halten B fest und betrachten $\det(B\,A)$ als Funktion in A:

$$D_B(A) = \det(B\,A) .$$

Setzen wir

$$B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}, \quad A = (a^1, \ldots, a^n) ,$$

so erhalten wir

$$B\,A = \begin{pmatrix} b_1 \cdot a^1 & b_1 \cdot a^2 & \cdots & b_1 \cdot a^n \\ \vdots & \vdots & & \vdots \\ b_n \cdot a^1 & b_n \cdot a^2 & \cdots & b_n \cdot a^n \end{pmatrix} .$$

Also ist $\det(B\,A)$ multilinear und alternierend in den Spalten von A, woraus sich

$$D_B(A) = \lambda_B \det A , \quad \lambda_B = D_B(E) = \det(B\,E) = \det B$$

ergibt, also die Behauptung $\det(B\,A) = D_B(A) = \det B \cdot \det A$ für alle A, B. $\square$

ad 10: a) A sei invertierbar $\Longrightarrow A^{-1} \cdot A = E \Longrightarrow \det A^{-1} \det A = \det E = 1 \Longrightarrow \det A \ne 0$ (und $\det A^{-1} = (\det A)^{-1}$).

b) Ist A nicht invertierbar, so ist $\operatorname{rang} A = r < n$, und es gibt invertierbare Matrizen S, T mit

$$S\,A\,T^{-1} = N = \left(\begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array}\right) .$$

Dann gilt $A = S^{-1}\,N\,T$ und damit $\det A = (\det S)^{-1} \cdot \det N \cdot \det T = 0$ wegen $\det N = 0$ (nach 7.). $\square$

Warnung. I. a. gilt $\det(A + B) \neq \det A + \det B$!

Bemerkung. Mit der Eigenschaft 10 läßt sich sehr einfach der Nachweis führen, daß die Menge der invertierbaren Matrizen eine Gruppe bzgl. Multiplikation bildet.

Die Formel $(*)$ in Satz 7 für die Determinante ist für große n viel zu unhandlich, so daß sich wieder die Frage nach *effektiven* Methoden zur Berechnung stellt. Ein möglicher Algorithmus sieht wie folgt aus: Spaltenumformungen vom Typ III und IV bringen A in Spaltenstufenform B. Dabei ist klar:

$$\det A = (-1)^k \det B \,, \quad k = \text{Anzahl der Umformungen vom Typ IV} .$$

Nun ist eine quadratische Matrix in Spaltenstufenform eine untere Dreiecksmatrix, also die Transponierte einer oberen Dreiecksmatrix, und damit gilt

$$\det B = \prod_{j=1}^{n} \lambda_j \,, \quad \text{wenn} \quad B = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{pmatrix} .$$

Eine weitere nützliche Formel ist die folgende: Ist die Matrix A von der Gestalt

$$A = \left(\begin{array}{c|c} A_1 & C \\ \hline 0 & A_2 \end{array}\right)$$

mit quadratischen Matrizen A_1 und A_2, so gilt

$$\det A = \det A_1 \cdot \det A_2 .$$

(Diese Aussage wird in der Übungsaufgabe 49 wesentlich verallgemeinert).

Wir kommen am Ende noch zu dem Begriff der *klassischen Adjungierten* und zur *Cramerschen Regel*: Zu der Matrix $A \in M(n \times n, \mathbb{K})$ assoziieren wir eine neue Matrix $\widetilde{A} = (\widetilde{a}_{jk})$ durch

$$\widetilde{a}_{jk} = (-1)^{k+j} \det A'_{kj} \,, \quad \text{wobei} \quad A'_{kj} \in M((n-1) \times (n-1), \mathbb{K})$$

wie früher aus A durch Streichen der k–ten Zeile und j–ten Spalte hervorgeht. Wir nennen $\widetilde{A}$ die *klassische Adjungierte* zu A. Aufgrund des Laplaceschen Entwicklungssatzes erhält man sofort die Darstellung

$$\widetilde{a}_{jk} = \det(a^1, \ldots, a^{j-1}, \underbrace{e^k}_{j}, a^{j+1}, \ldots, a^n) .$$

Damit bekommt man leicht:

Lemma 8.9

$$\widetilde{A}\,A = A\,\widetilde{A} = (\det A)\,E_n .$$

Beweis. Es ist

$$\begin{aligned}\sum_{k=1}^{n} \widetilde{a}_{jk}\, a_{k\ell} &= \sum_{k=1}^{n} a_{k\ell}\, \det(a^1, \ldots, a^{j-1}, \overbrace{e^k}^{j}, a^{j+1}, \ldots, a^n) \\ &= \det(a^1, \ldots, a^{j-1}, \sum_{k=1}^{n} a_{k\ell}\, e^k, a^{j+1}, \ldots, a^n) \\ &= \det(a^1, \ldots, a^{j-1}, a^\ell, a^{j+1}, \ldots, a^n) \\ &= \begin{cases} \det A\,, & \ell = j\,, \\ 0\,, & \ell \neq j\,, \end{cases}\end{aligned}$$

und folglich $\widetilde{A}\, A = (\det A)\, E_n$. Damit ist aber auch

$$A\, \widetilde{A} = {}^t({}^t\widetilde{A}\, {}^tA) = {}^t((\det {}^tA)\, E_n) = (\det A)\, E_n\,. \qquad \square$$

Wir kommen noch einmal zu linearen Gleichungssystemen im Spezialfall $m = n$ zurück. Es sei also $A \in M(n \times n, \mathbb{K})$ und $(*)\ A\,x = b \in \mathbb{K}^n$ das zugehörige Gleichungssystem. Dieses ist *universell lösbar*, d. h. lösbar für alle rechten Seiten b, wenn $\operatorname{rang} A = n$, d. h. wenn $\det A \neq 0$. Unter dieser Voraussetzung kann man die Lösung konkret durch Determinanten beschreiben.

Satz 8.10 (Cramersche Regel) *Es sei $A \in \mathrm{GL}(n, \mathbb{K})$. Dann hat die Gleichung $A\,x = b \in \mathbb{K}^n$ die eindeutig bestimmte Lösung*

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \textit{wobei} \quad x_k = \frac{\det(a^1, \ldots, \overbrace{b}^{k}, \ldots, a^n)}{\det(a^1, \ldots, a^n)}\,, \quad k = 1, \ldots, n\,.$$

Beweis. Mit $b = \sum_{\ell=1}^{n} b_\ell e^\ell$ gilt

$$\det(a^1, \ldots, \overbrace{b}^{k}, \ldots, a^n) = \sum_{\ell=1}^{n} b_\ell\, \det(a^1, \ldots, \overbrace{e^\ell}^{k}, \ldots, a^n) = \sum_{\ell=1}^{n} b_\ell\, \widetilde{a}_{k\ell}$$

und damit für alle $j = 1, \ldots, n$:

$$\begin{aligned}\sum_{k=1}^{n} a_{jk}\, x_k &= \frac{1}{\det A} \sum_{k=1}^{n} a_{jk}\, \det(a^1, \ldots, \overbrace{b}^{k}, \ldots, a^n) \\ &= \frac{1}{\det A} \sum_{k=1}^{n} a_{jk} \sum_{\ell=1}^{n} b_\ell\, \widetilde{a}_{k\ell} = \frac{1}{\det A} \sum_{\ell=1}^{n} b_\ell \sum_{k=1}^{n} a_{jk}\, \widetilde{a}_{k\ell} \\ &= \frac{1}{\det A} \sum_{\ell=1}^{n} b_\ell\, \delta_{j\ell}\, (\det A) = b_j\,. \qquad \square\end{aligned}$$

9 Dualräume

Im folgenden seien wie bisher V und W $\mathbb{K}$–Vektorräume. Eine wesentliche Rolle spielte für uns bisher der *Vektorraum der linearen Abbildungen*

$$\mathrm{Hom}\,(V,\,W) \,=\, \{F : V \longrightarrow W :\; F \;\text{ ist }\; \mathbb{K}\text{–linear}\,\}\;.$$

Einen Spezialfall stellt $W = \mathbb{K}$ dar:

$$V^* \,:=\, \mathrm{Hom}\,(V,\,\mathbb{K})$$

heißt der *Dualraum* von V, seine Elemente

$$\lambda : V \longrightarrow \mathbb{K}$$

heißen *Linearformen* auf V (oder *lineare Funktionale*). Sie erfüllen nach Definition die Beziehung

$$\lambda\,(av + bw) \,=\, a\,\lambda\,(v) + b\,\lambda\,(w) \quad \text{für alle} \quad v\,,\; w \in V\,,\; a\,,\; b \in \mathbb{K}\,.$$

Beispiele. 1. Für einen Zeilenvektor $a = (a_1, \ldots, a_n) \in \mathbb{K}^n$ sei λ_a definiert durch

$$\lambda_a : \left\{ \begin{array}{ccc} \mathbb{K}^n & \longrightarrow & \mathbb{K} \\[2ex] x \,=\, {}^t(x_1, \ldots, x_n) & \longmapsto & a \cdot x \,=\, \displaystyle\sum_{j=1}^{n} a_j\,x_j \;=:\; \langle\, a,\, x\,\rangle\;. \end{array} \right.$$

Offensichtlich kann man jede Gleichung in einem linearen Gleichungssystem mit Hilfe solcher λ_a ausdrücken.

2.

$$L : \left\{ \begin{array}{ccc} V \,=\, \mathrm{Konv}\,(\mathbb{N},\,\mathbb{R}) & \longrightarrow & \mathbb{R} \\[2ex] (a_j)_{j\in\mathbb{N}} & \longmapsto & \displaystyle\lim_{j\to\infty} a_j\;. \end{array} \right.$$

3. Ist $I = [\,a,\, b\,] \subset \mathbb{R}$ und $V = \mathcal{C}^0(I,\,\mathbb{R})$ oder gleich $\mathrm{Int}\,(I,\,\mathbb{R})$ im Riemannschen oder Lebesgueschen Sinne, so ist

$$V \ni f \longmapsto \int_I f\,(x)\,dx$$

eine Linearform.

Wir wissen von früher: Ist $\mathcal{B} = \{v_\iota : \iota \in I\}$ eine Basis von V, so gibt es zu jedem System $\{a_\iota : a_\iota \in \mathbb{K}\}_{\iota\in I}$ genau eine Linearform $\lambda \in V^*$ mit $\lambda\,(v_\iota) = a_\iota$. Ist speziell $\dim V = n < \infty$, so existieren bei fest gewählter Basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ eindeutig bestimmte Linearformen $v_k^* \in V^*$, $k = 1, \ldots, n$, mit

$$v_k^*(v_j) \,=\, \delta_{jk} \,=\, \begin{cases} 1\,, & j \,=\, k\,, \\ 0\,, & j \,\neq\, k\,. \end{cases}$$

Satz 9.1 *Ist* $\mathcal{B} = \{v_1, \ldots, v_n\}$ *eine Basis von* V, *so ist* $\mathcal{B}^* = \{v_1^*, \ldots, v_n^*\}$ *eine Basis von* V^*. *Insbesondere gilt* :

$$\dim V \,=\, n \,<\, \infty \quad\Longrightarrow\quad \dim V^* \,=\, \dim V\;.$$

Aus Satz 1 folgt sofort, daß die im Beispiel 1 angegebene lineare Abbildung $a \mapsto \lambda_a$ einen Isomorphismus $\mathbb{K}^n \longrightarrow (\mathbb{K}^n)^*$ induziert. Wegen des Bildungsgesetzes von λ_a identifiziert man oft $(\mathbb{K}^n)^*$ mit dem Raum der *Zeilenvektoren*.

Warnung. Aus $\dim V^* = \dim V$ folgt natürlich die Existenz von Isomorphismen $V \to V^*$. Aber keiner dieser Isomorphismen ist i.a. *basisunabhängig*, also kanonisch. Eine solche Isomorphie ist somit prinzipiell wertlos! (Eine Ausnahme bildet $V = \mathbb{K}^n$, da dieser Vektorraum eine *kanonische* Basis besitzt, der man die duale Basis eindeutig zuordnen kann, und z. B. jeder *euklidische* Vektorraum; siehe das nächste Kapitel). Im Unendlichdimensionalen gibt es dagegen nicht einmal *unkanonische* Isomorphismen zwischen V und V^*. So hat offensichtlich der Vektorraum $\mathbb{K}^{(\mathbb{N})}$ aller *endlichen* Folgen in $\mathbb{K}$ als Dualraum den Raum $\mathbb{K}^{\mathbb{N}}$ *aller* Folgen mit Werten in $\mathbb{K}$.

Beweis von Satz 1. 1. $\mathcal{B}^*$ ist linear unabhängig: Aus $\lambda = \sum_{j=1}^n a_j\, v_j^* = 0$ folgt unmittelbar

$$0 = \lambda\,(v_k) = (\sum_{j=1}^n a_j\, v_j^*)\,(v_k) = \sum_{j=1}^n a_j\, v_j^*(v_k) = \sum_{j=1}^n a_j\, \delta_{jk} = a_k$$

für alle $k = 1, \ldots, n$.

2. $\mathcal{B}^*$ erzeugt V^*: Ist $\lambda \in V^*$ und $\lambda\,(v_k) = a_k$ für $k = 1, \ldots, n$, so bilde man:

$$\lambda^* = \sum_{k=1}^n a_k\, v_k^* \in V^* .$$

Wegen $(\lambda^* - \lambda)\,(v_j) = (\sum_{k=1}^n a_k\, v_k^* - \lambda)\,(v_j) = \sum_{k=1}^n a_k\, \delta_{kj} - a_j = 0$ für alle j ergibt sich dann $\lambda = \lambda^*$ und damit $V^* \subset \operatorname{span}(v_1^*, \ldots, v_n^*)$. $\square$

Erstaunlicherweise gibt es eine völlig kanonische Abbildung von V in sein „Doppeldual“ oder „Bidual“. Dies hängt mit der kanonischen „Paarung“

$$(*) \qquad \begin{cases} V^* \times V \longrightarrow \mathbb{K} \\ (\lambda \ , \ v) \longmapsto \lambda\,(v) \end{cases}$$

zusammen, die offensichtlich linear in beiden Argumenten (also *bilinear*) ist.

Definition. Eine bilineare Abbildung

$$b : \begin{cases} V_1 \times V_2 \longrightarrow \mathbb{K} \\ (v_1 \ , \ v_2) \longmapsto b\,(v_1,\, v_2) \end{cases}$$

nennt man auch eine *Bilinearform*. Sie heißt *nicht ausgeartet im* 1. *Argument*, falls $b\,(v_1,\, v_2) = 0$ für alle v_2 die Bedingung $v_1 = 0$ nach sich zieht. (Entsprechend definiert man Nichtausgeartetheit im 2. Argument). Die Form heißt *nicht ausgeartet* schlechthin (oder auch *duale Paarung*), falls sie in beiden Argumenten nicht ausgeartet ist. Wir schreiben oft auch $\langle\, .\, ,\, .\, \rangle_b$ oder sogar $\langle\, .\, ,\, .\, \rangle$ anstelle von $b\,(\, .\, ,\, .\,)$, wenn aus dem Kontext heraus klar ist, um welche Bilinearform es sich handelt.

Offensichtlich gibt eine Bilinearform $b : V_1 \times V_2 \to \mathbb{K}$ Anlaß zu linearen Abbildungen

$$\begin{cases} V_2 \longrightarrow V_1^* \\ v_2 \longmapsto b\,(\, .\, ,\, v_2) \end{cases} \quad \text{und} \quad \begin{cases} V_1 \longrightarrow V_2^* \\ v_1 \longmapsto b\,(v_1,\, .) \end{cases}$$

(und jede solche Abbildung liefert eine Bilinearform; ist z. B. $g \in \operatorname{Hom}(V_2,\, V_1^*)$, so setze $b\,(v_1,\, v_2) = (g\,(v_2))\,(v_1)$). Selbstverständlich hat die Menge $\operatorname{Bil}_{\mathbb{K}}(V_1,\, V_2)$ aller Bilinearformen eine natürliche $\mathbb{K}$–Vektorraumstruktur; nach dem eben Bemerkten gibt es kanonische Isomorphien

$$\operatorname{Bil}_{\mathbb{K}}(V_1,\, V_2) \cong \operatorname{Hom}(V_2,\, V_1^*) \cong \operatorname{Hom}(V_1,\, V_2^*) .$$

Satz 9.2 *Eine Bilinearform $b : V_1 \times V_2 \to \mathbb{K}$ ist genau dann nicht ausgeartet im 1. Argument, wenn die zugeordnete Abbildung $g : V_1 \to V_2^*$ injektiv ist.*

Beweis. $g : V_1 \to V_2^*$ injektiv $\iff (g(v_1) = 0 \Longrightarrow v_1 = 0) \iff (b(v_1, v_2) = 0$ für alle $v_2 \in V_2 \Longrightarrow v_1 = 0)$. □

Bei einer *dualen Paarung* endlich dimensionaler Vektorräume folgt mit diesem Satz unter Zuhilfenahme der Ungleichungs–Kette

$$\dim V_1 \le \dim V_2^* = \dim V_2 \le \dim V_1^* = \dim V_1$$

die Existenz kanonischer Isomorphismen

$$V_2 \cong V_1^* , \quad V_1 \cong V_2^* .$$

Eine *Bilinearform auf dem Vektorraum* V ist per definitionem eine solche auf $V \times V$. Diese ist übrigens im endlich dimensionalen Fall nicht ausgeartet, wenn sie diese Eigenschaft für eines der beiden Argumente besitzt (noch allgemeiner ist dies richtig für beliebige Bilinearformen auf $V_1 \times V_2$ mit $\dim V_1 = \dim V_2 < \infty$). In diesem Fall vermittelt die duale Paarung i. a. *zwei* Isomorphismen $V \cong V^*$, die offensichtlich genau dann übereinstimmen, wenn $\langle\, . \,, . \,\rangle$ *symmetrisch* ist:

$$\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle \quad \text{für alle} \quad v_1 , v_2 \in V .$$

Die kanonische Paarung $V^* \times V \to \mathbb{K}$ ist trivialerweise nicht ausgeartet im 1. Argument: $\lambda(v) = 0$ für alle $v \in V \Longrightarrow \lambda = 0$. Dies entspricht der trivialen Einsicht, daß die kanonische Abbildung $g = \mathrm{id} : V^* \to V^*$ injektiv ist.

Satz 9.3 *Die kanonische Paarung* $V^* \times V \to \mathbb{K}$ *ist nicht ausgeartet im* 2. *Argument. Also existiert ein kanonischer injektiver Homomorphismus*

$$V \hookrightarrow V^{**} ,$$

der im Falle $\dim V = n < \infty$ *sogar ein Isomorphismus ist.*

Bemerkung. Die Abbildung $V \to V^{**}$ ist gegeben durch $v \mapsto L_v$, wobei

$$L_v(\lambda) = \lambda(v) , \quad \lambda \in V^* , \quad v \in V .$$

Beweis von Satz 3. Sei $v \neq 0$ in V; dann gibt es nach dem Basisergänzungssatz eine Basis $\mathcal{B}$ von V mit $v \in \mathcal{B}$ und nach einer früheren Bemerkung eine Linearform $\lambda \in V^*$ mit $\lambda(v) = 1$. Dies beweist den ersten Teil.

Der zweite Teil der Aussage folgt wegen $\dim V^{**} = \dim V^* = \dim V$. □

Im Beispiel $\mathbb{K}^n$ hat man eine natürliche Basis $(e_1, \ldots, e_n)$, so daß der durch $e_j \mapsto e_j^*$ vermittelte Isomorphismus

$$\mathbb{K}^n \longrightarrow (\mathbb{K}^n)^*$$

tatsächlich ausgezeichnet ist. Die Existenz dieses Isomorphismus hat aber eine tiefere Bedeutung: Es gibt die kanonische Bilinearform

$$\begin{cases} \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K} \\ (x\ ,\ y) \longmapsto \sum x_j\, y_j , \end{cases}$$

die offensichtlich nicht ausgeartet ist und damit $\mathbb{K}^n$ mit $(\mathbb{K}^n)^*$ identifiziert, und zwar genau auf die oben beschriebene Weise.

Wir wollen nun die Theorie der Linearformen auf das Problem der Lösung von (homogenen) linearen Gleichungssystemen

$$(+) \qquad A\,x = 0\,, \quad A \in M(m \times n, \mathbb{K})$$

anwenden. Zur Erinnerung: Fassen wir A auf als lineare Abbildung

$$F_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m\,, \quad F_A(x) = A \cdot x\,,$$

so ist der Lösungsraum von (+) gerade der Raum

$$\ker F_A\,.$$

F_A faktorisiert natürlich über $\operatorname{im} F_A$, und wählt man einen direkten Summanden V' von $\ker F_A$ in $\mathbb{K}^n$, so erhält man den folgenden Isomorphismus in der unteren Zeile:

$$\begin{tikzcd}
\ker F_A \oplus V' = \mathbb{K}^n \arrow[r, "F_A"] \arrow[dr] & \mathbb{K}^m \\
V' \arrow[u] \arrow[r, "\sim"'] & \operatorname{im} F_A \arrow[u]
\end{tikzcd}$$

so daß also

$$\begin{aligned}\dim \ker F_A &= \dim \mathbb{K}^n - \dim V' \\ &= n - \dim \operatorname{im} F_A\,,\end{aligned}$$

und $\dim \operatorname{im} F_A$ ist der *Rang* der Matrix A, d. h. genauer ihr *Spaltenrang* (Maximalzahl linear unabhängiger Spaltenvektoren).

Wir können aber (+) auch ganz anders interpretieren: Die Zeilen der Matrix A entsprechen m Linearformen auf $\mathbb{K}^n$. Da Lösungen von (+) invariant sind unter Linearkombinationen der Gleichungen, kommt es nur auf das lineare Erzeugnis der Zeilen, also auf den *Zeilenraum* an. Abstrakt heißt dies: Ein homogenes lineares Gleichungssystem entspricht der Vorgabe eines Untervektorraums

$$W \subset V^*\,,$$

und der Lösungsraum ist die zu W *orthogonale Menge*

$$W^{\perp} := \{v \in V : \mu(v) = 0 \text{ für alle } \mu \in W\}\,.$$

Zur weiteren Verfolgung dieses Ansatzes benötigen wir die Tatsache, daß jeder linearen Abbildung $F : V \longrightarrow W$ von Vektorräumen in kanonischer Weise eine *duale* Abbildung

$$F^* : \begin{cases} W^* \longrightarrow V^* \\ \mu \longmapsto \mu \circ F \end{cases}$$

zugeordnet ist. Man sieht unmittelbar, daß für eine Inklusion $V_1 \subset V_2$ die duale Abbildung $V_2^* \to V_1^*$ nichts anderes als die Einschränkungs–Abbildung $V_2^* \ni \lambda \mapsto \lambda_{|V_1} \in V_1^*$ ist. Mit dem Basisergänzungssatz läßt sich aber jede Linearform auf V_1 nach V_2 fortsetzen, so daß folglich die duale Abbildung in diesem Fall surjektiv ist.

Die spezielle Inklusion $j : W \hookrightarrow V^*$ oben gibt dann nach den früheren Ergebnissen Anlaß zu einer surjektiven Komposition

$$L : V \xrightarrow{\sim} V^{**} \xrightarrow{j^*} W^*\,,$$

so daß also nach der Rangformel

$$\dim \ker L + \dim W^* = \dim V = n$$

gilt.

Behauptung. $\ker L$ *ist gleich dem Lösungsraum* $W^\perp$ *des homogenen Systems* $A\,x = 0$.

Beweis. Nach Definition gilt für alle $\mu \in W$ die Beziehung $(L(v))(\mu) = \mu(v)$. Also ist

$$\ker L = \{v \in V : L(v) = 0\} = \{v \in V : \mu(v) = 0 \text{ für alle } \mu \in W\}\,. \qquad \square$$

Wir haben damit wegen $\dim W = \dim W^*$ eine neue Dimensionsformel bewiesen:

Satz 9.4 *Es sei* W *ein Unterraum von* V^*, $\dim V = n < \infty$, *und* $W^\perp$ *die orthogonale Menge zu* W *in* V. *Dann gilt*

$$\dim W + \dim W^\perp = \dim V\,.$$

Da offensichtlich im Falle (+) eines Gleichungssystems $\dim W$ gleich dem *Zeilenrang* von A ist, ergibt sich sofort wieder zusammen mit der Rangformel die

Folgerung 9.5 *Für jede Matrix* $A \in M(m \times n, \mathbb{K})$ *stimmen Spalten– und Zeilenrang überein.*

Da W und $W^\perp$ in verschiedenen Räumen leben, macht es eigentlich keinen Sinn, von *Orthogonalität* zu reden. Dies ändert sich aber in gewisser Weise, wenn man (wie auf $\mathbb{K}^n$) eine nicht ausgeartete Bilinearform hat, bzgl. der sich V und V^* identifizieren lassen. Es sei also

$$\begin{cases} V \times V \longrightarrow \mathbb{K} \\ (v_1\,,\,v_2) \longmapsto \langle\, v_1,\, v_2\,\rangle \end{cases}$$

eine solche Form und $W \subset V \cong V^*$ ein Unterraum (der Isomorphismus gegeben durch $w \longmapsto \langle\,.\,,\, w\,\rangle$). Dann ist

$$W^\perp = \{v \in V : \langle\, v,\, w\,\rangle = 0 \text{ für alle } w \in W\}\,.$$

Interpretiert man das Verschwinden $\langle\, v,\, w\,\rangle = 0$ als *Senkrecht–Stehen* von v und w (siehe das nächste Kapitel), so besteht $W^\perp$ also aus allen Vektoren v, die auf (jedem Vektor von) W senkrecht stehen.

Obwohl $W^\perp$ manchmal als *orthogonales Komplement* von W bezeichnet wird, muß man sich davor hüten, $W^\perp$ als Komplement von W anzusehen in dem Sinne, daß $V = W \oplus W^\perp$. Es kann bei einer allgemeinen nicht ausgearteten Bilinearform durchaus vorkommen, daß $W \cap W^\perp \neq \{0\}$. Ist z. B.

$$\langle\, x,\, y\,\rangle = x_1\, y_1 - x_2\, y_2$$

auf $\mathbb{R}^2$ gegeben und $W = \{{}^t(\xi, \xi) \in \mathbb{R}^2,\, \xi \in \mathbb{R}\}$, so sieht man sofort, daß $W \subset W^\perp$. Im *euklidischen* Fall, den wir im nächsten Kapitel studieren, können solche Pathologien aber nicht auftreten. Wir formulieren schon jetzt den richtigen Satz in diesem Zusammenhang.

Satz 9.6 *Es sei* V *ein endlich dimensionaler* $\mathbb{K}$*–Vektorraum mit Bilinearform* $\langle\,.\,,\,.\,\rangle$, *die die zusätzliche Eigenschaft*

$$(\times) \qquad \langle\, u,\, u\,\rangle = 0 \Longrightarrow u = 0$$

besitze (und damit automatisch nicht ausgeartet ist). Es seien $U_1, U_2 \subset V$ *Untervektorräume. Dann sind die folgenden Aussagen äquivalent* :

i) $U_2 = U_1^\perp$;

ii) $U_1 \perp U_2$ $\;$ (*d. h.* $\langle\, u_1,\, u_2\,\rangle = 0 \;\; \forall\, u_1 \in U_1,\, u_2 \in U_2$) *und* $V = U_1 + U_2$;

iii) $U_1 \perp U_2$ *und* $V = U_1 \oplus U_2$;

iv) $U_1 \perp U_2$ *und* $\dim U_1 + \dim U_2 = \dim V$.

Definition. Unter einer dieser Voraussetzungen heißt U_2 das *orthogonale Komplement* von U_1; i. Z.: $V = U_1 \operatorname{\bigcirc\!\!\!\!\perp} U_2$.

Beweis von Satz 6. $(\times)$ impliziert sofort, daß mit $U_1 \perp U_2$ auch $U_1 \cap U_2 = \{0\}$ ist. Folglich gilt ii) $\Longleftrightarrow$ iii). iii) $\Longleftrightarrow$ iv) ist dann ebenfalls nach früheren Sätzen richtig. Ferner ist i) $\Longrightarrow$ iv) klar nach dem vorigen Satz. Schließlich folgt aus iv) $U_2 \subset U_1^\perp$ und $\dim U_1^\perp = \dim V - \dim U_1 = \dim U_2$, also $U_2 = U_1^\perp$. □

Beispiele. 1. Das obige Beispiel in $\mathbb{R}^2$ erfüllt $(\times)$ nicht. Es ist hierbei W gerade eine der beiden Komponenten der Menge der *isotropen* Vektoren $w : \langle w, w \rangle = 0$. (Die andere besteht aus der Menge $\{(\xi, -\xi) : \xi \in \mathbb{R}\}$).

2. In $\mathbb{R}^n$ erfüllt das euklidische Skalarprodukt $\sum x_j\, y_j$ die Bedingung $(\times)$, *nicht* aber auf $\mathbb{C}^n$.

Wir geben am Ende dieses Abschnitts noch einen weiteren konzeptionellen Beweis für die Gleichheit von Zeilenrang und Spaltenrang. Wir erinnern daran, daß in der Übungsaufgabe 20 das Folgende zu zeigen ist: Wird eine lineare Abbildung relativ zu festen Basen beschrieben durch die Matrix A, so wird die duale Abbildung bzgl. der jeweiligen dualen Basen beschrieben durch die transponierte Matrix ${}^tA$. Also folgt die in Rede stehende Formel sofort aus dem folgenden Resultat:

Satz 9.7 *Es sei $F : V \to W$ eine lineare Abbildung endlich dimensionaler $\mathbb{K}$–Vektorräume, $F^* : W^* \to V^*$ bezeichne die duale Abbildung. Dann gilt :*

$$\ker F = (\operatorname{im} F^*)^\perp\,, \quad \operatorname{im} F = (\ker F^*)^\perp\,.$$

Insbesondere hat man

$$\dim \operatorname{im} F = \dim \operatorname{im} F^*\,.$$

Besitzt V eine nicht ausgeartete Bilinearform mit $(\times)$, so gilt

$$V = \ker F \operatorname{\bigcirc\!\!\!\!\perp} \operatorname{im} F^*\,,$$

und F induziert einen Isomorphismus

$$F : \operatorname{im} F^* \xrightarrow{\sim} \operatorname{im} F\,.$$

Beweis. Jeweils mit dem Basisergänzungssatz bzw. mit der Rangformel schließt man:

$$\begin{aligned}
(\operatorname{im} F^*)^\perp &= \{v \in V : \lambda(v) = 0 \quad \forall \lambda \in \operatorname{im} F^*\} \\
&= \{v \in V : \mu \circ F(v) = 0 \quad \forall \mu \in W^*\} \\
&= \ker F\,. \\
(\ker F^*)^\perp &= \{w \in W : \mu(w) = 0 \quad \forall \mu \in \ker F^*\} \\
&= \{w \in W : \mu(w) = 0 \quad \forall \mu \in W^* \quad \text{mit} \quad \mu_{|\operatorname{im} F} = 0\} \\
&= \operatorname{im} F\,. \\
\dim \operatorname{im} F^* &= \dim W^* - \dim \ker F^* \\
&= \dim (\ker F^*)^\perp = \dim \operatorname{im} F\,.
\end{aligned}$$

Die letzten Behauptungen sind klar. □

Zum Schluß geben wir noch ein Beispiel dafür, daß im Unendlichdimensionalen die Dinge wesentlich anders liegen. Durch

$$\langle f, g \rangle = \int_I f(x)\, g(x)\, dx\,, \quad f, g \in V = \mathcal{C}^0(I)\,, \quad I = [a, b] \subset \mathbb{R}\,,$$

wird eine nicht ausgeartete Bilinearform auf V definiert, liefert also eine kanonische Injektion $V \hookrightarrow V^*$, $g \mapsto \langle\, .\,,\, g \,\rangle$. Für $x_0 \in I$ ist $f \mapsto f\,(x_0)$ ebenfalls eine Linearform; aber offensichtlich gibt es keine Funktion $g \in \mathcal{C}^0(I)$ mit

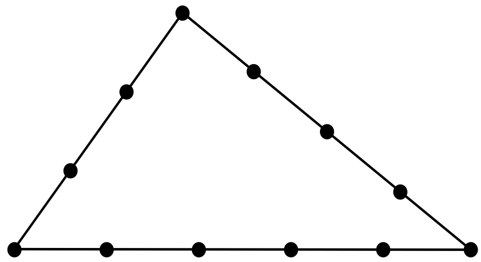
$$\int_I f\,(x)\, g\,(x)\, dx \;=\; f\,(x_0) \quad \text{für alle} \quad f \in \mathcal{C}^0(I)\ .$$

10 Euklidische und unitäre Vektorräume

In unserer Anschauung sind die Ebene und der uns umgebende Raum nicht nur mit einer linearen Struktur versehen; wir haben auch eine Vorstellung von Begriffen wie *Längen* von Strecken und dem *Senkrechtstehen (Orthogonalität)* von Strecken oder Geraden. Überdies reden wir auch von *Winkeln* zwischen sich schneidenden Geraden (obwohl dieser Begriff, wie wir sehen werden, aus den beiden anderen abgeleitet werden kann). Alle diese Begriffsbildungen gehören in den Bereich der *euklidischen* Geometrie.

Orthogonalität und Längenmessung sind in der euklidischen Ebene eng miteinander durch den *Satz des Pythagoras* verknüpft: Für je zwei Vektoren $x, y \in \mathbb{R}^2$ mit $x \perp y$ gilt:

$$\| x + y \|^2 = \| x \|^2 + \| y \|^2 .$$



Figur 10.1

Aber auch umgekehrt kann man mit elementarer Geometrie sofort zeigen, daß aus der Gültigkeit des Satzes von Pythagoras die Orthogonalität von x und y folgt.

Setzen wir nun allgemein im $\mathbb{R}^n$ voraus, daß wir eine *Längenmessung*

$$\| \cdot \| : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0}$$

besitzen und einen symmetrischen *Orthogonalitätsbegriff* $\perp$, so können wir leicht die übliche Form der *euklidischen Norm* und des *euklidischen Skalarproduktes* aus wenigen (unmittelbar einsichtigen) Axiomen ableiten. – Wir fordern:

1. Es gibt eine *orthonormierte* Basis $(e_1, \ldots, e_n)$, d. h. eine Basis mit $\| e_j \| = 1$ für alle j und $e_j \perp e_k$ für alle $j \neq k$.

2. Zwei Vektoren x und y stehen genau dann senkrecht aufeinander, wenn

$$\| x + y \|^2 = \| x \|^2 + \| y \|^2 .$$

3. Gilt $x \perp y_1$ und $x \perp y_2$, so auch $x \perp y$ für alle $y \in \mathrm{span}\,(y_1, y_2)$.

4. $\| a\, x \| = | a | \cdot \| x \|$ für alle $a \in \mathbb{R}$, $x \in \mathbb{R}^n$.

Schreibt man nun $x = x_1 e_1 + \cdots + x_n e_n$, so folgt aus 1. und 3. die Relation $x_n e_n \perp x_1 e_1 + \cdots + x_{n-1} e_{n-1}$ und damit

$$\| x \|^2 = \| x_1 e_1 + \cdots + x_{n-1} e_{n-1} \|^2 + \| x_n e_n \|^2 .$$

Per Induktion ergibt sich daraus sofort

$$\| x \|^2 = x_1^2 + \cdots + x_n^2 ,$$

und, da Längen nicht negativ sein sollten,

$$\| x \| = \Big(\sum_{j=1}^{n} x_j^2 \Big)^{1/2} \quad \text{(positive Wurzel)} .$$

Die einfachsten Eigenschaften dieser *euklidischen Norm* fassen wir in folgender Definition zusammen:

Definition. Es sei V ein $\mathbb{R}$– oder $\mathbb{C}$–Vektorraum. Eine Abbildung

$$\| \cdot \| : V \longrightarrow \mathbb{R}$$

heißt eine *Norm*, falls die folgenden Eigenschaften erfüllt sind:

i) $\| v \| \geq 0 \, ; \quad \| v \| = 0 \iff v = 0 \, .$

ii) $\| a\, v \| = | a | \cdot \| v \| \, , \quad v \in V \, , \; a \in \mathbb{R}$ (bzw. $\mathbb{C}$).

iii) $\| v_1 + v_2 \| \leq \| v_1 \| + \| v_2 \|$ (Dreiecksungleichung).

Man nennt dann das Paar $(V, \| \cdot \|)$ auch einen *normierten Vektorraum*.

Weitere *Beispiele* für normierte Vektorräume sind, wovon man sich leicht überzeugt, die folgenden:

1. $\mathbb{R}^n$ mit $\| x \|_\infty := \max_{j=1,\ldots,n} | x_j | \, .$ (*Maximum–Norm*).

2. $\mathcal{C}^0(I)$ mit $\| f \|_\infty = \sup_{x \in I} | f(x) | \, ,$ I ein kompaktes Intervall. (*Supremums–Norm*).

3. $\mathcal{C}^0(I)$ mit $\| f \|_2^2 = \int_I f^2(x)\, dx \, .$ (L^2*–Norm*).

Ausgehend von dem Axiom 2 kann man die *Orthogonalität* im euklidischen (Standard–) Raum $\mathbb{R}^n$ mit Hilfe einer *Bilinearform* analytisch fassen. Es gilt offensichtlich für $x, y \in \mathbb{R}^n$: $x \perp y \iff \| x + y \|^2 = \| x \|^2 + \| y \|^2 \iff$

$$\sum_{j=1}^{n} (x_j + y_j)^2 = \sum_{j=1}^{n} x_j^2 + \sum_{j=1}^{n} y_j^2 \iff \langle x, y \rangle := x_1 y_1 + \cdots + x_n y_n = 0 \, .$$

Wir haben also eine kanonische Bilinearform $\langle \, . \, , \, . \, \rangle$ auf $\mathbb{R}^n$ zur Verfügung, so daß

$$\| x \| = \langle x, x \rangle^{1/2} \quad \text{und} \quad x \perp y \quad \iff \quad \langle x, y \rangle = 0 \, .$$

Dieses Beispiel verallgemeinern wir sofort zu der folgenden generellen

Definition. Es sei V ein $\mathbb{R}$–Vektorraum. Ein *euklidisches Skalarprodukt* auf V ist eine symmetrische Bilinearform

$$\langle \, . \, , \, . \, \rangle : V \times V \longrightarrow \mathbb{R}$$

mit $\langle v, v \rangle > 0$ für $v \neq 0$ (also eine sogenannte *positiv–definite Form*). Das Paar $(V, \langle \, . \, , \, . \, \rangle)$ heißt ein *euklidischer Vektorraum*.

Beispiele für euklidische Vektorräume sind u. a.:

1. $\mathbb{R}^n$ mit $\langle x, y \rangle = \sum_{j=1}^{n} x_j \, y_j \, .$

2. $\mathcal{C}^0(I)$ mit $\langle f, g \rangle = \int_I f(x)\, g(x)\, dx \, .$

Die Positiv–Definitheit ist natürlich notwendig, um überhaupt eine Norm definieren zu können: Der Ausdruck

$$\| v \| = \langle v, v \rangle^{1/2}$$

erfüllt Bedingung i) einer Norm, aber auch ii):

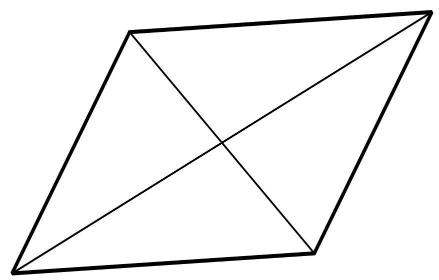
$$\| av \|^2 = \langle av, av \rangle = a^2 \langle v, v \rangle \implies \| av \| = | a | \cdot \| v \| .$$

Die Eigenschaft iii) beweisen wir später.

Definition. $\| v \| := \langle v, v \rangle^{1/2}$ heißt die dem Skalarprodukt $\langle \,.\, , \,.\, \rangle$ *zugeordnete Norm.*

Lemma 10.1 *Es sei* V *ein euklidischer Vektorraum. Dann gilt* :

1. $\| v + w \|^2 = \| v \|^2 + \| w \|^2 + 2 \langle v, w \rangle$ (Satz des Pythagoras).
2. $\| v + w \|^2 + \| v - w \|^2 = 2(\| v \|^2 + \| w \|^2)$ (Parallelogramm–Gleichung).



Figur 10.2

Wegen 1. definiert man allgemein für einen euklidischen Vektorraum:

$$v \perp w \iff \langle v, w \rangle = 0 \quad (\iff \| v + w \|^2 = \| v \|^2 + \| w \|^2).$$

Außerdem erhält man leicht die sogenannte *Polarisations–Gleichung*:

$$\begin{aligned} \langle v, w \rangle &= \frac{1}{2} (\| v + w \|^2 - \| v \|^2 - \| w \|^2) \\ &= \frac{1}{4} (\| v + w \|^2 - \| v - w \|^2) . \end{aligned}$$

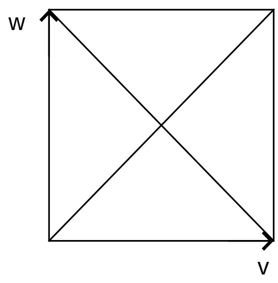
D. h.: Im euklidischen Fall ist $\langle \,.\, , \,.\, \rangle$ schon durch die Norm $\| \cdot \|$ bestimmt. – In bezug auf die Umkehrung gilt der folgende

Satz 10.2 *Eine Norm* $\| \cdot \|$ *kommt genau dann von einem Skalarprodukt her, wenn sie der Parallelogramm–Gleichung genügt.*

Beweis. Übungsaufgabe 59.

Folgerung 10.3 *Die Maximum–Norm kommt nicht von einem Skalarprodukt her.*

Beweis. Man wähle z. B. $v = (1, 0)$, $w = (0, 1)$. (Siehe die folgende Zeichnung).



Figur 10.3

Wir wollen im euklidischen $\mathbb{R}^n$ auch *Winkel* zwischen Vektoren v, w definieren. Das geht natürlich nur, wenn $v \neq 0$, $w \neq 0$. Der Winkel sollte unabhängig von der Länge der Vektoren sein. Wir normieren sie deshalb durch

$$v_0 = \frac{v}{\| v \|}, \quad w_0 = \frac{w}{\| w \|}$$

auf die Länge 1 und versuchen eine Zerlegung

$$w_0 = w_1 + w_2 , \quad w_1 \in \operatorname{span}(v_0) , \quad w_2 \perp v_0 ,$$

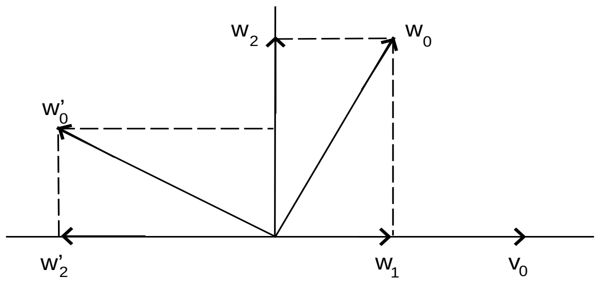
was tatsächlich mit $w_1 := \langle w_0, v_0 \rangle v_0$, $w_2 := w_0 - w_1$ gelingt:

$$\begin{aligned} \langle w_2, v_0 \rangle &= \langle w_0, v_0 \rangle - \langle w_1, v_0 \rangle \\ &= \langle w_0, v_0 \rangle - \langle w_0, v_0 \rangle \langle v_0, v_0 \rangle = 0 . \end{aligned}$$

Wir haben hierbei nichts anderes getan, als den Vektor w_0 auf v_0 *orthogonal* zu projizieren (siehe Figur 5). Ferner ist $\| w_1 \|^2 + \| w_2 \|^2 = \| w_1 + w_2 \|^2 = \| w_0 \|^2 = 1$, so daß also aufgrund der Zeichnung der Winkel α durch

$$\cos \alpha = \pm \| w_1 \| = \langle w_0, v_0 \rangle = \frac{\langle w, v \rangle}{\| w \| \cdot \| v \|}$$

bestimmt ist.



Figur 10.4

Daß man auf diese Weise in euklidischen Vektorräumen immer einen Winkel definieren kann, liegt an der *Cauchy–Schwarzschen Ungleichung*, manchmal auch kurz CSU genannt.

Satz 10.4 (Cauchy - Schwarzsche Ungleichung) *Es sei V ein euklidischer Vektorraum. Dann gilt stets*

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

und Gleichheit besteht hier genau dann, wenn v und w linear abhängig sind.

Beweis. Seien v, w abhängig, also ohne Einschränkung $w = av$. Dann ist

$$|\langle v, w \rangle| = |\langle v, av \rangle| = |a| \cdot \| v \|^2 = \| v \| \cdot \| w \| .$$

Seien umgekehrt v, w linear unabhängig, so daß also insbesondere $v - aw \neq 0$ für alle $a \in \mathbb{R}$ und $w \neq 0$. Dann ergibt sich

$$0 < \langle v - aw, v - aw \rangle = \| v \|^2 - 2a \langle v, w \rangle + |a|^2 \| w \|^2 .$$

Wählt man hier speziell

$$a = \frac{\langle v, w \rangle}{\| w \|^2} ,$$

so folgt sofort

$$0 < \| v \|^2 - \frac{\langle v, w \rangle^2}{\| w \|^2}$$

und damit die Behauptung

$$|\langle v, w \rangle| < \| v \| \cdot \| w \| .$$ □

Bemerkung. Mit Hilfe dieses Satzes kann man in euklidischen Vektorräumen durch

$$\cos \alpha = \frac{\langle v, w \rangle}{\| v \| \cdot \| w \|}$$

stets einen Winkel α zwischen zwei Vektoren mit $0 \le \alpha \le \pi$ eindeutig definieren.

Folgerung 10.5 *Ist $\langle \, . \, , \, . \, \rangle$ ein euklidisches Skalarprodukt, so wird durch $\| v \| = \sqrt{\langle v, v \rangle}$ eine Norm auf V definiert.*

Beweis. Es fehlt nur noch die Dreiecksungleichung:

$$\begin{aligned} \| v + w \|^2 &= \langle v + w, v + w \rangle = \| v \|^2 + 2 \langle v, w \rangle + \| w \|^2 \\ &\le \| v \|^2 + 2 \| v \| \cdot \| w \| + \| w \|^2 \qquad \text{(CSU)} \\ &= (\| v \| + \| w \|)^2 . \end{aligned}$$

Wir bemerken schließlich, daß offensichtlich jedes euklidische Skalarprodukt die Bedingung $(\times)$ aus Kapitel 9 erfüllt. Damit existieren in diesem Fall orthogonale Komplemente.

Satz 10.6 *Es sei V ein endlich–dimensionaler euklidischer Vektorraum mit Skalarprodukt $\langle \, . \, , \, . \, \rangle$, und $W \subset V$ sei ein Untervektorraum. Dann gilt*

$$V = W \perp W^\perp , \quad W^\perp = \{ v \in V : v \perp w \ \textit{für alle} \ w \in W \} .$$

Insbesondere ist für jede lineare Abbildung

$$F : V \longrightarrow V'$$

der Raum V die orthogonale Summe

$$\ker F \perp \operatorname{im} F^* .$$

Der $\mathbb{R}^n$ mit euklidischem Skalarprodukt $\langle \, . \, , \, . \, \rangle$ besitzt die *Orthonormalbasis* $(e_1, \ldots, e_n)$:

$$\langle e_j, e_k \rangle = \delta_{jk} .$$

Wir wollen zeigen, daß jeder euklidische (endlich–dimensionale) Vektorraum solche Basen besitzt. Genauer gilt:

Satz 10.7 (Schmidtsches Orthonormalisierungs - Verfahren) *Es sei V ein euklidischer Vektorraum mit Skalarprodukt $\langle\,.\,,\,.\,\rangle$ und Basis $(u_1,\ldots,u_n)$. Dann gibt es eine Orthonormalbasis $(v_1,\ldots,v_n)$ mit*

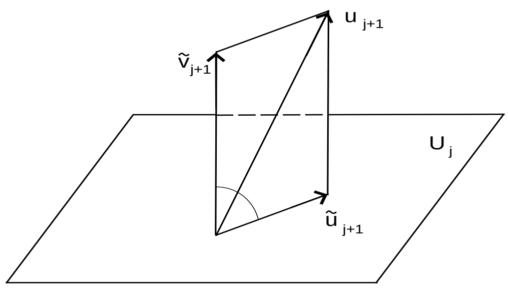
$$\mathit{span}\,(u_1,\ldots,u_j) \;=\; \mathit{span}\,(v_1,\ldots,v_j)\,, \quad j=1,\ldots,n\,.$$

Beweis. Setze $U_j := \operatorname{span}(u_1,\ldots,u_j)\,,\ j=1,\ldots,n$. Im Falle $j=1$ braucht man nur $v_1 = \dfrac{u_1}{\|\,u_1\,\|}$ zu wählen. Sei die Aussage für $j<n$ schon bewiesen. Dann ist

$$u_{j+1} \notin U_j \;=\; \operatorname{span}(v_1,\ldots,v_j)\,.$$

Es sei nun $\widetilde{u}_{j+1}$ die *orthogonale Projektion* von u_{j+1} nach U_j:

$$\widetilde{u}_{j+1} \;=:\; \sum_{k=1}^{j} \langle\, u_{j+1},\, v_k\,\rangle\, v_k\,.$$



Figur 10.5

Dann steht

$$\widetilde{v}_{j+1} \;=\; u_{j+1} - \widetilde{u}_{j+1}$$

(tatsächlich) senkrecht auf U_j:

$$\begin{aligned}\langle\, \widetilde{v}_{j+1},\, v_\ell\,\rangle \;&=\; \langle\, u_{j+1},\, v_\ell\,\rangle \;-\; \sum_{k=1}^{j} \langle\, u_{j+1},\, v_k\,\rangle\,\langle\, v_k,\, v_\ell\,\rangle \\ &=\; \langle\, u_{j+1},\, v_\ell\,\rangle \;-\; \langle\, u_{j+1},\, v_\ell\,\rangle \;=\; 0\,, \quad \ell=1,\ldots,j\,.\end{aligned}$$

Es ist ferner $\widetilde{v}_{j+1} \neq 0$, da sonst $u_{j+1} = \widetilde{u}_{j+1} \in \operatorname{span}(v_1,\ldots,v_j) = U_j$ wäre. Also ist

$$(v_1,\ldots,v_{j+1}) \quad \text{mit} \quad v_{j+1} \;=\; \frac{\widetilde{v}_{j+1}}{\|\,\widetilde{v}_{j+1}\,\|}$$

orthonormiert, und

$$\begin{aligned}\operatorname{span}(v_1,\ldots,v_{j+1}) \;&=\; \operatorname{span}(v_1,\ldots,v_j,\widetilde{v}_{j+1}) \\ &=\; \operatorname{span}(v_1,\ldots,v_j,\, u_{j+1}) \\ &=\; \operatorname{span}(u_1,\ldots,u_j,\, u_{j+1}) \;=\; U_{j+1}\,.\end{aligned}$$

Für $j=n$ ergibt sich $V = \operatorname{span}(v_1,\ldots,v_n)$, und wegen $\dim V = n$ muß $(v_1,\ldots,v_n)$ eine Basis von V bilden. □

Das letzte Argument ist bei unendlicher Dimension nicht anwendbar. Deswegen ist eine andere Aussage nützlich, die aber auch für endliche Dimension von Interesse ist.

Satz 10.8 *Es sei* $(v_\iota)_{\iota \in I}$ *ein orthogonales System von Vektoren* $v_\iota \neq 0$. *Dann ist* $(v_\iota)_{\iota \in I}$ *linear unabhängig.*

Beweis . Die Voraussetzung bedeutet natürlich

$$v_\iota \perp v_\kappa$$

für alle $i \neq \kappa$ in I. Ist nun für eine endliche Teilmenge I_0

$$\sum_{\iota \in I_0} a_\iota \, v_\iota \;=\; 0 \; ,$$

so multipliziert man mit einem beliebigen v_κ und erhält

$$0 \;=\; \langle \sum_{\iota \in I_0} a_\iota \, v_\iota, \, v_\kappa \rangle \;=\; \sum_{\iota \in I_0} a_\iota \langle \, v_\iota, \, v_\kappa \, \rangle \;=\; a_\kappa \, \| \, v_\kappa \, \|^2 \; ,$$

d. h. $a_\kappa = 0, \; \kappa \in I_0$. □

Am Ende dieses Abschnitts wollen wir die erzielten Ergebnisse noch in Matrizenform umgestalten. Betrachten wir zunächst $\mathbb{K}^n$ allgemein (als Raum der Spaltenvektoren x, y etc.) und eine beliebige Matrix $A \in M(n \times n, \mathbb{K})$, so wird offensichtlich durch

$$\langle \, x, \, y \, \rangle_A \; := \; {}^t x \, A \, y$$

eine Bilinearform auf $\mathbb{K}^n$ erklärt. Die Zuordnung

$$\begin{cases} M(n \times n, \, \mathbb{K}) & \longrightarrow \quad \mathrm{Bil}_{\mathbb{K}}\,(\mathbb{K}^n) \\ \qquad\qquad A & \longmapsto \quad \langle \, . \, , \, . \, \rangle_A \end{cases}$$

ist $\mathbb{K}$–linear. Man kann leicht eine Umkehrabbildung angeben: Für die kanonische Basis $(e_1, \ldots, e_n)$ von $\mathbb{K}^n$ setze man mit Hilfe der gegebenen Form $\langle \, . \, , \, . \, \rangle$:

$$a_{jk} \;=\; \langle \, e_j, \, e_k \, \rangle \quad \text{und} \quad A \;=\; (a_{jk}) \; .$$

Dann ist tatsächlich, wie gewünscht,

$$\begin{aligned} \langle \, x, \, y \, \rangle \;&=\; \langle \textstyle\sum_j \, x_j \, e_j, \, \sum_k \, y_k \, e_k \, \rangle \;=\; \sum_{j,\,k} \, x_j \, y_k \, \langle \, e_j, \, e_k \, \rangle \\ &=\; \textstyle\sum_j \, x_j \, (\sum_k \, a_{jk} \, y_k) \;=\; {}^t x \, A \, y \; . \end{aligned}$$

Es besteht also eine kanonische Isomorphie

$$M(n \times n, \, \mathbb{K}) \;\cong\; \mathrm{Bil}_{\mathbb{K}}\,(\mathbb{K}^n) \; .$$

Für einen beliebigen $\mathbb{K}$–Vektorraum V der Dimension n hat man nach Wahl einer Basis $\mathcal{A} = (v_1, \ldots, v_n)$ eine Isomorphie $V \cong \mathbb{K}^n$ und damit auch eine, von $\mathcal{A}$ abhängige, Isomorphie

$$\begin{cases} \mathrm{Bil}_{\mathbb{K}}(V) \;\;\cong\;\; \mathrm{Bil}_{\mathbb{K}}(\mathbb{K}^n) \;\;\cong\;\; M(n \times n, \, \mathbb{K}) \\ \;\;\langle \, . \, , \, . \, \rangle \qquad\qquad \longmapsto \qquad\quad A \;=\; (a_{jk}) \; , \quad a_{jk} \;=\; \langle \, v_j, \, v_k \, \rangle \; . \end{cases}$$

Lemma 10.9 $\langle \, . \, , \, . \, \rangle$ *ist genau dann symmetrisch, wenn* A *symmetrisch ist, d. h. wenn* $a_{jk} = a_{kj}$ *für alle* $j, \, k$ *gilt.*

Beweis. Straightforward. □

Bemerkung. Wie wir für $\mathbb{K} = \mathbb{R}$ die *Positiv–Definitheit* einer Bilinearform durch eine sie beschreibende Matrix A fassen können, werden wir erst in Kapitel 14 sehen.

Die Orthogonalisierung bedeutet natürlich, daß wir die $\langle\,.\,,\,.\,\rangle$ zugeordnete Matrix auf „Diagonalgestalt gebracht haben“. Wir müssen aber genau nachprüfen, was dies in unserem Zusammenhang bedeutet. Es sei also neben $\mathcal{A} = (v_1, \ldots, v_n)$ eine weitere Basis $\mathcal{B} = (w_1, \ldots, w_n)$ gegeben und $B = (b_{jk})$, $b_{jk} = \langle\, w_j,\, w_k \,\rangle$. Ferner sei

$$S \,=\, M_{\mathcal{B}}^{\mathcal{A}}\,(\mathrm{id}_V) \in GL(n,\, \mathbb{K})$$

die Transformationsmatrix des Basiswechsels $\mathcal{A} \to \mathcal{B}$, die das folgende Diagramm kommutativ macht:

$$\begin{tikzcd}
V \arrow[r, "\mathrm{id}_V"] \arrow[d, "\Phi_{\mathcal{A}}^{-1}"'] & V \arrow[d, "\Phi_{\mathcal{B}}^{-1}"] \\
\mathbb{K}^n \arrow[r, "S"'] & \mathbb{K}^n
\end{tikzcd}$$

Zur Erinnerung: Ist $v \,=\, \sum a_j\, v_j \,=\, \sum b_j\, w_j$, so ist $\Phi_{\mathcal{A}}^{-1}(v) \,=\, {}^t(a_1, \ldots, a_n) \,=:\, {}^t a$, $\Phi_{\mathcal{B}}^{-1}(v) \,=\, {}^t(b_1, \ldots, b_n) \,=:\, {}^t b$ und $b \,=\, S\, a$. Entsprechend betrachten wir $v' \in V$ etc. Es ist dann nach Definition von A und B:

$${}^t a \cdot A \cdot a' \,=\, \langle\, v,\, v' \,\rangle \,=\, {}^t b \cdot B \cdot b' \,=\, {}^t(Sa) \cdot B \cdot (Sa') \,=\, {}^t a({}^t S\, B\, S)a' \;,$$

und diese Identität ist nur dann richtig für alle a, $a' \in \mathbb{K}^n$ (setze $a = e_j$, $a' = e_k$), wenn gilt:

Satz 10.10 (Transformationsformel) *Unter den obigen Voraussetzungen gilt*

$$A \,=\, {}^t S \cdot B \cdot S \;.$$

Warnung. Matrizen zu *Endomorphismen* transformieren sich nach der Formel $A = S^{-1}\, B\, S$. Dies ist ein wesentlich anderes Gesetz!

Der Satz über die Orthonormalisierung kann auch folgendermaßen *matrizen–theoretisch* formuliert werden:

Satz 10.11 *Es sei* $A \in M(n \times n,\, \mathbb{R})$ *eine symmetrische, positiv–definite Matrix (d. h. es gelte* ${}^t x\, A\, x > 0$ *für alle* $x \neq 0$ *). Dann gibt es eine invertierbare Matrix* $S \in \mathrm{GL}\,(n,\, \mathbb{R})$ *mit*

$${}^t S \cdot S\, (=\, {}^t S\, E_n\, S) \,=\, A \;.$$

In diesem Sinne besitzen also positiv–definite symmetrische Matrizen stets eine *Quadratwurzel*.

Bei *komplexen* Vektorräumen muß man notgedrungen die „euklidischen“ Axiome modifizieren. Z. B. hat die naheliegende Form

$$\langle\, z,\, w \,\rangle \,=\, \sum_{j=1}^{n} z_j\, w_j$$

auf $\mathbb{C}^n$ isotrope Vektoren (für $n \geq 2$), ist also ausgeartet (das Produkt von ${}^t(1,\, i,\, 0, \ldots, 0)$ mit sich selbst ist gleich Null). Außerdem ist diese Form (auch schon für $n = 1$) nicht positiv–definit. Aber nur unter dieser Bedingung können wir jedem $z \in \mathbb{C}^n$ eine „vernünftige“ (sprich *reelle*) Länge zuordnen. Dieses Malheur passiert übrigens für *jede* komplexe Bilinearform, denn es ist ja stets für alle $v \in V$:

$$\langle\, iv,\, iv \,\rangle \,=\, i^2\, \langle\, v,\, v \,\rangle \,=\, -\, \langle\, v,\, v \,\rangle \;,$$

und nur einer der beiden Ausdrücke kann positiv sein.

Man muß sich daher von der Linearität in beiden Argumenten trennen und definiert in leichter Modifikation:

Definition. Eine *Sesquilinearform* auf einem komplexen Vektorraum V ist eine Abbildung

$$\langle\, .\,,\, .\,\rangle : V \times V \to \mathbb{C}\ ,$$

die linear im 2. und antilinear (bzgl. Konjugation) im 1. Argument ist, d. h.:

$$\begin{aligned} \langle\, v,\, aw_1 + bw_2 \,\rangle &= a \,\langle\, v,\, w_1 \,\rangle + b \,\langle\, v,\, w_2 \,\rangle\ , \\ \langle\, av_1 + bv_2,\, w \,\rangle &= \overline{a} \,\langle\, v_1,\, w \,\rangle + \overline{b} \,\langle\, v_2\,,\, w \,\rangle\ . \end{aligned}$$

Hinweis. Manchmal wird in der Literatur auch die Antilinearität im zweiten Argument und die Linearität im ersten Argument verlangt. „Sesqui“ heißt in etwa „anderthalbfach“.

Beispiele für solche Vektorräume mit Sesquilinearform sind:

1. $\mathbb{C}^n$ mit $\langle\, z,\, w \,\rangle = \displaystyle\sum_{j=1}^{n} \overline{z_j}\, w_j\ .$

2. $\mathcal{C}^0(I,\, \mathbb{C})$ mit $\langle\, f,\, g \,\rangle = \displaystyle\int_I \overline{f(x)}\, g(x)\, dx\ .$

3. $\ell^2_{\mathbb{C}} = \{(a_j)_{j\in\mathbb{N}} : a_j \in \mathbb{C}\,,\ \displaystyle\sum_{j=1}^{\infty} |\, a_j\,|^2 < \infty\}$ mit $\langle\, (a_j),\, (b_j) \,\rangle = \displaystyle\sum_{j=0}^{\infty} \overline{a_j}\, b_j\ .$

Konzeptionell kann man die Antilinearität im 1. Argument noch eleganter fassen: Ist V ein $\mathbb{C}$–Vektorraum, so bezeichne $\overline{V}$ als *Menge* V selbst. Die *Addition* sei ebenfalls genau wie in V erklärt, die *Multiplikation* mit Skalaren $c \in \mathbb{C}$ aber durch

$$\mathbb{C} \times \overline{V} \ni (c,\, v) \longmapsto \overline{c}v \in \overline{V}\ .$$

$\overline{V}$ ist dann ebenfalls ein $\mathbb{C}$–Vektorraum, und eine Sesquilinearform auf V ist nichts anderes als eine Bilinearform

$$\overline{V} \times V \longrightarrow \mathbb{C}\ .$$

So betrachtet ist es klar, daß wir auch die Symmetrie aus der Definition von euklidischen Skalarprodukten im komplexen Fall nicht einfach übernehmen können.

Definition. Eine Sesquilinearform heißt eine *hermitesche Form*, wenn

$$\langle\, w,\, v \,\rangle = \overline{\langle\, v,\, w \,\rangle} \quad \text{für alle} \quad v,\, w \in V\ .$$

Diese Bedingung ist jedenfalls verträglich mit den Linearitätsbedingungen und wird von unseren Standardformen erfüllt. Man beachte, daß man bei Voranstellung dieses Axioms auf die Antilinearitäts–Bedingung im 1. Argument vollständig verzichten kann.

Satz 10.12 *Eine Abbildung* $\langle\, .\,,\, .\,\rangle : V \times V \to \mathbb{C}$ *ist genau dann hermitesch, wenn*

i) $\langle\, w,\, v \,\rangle = \overline{\langle\, v,\, w \,\rangle}\ ,$

ii) $\langle\, v,\, a\,w_1 + b\,w_2 \,\rangle = a\,\langle\, v,\, w_1 \,\rangle + b\,\langle\, v,\, w_2 \,\rangle\ .$

Insbesondere ist für hermitesche Formen das Produkt von v mit sich selbst stets reell. Es ist daher sinnvoll, auch für solche Formen den Begriff der *Positiv–Definitheit* einzuführen:

$$\langle v, v \rangle > 0 , \quad v \neq 0 .$$

Definition. Ein komplexer Vektorraum V heißt *unitär*, wenn er mit einer positiv–definiten hermiteschen Form $\langle\, . \,, \, . \,\rangle$ versehen ist.

Die meisten Aussagen über euklidische Vektorräume übertragen sich jetzt fast wörtlich auf unitäre Vektorräume.

1. $\| v \| := \langle v, v \rangle^{1/2}$ ist eine Norm auf V.

2. Es gilt die Cauchy–Schwarzsche Ungleichung:

$$| \langle v, w \rangle | \leq \| v \| \cdot \| w \| .$$

3. Orthonormalisierung kann im Endlich–Dimensionalen durchgeführt werden.

Hermitesche Formen werden relativ zu Basen per definitionem durch *hermitesche Matrizen* beschrieben:

$$\langle v, w \rangle = {}^t\overline{v}\, H\, w , \quad {}^tH = \overline{H} ,$$

wobei stets $\overline{H} = (\overline{h_{jk}})$ für $H = (h_{jk})$ gesetzt wird und $\overline{v} \in \mathbb{C}^n$ natürlich den Vektor ${}^t(\overline{z_1}, \ldots, \overline{z_n})$ bezeichnet, wenn $v = {}^t(z_1, \ldots, z_n)$. Hermitesche Matrizen transformieren sich demgemäß nach der Formel

$$A = {}^t\overline{S}\, B\, S .$$

In unitären Vektorräumen hat man einige wenige Modifikationen bei der Formulierung der bekannten Sätze im Euklidischen anzubringen. So lautet der *Satz von Pythagoras* in unitären Räumen:

$$\| v + w \|^2 = \| v \|^2 + \| w \|^2 + 2 \operatorname{Re} \langle v, w \rangle .$$

Die *Parallelogramm–Gleichung* ist wie im euklidischen Fall gültig (und folgt auch hier schlicht und einfach durch Addition des Phythagoras für $v + w$ und $v - w$):

$$\| v + w \|^2 + \| v - w \|^2 = 2(\| v \|^2 + \| w \|^2) .$$

Hieraus ergibt sich sofort:

$$4 \operatorname{Re}\langle v, w \rangle = \| v + w \|^2 - \| v - w \|^2 .$$

Erinnern wir uns, daß im euklidischen Fall das Skalarprodukt durch die Norm bestimmt ist, so verwundert uns vielleicht diese und die vorvorletzte Formel, die *nur* den Realteil von $\langle v, w \rangle$ durch Normen auszudrücken scheint. Nun ist aber im unitären Fall $\langle w, v \rangle = \overline{\langle v, w \rangle}$ und damit

$$\operatorname{Im} \langle v, w \rangle = - \operatorname{Im} \overline{\langle v, w \rangle} = - \operatorname{Im} \langle w, v \rangle = \operatorname{Re} i \langle w, v \rangle ,$$

so daß wir leicht die folgende Formel erhalten:

$$\begin{aligned} 4 \langle v, w \rangle &= 4 \operatorname{Re} \langle v, w \rangle + 4i \operatorname{Im} \langle v, w \rangle \\ &= 4 \operatorname{Re} \langle v, w \rangle + 4i \operatorname{Re} i \langle w, v \rangle \\ &= 4 \operatorname{Re} \langle v, w \rangle + 4i \operatorname{Re} \langle -iw, v \rangle \\ &= \| v + w \|^2 - \| v - w \|^2 + i \| v - iw \|^2 - i \| v + iw \|^2 , \end{aligned}$$

die man wie im Reellen die *Polarisationsformel* nennt.

Zum Schluß wollen wir noch einige Bemerkungen über *euklidische Geometrie* anschließen. Diese verhält sich zur affinen Geometrie wie der Begriff des euklidischen Vektorraumes zu dem des allgemeinen Vektorraums.

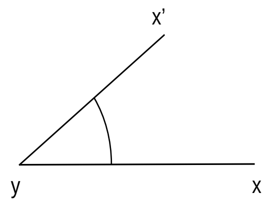
Definition. Ein *euklidischer Raum* ist ein Quadrupel $(X, T(X), \langle\,.\,,\,.\,\rangle, \tau)$, wobei $(X, T(X), \tau)$ ein affiner Raum über $\mathbb{R}$ und $T(X)$ ein euklidischer Vektorraum mit Skalarprodukt $\langle\,.\,,\,.\,\rangle$ ist.

Auf einem solchen euklidischen Raum gibt es eine *Abstandsfunktion* (oder *Metrik*):

$$d : \begin{cases} X \times X & \longrightarrow \quad \mathbb{R}_{\geq 0} \\ (x_1, x_2) & \longmapsto \quad \| \overrightarrow{x_1 x_2} \| = \sqrt{\langle \overrightarrow{x_1 x_2}, \overrightarrow{x_1 x_2} \rangle} \, . \end{cases}$$

Außerdem kann man Winkel erklären:

$$\angle(xyx') = \arccos \left(\frac{\langle \overrightarrow{yx}, \overrightarrow{yx'} \rangle}{\| \overrightarrow{yx} \| \cdot \| \overrightarrow{yx'} \|} \right) .$$



Figur 10.6

Als typisches Problem der euklidischen Geometrie behandeln wir die Aufgabe, den Abstand eines Punktes $\widetilde{x}$ von einer gegebenen Hyperebene H zu bestimmen. Es sei z. B.

$$H = \{ \sum_{j=1}^{n} a_j \, x_j = c \}$$

und $x^{(0)}$ ein fester Punkt auf H mit den Koordinaten $x_j^{(0)}$. Dann gilt also

$$H = \{ x \in \mathbb{R}^n : \langle a, x - x^{(0)} \rangle = 0 \}$$

mit dem Vektor $a = {}^t(a_1, \ldots, a_n) \neq 0$. Ohne Einschränkung können wir den Vektor a normieren: $\| a \| = 1$. Er ist dann (bis auf den Faktor ± 1) durch H eindeutig bestimmt und heißt *Normalen(einheits)vektor* zu H. Man bezeichnet eine solche „normierte“ Form der Hyperebenengleichung

$$\langle a, x \rangle = c \, , \quad \| a \| = 1$$

auch als *Hessesche Normalform* der Hyperebene H.

Man kann nun diese Normalform sehr günstig benutzen, um den gesuchten Abstand des Punktes $\widetilde{x}$ von H zu bestimmen. Ist nämlich

$$\widetilde{x} - x^{(0)} = \lambda \, a - \mu \, (x - x^{(0)}) \quad \text{mit} \quad x \in H \, ,$$

so ist $\lambda \in \mathbb{R}$ (bis auf das Vorzeichen) gerade die gesuchte Größe. Andererseits ist aber

$$\lambda = \lambda \langle a, a \rangle = \langle a, \lambda a \rangle = \langle a, \widetilde{x} - x^{(0)} \rangle$$

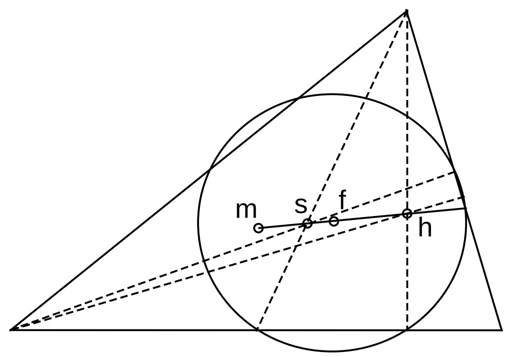
und damit

$$| \lambda | = | \langle a, \widetilde{x} \rangle - c | .$$

In der euklidischen Geometrie können wir auch von *Mittelsenkrechten, Winkelhalbierenden* und *Loten* in Dreiecken sprechen. Wir bezeichnen mit s den *Schwerpunkt* eines Dreiecks, d. h. den Schnittpunkt seiner Seitenhalbierenden, mit m den der Mittelsenkrechten und mit h den der *Höhen*, d. h. der Lote von den Ecken auf die gegenüberliegenden Seiten. Ferner können wir auch von *Kreisen* reden. Jeder Kreis ist durch die Vorgabe von drei nichtkollinearen Punkten bestimmt. Also gibt es zum Beispiel stets einen Kreis durch die Seitenmittelpunkte eines Dreiecks, der oft als *Feuerbach–Kreis* bezeichnet wird. Seinen Mittelpunkt nennen wir f.

Ein charakteristischer Satz in diesem Zusammenhang ist die folgende Zusammenfassung zweier Resultate von Euler und Feuerbach (ein Teil wird in den Übungen behandelt):

Satz 10.13 *Für jedes Dreieck in der euklidischen Ebene liegen die drei Punkte* m, s, f *und* h *auf einer Geraden, und die Punkte* s, f *zerlegen die Strecke von* m *nach* h *im Verhältnis* $2 : 1 : 3$.



Figur 10.7

Im euklidischen Fall sind die *Bewegungen* (oder *Isometrien*) die „relevanten“ Transformationen im Sinne von Felix Kleins Geometrie–Entwurf; sie werden gekennzeichnet durch die Bedingung

$$d\,(f(x_1),\, f(x_2)) = d\,(x_1,\, x_2) \quad \text{für alle} \quad x_1, x_2 \in X .$$

Alle Bewegungen zusammen bilden natürlich eine Gruppe.

Man überzeugt sich leicht davon, daß im Falle des euklidischen Standard–Raumes, den man manchmal auch mit $\mathbb{E}_n$ bezeichnet, die Translationen zu der Gruppe der Bewegungen gehören. Hat man nun in $\mathbb{E}_n$ eine beliebige Bewegung gegeben, so kann man durch Hinterschaltung einer geeigneten Translation erreichen, daß ein vorgegebener Punkt x_0 fest bleibt, so daß die Bestimmung der Bewegungsgruppe reduziert wird auf die Frage, welches die Bewegungen mit einem Fixpunkt sind. Wählt man diesen zum Ursprung des unterliegenden euklidischen Vektorraumes V, so müssen diese Bewegungen zumindest die Norm festlassen. Wir werden in Kapitel 13 zeigen, daß diese dann notwendig linear sind (und auch das Skalarprodukt in V respektieren). Es handelt sich also um *orthogonale* Endomorphismen. Man sieht damit, daß die Gruppe der Bewegungen ein (sogenanntes „verschränktes“) Produkt der abelschen Gruppe der Translationen mit der orthogonalen Gruppe $\mathrm{O}\,(n)$ ist. Insbesondere sind alle Bewegungen affin!

Wir schließen mit ein paar Bemerkungen zu *Symmetriegruppen*. Ist $K \subset \mathbb{E}_n$ eine Teilmenge, so bezeichnen wir als *Symmetrie* von K jede Bewegung F des umgebenden euklidischen Raumes, die K festläßt: $F\,(K) = K$. Die Gesamtheit der Symmetrien bildet eine Gruppe, die *Symmetriegruppe*

$\mathrm{Sym}\,(K)$ von K. Man kann mit etwas Analysis leicht zeigen, daß alle Symmetrien einer *kompakten* Menge K einen gemeinsamen Fixpunkt besitzen müssen. Wählt man diesen wieder als Ursprung, so erkennt man, daß die Symmetriegruppe einer kompakten Menge in dem euklidischen Raum $\mathbb{E}_n$ stets als Untergruppe von $\mathrm{O}\,(n)$ realisiert werden kann. Besonders interessant sind in diesem Zusammenhang die Fälle von *endlichen* Symmetriegruppen im *dreidimensionalen* Raum. Diese werden realisiert durch die *Platonischen Körper* (und zwei weitere Serien mit *zyklischer* bzw. *Dieder*–Symmetrie).

11 Diagonalisierbarkeit und Trigonalisierbarkeit von Endomorphismen

Wir wollen in diesem und in dem folgenden Kapitel die beschreibenden Matrizen von Endomorphismen auf endlich–dimensionalen Vektorräumen auf *bestmögliche* Gestalt bringen; dies ist das sogenannte *Normalformen–Problem* für $n \times n$–Matrizen.

Das Einfachste (aber auch gleichzeitig das Langweiligste), was wir evtl. erreichen können, ist sicherlich Diagonalgestalt. Wir stellen also als erstes die Frage: Wann kann ein Endomorphismus $F \in \mathrm{End}\,(V)$ bzgl. einer geeigneten Basis von V in Diagonalgestalt dargestellt werden, wann ist er *diagonalisierbar*? Diese Frage ist gleichbedeutend mit: Wann ist eine Matrix $A \in M(n \times n, \mathbb{K})$ *ähnlich* zu einer Diagonalmatrix

$$\mathrm{diag}\,(\lambda_1, \ldots, \lambda_n) \;=\; \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} ,$$

d. h.: Wann gibt es eine invertierbare Matrix $S \in \mathrm{GL}(n, \mathbb{K})$ mit $S^{-1}AS \;=\; \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n)$, oder schließlich: wann gibt es eine Basis $(v_1, \ldots, v_n)$ von V mit $F(v_j) = \lambda_j v_j$ für alle j?

Die letzte Umformulierung der Fragestellung gibt sofort Anlaß zu der folgenden

Definition. Es sei $F \in \mathrm{End}_{\mathbb{K}}\, V$; dann heißt ein Vektor $v \in V$, $v \neq 0$ ein *Eigenvektor* von F, falls es ein $\lambda \in \mathbb{K}$ gibt, so daß $F(v) = \lambda v$. Die zu Eigenvektoren v gehörenden Skalare λ heißen *Eigenwerte* von F.

Warnung: Ein *Eigenvektor* v ist per definitionem stets ungleich 0, ein *Eigenwert* kann aber durchaus Null sein.

Folgerung 11.1 *Eine $n \times n$–Matrix A ist genau dann diagonalisierbar, wenn es eine Basis von $\mathbb{K}^n$ gibt, die nur aus Eigenvektoren von A besteht.*

Wir betrachten nun die Menge aller Eigenvektoren zu einem festen Eigenwert λ. Diese kann keinen Untervektorraum von V bilden, da der Nullvektor nicht dazugehört. Wir definieren deshalb den *Eigenraum* von F zum Wert λ durch

$$\mathrm{Eig}\,(F;\, \lambda) \;:=\; \{v \in V : \, F(v) \;=\; \lambda\, v\} \; .$$

Es gilt dann der folgende

Satz 11.2 *Es sei F ein Endomorphismus des $\mathbb{K}$–Vektorraums V, und λ sei ein Wert in dem Grundkörper $\mathbb{K}$. Dann gilt:*

a) $\mathrm{Eig}\,(F;\, \lambda)$ *ist ein Untervektorraum von* V;

b) λ *ist ein Eigenwert von F genau dann, wenn* $\mathrm{Eig}\,(F;\, \lambda) \neq \{0\}$;

c) $\mathrm{Eig}\,(F;\, \lambda) \setminus \{0\}$ *ist die Menge der zu λ gehörenden Eigenvektoren*;

d) $\mathrm{Eig}\,(F;\, \lambda) \;=\; \ker\,(F \,-\, \lambda\,\mathrm{id})$;

e) *Ist $\lambda_1 \neq \lambda_2$, so gilt* $\mathrm{Eig}\,(F;\, \lambda_1) \,\cap\, \mathrm{Eig}\,(F;\, \lambda_2) \;=\; \{0\}$.

Beweis. Die Behauptungen a), b) und c) sollten unmittelbar klar sein. Auch d) und e) sind sehr einfach:

d) $\qquad v \in \mathrm{Eig}\,(F;\, \lambda) \Longleftrightarrow F(v) = \lambda v \Longleftrightarrow (F - \lambda\,\mathrm{id})(v) = 0 \Longleftrightarrow v \in \ker\,(F - \lambda\,\mathrm{id})$.

e) $\qquad \lambda_1 v \;=\; F(v) \;=\; \lambda_2 v \Longrightarrow (\lambda_1 \,-\, \lambda_2) v \;=\; 0 \Longrightarrow v \;=\; 0$. $\qquad \square$

Zusatz. *Ist* $\dim V < \infty$ *, so gilt außerdem*

$$\mathrm{Eig}\,(F;\,\lambda) \, \neq \, \{0\} \iff \det(F \, - \, \lambda\, \mathrm{id}) \, = \, 0 \; .$$

Hierbei ist $\det G$ *für einen Endomorphismus* G *definiert durch* $\det B$ *, wenn* B *die Matrix von* G *bezgl. irgendeiner Basis ist. (Daß diese Definition unabhängig von Basen ist, sehen wir noch weiter unten).*

Denn: $\mathrm{Eig}\,(F;\,\lambda) \neq \{0\} \iff \lambda$ ist Eigenwert zu $F \iff \exists v \neq 0$ so daß $v \in \ker(F - \lambda\,\mathrm{id}) \iff F - \lambda\,\mathrm{id}$ ist nicht injektiv $\iff F - \lambda\,\mathrm{id}$ ist nicht bijektiv $\iff \det(F - \lambda\,\mathrm{id}) = 0$. □

Lemma 11.3 *Es sei* $\dim V = n < \infty$ *, und der Endomorphismus* F *besitze Eigenvektoren* $v_1, \ldots, v_n$ *zu den paarweise verschiedenen Eigenwerten* $\lambda_1, \ldots, \lambda_n$ *. Dann ist* $(v_1, \ldots, v_n)$ *eine Basis von* V *.*

Beweis. Es ist nur zu zeigen, daß die Vektoren $v_1, \ldots, v_m$ linear unabhängig sind für $1 \leq m \leq n$. Dazu führen wir Induktion nach m, wobei wegen $v_1 \neq 0$ im Falle $m = 1$ nichts zu zeigen ist. Es sei also die Unabhängigkeit der ersten $m - 1$ Vektoren schon gezeigt, und es gelte $v = a_1 v_1 + \cdots + a_m v_m = 0$. Dann folgt

$$a_1(\lambda_1 - \lambda_m)\, v_1 \, + \, \cdots \, + \, a_{m-1}(\lambda_{m-1} - \lambda_m) v_{m-1} \, = \, F(v) \, - \, \lambda_m v \, = \, 0$$

und damit $a_1 = \cdots = a_{m-1} = 0$, was unmittelbar auch $a_m = 0$ wegen $v_m \neq 0$ nach sich zieht. □

Wie kann man die Eigenwerte eines Endomorphismus F tatsächlich bestimmen? Nach den obigen Ergebnissen läuft dies jedenfalls darauf hinaus, daß bzgl. einer beschreibenden Matrix A von F die Determinante $\det(A - \lambda E_n)$ verschwinden muß. Aus diesem Grunde betrachten wir für jede Matrix $A \in M(n \times n,\, \mathbb{K})$ die Abbildung

$$\mathbb{K} \ni t \longmapsto \det(A - tE_n) \in \mathbb{K} \; .$$

Wie man sich leicht durch Determinantenentwicklung mittels der expliziten Form

$$\chi_A \; : \; t \longmapsto \det \begin{pmatrix} a_{11} - t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - t & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - t \end{pmatrix}$$

überlegt, wird dadurch eine *polynomiale* Abbildung χ_A gegeben, die von der Gestalt

$$\chi_A(t) \; = \; (-1)^n t^n \; + \; (-1)^{n-1} \alpha_{n-1} t^{n-1} \; + \cdots + \; \alpha_0$$

ist, wobei durch Einsetzen von $t = 0$ sofort

$$\alpha_0 \; = \; \det A$$

kommt. Außerdem überzeugt man sich leicht, daß

$$\alpha_{n-1} \; = \; a_{11} \; + \cdots + \; a_{nn} \; =: \; \mathrm{spur} A \; .$$

Wir schreiben im folgenden $\pm\chi_A$, wenn wir das *normierte* Polynom $(-1)^n \chi_A(t) = \det\,(tE - A)$ meinen.

Wenn man berücksichtigt, daß man Determinanten statt mit Einträgen in dem Grundkörper $\mathbb{K}$ auch ohne Schwierigkeiten für Matrizen mit Einträgen in dem kommutativen Ring $\mathbb{K}\,[\,t\,]$ der Polynome in t erklären kann, so kann man χ_A auch als *echtes* Polynom in der Variablen t über dem Grundkörper $\mathbb{K}$ vom Grad n auffassen:

$$\chi_A \; \in \; \mathbb{K}\,[\,t\,] \; .$$

Folgerung 11.4 λ *ist genau dann ein Eigenwert der Matrix* $A \in M(n \times n,\, \mathbb{K})$ *, wenn* λ *eine Nullstelle des charakteristischen Polynoms* χ_A *in* $\mathbb{K}$ *ist.*

Wir wollen noch zeigen, daß das charakteristische Polynom $\chi_F := \chi_A$ jedem Endomorphismus F *intrinsisch* (d. h. unabhängig von jeder Basis) zugeordnet ist. Damit haben natürlich auch alle Koeffizienten von χ_F intrinsische Bedeutung für F. Wichtig ist aber außer $\alpha_0 = \det F$ i. a. nur die *Spur* von F:

$$\operatorname{spur} F = \operatorname{spur} A .$$

Um dieses einzusehen, reicht es, das folgende Lemma zu beweisen:

Lemma 11.5 *Sind A und B ähnlich, so gilt $\chi_A(t) = \chi_B(t)$.*

Beweis. Aus $A = S^{-1}BS$ folgt unmittelbar

$$\begin{aligned}\chi_A(t) &= \det(A - t\,E_n) = \det(S^{-1}\,B\,S - t\,E_n)\\ &= \det(S^{-1}\,B\,S - t\,S^{-1}\,S) = \det(S^{-1}(B - t\,E_n)\,S)\\ &= \det S^{-1}\,\det S\,\det(B - t\,E_n) = \chi_B(t)\,. \end{aligned}\qquad\square$$

Beim Bestimmen der Eigenwerte eines Endomorphismus entsteht ein völlig neuartiges Phänomen, was die effektive Berechenbarkeit anbelangt. Während in dem eigentlichen Bereich der *linearen* Algebra praktisch alle Fragen im Prinzip *endlich algorithmisch* lösbar sind, handelt es sich jetzt um das Auffinden der Wurzeln eines Polynoms vom Grade $n = \dim V$. Dieses Problem gehört aber der *Algebra* an, in der explizite Formeln für $n = 2, 3, 4$ angegeben werden, die aber außer den üblichen arithmetischen Operationen auch das Ziehen von Wurzeln verlangen, so daß man schon hier nur näherungsweise vorgehen kann. Für $n \geq 5$ ist die Situation sogar theoretisch noch viel unbefriedigender: ABEL hat als erster bewiesen, daß es in diesen Fällen keine allgemeingültige Formel von dem eben beschriebenen Typ geben kann, die es erlaubt, aus den Koeffizienten eines Polynoms seine Wurzeln zu bestimmen.

Kommen wir noch einmal zu der Charakterisierung der Eigenwerte zurück, die wir etwas umformulieren wollen. Wir wissen: λ ist Eigenwert von $F \iff \lambda$ ist Nullstelle von $\chi_F \iff \chi_F(t) = (t-\lambda)Q(t)$ mit einem Polynom $Q(t) \in \mathbb{K}[t]$ vom Grade $n-1$, d. h. $(t-\lambda)$ ist ein sogenannter *Linearfaktor* von χ_F. Die letzte Umformulierung ergibt sich aus der Existenz der Division mit Rest in jedem Polynomring $\mathbb{K}[t]$ (siehe z. B. Anhang FISCHER):

$$\chi_F(t) = (t-\lambda)\,Q(t) + R(t)\,, \quad \deg R < \deg(t-\lambda) = 1\,.$$

Also ist $R \in \mathbb{K}$, und mit $t = \lambda$ folgt $\chi_F(\lambda) = 0$, also $R = 0$.

Wir können nun aus dem charakteristischen Polynom für jeden Eigenwert λ die maximale Potenz des Linearfaktors $(t-\lambda)$ herausziehen:

$$\chi_F(t) = (t-\lambda)^m\,Q(t)\,, \quad Q(\lambda) \neq 0\,.$$

Wir schreiben $\mu(F;\,\lambda) := m$.

Definition. $\mu(F;\,\lambda)$ heißt die (algebraische) *Multiplizität* oder *Vielfachheit* des Eigenwertes λ von F. Die Zahl $\dim_{\mathbb{K}} \operatorname{Eig}(F;\,\lambda)$ nennt man die *geometrische* Multiplizität.

Klar ist: Kennt man λ, so ist der Eigenraum $\operatorname{Eig}(F;\,\lambda)$ und damit auch seine Dimension einfach zu bestimmen, denn es ist ja

$$\operatorname{Eig}(F;\,\lambda) = \ker(F - \lambda\,\mathrm{id})$$

der Lösungsraum eines homogenen Gleichungssystems.

Lemma 11.6 $$\mu(F;\,\lambda) \geq \dim_{\mathbb{K}} \operatorname{Eig}(F;\,\lambda)\,.$$

Beweis. Sei $v_1, \ldots, v_r$ eine Basis von $\mathrm{Eig}\,(F;\, \lambda)$; ergänze diese durch $v_{r+1}, \ldots, v_n$ zu einer Basis von V. Dann wird F relativ zu dieser Basis beschrieben durch eine Matrix der Gestalt

$$A = \left(\begin{array}{c|c} \lambda E_r & * \\ \hline 0 & A' \end{array} \right) ,$$

wobei $r = \dim \mathrm{Eig}\,(F;\, \lambda)$. Hieraus ergibt sich sofort, daß

$$\chi_A(t) = \det\,(A - t\,E) = (\lambda - t)^r \det\,(A' - t\,E)$$

und $r \le \mu(F;\, \lambda)$. □

Satz 11.7 *Für einen Endomorphismus F sind die folgenden Bedingungen äquivalent:*

1. *F ist diagonalisierbar* ;

2. i) *$\chi_F(t)$ zerfällt vollständig in Linearfaktoren* :

$$\pm\chi_F(t) = (t - \lambda_1)^{m_1} \cdot \ldots \cdot (t - \lambda_\ell)^{m_\ell} , \quad \lambda_j \neq \lambda_k ,\ 1 \le j < k \le \ell ;$$

und

ii) *für jeden Eigenwert λ ist*

$$\mu(F;\, \lambda) = \dim \mathrm{Eig}\,(F;\, \lambda) ;$$

3. $$\sum_{j=1}^{\ell} \dim \mathrm{Eig}\,(F;\, \lambda_j) = \dim V ;$$

4. $$V = \bigoplus_{j=1}^{\ell} \mathrm{Eig}\,(F;\, \lambda_j) .$$

Bevor wir diesen Satz beweisen, machen wir noch die folgende Bemerkung, die eine Verschärfung von Satz 2.e) ist.

Lemma 11.8 *Sind $\lambda_1, \ldots, \lambda_k$ paarweise verschiedene Eigenwerte des Endomorphismus F, so ist die Summe*

$$\sum_{j=1}^{k} \mathrm{Eig}\,(F;\, \lambda_j)$$

direkt und insbesondere

$$\dim \sum_{j=1}^{k} \mathrm{Eig}\,(F;\, \lambda_j) = \sum_{j=1}^{k} \dim \mathrm{Eig}\,(F;\, \lambda_j) .$$

Beweis. Wir führen Induktion nach k, wobei für $k = 1$ nichts zu beweisen ist. Sei also die obige Summe für ein $k \ge 1$ direkt, und sei λ_{k+1} ein weiterer Eigenwert. Wir brauchen dann nach Übungsaufgabe 28 nur noch zu zeigen, daß

$$\Big(\sum_{j=1}^{k} \mathrm{Eig}\,(F;\, \lambda_j)\Big) \cap \mathrm{Eig}\,(F;\, \lambda_{k+1}) = \{0\} .$$

Es sei also $v \in \mathrm{Eig}\,(F;\, \lambda_{k+1})$, $v = \sum_{j=1}^{k} v_j$, $v_j \in \mathrm{Eig}\,(F;\, \lambda_j)$. Dann gilt:

$$\begin{aligned} \lambda_{k+1}\, v_1 + \cdots + \lambda_{k+1}\, v_k &= \lambda_{k+1}\, v = F(v) = F(v_1 + \cdots + v_k) \\ &= F(v_1) + \cdots + F(v_k) = \lambda_1\, v_1 + \cdots + \lambda_k\, v_k , \end{aligned}$$

und daraus folgt wegen $\lambda_j \neq \lambda_{k+1}$, $j \leq k$, sofort $v_1 = \cdots = v_k = 0$ und damit $v = 0$. □

Beweis (von Satz 7). 1. $\Longrightarrow$ 2. F ist diagonalisierbar $\Longleftrightarrow$ F wird in geeigneter Basis durch eine Diagonalmatrix

$$A = \mathrm{diag}\,(\underbrace{\lambda_1, \ldots, \lambda_1}_{m_1}, \ldots, \underbrace{\lambda_\ell, \ldots, \lambda_\ell}_{m_\ell})$$

dargestellt mit $\lambda_j \neq \lambda_k$ für $1 \leq j < k \leq \ell$. Daraus folgt

$$\chi_A(t) = (\lambda_1 - t)^{m_1} \cdot \ldots \cdot (\lambda_\ell - t)^{m_\ell} \ ,$$

und $\mathrm{Eig}\,(F;\, \lambda_j) \;=\; \ker\,(F - \lambda_j\,\mathrm{id})$ besitzt die Dimension m_j.

2. $\Longrightarrow$ 3.

$$\sum_{j=1}^{\ell} \dim \mathrm{Eig}\,(F;\, \lambda_j) \;=\; \sum_{j=1}^{\ell} \mu\,(F;\, \lambda_j) \;=\; \sum_{j=1}^{\ell} m_j \;=\; n \;=\; \dim V \ .$$

3. $\Longrightarrow$ 4. Aus 3. folgt mit Lemma 8 sofort

$$\dim V \;=\; \sum_{j=1}^{\ell} \dim \mathrm{Eig}\,(F;\, \lambda_j) \;=\; \dim \sum_{j=1}^{\ell} \mathrm{Eig}\,(F;\, \lambda_j) \ ,$$

also

$$V \;=\; \sum_{j=1}^{\ell} \mathrm{Eig}\,(F;\, \lambda_j) \;=\; \bigoplus_{j=1}^{\ell} \mathrm{Eig}\,(F;\, \lambda_j) \ .$$

4. $\Longrightarrow$ 1. Aus $V = \bigoplus\limits_{j=1}^{\ell} \mathrm{Eig}\,(F;\, \lambda_j)$ ergibt sich sofort Diagonalgestalt bzgl. der Vereinigung von Basen der Eigenräume. □

Als Folgerung erhält man hieraus erneut das frühere Lemma 3.

Beispiele. 1. Die Matrix

$$A_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \in M(2 \times 2,\, \mathbb{R})$$

besitzt das charakteristische Polynom

$$\chi_{A_\alpha}(t) \;=\; (\cos\alpha - t)^2 \;+\; \sin^2\alpha \;=\; t^2 - 2\cos\alpha \cdot t \;+\; 1 \ ,$$

das genau dann reelle Wurzeln besitzt, wenn

$$\cos^2\alpha - 1 \;\geq\; 0 \ .$$

Die letzte Bedingung ist gleichbedeutend mit $\cos^2\alpha \geq 1$, also $\cos^2\alpha = 1$ und damit $\alpha = 0$ oder π modulo 2π . Somit sind die einzigen Matrizen A_α mit reellen Eigenwerten die Matrizen

$$A_0 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad A_\pi \;=\; \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \ .$$

I. a. hat also eine reelle Matrix A *keine* (reelle) Eigenwerte.

2. Die Matrix

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

besitzt das charakteristische Polynom

$$\chi_A(t) \;=\; \det \begin{pmatrix} 1-t & 1 \\ 0 & 1-t \end{pmatrix} \;=\; (1-t)^2\;,$$

so daß $\mu(A;\,1) = 2$. Andererseits ist

$$\mathrm{Eig}\,(A;\,1) \;=\; \ker\,(A - E_2) \;=\; \{(x_1,\,x_2) : \,x_2 = 0\}\;,$$

so daß $\dim \mathrm{Eig}\,(A;\,1) = 1 \neq 2 = \mu(A;\,1)$. A ist also *nicht* diagonalisierbar.

Wir wollen jetzt noch die Bedingung 2 i) in Satz 7 weiter untersuchen.

Satz 11.9 *Die folgenden Aussagen über einen Endomorphismus* $F \in \mathrm{End}\,(V)$ *sind äquivalent:*

a) F *ist* trigonalisierbar*, d. h. es existiert eine Basis, s. d.* F *durch eine obere Dreiecksmatrix* A *dargestellt wird.*

b) *Es existiert eine Fahne von Untervektorräumen*

$$\{0\} = V_0 \subset V_1 \subset \cdots \subset V_n = V\;, \quad \dim V_j = j\;,$$

die F*–invariant ist, d. h. für die* $F(V_j) \subset V_j$ *für alle* j *gilt.*

c) χ_F *zerfällt vollständig in Linearfaktoren:*

$$\pm\chi_F \;=\; \prod_{j=1}^{n} (t\,-\,\lambda_j)\;.$$

Beweis. b) $\Longrightarrow$ a) Aus der Existenz einer solchen Fahne folgt die Existenz einer Basis $\mathcal{B} = (v_1, \ldots, v_n)$, s. d. $V_j = \mathrm{span}\,(v_1, \ldots, v_j)$. Bzgl. dieser Basis besitzt die beschreibende Matrix wegen $F(v_j) \in V_j$, $j = 1, \ldots n$, obere Dreiecksgestalt.

a) $\Longrightarrow$ c): Ohne Einschränkung sei A eine obere Dreiecksmatrix mit Einträgen $(a_{11}, \ldots, a_{nn})$ in der Hauptdiagonalen. Dann ist

$$\chi_A(t) = \det\,(A - tE) = (a_{11} - t) \cdot \ldots \cdot (a_{nn} - t)\;.$$

Insbesondere sind die Eigenwerte von A genau die Diagonalelemente.

c) $\Longrightarrow$ b) Induktion nach $n \geq 1$. Der Fall $n = 1$ ist trivial. Es sei also $n \geq 2$. Dann hat man

$$\pm\chi_F(t) \;=\; (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_n)\;.$$

Es sei v_1 Eigenvektor zu λ_1, $V_1 = \mathrm{span}\,(v_1) \Longrightarrow F(V_1) \subset V_1$, und die beschreibende Matrix A hat die folgende Gestalt bzgl. einer beliebigen Basis–Ergänzung $(v_1,\, w_2, \ldots, w_n)$:

$$A \;=\; \left(\begin{array}{c|cc} \lambda_1 & * & * \\ \hline 0 & & \\ \vdots & & A' \\ 0 & & \end{array}\right)\;.$$

Es sei $V' = \mathrm{span}\,(w_2, \ldots, w_n)$, so daß $V = V_1 \oplus V'$. Setze dann den Homomorphismus $F' : V' \to V'$ an als die Komposition

$$V' \xrightarrow{F|V'} V \xrightarrow{\mathrm{pr}} V'\;,$$

wobei $\mathrm{pr}(a\,v_1 + \sum b_j\,w_j) = \sum b_j\,w_j$. F' wird offensichtlich beschrieben durch die oben eingezeichnete Matrix A' , und aus

$$(\lambda_1 - t) \cdot \ldots \cdot (\lambda_n - t) = \chi_A(t) = (\lambda_1 - t) \cdot \det(A' - t\,E_{n-1})$$

folgt unmittelbar

$$\chi_{F'}(t) = \chi_{A'}(t) = \pm(t - \lambda_2) \cdot \ldots \cdot (t - \lambda_n) .$$

Nach Induktionsvoraussetzung gibt es eine F'–invariante Fahne

$$\{0\} = V_0' \subset V_1' \subset \cdots \subset V_{n-1}' = V' .$$

Setzt man nun $V_2 = V_1' + V_1 \subset V_3 = V_2' + V_1 \subset \cdots \subset V_n = V' + V_1 = V$, so ist auch $\{0\} = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_n = V$ eine Fahne, und $V_j \ni u_j = u_{j-1}' + u_1$ impliziert

$$F(u_j) = F(u_{j-1}') + F(u_1) \equiv F'(u_{j-1}') \text{ modulo } V_1$$

und damit

$$F(u_j) \in V_{j-1}' + V_1 = V_j . \qquad \square$$

Definition. Ein Körper $\mathbb{K}$ heißt *algebraisch abgeschlossen*, wenn eine der beiden folgenden äquivalenten Bedingungen erfüllt ist:

1. Jedes Polynom $P \in \mathbb{K}[\,t\,]$ vom Grad $\deg P \geq 1$ hat mindestens eine Nullstelle in $\mathbb{K}$.
2. Jedes Polynom zerfällt vollständig in Linearfaktoren.

Folgerung 11.10 *Über einem algebraisch abgeschlossenen Körper ist jede $n \times n$–Matrix trigonalisierbar.*

Damit ist über dem Grundkörper der *komplexen Zahlen* jeder Endomorphismus trigonalisierbar wegen

Satz 11.11 (Fundamentalsatz der Algebra) *Der Körper $\mathbb{C}$ ist algebraisch abgeschlossen.*

Einen *Beweis* für diesen Fundamentalsatz findet man in jedem Lehrbuch der Funktionentheorie.

Für den *Satz von Cayley–Hamilton* gebe ich einen Beweis, der den Vorteil hat, über beliebigen Grundkörpern gültig zu sein.

Satz 11.12 (Cayley–Hamilton) *Es sei $F \in \mathrm{End}(V)$ ein Endomorphismus mit charakteristischem Polynom*

$$\chi_F(t) = \pm(t^n - \alpha_1 t^{n-1} + \cdots \pm \alpha_n) \in \mathbb{K}[\,t\,] .$$

Dann verschwindet der Endomorphismus

$$\chi_F(F) := \pm(F^n - \alpha_1 F^{n-1} + \cdots \pm \alpha_n F^0) .$$

Bemerkungen. 1. Für eine Matrix $A \in M(n \times n,\, \mathbb{K})$ bedeutet dieser Satz das Bestehen der polynomialen Identität

$$A^n - \alpha_1 A^{n-1} + \cdots \pm \alpha_n A^0 = 0 ,$$

die man *nicht* durch bloßes Einsetzen von A in $\chi_A = \det(A - tE)$ beweisen kann!

2. Daß A jedoch irgendeiner polynomialen Identität genügen muß ist klar, da wegen $\dim M(n \times n,\, \mathbb{K}) = n^2$ die Matrizen A^0 , $A^1, \ldots,$ A^{n^2} linear abhängig über $\mathbb{K}$ sind.

3. Es besteht somit die Menge $\operatorname{Ann} A = \{P \in \mathbb{K}[t] : P(A) = 0\}$ nicht nur aus dem Nullpolynom. Offensichtlich ist $\operatorname{Ann} A$ ein *Ideal* in $\mathbb{K}[t]$, d. h.: mit $P_1, P_2 \in \operatorname{Ann} A$ ist auch $P_1 + P_2 \in \operatorname{Ann} A$ und $P \in \operatorname{Ann} A$, $Q \in \mathbb{K}[t]$ impliziert $QP \in \operatorname{Ann} A$. Nun kann man zeigen, daß $\mathbb{K}[t]$ ein *Hauptidealring* ist, d. h. jedes Ideal $\mathfrak{a}$ in $\mathbb{K}[t]$ wird erzeugt von (genau) einem (normierten) Polynom $\mu(t)$:

$$\mathfrak{a} = \{Q(t)\mu(t) : Q(t) \in \mathbb{K}[t]\} .$$

(Der Ring $\mathbb{Z}$ hat übrigens die gleiche Eigenschaft). Folglich gibt es zu jeder Matrix $A \in M(n \times n, \mathbb{K})$ ein sogenanntes *Minimalpolynom* $\mu_A(t)$ mit $\mu_A(A) = 0$, das wegen des Satzes von Cayley–Hamilton das charakteristische Polynom teilt (insbesondere einen Grad $\leq n$ besitzt). Zerfällt speziell χ_A in Linearfaktoren

$$\chi_A(t) = \pm(t - \lambda_1)^{m_1} \cdot \ldots \cdot (t - \lambda_r)^{m_r}$$

so ist notwendig

$$\mu_A(t) = \pm(t - \lambda_1)^{m_1'} \cdot \ldots \cdot (t - \lambda_r)^{m_r'} , \quad m_\rho' \leq m_\rho ,$$

so daß man in jedem Einzelfall μ_A durch Einsetzen von A in die Teiler von χ_A bestimmen kann. (Man kann übrigens a priori zeigen, daß alle $m_\rho' \geq 1$ sein müssen, d. h. daß μ_A dieselben Nullstellen wie χ_A besitzt).

Nun also zum *Beweis* des Satzes von Cayley–Hamilton: Ist $v \neq 0$ ein beliebiger Vektor in V, so gibt es eine eindeutig bestimmte *maximale* Zahl $r \leq n$, s. d. $v_1 = v$, $v_2 = F(v_1), \ldots, v_r = F(v_{r-1}) = F^{r-1}(v_1)$ linear unabhängig sind. Dann gilt $F(v_r) = a_1 v_1 + \cdots + a_r v_r$, und $U = \operatorname{span}(v_1, \ldots, v_r)$ ist ein F–invarianter Unterraum. Hieraus folgt, wie schon oft verwendet, daß $\chi_F(t) = Q(t) \cdot \chi_{F|U}(t)$.

Wir wollen zeigen: $\chi_F(F) = 0 \in \operatorname{End} V$, d. h. $\chi_F(F)(v) = 0$ für alle $v \in V$. Dazu genügt offensichtlich $\chi_{F|U}(F)(v) = 0$ für alle $v \neq 0$, wobei U wie oben mit Hilfe des Vektors v konstruiert wird. Nun hat $F|U$ eine beschreibende Matrix von der Form

$$B = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_1 \\ 1 & 0 & \cdots & 0 & a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_r \end{pmatrix} ,$$

aus der man leicht ableitet, daß $\pm\chi_{F|U}(t) = \det(tE_r - B) = t^r - a_r t^{r-1} - \cdots - a_1$ und

$$\begin{aligned} \pm\chi_{F|U}(F)(v) &= (F^r - a_r F^{r-1} - \cdots - a_1 F^0)(v) \\ &= F(v_r) - (a_1 v_1 + \cdots + a_r v_r) = 0 . \end{aligned}$$

□

12 Jordansche Normalform

Ist $F \in \operatorname{End} V$ und $V_1 \subset V$ ein F–invarianter Unterraum, so kann man fragen, ob es einen direkten Summanden V_2 von V_1 in V gibt, der selbst F–invariant ist. Dies wäre natürlich gleichbedeutend mit einer Blockstruktur der beschreibenden Matrix in einer geeigneten Basis:

$$\left(\begin{array}{c|c} * & 0 \\ \hline 0 & * \end{array}\right) .$$

Eine solche Zerlegung ist jedoch i. a. nicht möglich. Anderenfalls würde der Beweis für die Trigonalisierbarkeit von Endomorphismen über algebraisch abgeschlossenen Körpern stets die Diagonalisierbarkeit zur Folge haben, und dies ist aufgrund eines Gegenbeispiels im letzten Kapitel nicht richtig! – Unter gewissen Voraussetzungen gibt es aber invariante Zerlegungen, und diese sind das Herzstück des Beweises über die Existenz der Jordanschen Normalform.

Es sei i. f. V ein $\mathbb{K}$–Vektorraum der endlichen Dimension n, und $F \in \operatorname{End} V$ sei ein Endomorphismus. Über den Körper $\mathbb{K}$ wird vorerst noch nichts vorausgesetzt. Wir betrachten zuerst die Folge aller Potenzen F^k und die Unterräume $\ker F^k$, $\operatorname{im} F^k$. Dann gilt offensichtlich

$$\begin{cases} \ker F \subset \ker F^2 \subset \ker F^3 \subset \cdots , \\ \operatorname{im} F \supset \operatorname{im} F^2 \supset \operatorname{im} F^3 \supset \cdots , \end{cases}$$

wobei alle auftretenden Vektorräume F–invariant sind.

Wegen $\dim V = n < \infty$ muß die Kette der $\ker F^j$ „stationär“ werden. Insbesondere gibt es eine kleinste positive Zahl q, so daß $\ker F^q = \ker F^{q+1}$. Dann ist aber auch

$$\ker F^{q+j} = \ker F^q \quad \text{für alle} \quad j \geq 1 .$$

Denn ist $F^{q+j}(v) = 0$, $j \geq 2$, so folgt $F^{q+1}(F^{j-1}(v)) = 0$ und damit $F^{q+j-1}(v) = F^q(F^{j-1}(v)) = 0$, so daß sich der Beweis leicht durch Induktion beenden läßt.

Satz 12.1 *Mit den obigen Bezeichnungen gilt*

$$(*) \qquad V = \ker F^q \oplus \operatorname{im} F^q .$$

$F|\ker F^q$ ist nilpotent von der Ordnung q (d. h. die q–te Potenz dieses Endomorphismus verschwindet, aber keine niedrigere Potenz), und $F|\operatorname{im} F^q$ ist invertierbar.

Die Zerlegung $(*)$ ist überdies eindeutig: Zerfällt V in eine direkte Summe $V_1 \oplus V_2$ von F–invarianten Unterräumen mit $F|V_1$ nilpotent, $F|V_2$ invertierbar, so ist notwendig $V_1 = \ker F^q$, $V_2 = \operatorname{im} F^q$.

Beweis (aus dem Buch von Halmos). Wegen der Rangformel ist $\dim \ker F^j + \dim \operatorname{im} F^j = n = \dim V$ für alle $j \geq 1$. Um $(*)$ zu zeigen, benötigt man deshalb nur noch

$$\ker F^q \cap \operatorname{im} F^q = \{0\} .$$

Es sei also v in dem Durchschnitt auf der linken Seite enthalten; dann folgt $F^q(v) = 0$ und $v = F^q(w)$, also $F^{2q}(w) = 0$ und damit wegen der Bemerkung unmittelbar vor Satz 1 auch $v = F^q(w) = 0$. F^q verschwindet selbstverständlich auf $\ker F^q$, nicht aber eine kleinere Potenz wegen $\ker F^j \neq \ker F^q$, $j < q$.

Ist schließlich $v \in \operatorname{im} F^q$, d. h. $v = F^q(w)$, und $F(v) = 0$, so folgt wieder $F^{q+1}(w) = 0$ und $v = F^q(w) = 0$. Also ist $F|\operatorname{im} F^q$ injektiv und damit auch invertierbar.

Zur Eindeutigkeitsaussage: Hat man eine Zerlegung $V = V_1 \oplus V_2$ mit den geforderten Eigenschaften, so ist notwendig wegen $F^k|V_1 = 0$ für ein geeignetes k:

$$V_1 \subset \ker F^k \subset \ker F^{\max(k,q)} = \ker F^q .$$

Da F auf V_2 invertierbar ist, gilt $V_2 \subset \operatorname{im} F^k$ für alle k. Insbesondere ist $V_2 \subset \operatorname{im} F^q$. Aus (∗) folgt dann sofort $V_1 = \ker F^q$, $V_2 = \operatorname{im} F^q$. □

Bevor wir uns dem Jordanschen Normalformen–Problem zuwenden (das die algebraische Abgeschlossenheit von $\mathbb{K}$ oder zumindest die Linearfaktor–Zerlegung von χ_F des betreffenden Endomorphismus erfordert), bringen wir nilpotente Endomorphismen in Normalgestalt. Dies erklärt die Einsen auf der Nebendiagonalen in der Jordanschen Normalform (siehe Folgerung 4). Bei diesem Beweis folgen wir FISCHER: *Lineare Algebra.*

Satz 12.2 *Es sei W ein endlich–dimensionaler $\mathbb{K}$–Vektorraum, und $N \in \operatorname{End} W$ sei nilpotent. Dann gibt es eine Basis $(w_1, \ldots, w_n)$ von W mit*

$$N(w_j) = w_{j-1} \quad \textit{oder} \quad = 0 \, .$$

In einer solchen Basis hat also N eine Matrix–Darstellung, die aus lauter Blöcken der Gestalt

$$\left.\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}\right\} r_\ell$$

entlang der Hauptdiagonalen besteht. Die Darstellung ist eindeutig, wenn man die r_ℓ monoton fallend anordnet : $r_1 \ge r_2 \ge \cdots$. Dabei ist r_1 die Nilpotenzordnung q von N.

Beweis. Wir betrachten wie vorhin die Folge $\{0\} = V_0 = \ker N^0 \subset V_1 = \ker N^1 \subset \cdots \subset V_q = \ker N^q = W$. Es gilt dann

a) $V_{j-1} \ne V_j$,

b) $N^{-1}(V_{j-1}) = V_j$,

c) ist U ein Untervektorraum von W mit $U \cap V_j = \{0\}$, so ist $N|U$ injektiv.

Zu a). Beweis des vorigen Satzes.

Zu b).

$$v \in N^{-1}(V_{j-1}) \iff N(v) \in V_{j-1} \iff N^{j-1}N(v) = 0 \iff N^j(v) = 0 \iff v \in V_j \, .$$

Zu c). $V_1 = \ker N \subset V_j \Longrightarrow U \cap V_1 \subset U \cap V_j = \{0\} \Longrightarrow N|U$ ist injektiv.

Als nächstes zeigen wir: Es gibt Unterräume $U_1, \ldots, U_q$ von W mit

d) $V_j = V_{j-1} \oplus U_j$, $j \ge 1$,

e) $N : U_j \hookrightarrow U_{j-1}$, $j \ge 2$.

Zu d) und e). U_q sei ein beliebiger direkter Summand von V_{q-1} in $V_q = W$. Aus $N^{-1}(V_{j-2}) = V_{j-1}$ folgt, wenn die Zerlegung $V_j = V_{j-1} \oplus U_j$ schon gegeben ist, $N(U_j) \cap V_{j-2} = \{0\}$. Also gibt es wegen $N(U_j) \subset N(V_j) \subset V_{j-1}$ eine Zerlegung

$$V_{j-1} = V_{j-2} \oplus U_{j-1} \quad \text{mit} \quad N(U_j) \subset U_{j-1} \, .$$

Wegen $U_j \cap V_{j-1} = \{0\}$ ist $N|U_j$ injektiv, $j \ge 2$ (Teil c)).

Wir können uns nun eine geeignete Basis verschaffen. Es ist

$$W = V_q = V_{q-1} \oplus U_q = V_{q-2} \oplus U_{q-1} \oplus U_q = \cdots = U_1 \oplus \cdots \oplus U_q \, .$$

Wähle dann eine Basis $u_1^{(q)}, \ldots, u_{\ell_q}^{(q)}$ von U_q. Da $N : U_q \hookrightarrow U_{q-1}$ injektiv ist, kann man $N(u_1^{(q)}), \ldots, N(u_{\ell_q}^{(q)})$ durch $u_1^{(q-1)}, \ldots, u_{\ell_{q-1}}^{(q-1)}$ zu einer Basis von U_{q-1} ergänzen. Im letzten Schritt erhält man dann eine Basis von U_1 der Gestalt:

$$N^{q-1}(u_1^{(q)}), \ldots, N^{q-1}(u_{\ell_q}^{(q)})\,, \quad N^{q-2}(u_1^{(q-1)}), \ldots, N^{q-2}(u_{\ell_{q-1}}^{(q-1)})\,, \quad \ldots$$

$$\ldots\,, \quad N(u_1^{(2)}), \ldots, N(u_{\ell_2}^{(2)})\,, \quad u_1^{(1)}, \ldots, u_{\ell_1}^{(1)}\,.$$

Wegen $U_1 = V_1 = \ker N$ bildet N diese Vektoren auf Null ab, und der Satz ist bewiesen, wenn man die gesamte Basis wie folgt anordnet:

$$\begin{array}{cccc}
N^{q-1}(u_1^{(q)})\,, & N^{q-2}(u_1^{(q)})\,, & \ldots, & u_1^{(q)} \\
N^{q-1}(u_2^{(q)})\,, & N^{q-2}(u_2^{(q)})\,, & \ldots, & u_2^{(q)} \\
\vdots & & & \\
N^{q-2}(u_1^{(q-1)})\,, & N^{q-3}(u_1^{(q-1)})\,, & \ldots, & u_1^{(q-1)} \\
\vdots & & & \\
\vdots & & &
\end{array}$$

$\square$

Um den Satz über die *Jordan–Zerlegung*, die eine Verfeinerung der Trigonalisierung darstellt, beweisen zu können, müssen wir jetzt voraussetzen, daß χ_F in Linearfaktoren zerfällt:

$$\chi_F(t) = \pm(t - \lambda_1)^{m_1} \cdot \ldots \cdot (t - \lambda_k)^{m_k}\,, \quad \lambda_j \text{ paarweise verschieden}\,.$$

Erinnern wir uns an das Beispiel

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\,,$$

so wissen wir $\mathrm{Eig}(A;\,1) = \ker(A - E) \neq \mathbb{K}^2$. Offensichtlich ist aber $\ker(A - E)^2 = \ker 0 = \mathbb{K}^2$.

Dies führt uns zu der Definition der *Haupträume* oder *verallgemeinerten Eigenräume*

$$\mathrm{Eig}^\infty(F;\,\lambda) = \bigcup_{k=0}^{\infty} \ker(F - \lambda\,\mathrm{id})^k \subset V$$

zu einem Eigenwert λ von $F \in \mathrm{End}\,V$. Jeder Raum $\ker(F - \lambda\,\mathrm{id})^k$ ist $(F - \lambda\,\mathrm{id})$–invariant, und damit sind die Haupträume F–invariant.

Satz 12.3 (Jordan - Zerlegung) *Unter den obigen Voraussetzungen hat man eine F–invariante Zerlegung*

$$V = \bigoplus_{j=1}^{k} \mathrm{Eig}^\infty(F;\,\lambda_j)\,.$$

Es ist $\mathrm{Eig}^\infty(F;\lambda_j) = \ker(F - \lambda_j)^{m_j}$ *und* $\dim \mathrm{Eig}^\infty(F;\,\lambda_j) = m_j$. *Auf jedem Hauptraum hat F die Gestalt:*

$$F = \lambda_j\,\mathrm{id} + N_j\,,$$

wobei N_j nilpotent von der Ordnung $\leq m_j$ ist.

Beweis. Wir setzen $F_j = F - \lambda_j \,\mathrm{id}$ für jeden Eigenwert λ_j, s. d.

$$V_j := \mathrm{Eig}^\infty(F;\, \lambda_j) = \bigcup_{\ell=0}^{\infty} \ker F_j^\ell \,.$$

Nach Satz 1 hat man dann eine direkte F–invariante Summenzerlegung

$$V = V_j \oplus W_j \,, \quad V_j = \ker F_j^{q_j} \,, \quad W_j = \mathrm{im}\, F_j^{q_j} \,, \quad q_j \text{ geeignet} \,,$$

s. d. F_j auf V_j nilpotent und auf W_j invertierbar ist. Wir setzen ohne Einschränkung $j = 1$ und betrachten F bzgl. einer Basis, in der F obere Dreiecksgestalt hat mit λ_1 am Anfang der Hauptdiagonalen m_1–mal aufgereiht. Dann wird $F_1 = F - \lambda_1 \,\mathrm{id}$ repräsentiert durch die Matrix

$$A = \left(\begin{array}{c|c} N & B \\ \hline 0 & D \end{array}\right) \,,$$

wobei N eine obere $m_1 \times m_1$–Dreiecksmatrix mit Nullen in der Hauptdiagonalen und D eine weitere obere Dreiecksmatrix mit den Werten $\lambda_j - \lambda_1 \,,\ j > 1$, in der Hauptdiagonalen ist. Es folgt leicht

$$A^s = \left(\begin{array}{c|c} N^s & B_s \\ \hline 0 & D^s \end{array}\right) \,,$$

so daß sich wegen $\det D^s = (\det D)^s \neq 0$, $\det N^s = (\det N)^s = 0$ stets $\mathrm{rang}\, A^s \geq n - m_1$ ergibt. Da außerdem $N^s = 0$ für $s \geq m_1$ ist, ergibt sich $\mathrm{rang}\, A^s = n - m_1$ für alle $s \geq m_1$, und damit ist $\dim \ker (F - \lambda_1 \,\mathrm{id})^s$ konstant für diese s. Es folgt wie behauptet

$$\mathrm{Eig}^\infty(F;\, \lambda_1) = \ker (F - \lambda_1 \,\mathrm{id})^{m_1} \,,$$

und

$$\dim \mathrm{Eig}^\infty(F;\, \lambda_1) = m_1 \,.$$

Weiter folgt leicht durch Induktion, daß

$$\mathrm{Eig}^\infty(F;\, \lambda_j) \cap \sum_{\ell \neq j} \mathrm{Eig}^\infty(F;\, \lambda_\ell) = \{0\} \,;$$

also ist die Summe

$$\sum_j \mathrm{Eig}^\infty\,(F;\, \lambda_j)$$

direkt, und wegen $\sum_j \dim \mathrm{Eig}^\infty(F;\, \lambda_j) = \sum_j m_j = n$ ergibt sich schließlich

$$V = \bigoplus_{j=1}^{k} \mathrm{Eig}^\infty\,(F;\, \lambda_j) \,. \qquad \square$$

Zusammen mit dem vorigen Satz liefert diese Aussage natürlich den Beweis für die *Jordansche Normalform*:

Folgerung 12.4 *Es sei V ein endlich–dimensionaler $\mathbb{K}$–Vektorraum, und $F \in \mathrm{End}\, V$ sei ein Endomorphismus. Dann gibt es eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V, so daß F bzgl. $\mathcal{B}$ eine Matrix–Darstellung besitzt, die aus lauter* Jordan–Blöcken *der Gestalt*

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

entlang der Hauptdiagonalen besteht. Dabei ist λ ein Eigenwert von F. Bei festem λ kann es (selbstverständlich) mehrere solche Jordan–Blöcke geben.

Noch struktureller ist die

Folgerung 12.5 *Jeder Endomorphismus F schreibt sich (eindeutig) in der Form $F = D + N$ mit*

i) *D diagonalisierbar (manchmal wird D auch* halbeinfach *genannt),*

ii) *N nilpotent,*

iii) *D und N vertauschen.*

Beweisidee. Die Existenz folgt sofort aus der Jordanschen Normalform. Ist $F = D + N = D' + N'$, also $D - D' = N' - N$, so gilt $D'F = D'^2 + D'N' = D'^2 + N'D' = FD'$. D' vertauscht also mit F und damit auch mit jedem Polynom in F.

Man kann zeigen (siehe z. B. Brieskorn, Bd. 2, oder Halmos), daß in der obigen Konstruktion D ein Polynom in F ist. Also ergibt sich die Vertauschbarkeit von D und D'. In einer Übungsaufgabe wird gezeigt, daß D und D' dann auch *simultan* diagonalisiert werden können. Ferner vertauschen aus denselben Gründen N und N', woraus sofort die Nilpotenz von $N' - N$ folgt. Damit reduziert sich die Eindeutigkeitsaussage auf die folgende triviale Bemerkung: Jede nilpotente Diagonalmatrix ist notwendig die Nullmatrix. □

13 Normalformen für Endomorphismen von Skalarprodukträumen

Im folgenden sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$. Wir nennen einen $\mathbb{K}$–Vektorraum V kurz einen *Skalarproduktraum*, wenn er euklidisch ($\mathbb{K} = \mathbb{R}$) bzw. unitär ($\mathbb{K} = \mathbb{C}$) ist. Das Skalarprodukt werde wie immer mit $\langle\, .\,,\, .\,\rangle$, die zugehörige Norm mit $\|\, .\, \|$ bezeichnet. Wir sind jetzt vornehmlich an Abbildungen interessiert, die diese zusätzliche Struktur respektieren. Z. B. sollten wir nur Basen betrachten, die aus Orthonormalsystemen bestehen.

Wir untersuchen daher zunächst die Frage: Wann beschreibt eine Matrix $S \in \mathrm{GL}\,(n, \mathbb{K})$ einen Wechsel von Orthonormalbasen $\mathcal{A}$ bzw. $\mathcal{B}$?

$$\begin{tikzcd}
\mathbb{K}^n \arrow[r, "S"] \arrow[d, "\Phi_{\mathcal{A}}"'] & \mathbb{K}^n \arrow[d, "\Phi_{\mathcal{B}}"] \\
V \arrow[r, "\mathrm{id}"'] & V
\end{tikzcd}$$

Offensichtlich ist dies genau dann der Fall, wenn

$$\langle\, Se_j,\, Se_k\,\rangle \;=\; v_j,\, v_k\,\rangle \;=\; \delta_{jk}\ ,$$

d. h. wenn die Spalten von S ein Orthonormalsystem in $\mathbb{K}^n$ bzgl. der kanonischen Skalarproduktstruktur darstellen, d. h. wenn ${}^t\overline{S}\, S = E_n$ (wobei im euklidischen Fall $\mathbb{K} = \mathbb{R}$ die Matrix ${}^t\overline{S}$ schlicht durch ${}^tS$ zu ersetzen ist).

Definition. 1. Eine $n \times n$–Matrix $S \in M\,(n \times n\,,\ \mathbb{K})$ heißt *unitär* (im Fall $\mathbb{K} = \mathbb{R}$ auch *orthogonal*), falls

$${}^t\overline{S}\, S \;=\; E_n\ .$$

2. Zwei Matrizen A, B heißen *unitär (orthogonal) ähnlich*, falls es eine unitäre (orthogonale) Matrix S gibt, s. d.

$$B \;=\; {}^t\overline{S}\, A\, S\ .$$

Bemerkungen. 1. Aus ${}^t\overline{S}\, S = E_n$ folgert man $|\det S\,|^2 = \overline{\det S} \cdot \det S = \det {}^t\overline{S}\, \det S = 1$, also $|\det S\,| = 1$. Insbesondere ist $\det S \neq 0$ und damit S invertierbar mit $S^{-1} = {}^t\overline{S}$, so daß sich auch automatisch $S\,{}^t\overline{S} = E_n$ ergibt. Im Reellen schließt man entsprechend $\det S = \pm 1$ und $S\,{}^tS = E_n$.

2. Zwei Matrizen sind genau dann unitär ähnlich, wenn sie denselben Endomorphismus bzgl. zweier Orthonormalbasen darstellen.

3. Die Mengen

$$\mathrm{U}\,(n) \;:=\; \{S \in M\,(n \times n, \mathbb{C}) \,:\, {}^t\overline{S}\, S \;=\; E_n\}$$

und

$$\mathrm{O}\,(n) \;:=\; \{S \in M\,(n \times n, \mathbb{R}) \,:\, {}^tS\, S \;=\; E_n\}$$

bilden Untergruppen von $\mathrm{GL}\,(n, \mathbb{C})$ bzw. $\mathrm{GL}\,(n, \mathbb{R})$: die *unitäre* bzw. *orthogonale Gruppe*. Ihre Elemente heißen *unitäre* bzw. *orthogonale Matrizen.* Es gilt

$$\begin{aligned} \mathrm{U}\,(n) \ni S \quad &\Longrightarrow \quad |\det S| \;=\; 1 \\ \mathrm{O}\,(n) \ni S \quad &\Longrightarrow \quad \det S \;=\; \pm 1\ . \end{aligned}$$

Nun ist die Determinantenbildung $\det : \mathrm{U}\,(n) \to \mathbb{C}^*$ bzw. $\mathrm{O}\,(n) \to \mathbb{R}^*$ ein Gruppenhomomorphismus. Infolgedessen bilden auch die *speziellen* unitären (orthogonalen) Matrizen je eine Gruppe:

$$\begin{aligned} \mathrm{SU}\,(n) \;&=\; \{S \in \mathrm{U}\,(n) \,:\, \det S \;=\; 1\} \\ \mathrm{SO}\,(n) \;&=\; \{S \in \mathrm{O}\,(n) \,:\, \det S \;=\; 1\}\ . \end{aligned}$$

Als nächstes wollen wir diejenigen Endomorphismen charakterisieren, die sich bzgl. Orthonormalbasen diagonalisieren lassen. Dazu benötigen wir noch eine allgemeine Konstruktion, nämlich die der *adjungierten* Abbildung (die nur eine Variante der dualen Abbildung ist). Ist allgemein $F : V \to W$ ein Homomorphismus, so wird $F^* : W^* \to V^*$ definiert durch

$$F^*(\mu) = \mu \circ F \,, \quad \mu \in W^* = \mathrm{Hom}_{\mathbb{K}}(W, \mathbb{K}) \,.$$

Mit den kanonischen Paarungen $V \times V^* \to \mathbb{K}$ und $W \times W^* \to \mathbb{K}$ schreibt sich diese Definition als

$$\langle\, v,\, F^*(\mu) \,\rangle = (F^*(\mu))(v) = (\mu \circ F)\,(v) = \mu(F\,(v)) = \langle\, F\,(v),\, \mu \,\rangle \,.$$

Sind nun V, W Skalarprodukträume, so können wir V mit V^* resp. W mit W^* kanonisch identifizieren und F^* zu einer Abbildung $F^{ad} : W \to V$ liften:

$$\begin{tikzcd}
V \arrow[d, "\cong"'] & W \arrow[l, "F^{ad}"'] \arrow[d, "\cong"] \\
V^* & W^* \arrow[l, "F^*"]
\end{tikzcd}$$

Satz 13.1 *Für Skalarprodukträume* V, W *gibt es zu jedem Homomorphismus* $F \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ *genau einen Homomorphismus* $F^{ad} \in \mathrm{Hom}_{\mathbb{K}}(W, V)$ *mit*

$$\langle\, v,\, F^{ad}(w) \,\rangle = \langle\, F\,(v),\, w \,\rangle \,, \quad v \in V \,,\ w \in W \,.$$

Definition. F^{ad} heißt die zu F *adjungierte* Abbildung.

Sind $\mathcal{A}$ und $\mathcal{B}$ Orthonormalbasen von V bzw. W, und wird F durch die Matrix A relativ zu diesen Basen dargestellt: $A = M_{\mathcal{B}}^{\mathcal{A}}(F)$, so wird F^{ad} dargestellt durch ${}^t\overline{A}$:

$$M_{\mathcal{A}}^{\mathcal{B}}(F^{ad}) = {}^t\overline{(M_{\mathcal{B}}^{\mathcal{A}}(F))} \,.$$

Bemerkung. $(F^{ad})^{ad} = F$.

Ab jetzt sei für längere Zeit $\mathbb{K} = \mathbb{C}$. Entscheidend sind die beiden folgenden Sätze:

Satz 13.2 *Jeder Endomorphismus* F *eines endlich–dimensionalen unitären Vektorraumes* V *kann unitär trigonalisiert werden.*

Denn: Konstruiere eine F–invariante Fahne $U_0 \subset U_1 \subset \cdots \subset U_n = V$ und wähle nach Schmidt eine Orthonormalbasis $v_1, \ldots, v_n$ mit $U_j = \mathrm{span}(v_1, \ldots, v_j)$. □

Satz 13.3 *Sei* $F \in \mathrm{End}\, V$. *Dann läßt sich* F *genau dann unitär diagonalisieren, wenn*

$$(*) \qquad F\,F^{ad} = F^{ad}\,F \,.$$

Definition. Endomorphismen mit der Eigenschaft $(*)$ nennt man auch *normal*. Eine Matrix $A \in M\,(n \times n, \mathbb{C})$ heißt entsprechend normal, falls ${}^t\overline{A}\,A = A\,{}^t\overline{A}$.

Folgerung 13.4

1. *Zu jeder komplexen* $n \times n$*–Matrix gibt es eine unitäre Matrix* $S \in \mathrm{U}\,(n)$, *so daß* ${}^t\overline{S}\,A\,S$ *eine obere Dreiecksmatrix ist.*
2. *Eine Matrix* $A \in M\,(n \times n, \mathbb{C})$ *ist genau dann normal, wenn es eine unitäre Matrix* $S \in \mathrm{U}(n)$ *gibt, so daß* ${}^t\overline{S}\,A\,S$ *Diagonalgestalt hat.*

Bevor wir den Beweis von Satz 3 in Angriff nehmen können, benötigen wir die folgenden zwei Lemmata.

Lemma 13.5 *Es sei* $F \in \operatorname{End} V$ *normal. Dann gilt:*

1. $\ker F = \ker F^{ad}$;
2. $\operatorname{Eig}(F;\, \lambda) = \operatorname{Eig}(F^{ad};\, \overline{\lambda})$ *für alle* $\lambda \in \mathbb{C}$;
3. $\operatorname{Eig}(F;\, \lambda_j) \perp \operatorname{Eig}(F;\, \lambda_k)\,, \quad \lambda_j \neq \lambda_k\,.$

Beweis. 1. Dies folgt sofort aus

$$\langle F(v),\, F(v) \rangle = \langle F^{ad} F(v),\, v \rangle = \langle F\, F^{ad}(v),\, v \rangle = \langle F^{ad}(v),\, F^{ad}(v) \rangle\,.$$

2. Setze $G = F - \lambda\, \mathrm{id}$. Aus der definierenden Gleichung für G^{ad} erhält man unmittelbar $G^{ad} = F^{ad} - \overline{\lambda}\, \mathrm{id}$ und $G\, G^{ad} = G^{ad}\, G$; G ist also ebenfalls normal, und mit 1. folgt

$$\operatorname{Eig}(F;\, \lambda) = \ker G = \ker G^{ad} = \operatorname{Eig}(F^{ad};\, \overline{\lambda})\,.$$

3. Mit $v \in \operatorname{Eig}(F;\, \lambda_j)$, $w \in \operatorname{Eig}(F;\, \lambda_k)$ ist $\lambda_j \langle v,\, w \rangle = \langle \overline{\lambda_j}\, v,\, w \rangle =$ $\langle F^{ad}(v),\, w \rangle = \langle v,\, F(w) \rangle = \langle v,\, \lambda_k\, w \rangle = \lambda_k \langle v,\, w \rangle$ und folglich $\langle v,\, w \rangle = 0$. □

Folgerung 13.6 *Für jeden normalen Endomorphismus* $F \in \operatorname{End} V$ *gilt*

$$V = \bigoplus_{j=1}^{\ell} \operatorname{Eig}(F;\, \lambda_j)\,.$$

Denn wegen des obigen Satzes ist V die direkte Summe der Eigenräume, und diese stehen wegen 3. im obigen Lemma 5 paarweise senkrecht aufeinander. □

Lemma 13.7 *Ist* $F \in \operatorname{End} V$, V *ein Skalarproduktraum und* $U \subset V$ *ein* F*–invarianter Untervektorraum. Dann ist* $U^{\perp}$ *invariant unter* F^{ad}.

Beweis. Es sei $v \in U$, $w \in U^{\perp}$. Dann gilt $\langle v,\, F^{ad}(w) \rangle = \langle F(v),\, w \rangle = 0$, da $F(v) \in U$. Da dies für alle $v \in U$ richtig ist, ergibt sich $F^{ad}(w) \in U^{\perp}$. □

Wir kommen nun zu dem *Beweis* von Satz 3. Wir wollen zeigen: F ist genau dann normal, wenn V eine orthonormale Basis aus Eigenvektoren von F besitzt. Sei also eine solche Basis $\mathcal{B}$ gegeben, so daß

$$v \in \mathcal{B} \Longrightarrow F(v) = \lambda\, v\,,\ F^{ad}(v) = \overline{\lambda}\, v \Longrightarrow F\, F^{ad}(v) = F(\overline{\lambda}\, v) = \overline{\lambda}\, F(v) = \lambda\, \overline{\lambda}\, v$$

und genauso $F^{ad} F(v) = \overline{\lambda}\, \lambda\, v$. Damit ist aber $F\, F^{ad} = F^{ad}\, F$.

Sei umgekehrt F normal. Da $\mathbb{K} = \mathbb{C}$ ist, gibt es einen Eigenvektor v von F : $F(v) = \lambda\, v$. Ohne Einschränkung sei $\|\, v \,\| = 1$. Bilde nun $V_1 = \langle v \rangle$ und

$$W = V_1^{\perp} = \{ w \in V : \langle v,\, w \rangle = 0 \}\,.$$

Dann ist für alle $w \in W$:

$$\langle v,\, F(w) \rangle = \langle F^{ad}(v),\, w \rangle = \langle \overline{\lambda}\, v,\, w \rangle = \lambda \langle v,\, w \rangle = 0\,,$$

d. h. $F(w) \in W$. M. a. W.: W ist F–invariant. Nach dem vorigen Lemma 7 ist W als orthogonales Komplement des F–invarianten Raumes V_1 auch F^{ad}–invariant. Daraus folgt unmittelbar, daß die Einschränkung $F' = F|W$ normal ist mit $F'^{ad} = F^{ad}|W$, und man beendet den Beweis durch Induktion. □

Wir werden diesen Satz nun in zwei Spezialfällen bei unitären Vektorräumen anwenden, nämlich bei *unitären* und *selbstadjungierten* Abbildungen. Es wird dann am Ende des Kapitels unsere Aufgabe sein, daraus Folgerungen im euklidischen Fall abzuleiten.

Definition. Ein Endomorphismus $F \in \operatorname{End} V$, V ein (endlich–dimensionaler) Skalarproduktraum, heißt *unitär* ($\mathbb{K} = \mathbb{C}$) bzw. *orthogonal* ($\mathbb{K} = \mathbb{R}$), falls

$$\langle F(v), F(w) \rangle = \langle v, w \rangle \quad \text{für alle} \quad v, w \in V .$$

Bemerkungen. 1. Für solche Endomorphismen gilt notwendig

$$\| F(v) - F(w) \| = \| v - w \| , \quad \text{insbesondere} \quad \| F(v) \| = \| v \| ,$$

sie sind also *Isometrien* und *normerhaltend*. Wegen der Polarisierungsgleichung ist die Normerhaltung gleichwertig zu der Definition der Unitarität. Wir werden überdies am Ende des Kapitels zeigen, daß die Isometrieeigenschaft einer *beliebigen* Selbstabbildung eines euklidischen Raumes, die den Ursprung festhält, automatisch deren Linearität nach sich zieht. Dies beendet dann auch den Nachweis der in Kapitel 10 formulierten Tatsache, daß Isometrien in euklidischen Räumen affin sind.

2. Unitäre (orthogonale) Endomorphismen werden *relativ zu Orthonormalbasen* durch unitäre (orthogonale) Matrizen beschrieben.

Folgerung 13.8 *Es sei* $F \in \operatorname{End} V$ *unitär oder orthogonal. Dann gilt:*

i) *Jeder Eigenwert* $\lambda \in \mathbb{C}$ *in* F *ist unimodular, d. h.* $|\lambda| = 1$.

ii) $v \perp w$ *impliziert* $F(v) \perp F(w)$.

iii) F *ist injektiv (und damit bijektiv für* $\dim V < \infty$ *mit unitärem bzw. orthogonalem Inversen).*

Beweis. i) Sei $v \neq 0$ ein Eigenvektor zu λ. Dann ist

$$\| v \| = \| F(v) \| = | \lambda | \, \| v \| ,$$

also $|\lambda| = 1$. Die Aussage ii) folgt ebenso leicht, und iii) ist eine Konsequenz der Normerhaltung. Ist F bijektiv, so setze man $v = F(v')$, $w = F(w')$ und folgere:

$$\begin{aligned} \langle F^{-1}(v), F^{-1}(w) \rangle &= \langle F^{-1} F(v'), F^{-1} F(w') \rangle = \langle v', w' \rangle \\ &= \langle F(v'), F(w') \rangle = \langle v, w \rangle . \end{aligned} \qquad \Box$$

Lemma 13.9 *Unitäre Endomorphismen sind normal.*

Beweis. Mit $\langle (F^{ad} F)(v), w \rangle = \langle F(v), F(w) \rangle = \langle v, w \rangle$ ist $(F^{ad} F)(v) = v$ für alle $v \in V$, also $F^{ad} F = \mathrm{id}$. Entsprechend ergibt sich $F F^{ad} = \mathrm{id}$. $\Box$

Als Folgerung erhalten wir

Satz 13.10 1. *Jeder unitäre Endomorphismus* $F \in \operatorname{End} V$ *ist unitär diagonalisierbar:*

$$V = \bigoplus_{j=1}^{\ell} \operatorname{Eig}(F; \lambda_j) .$$

2. *Für jede Matrix* $A \in \mathrm{U}(n)$ *gibt es ein* $S \in \mathrm{U}(n)$ *mit*

$$^t\overline{S} \, A \, S = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) , \quad |\lambda_j| = 1 , \; j = 1, \ldots, n .$$

Das Analogon dieses Satzes im Reellen ist *nicht* richtig. Dies liegt daran, daß orthogonale Matrizen i. a. keine *reelle* Eigenwerte besitzen, wie die Beispiele

$$A_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \quad a \not\equiv 0 \bmod \pi$$

zeigen. Gerade diese Beispiele sind aber die charakteristischen; es gilt der folgende Satz, den wir weiter unten beweisen werden:

Satz 13.11 *Es sei $F \in \operatorname{End} V$ ein orthogonaler Endomorphismus eines endlich–dimensionalen euklidischen Vektorraumes V. Dann zerfällt V in eine orthogonale Summe F–invarianter Unterräume:*

$$V = V_0 \oplus V_1 \oplus \cdots \oplus V_s ,$$

so daß $F|V_0$ orthogonal diagonalisierbar (mit Eigenwerten ± 1) ist und die V_j 2–dimensionale Vektorräume sind, auf denen F bzgl. einer Orthonormalbasis wie eine der oben angegebenen Matrizen A_α operiert.

Wir wenden uns zunächst erst einem anderen Spezialfall zu.

Definition. Es sei V ein Skalarproduktraum, und $F \in \operatorname{End} V$ sei ein Endomorphismus. F heißt *selbstadjungiert*, falls $F = F^{ad}$.

Bemerkungen. 1. Im unitären Fall sind selbstadjungierte Endomorphismen normal. Sie werden bzgl. Orthonormalbasen durch *hermitesche* Matrizen A beschrieben: ${}^t\overline{A} = A$.

2. Im euklidischen Fall sind die beschreibenden Matrizen bzgl. Orthonormalbasen genau die *symmetrischen* Matrizen: ${}^tA = A$.

Lemma 13.12 *Selbstadjungierte Endomorphismen besitzen (im unitären und im euklidischen Fall) nur reelle Eigenwerte.*

Beweis. Es sei zunächst $\mathbb{K} = \mathbb{C}$; v sei ein Eigenvektor zu dem Eigenwert λ. Dann gilt

$$\lambda v = F(v) = F^{ad}(v) = \overline{\lambda} v , \quad \text{also} \quad \lambda = \overline{\lambda} \in \mathbb{R} .$$

Ist $\mathbb{K} = \mathbb{R}$ und A eine beschreibende Matrix bzgl. einer Orthonormalbasis, dann ist die reelle symmetrische Matrix A, aufgefaßt als komplexe Matrix, hermitesch und damit normal. Also gibt es eine Matrix $S \in \mathrm{U}(n)$ mit $S^{-1}AS = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, wobei die Werte $\lambda_1, \ldots, \lambda_n$ nach dem ersten Teil reell sind. Damit ergibt sich

$$\chi_F(t) = \chi_A(t) = \pm(t - \lambda_1) \cdot \ldots \cdot (t - \lambda_n) \in \mathbb{R}[t] ,$$

also eine *reelle* Linearfaktorzerlegung des charakteristischen Polynoms. □

Satz 13.13 1. *Jeder selbstadjungierte Endomorphismus F eines endlich–dimensionalen unitären Vektorraumes V läßt sich unitär diagonalisieren:*

$$V = \bigoplus_{j=1}^{\ell} \operatorname{Eig}(F; \lambda_j) , \quad \lambda_j \in \mathbb{R} .$$

2. *Zu jeder hermiteschen Matrix $A \in M(n \times n, \mathbb{C})$ gibt es eine unitäre Matrix $S \in \mathrm{U}(n)$, s. d.*

$${}^t\overline{S}AS = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) , \quad \lambda_j \in \mathbb{R} , \quad j = 1, \ldots, n .$$

Da die Eigenwerte im reellen Fall tatsächlich reell sind, dürfte es nicht überraschen, daß in dieser Situation das reelle Analogon wirklich richtig bleibt:

Satz 13.14 (Hauptachsentransformation, Spektralsatz) 1. *Jeder selbstadjungierte Endomorphismus F eines endlich–dimensionalen euklidischen Vektorraumes V ist orthogonal diagonalisierbar* :

$$V = \bigoplus_{j=1}^{\ell} \mathrm{Eig}\,(F;\, \lambda_j)\ .$$

2. *Zu jeder symmetrischen Matrix $A \in M\,(n \times n, \mathbb{R})$ gibt es eine orthogonale Matrix $S \in \mathrm{O}\,(n)$, so daß*

$$^{t}SAS = \mathrm{diag}\,(\lambda_1, \ldots \lambda_n)\ .$$

Man kann den obigen Satz wie für normale Endomorphismen beweisen: Da F selbstadjungiert ist, besitzt F einen reellen Eigenwert λ und damit einen reellen Eigenvektor $v \in V$. Es sei $U_1 = \mathrm{span}\,(v)$ und $U_2 = U_1^{\perp}$. Da F selbstadjungiert ist, bildet F den Raum U_2 in sich ab; denn aus $v \in U_1$, $w \in U_2 = U_1^{\perp}$ folgt

$$\langle\, v,\, F\,(w)\,\rangle = \langle\, F\,(v),\, w\,\rangle = \langle\, \lambda v,\, w\,\rangle = \lambda\,\langle\, v,\, w\,\rangle = 0\ .$$

Wie früher sieht man $(F|U_2)^{ad} = F^{ad}|U_2 = F|U_2$. Also ist $F|U_2$ selbstadjungiert, und nach Induktionsvoraussetzung gibt es eine Orthonormalbasis von U_2 , in der $F' = F|U_2$ Diagonalgestalt hat. □

Man kann den gesamten Themenkreis der Hauptachsentransformation auch analytisch angehen: Es sei $A \in M\,(n \times n, \mathbb{R})$ eine symmetrische Matrix, und

$$q = q\,(x_1, \ldots, x_n) = {}^{t}x\,A\,x = \sum_{i,j=1}^{n} a_{ij}\,x_i\,x_j\ .$$

Man betrachtet dann $q|S^{n-1}$, wobei S^{n-1} die Einheitskugel

$$\{x \in \mathbb{R}^n :\, \|\,x\,\|^2 = 1\}$$

bezeichnet. Da S^{n-1} kompakt ist, nimmt q auf S^{n-1} ein Maximum, sagen wir bei x_0 , an. Dort ist dann nach dem *Satz über Extrema mit Nebenbedingungen*:

$$\mathrm{grad}\,q\,(x_0) = \lambda\,\mathrm{grad}\,f(x_0)\ , \quad f(x) = \|x\|^2 - 1\ ,$$

λ eine reelle Zahl. Nun ist aber

$$\mathrm{grad}\,f\,(x_0) = 2\,x_0$$

und (wegen der Symmetrie der Matrix A):

$$\mathrm{grad}\,q\,(x_0) = 2\,A\,x_0\ .$$

Also ist x_0 gerade ein Eigenvektor von A . Man wählt dann eine Hyperebene senkrecht zu x_0 und fährt entsprechend fort. □

Es muß noch der Beweis für die Normalform *orthogonaler* Matrizen nachgeholt werden. Dazu benötigen wir einen kurzen Exkurs über *reelle* Polynome $P\,(t) \in \mathbb{R}\,[\,t\,]$. Hat P eine *komplexe* Nullstelle λ , so folgt aus

$$P\,(\overline{\lambda}) = \overline{P(\lambda)} = \overline{0} = 0\ ,$$

daß auch $\overline{\lambda}$ eine Nullstelle von P ist. M. a. W.: eine Nullstelle λ von P ist entweder reell, oder sie ist nicht reell und kommt automatisch als Paar mit ihrer Konjugiert–Komplexen $\overline{\lambda}$ vor. Die Linearfaktor–Zerlegung von P über dem komplexen Zahlkörper $\mathbb{C}$ läßt sich deshalb in die folgende Form bringen:

$$P(t) = c(t - \lambda_1) \cdot \ldots \cdot (t - \lambda_s) \cdot (t - \mu_1)(t - \overline{\mu_1}) \cdot \ldots \cdot (t - \mu_r)(t - \overline{\mu_r})\ ,$$

wobei $\lambda_1, \ldots, \lambda_s \in \mathbb{R}$, $\mu_1, \ldots, \mu_r \in \mathbb{C} \setminus \mathbb{R}$, $c \in \mathbb{R}^*$. Insbesondere ergibt sich sofort:

Jedes reelle Polynom ungeraden Grades hat mindestens eine reelle Nullstelle.

(Dies folgt natürlich auch mit rein reellen Methoden aus dem Zwischenwertsatz und der Tatsache, daß $\lim_{t\to\infty} P(t) = \pm\infty$ und $\lim_{t\to-\infty} P(t) = -\lim_{t\to\infty} P(t)$.) Faßt man die Terme $(t-\mu)(t-\overline{\mu})$ zu $t^2 - (\mu + \overline{\mu})\,t + \mu\,\overline{\mu} \in \mathbb{R}[t]$ zusammen und beachtet man die Eindeutigkeit der Linearfaktorzerlegung über $\mathbb{C}$, so ergibt sich der folgende

Satz 13.15 *Jedes reelle Polynom $P(t)$ hat eine (bis auf Reihenfolge eindeutige) Produktzerlegung*

$$P(t) = c\,(t-\lambda_1)\cdot\ldots\cdot(t-\lambda_s)\cdot Q_1(t)\cdot\ldots\cdot Q_r(t)\,,$$

$c \in \mathbb{R}^*$, $\lambda_1,\ldots,\lambda_s \in \mathbb{R}$, $Q_\rho(t) = t^2 + a_\rho\,t + b_\rho \in \mathbb{R}[t]$, $\rho = 1,\ldots,r$, *wobei die quadratischen Polynome Q_ρ keine reellen Nullstellen besitzen (d. h. $a_\rho^2 - 4b_\rho < 0$).*

Als Folgerung erhält man folgendes

Korollar 13.16 *Es sei V ein endlich dimensionaler $\mathbb{R}$–Vektorraum und $F \in \operatorname{End} V$. Dann besitzt V einen F–invarianten Unterraum U der Dimension* 1 *oder* 2.

Beweis. Hat χ_F eine reelle Nullstelle, so ist der von dem entsprechenden Eigenvektor v erzeugte Unterraum eindimensional und F–invariant. Wir können daher annehmen, daß χ_F nur quadratische „Primfaktoren“ besitzt: $\chi_F = Q_1\cdot\ldots\cdot Q_s$. Nach dem Satz von Cayley–Hamilton ist für alle $v \in V\setminus\{0\}$:

$$\chi_F(F)\,(v) = 0\,,$$

d. h.

$$(Q_s(F)\circ\cdots\circ Q_1(F))\,(v) = 0\,.$$

Sei nun $j \geq 1$ der größte Index mit

$$u := (Q_{j-1}(F)\circ\cdots\circ Q_1(F))\,(v) \neq 0\,;$$

dann ist $Q_j(F)\,(u) = (F^2 + a_j\,F + b_j\,\mathrm{id})\,(u) = 0$, $u \neq 0$, und der Vektorraum

$$U = \operatorname{span}(u,\, F(u))$$

hat die gewünschte Eigenschaft. □

Der *Beweis* des Normalformensatzes ist nun wie in den anderen Fällen zum Abschluß zu bringen: Ist $U \subset V$ irgendein F–invarianter Unterraum, so ist sogar $F(U) = U$, da F als orthogonale Abbildung bijektiv ist. Da aber F^{-1} ebenfalls orthogonal ist, muß auch $U^\perp$ invariant unter F sein; denn aus $v \in U$, $u \in U^\perp$ folgt

$$\langle\, F(u),\, v\,\rangle = \langle\, F^{-1}F(u),\, F^{-1}(v)\,\rangle = \langle\, u,\, F^{-1}(v)\,\rangle = 0\,.$$

Mittels des obigen Korollars und Induktion zerfällt daher V in eine orthogonale Summe

$$U_1 \perp \cdots \perp U_\ell$$

F–invarianter Unterräume der Dimension 1 oder 2. Da die Eigenwerte von F nur $= \pm 1$ sein können, operiert F auf den eindimensionalen Räumen durch Multiplikation mit 1 oder -1. Was überbleibt, ist die Klassifikation der zweidimensionalen orthogonalen Abbildungen. Das Ergebnis fassen wir in dem folgenden Lemma zusammen.

Lemma 13.17 *Es sei* $\dim_{\mathbb{R}} V = 2$, $F \in \operatorname{End} V$ *sei ein orthogonaler Endomorphismus. Dann gibt es die folgenden Fälle :*

i) $\det F = 1$. *In diesem Fall ist* F *eine Drehung mit Normalform*

$$A_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \quad 0 \le \alpha < 2\pi .$$

ii) $\det F = -1$. *Dann ist* F *eine Spiegelung mit Normalform*

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Beweis. F wird beschrieben durch eine orthogonale Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2, \mathbb{R}) ,$$

d. h. $a^2 + c^2 = b^2 + d^2 = 1$ und $a\,b + c\,d = 0$. Aus den beiden ersten Bedingungen folgt die Existenz von Winkeln α, α', so daß

$$a = \cos\alpha\,,\ c = \sin\alpha\,,\ b = \sin\alpha'\,,\ d = \cos\alpha' .$$

Wegen der dritten Bedingung ist

$$\sin(\alpha + \alpha') = \sin\alpha\cos\alpha' + \cos\alpha\sin\alpha' = c\,d + a\,b = 0 ,$$

d. h. $\alpha + \alpha' \equiv 0 \bmod \pi$. Je nachdem, ob in $\alpha + \alpha' = n\pi$ die Zahl n gerade oder ungerade ist, ergeben sich die beiden Normalformen

$$A_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \quad \text{mit} \quad \det A_\alpha = 1$$

bzw.

$$B_\alpha = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \quad \text{mit} \quad \det B_\alpha = -1 .$$

Im zweiten Fall ist das charakteristische Polynom aber gleich

$$\det \begin{pmatrix} \cos\alpha - t & \sin\alpha \\ \sin\alpha & -\cos\alpha - t \end{pmatrix} = t^2 - (\cos^2\alpha + \sin^2\alpha) = t^2 - 1$$

mit den Eigenwerten $+1$ und -1. In der Tat sind die B_α *Spiegelungen* an den Geraden in $\mathbb{R}^2$ mit Steigungswinkel $\alpha/2$. □

Wir wollen uns noch einen Überblick über *dreidimensionale* orthogonale Abbildungen verschaffen. Wegen des Normalformensatzes hat man die folgenden Typen.

$\det F = 1$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} : \quad \text{Identität} .$$

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} : \quad \begin{array}{l} \text{Drehung senkrecht zur } x_3\text{–Achse} \\ \text{um den Winkel } \alpha . \end{array}$$

$\det F = -1$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} : \quad \text{Spiegelung an der } (x_1, x_2)\text{–Ebene} .$$

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & -1 \end{pmatrix} :$$

Drehung senkrecht zur x_3–Achse um den Winkel α mit anschließender Spiegelung an der (x_1, x_2)–Ebene .

Es sind also alle *eigentlichen* orthogonalen Abbildungen ($\det F = 1$) Drehungen um Achsen. Der Drehwinkel bestimmt sich (im wesentlichen, d. h. bis auf das Vorzeichen) aus der Formel

$$\cos\alpha = \frac{1}{2}(\operatorname{Spur} F - 1) .$$

Man kann übrigens auch die Drehachse bestimmen, ohne die Normalform herzustellen (außer wenn $\sin\alpha = 0$; siehe Übungsaufgaben): Ist $Q = (q_{jk}) \in \mathrm{SO}(3)$ und Q nicht symmetrisch (was zu $\sin\alpha \neq 0$ oder $Q^2 \neq E$ äquivalent ist), so ist

$$^t(q_{32} - q_{23}\, ,\ q_{13} - q_{31}\, ,\ q_{21} - q_{12})$$

ein Eigenvektor von Q zu dem Eigenwert 1, also ein Fixvektor.

Wir fügen jetzt noch den Beweis dafür an, daß Isometrie die Affinität impliziert.

Satz 13.18 *Es sei V ein euklidischer Vektorraum, und $F : V \to V$ sei eine Isometrie* :

$$\| F(v) - F(w) \| = \| v - w \| \quad \textit{für alle} \quad v, w \in V ,$$

die den Ursprung festlassen : $F(0) = 0$. *Dann ist F linear, also eine orthogonale Abbildung.*

Beweis. Wegen $F(0) = 0$ impliziert die Isometriebedingung auch die Normerhaltung und dann mit der Polarisationsgleichung in der Form

$$\langle v, w \rangle = \frac{1}{2}\left(\| v \|^2 + \| w \|^2 - \| v - w \|^2\right)$$

auch die Invarianz des Skalarprodukts unter der Abbildung F:

$$\langle F(v), F(w) \rangle = \langle v, w \rangle .$$

Es folgt damit

$$\begin{aligned} \| F(v + w) - F(v) - F(w) \|^2 &= \langle F(v + w), F(v + w) \rangle + \langle F(v), F(v) \rangle \\ &\quad + \langle F(w), F(w) \rangle - 2\langle F(v + w), F(w) \rangle - \cdots \\ &= \langle v + w, v + w \rangle + \langle v, v \rangle \\ &\quad + \langle w, w \rangle - 2\langle v + w, w \rangle - \cdots \\ &= \| (v + w) - v - w \|^2 = 0 , \end{aligned}$$

also $F(v + w) = F(v) + F(w)$. Entsprechend beweist man $F(\lambda v) = \lambda F(v)$. □

Anhang: Orientierung von Basen

Es seien $\mathcal{A} = (v_1, \ldots, v_n)$ und $\mathcal{B} = (w_1, \ldots, w_n)$ zwei Basen des reellen Vektorraums $V \neq \{0\}$, und S bezeichne die Transformationsmatrix des Basiswechsels von $\mathcal{A}$ nach $\mathcal{B}$. Dann ist $\det S \in \mathbb{R}^* = \mathbb{R} \setminus \{0\}$.

Definition. Man nennt $\mathcal{A}$ und $\mathcal{B}$ *von gleicher Orientierung* (oder *gleichorientiert*), wenn $\det S > 0$. Im anderen Fall heißen $\mathcal{A}$ und $\mathcal{B}$ *entgegengesetzt* orientiert.

Satz 13.19 *Die Eigenschaft, gleichorientiert zu sein, ist eine Äquivalenzrelation.*

Beweis. Elementar. □

Definition. Eine *Orientierung* auf V ist eine Äquivalenzklasse von gleichorientierten Basen.

Es ist klar, daß es stets nur zwei solche Äquivalenzklassen, also zwei Orientierungen gibt. Man nennt die eine positiv, die andere negativ. Welche man als positiv bezeichnet, ist nicht kanonisch vorgegeben, sondern reine Konvention.

Wir wollen als erstes beweisen, daß zwei Basen gleichorientiert sind, wenn sie *stetig ineinander deformiert* werden können (nachher zeigen wir die Umkehrung). Zuerst müssen wir diese Aussage präzisieren. Wir stellen uns vor, daß wir eine *Schar* von Übergangsmatrizen

$$\gamma(t) \in \mathrm{GL}(n, \mathbb{R}), \quad t \in [a, b]$$

besitzen mit $\gamma(a) = E$ und $\gamma(b) = S$. Fassen wir $\mathrm{GL}(n, \mathbb{R})$ einfach als offenen Teil von $\mathbb{R}^{n^2}$ auf:

$$\mathrm{GL}(n, \mathbb{R}) = \{A \in M(n \times n, \mathbb{R}) = \mathbb{R}^{n^2} : \det A \neq 0\},$$

so können wir davon reden, daß die Abbildung

$$\gamma : [a, b] \longrightarrow \mathrm{GL}(n, \mathbb{R})$$

stetig ist.

Definition. Zwei Basen $\mathcal{A}$ und $\mathcal{B}$ können *stetig ineinander deformiert* werden, falls es eine stetige Abbildung

$$\gamma : [a, b] \longrightarrow \mathrm{GL}(n, \mathbb{R})$$

gibt mit $\gamma(a) = E$, $\gamma(b) = S$, wobei S die Transformationsmatrix $\mathcal{A} \to \mathcal{B}$ bezeichnet.

Bemerkung. Es ist leicht zu sehen, daß diese Bedingung ebenfalls eine Äquivalenzrelation darstellt.

Nun zu der obigen Behauptung: Es sei $\gamma = \gamma(t)$ eine solche stetige Familie von invertierbaren Matrizen. Dann ist

$$(\det \gamma)(t) = \det \gamma(t) : [a, b] \longrightarrow \mathbb{R}^*$$

eine stetige Funktion mit $(\det \gamma)(a) = 1$. Dann muß aber wegen des Zwischenwertsatzes auch $\det S = (\det \gamma)(b)$ positiv sein!

Was bedeutet nun die Umkehrung? Diese ist doch offensichtlich erfüllt, wenn wir zeigen können, daß *sich jede Matrix A mit* $\det A > 0$ *durch eine stetige Schar in die Einheitsmatrix E innerhalb* $\mathrm{GL}(n, \mathbb{R})$ *deformieren läßt.* Unsere Behauptung lautet daher in anderen Worten:

Satz 13.20 *Die Teilmenge*

$$\mathrm{GL}^+(n, \mathbb{R}) = \{A \in \mathrm{GL}(n, \mathbb{R}) : \det A > 0\}$$

ist ein (offener und) wegzusammenhängender *Teil von* $\mathrm{GL}(n, \mathbb{R})$. *Entsprechendes gilt für* $\mathrm{GL}^-(n, \mathbb{R})$, *und* $\mathrm{GL}(n, \mathbb{R})$ *ist die disjunkte Vereinigung dieser beiden Zusammenhangskomponenten.*

Zum *Beweis* wollen wir uns zunächst mit der entsprechenden Frage für *Orthonormalbasen in euklidischen Vektorräumen* beschäftigen, auf den wir dann am Ende auch den allgemeinen Fall zurückführen werden. Im Fall der Orthonormalbasen sind die Transformationsmatrizen A natürlich orthogonal: $A \in \mathrm{O}(n)$, und wir würden selbstverständlich gern wissen, daß gleichorientierte Orthonormalbasen durch eine *stetige Schar von Orthonormalbasen* ineinander deformiert werden können. Dies ist tatsächlich richtig, d. h. es gilt der

Satz 13.21 *Zu jeder Matrix* $A \in \mathrm{SO}(n)$ *gibt es eine stetige Abbildung*

$$\gamma : [a, b] \longrightarrow \mathrm{SO}(n) \subset \mathbb{R}^{n^2}$$

mit $\gamma(a) = E$, $\gamma(b) = A$. *Insbesondere ist* $\mathrm{SO}(n)$ *wegzusammenhängend, und* $\mathrm{O}(n)$ *zerfällt in die beiden Zusammenhangskomponenten* $\mathrm{SO}(n)$ *und* $\mathrm{O}^-(n)$ *der orthogonalen Matrizen mit Determinante* -1.

Beweis. Es sei ${}^tS\,A\,S = N$ in Normalform, $S \in \mathrm{O}(n)$. Wegen $\det A = 1$ kann es nur eine *gerade* Anzahl von dem Eigenwert -1 geben, so daß sich je zwei von solchen zusammenfassen lassen zu

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = A_\pi \, .$$

Nun lassen sich aber alle A_α stetig in $\mathrm{SO}(2)$ in die Einheitsmatrix deformieren durch

$$\gamma_\alpha(t) = \begin{pmatrix} \cos\alpha\, t & -\sin\alpha\, t \\ \sin\alpha\, t & \cos\alpha\, t \end{pmatrix}, \quad t \in [0, 1] \, .$$

Also gibt es eine stetige Schar $\widetilde{\gamma} : [0, 1] \to \mathrm{SO}(n)$ mit

$$\widetilde{\gamma}(0) = E \, , \quad \widetilde{\gamma}(1) = N \, .$$

Definiert man noch $\gamma(t) := {}^tS\,\widetilde{\gamma}(t)\,S$, so ist man fertig.

Der letzte Teil folgt aus der leicht einzusehenden Tatsache, daß $O^-(n) = T\,\mathrm{SO}(n)$ mit einer beliebigen orthogonalen Matrix T der Determinante -1, also einer fest gewählten Spiegelung gilt. □

Zum *Beweis* des Satzes für $\mathrm{GL}(n, \mathbb{R})$ kann man wie oben eine Reduktion auf spezielle Matrizen erreichen. Man mache sich zuerst klar, daß das Schmidtsche Orthonormalisierungsverfahren die folgende matrizentheoretische Konsequenz hat.

Satz 13.22 *Zu jeder invertierbaren Matrix* $A \in \mathrm{GL}(n, \mathbb{R})$ *gibt es Matrizen* $K \in \mathrm{O}(n)$, $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, $\lambda_j > 0$, *und eine obere Dreiecksmatrix* N *mit Einsen in der Hauptdiagonalen, s. d.*

$$A = K\,D\,N$$

*(*Iwasawa–Zerlegung, *auch* KAN–Zerlegung *genannt für kompakt, abelsch, nilpotent (oder besser unipotent)).*

Nun ist klar, daß $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ stetig in $\mathrm{GL}(n, \mathbb{R})$ nach $E_n = \mathrm{diag}(1, \ldots, 1)$ deformiert werden kann vermöge der Familie

$$\mathrm{diag}((1 - t) + t\,\lambda_1, \ldots, (1 - t) + t\,\lambda_n) \, .$$

Entsprechend kann man N vermöge

$$\begin{pmatrix} 1 & t\,a_{12} & \cdots & t\,a_{1n} \\ 0 & 1 & \cdots & t\,a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

in E_n deformieren. Ist nun $\det A > 0$, so muß auch $\det K > 0$, d. h. $\det K = 1$ und damit $K \in \mathrm{SO}(n)$ sein. Der Rest folgt dann aus den Sätzen 21 und 22. $\square$

Bemerkung. Wie im Falle $\mathrm{SO}(n)$ kann man durch Zurückführung auf die Jordansche Normalform zeigen, daß auch die komplexe Gruppe

$$\mathrm{GL}(n, \mathbb{C})$$

zusammenhängend ist.

Einen Überblick über die sogenannten *klassischen Gruppen* gibt Brieskorn im 2. Band seiner Linearen Algebra.

14 Quadratische Formen und Quadriken

Eine *quadratische Form* auf $\mathbb{R}^n$ ist eine Funktion $\mathbb{R}^n \to \mathbb{R}$ der Gestalt

$$q\,(x) \;=\; \sum_{j,k=1}^{n} a_{jk}\, x_j x_k \;.$$

Wegen

$$\sum_{j,k} a_{jk}\, x_j x_k \;=\; \sum_{j,k} a_{kj}\, x_k x_j \;=\; \sum_{j,k} \frac{a_{jk}+a_{kj}}{2}\, x_j x_k$$

können und werden wir stets annehmen, daß $a_{jk} \,=\, a_{kj}$ für alle $j,\,k$ gilt. Damit ist jeder quadratischen Form q eine *symmetrische* Matrix A zugeordnet. Ist umgekehrt $A \in M\,(n \times n,\, \mathbb{R})$ symmetrisch, so ist

$$\langle\, x,\, y\,\rangle_A \;:=\; {}^t x\, A\, y$$

eine symmetrische Bilinearform auf $\mathbb{R}^n$ und

$$q\,(x) \;=\; q_A(x) \;=\; \langle\, x,\, x\,\rangle_A$$

eine quadratische Form. Mit dieser Beziehung kann man allgemein quadratische Formen q auf reellen Vektorräumen mit Hilfe von symmetrischen Bilinearformen b erklären durch $q\,(x) \;=\; b\,(x,\,x)$, wobei umgekehrt die quadratische Form q ihrerseits die sie definierende symmetrische Bilinearform b bestimmt durch die sogenannte *Polarisationsformel*:

$$b\,(x,\,y) \;=\; \frac{1}{2}\,(q(x+y) - q(x) - q(y))\;.$$

Quadratische Formen treten in natürlicher Weise in der Analysis auf. Es gilt bekanntlich: *Ist $f \,:\, I \to \mathbb{R}$ zweimal stetig differenzierbar, $I \subset \mathbb{R}$ offen, $f'(x_0) \,=\, 0$ an einer Stelle $x_0 \in I$ und $f''(x_0) \,>\, 0$, so besitzt f an der Stelle x_0 ein lokales Minimum.* Dies verallgemeinert sich in mehreren Veränderlichen zu der folgenden Aussage:

Ist $G \subset \mathbb{R}^n$ ein Gebiet, $x_0 \in G$ und $f : G \to \mathbb{R}$ zweimal stetig differenzierbar, gilt ferner

$$\operatorname{grad} f(x_0) \;=\; 0$$

und ist die Hessesche *in* x_0

$$\sum_{j,k} \frac{\partial^2 f}{\partial x_j \partial x_k}\,(x_0)\, x_j\, x_k$$

positiv–definit*, so besitzt f in x_0 ein lokales Minimum.*

Wir brauchen also einfache Kriterien dafür, wann eine quadratische Form q positiv definit ist.

Satz 14.1 *Es sei q eine quadratische Form auf dem (euklidischen) Vektorraum V. Dann sind die folgenden Aussagen äquivalent:*

i) *q ist positiv definit*;

ii) *es gibt eine (Orthonormal–)Basis $(v_1, \ldots, v_n)$ und positive Zahlen $\lambda_j > 0$, s. d.*

$$q\,(\sum_{j=1}^{n} a_j\, v_j) \;=\; \sum_{j=1}^{n} \lambda_j\, a_j^2\;;$$

iii) *es gibt eine (Orthogonal–)Basis $(u_1, \ldots, u_n)$, so daß*

$$q\,(\sum_{j=1}^{n} a_j\, u_j) \;=\; \sum_{j=1}^{n} a_j^2\;.$$

D. h. mit anderen Worten:

Es sei A eine symmetrische Matrix. Dann sind äquivalent:

i) ${}^t x A x > 0 \quad \forall x \neq 0$;

ii) *es existiert $S \in \mathrm{O}(n)$, s. d.* ${}^t S A S = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, $\lambda_j > 0$, $j = 1, \ldots, n$;

iii) *es existiert $S \in \mathrm{GL}(n, \mathbb{R})$, s. d. ${}^t S A S = E_n$, d. h. es existiert $T \in \mathrm{GL}(n, \mathbb{R})$, so daß $A = {}^t T T$.*

Beweis. i) $\Longrightarrow$ ii). Wegen des Satzes über die Hauptachsentransformation kann man A diagonalisieren: es gibt $S \in \mathrm{O}(n)$, so daß

$$ {}^t S A S = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) . $$

Mit $x \neq 0$ ist auch $S x \neq 0$, also

$$ 0 < {}^t(S x) A (S x) = {}^t x {}^t S A S x = \sum_{j=1}^{n} \lambda_j x_j^2 . $$

Daraus folgt aber $\lambda_j > 0$ für $j = 1, \ldots, n$.

ii) $\Longrightarrow$ iii) Man ersetze S durch

$$ S \cdot \mathrm{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_n}}\right) . $$

iii) $\Longrightarrow$ i) Es sei $x \neq 0$ und $y = S^{-1} x$. Dann folgt $y \neq 0$ und

$$ {}^t x A x = {}^t y {}^t S A S y = {}^t y E_n y = \sum_{j=1}^{n} y_j^2 > 0 . \qquad \square $$

Dieser Satz (und ganz allgemein das Transformationsgesetz $A \mapsto {}^t S A S$) zeigen, daß man nicht von den Eigenwerten der symmetrischen Matrix A als von *den Eigenwerten* der zugehörigen quadratischen Form reden kann. Die Eigenwerte hängen wesentlich von den Basen ab, ändern sich aber nicht, wenn man eine Basis *orthogonal* in eine andere transformiert. Der sogenannte *Trägheitssatz von Sylvester* gibt an, welche Größen tatsächlich unter beliebigen Basiswechseln invariant bleiben.

Satz 14.2 *Es sei q eine quadratische Form und A die darstellende symmetrische Matrix bzgl. einer Basis. Dann sind die folgenden Zahlen Invarianten von q (also unabhängig von der Basis):*

$$ \begin{aligned} k &= \textit{Anzahl der positiven Eigenwerte von } A , \\ \ell &= \textit{Anzahl der negativen Eigenwerte von } A , \\ r &= \textit{Rang von } A . \end{aligned} $$

Beweis. Siehe z. B. Fischer, Lineare Algebra, Satz 6.7.4.

Man nennt dann auch r den *Rang* von q, k den *Index* von q und $k - \ell$ die *Signatur* von q. Hat

q den Rang r, den Index k und die Signatur s, so besitzt q eine beschreibende Matrix der Gestalt

$$\left(\begin{array}{cc|cc|cc} 1 & 0 & & & & \\ & \ddots & & & & \\ 0 & 1 & & & & \\ \hline & & -1 & 0 & & \\ & & & \ddots & & \\ & & 0 & -1 & & \\ \hline & & & & 0 & 0 \\ & & & & & \ddots \\ & & & & 0 & 0 \end{array}\right) \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\ \ddots\\ 0\end{array}}\right\} k \\ \left.\vphantom{\begin{array}{c}1\\ \ddots\\ 0\end{array}}\right\} \ell = r - k \\ \left.\vphantom{\begin{array}{c}1\\ \ddots\\ 0\end{array}}\right\} n - r \end{array}$$

und $s = k - (r - k) = 2k - r$. Ebenso ist der Wert $\det A$ der quadratischen Form q *nicht eindeutig zugeordnet*. Wegen

$$\det({}^t S\, A\, S) = (\det S)^2 \cdot \det A$$

ist aber das *Vorzeichen* von $\det A$ eindeutig bestimmt, sofern nur $\det A \neq 0$, d. h. wenn die Form q *nicht ausgeartet* ist. Wir sagen kurz, die quadratische Form q *habe positive Determinante*, wenn $\det A > 0$ für alle darstellenden Matrizen, d. h. wenn $r = n$ und $\ell \equiv 0 \mod 2$.

Das *Hurwitz–Kriterium* für positive Definitheit läßt sich dann wie folgt formulieren:

Satz 14.3 *Es sei q eine quadratische Form auf dem endlich–dimensionalen $\mathbb{R}$–Vektorraum V der Dimension n. Dann sind die folgenden Aussagen äquivalent:*

i) *q ist positiv definit* ;

ii) *für jeden Unterraum $U \subset V$ hat $q|U$ positive Determinante* ;

iii) *es gibt eine Fahne*

$$\{0\} = U_0 \subset U_1 \subset \cdots \subset U_{n-1} \subset U_n = V\,,$$

so daß $q|U_j$ positive Determinante hat für $j = 1, \ldots, n$.

Die klassische Form dieses Satzes ist nur eine Umformulierung (es ist im übrigen strittig, ob man diese Hurwitz zuschreiben soll):

Folgerung 14.4 *Es sei $A \in M(n \times n, \mathbb{R})$ eine symmetrische Matrix. Dann ist*

$$q_A(x) = {}^t x\, A\, x$$

positiv definit genau dann, wenn alle Hauptunterdeterminanten

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1j} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jj} \end{pmatrix}, \quad j = 1, \ldots, n$$

positiv sind.

Beweis des Satzes. i) $\Longrightarrow$ ii) $q|U$ ist nach Definition auch positiv definit. Man braucht also ii) nur für $U = V$ zu beweisen. Wegen Satz 1 ist aber E_n eine beschreibende Matrix für q; somit hat q positive Determinante.

ii) $\Longrightarrow$ iii) ist trivial.

iii) $\Longrightarrow$ i) per Induktion nach n. Für $n = 1$ ist die Aussage trivialerweise richtig. Sei sie also für Vektorräume der Dimension $n-1$ schon bewiesen. Dann können wir sie auch auf $q|U_{n-1}$ anwenden und finden eine Basis $(u_1, \ldots, u_{n-1})$ von U_{n-1}, so daß

$$q\,(a_1\,u_1 + \cdots + a_{n-1}\,u_{n-1}) \,=\, a_1^2 + \cdots + a_{n-1}^2 \,.$$

In dieser Basis gilt dann für die zugehörige Bilinearform b:

$$b\,(u_j,\, u_k) \,=\, \delta_{jk}\,, \quad 1 \le j,\, k \le n-1\,.$$

Es sei nun $(u_1, \ldots, u_{n-1},\, v_n)$ eine Basis von V. Setze

$$u_n \,=\, v_n \,-\, b\,(u_1,\, v_n)\,u_1 \,-\, \cdots \,-\, b\,(u_{n-1},\, v_n)\,u_{n-1}\,.$$

Dann ist auch $(u_1, \ldots, u_n)$ eine Basis von V, und es gilt

$$\begin{aligned} b\,(u_j,\, u_n) \,&=\, b\,(u_j,\, v_n) \,-\, \sum_{k=1}^{n} b\,(u_k,\, v_n)\,b\,(u_j,\, u_k) \\ &=\, b\,(u_j,\, v_n) \,-\, \sum_{k=1}^{n} \delta_{jk}\,b\,(u_k,\, v_n) \,=\, 0\,. \end{aligned}$$

Infolgedessen hat b in dieser Basis Diagonalgestalt:

$$\operatorname{diag}\,(1, \ldots, 1, b\,(u_n,\, u_n))\,.$$

Da q nach Voraussetzung positive Determinante hat, muß auch $b\,(u_n,\, u_n) \,>\, 0$ gelten, und folglich ist q positiv definit. $\square$

Im zweiten Teil dieses Abschnitts wollen wir uns mit *Kegelschnitten*, allgemeiner mit sogenannten *Quadriken* beschäftigen.

Man kann einen Kegel K in $\mathbb{R}^{n+1}$ mit den Koordinaten $\widetilde{x} \,=\, {}^t(x_0, \ldots, x_n)$ beschreiben durch die Gleichung:

$$K \,=\, \{x_0^2 \,=\, x_1^2 + \cdots + x_n^2\}\,.$$

Oder in Matrixschreibweise:

$$K \,=\, \{q_T\,(\widetilde{x}) \,=\, {}^t\widetilde{x}\,T\,\widetilde{x} \,=\, 0\}\,,$$

wobei

$$T \,=\, \operatorname{diag}\,(-1,\, 1, \ldots, 1)\,.$$

(Die indefinite *Minkowski–Form* q_T spielt übrigens in der *speziellen Relativitätstheorie* dieselbe grundlegende Rolle wie das (positiv–definite) Skalarprodukt $\langle\,.\,,\,.\,\rangle$ im euklidischen Raum).

Ein *Kegelschnitt* ist (zumindest im klassischen Sprachgebrauch für den Fall $n \,=\, 2$) nichts anderes als der Schnitt von K mit einer *affinen* (Hyper–)Ebene E:

$$K \cap \{\lambda\,(x_0, \ldots, x_n) \,=\, \alpha \,:\, \lambda \in (\mathbb{R}^{n+1})^* \setminus \{0\} \quad \text{und } \alpha \in \mathbb{R} \text{ fest }\}\,.$$

Da man die gegebene Hyperebene durch affine Transformation in die spezielle Hyperebene $x_0 \,=\, 0$ bewegen kann, ist dazu äquivalent, daß man den Kegel selbst *affin* bewegt:

$$\widetilde{x} \longmapsto A\,\widetilde{x} + \alpha\,, \quad A \in \mathrm{GL}\,(n+1, \mathbb{R})\,, \quad \alpha \in \mathbb{R}^{n+1}\,,$$

und mit der Hyperebene $\{x_0 \,=\, 0\}$ schneidet, also $x_0 \,=\, 0$ setzt. Schreiben wir also

$$x \,=\, {}^t(0,\, x_1, \ldots, x_n)\,,$$

wobei $\widetilde{x} = {}^t(x_0, x_1, \ldots, x_n)$, so wird jeder Kegelschnitt beschrieben durch eine Gleichung der Form

$$\begin{aligned} &{}^t x\,({}^t A\,T\,A)\,x + 2\,({}^t \alpha\,T\,A)\,x + {}^t \alpha\,T\,\alpha = \\ &{}^t x\,({}^t A\,T\,A)\,x + {}^t \alpha\,T\,A\,x + {}^t x\,{}^t A\,T\,\alpha + {}^t \alpha\,T\,\alpha = {}^t(A\,x + \alpha)\,T\,(A\,x + \alpha) = 0\,. \end{aligned}$$

Ist dann

$${}^t A\,T\,A = \begin{pmatrix} * & * & \cdots & * \\ * & s_{11} & \cdots & s_{1n} \\ \vdots & \vdots & & \vdots \\ * & s_{n1} & \cdots & s_{nn} \end{pmatrix}, \quad {}^t \alpha\,T\,A = (*, s_{01}, \ldots, s_{0n})$$

und ${}^t\alpha\,T\,\alpha = s_{00}$, setzt man ferner $s_{j0} = s_{0j}$ und

$$S' = (s_{jk})_{0 \le j,k \le n}\,, \quad S = (s_{jk})_{1 \le j,k \le n}\,,$$

so kann man also jeden Kegelschnitt schreiben in der Form

$$(\,*\,) \qquad \sum_{j,k=1}^{n} s_{jk}\,x_j\,x_k + 2\sum_{j=1}^{n} s_{0j}\,x_j + s_{00} = 0$$

oder

$$(\,**\,) \qquad K = \{x = {}^t(x_1, \ldots, x_n) \in \mathbb{R}^n : {}^t x'\,S'\,x' = 0\}\,,$$

wobei

$$S' = {}^t S' = \left(\begin{array}{c|ccc} s_{00} & s_{01} & \cdots & s_{0n} \\ \hline s_{10} & s_{11} & \cdots & s_{1n} \\ \vdots & \vdots & & \vdots \\ s_{n0} & s_{n1} & \cdots & s_{nn} \end{array}\right) \in M\,((n+1) \times (n+1), \mathbb{R})$$

und $x' = {}^t(1, x_1, \ldots, x_n)$.

Nach geeigneter linearer Transformation kann man S in Diagonalgestalt bringen:

$${}^t Q\,S\,Q = \operatorname{diag}(\underbrace{1, \ldots, 1}_{k}, \underbrace{-1, \ldots, -1}_{r-k}, 0, \ldots, 0)\,,$$

wobei r der Rang von S ist. Somit wird $(*)$ zu

$$\sum_{j=1}^{k} y_j^2 - \sum_{j=k+1}^{r} y_j^2 + 2\sum_{j=1}^{n} c_{0j}\,y_j + s_{00} = 0\,.$$

Durch quadratische Ergänzungen

$$(y_j + c_{0j})^2 = y_j^2 + 2\,c_{0j}\,y_j + c_{0j}^2 \quad \text{etc.}$$

und geeigneter Translation erhält man dann die Gestalt

$$(+) \qquad \sum_{j=1}^{k} y_j^2 - \sum_{j=k+1}^{r} y_j^2 + 2\sum_{j=r+1}^{n} c_{0j}\,y_j + c_{00} = 0\,.$$

Nun gibt es drei Fälle:

a) $c_{0j} = 0\,, \quad j = 0, r+1, \ldots, n\,;$

b) $c_{00} \neq 0\,, \quad c_{0j} = 0\,, \quad j = r+1, \ldots, n\,;$

c) es gibt ein $j > 0$ mit $c_{0j} \neq 0\,.$

Satz 14.5 *Nach geeigneter affiner Transformation hat jede Quadrik eine der folgenden Normalformen* :

a) $w_1^2 + \cdots + w_k^2 - w_{k+1}^2 - \cdots - w_r^2 = 0\,;$

b) $w_1^2 + \cdots + w_k^2 - w_{k+1}^2 - \cdots - w_r^2 = 1\,;$

c) $w_1^2 + \cdots + w_k^2 - w_{k+1}^2 - \cdots - w_r^2 + 2\,w_{r+1} = 0\,.$

Hat man die Darstellung $(**)$ für die Quadrik, ist r der Rang von S und r' der Rang von S', so entsprechen die Fälle a), b), c) genau den Fällen

$$a') \quad r = r'\,, \quad b') \quad r+1 = r'\,, \quad c') \quad r+2 = r'\,.$$

Um Eindeutigkeit zu erhalten, müssen wir uns bei a) und c) zwischen k und $r-k$ entscheiden (da man beide Gleichungen mit -1 multiplizieren und $-w_{r+1}$ durch w_{r+1} ersetzen kann). In der Tat gilt dann der folgende

Zusatz *Nach affiner Transformation ist jede nichtleere Quadrik im* $\mathbb{R}^n$ *beschreibbar durch genau eine der obigen Gleichungen, wobei* $k \geq r-k$ *in den Fällen* a) *und* c).

Man kann im übrigen anhand der erweiterten Matrizen S' von vornherein entscheiden, wann zwei Quadriken in diesem Sinne geometrisch äquivalent sind. Dazu benötigt man die Ränge von S und S' und die Beträge der Signaturen von S und S'. (Siehe Fischer, Analytische Geometrie, Satz 1. 4. 7).

Will man die Quadriken genauer bis auf *euklidische* Bewegungen klassifizieren (man spricht dann auch von *metrischer* Normalform), so besteht der entscheidende Unterschied zur affinen Klassifikation darin, daß wir bei der Hauptachsentransformation für S nur *orthogonale* Matrizen benutzen dürfen. Wir kommen dann auf die Gestalt

$${}^tQ\,S\,Q = \operatorname{diag}(\lambda_1, \ldots, \lambda_k, -\lambda_{k+1}, \ldots, -\lambda_r, 0, \ldots, 0)$$

mit positiven $\lambda_j \in \mathbb{R}$, $Q \in \mathrm{SO}(n)$, und schließlich zur Existenz positiver Zahlen $\alpha_1, \ldots, \alpha_r$, s. d. die metrischen Normalformen wie folgt aussehen:

a) $\left(\dfrac{y_1}{\alpha_1}\right)^2 + \cdots + \left(\dfrac{y_k}{\alpha_k}\right)^2 - \left(\dfrac{y_{k+1}}{\alpha_{k+1}}\right)^2 - \cdots - \left(\dfrac{y_r}{\alpha_r}\right)^2 = 0\,;$

b) $\left(\dfrac{y_1}{\alpha_1}\right)^2 + \cdots + \left(\dfrac{y_k}{\alpha_k}\right)^2 - \left(\dfrac{y_{k+1}}{\alpha_{k+1}}\right)^2 - \cdots - \left(\dfrac{y_r}{\alpha_r}\right)^2 = 1\,;$

c) $\left(\dfrac{y_1}{\alpha_1}\right)^2 + \cdots + \left(\dfrac{y_k}{\alpha_k}\right)^2 - \left(\dfrac{y_{k+1}}{\alpha_{k+1}}\right)^2 - \cdots - \left(\dfrac{y_r}{\alpha_r}\right)^2 + 2y_{r+1} = 0\,.$

Es ist unbedingt erforderlich, sich diese Normalformen in den Dimensionen $n = 2$ und 3 zu veranschaulichen und ihre Namen parat zu haben.

Wir geben eine vollständige Liste im Falle $n = 2$ auf der nächsten Seite. Hierbei ist der Fall $r = 0$ natürlich der lineare und sollte daher nicht ernsthaft zu den Quadriken gerechnet werden. Schließt man dann noch den Fall der leeren Menge aus („leere Quadrik“), so teilt man die verbleibenden Fälle ein in die *nicht–entarteten Quadriken* : Ellipse, Hyperbel, Parabel;

entarteten Quadriken : Doppelpunkt, Doppelgerade, zwei parallele Geraden, zwei sich schneidende Geraden.

Hiervon ist nur der Fall paralleler Geraden kein Kegelschnitt.

r	k	a)	b)	c)
[0	0	($\mathbb{R}^2$)	($\emptyset$)	Gerade]
1	0		($\emptyset$)	
	1	(Doppel-) Gerade	2 parallele Geraden	Parabel
2	0		($\emptyset$)	
	1	2 sich schneidende Geraden	Hyperbel	
	2	(Doppel-) Punkt	Ellipse (Kreis)	

In der Dimension $n = 3$ gibt es eine ganze Reihe von trivialen und entarteten Quadriken, z. B. alle Mengen der Form $K_2 \times \mathbb{R}$, K_2 eine zwei–dimensionale Quadrik (also z. B. „Zylinder“ über Kreisen, zwei sich schneidende Ebenen, etc.). Wir listen nur die nicht–entarteten, nicht–zylindrischen Formen auf mit ihren affinen Normalformen:

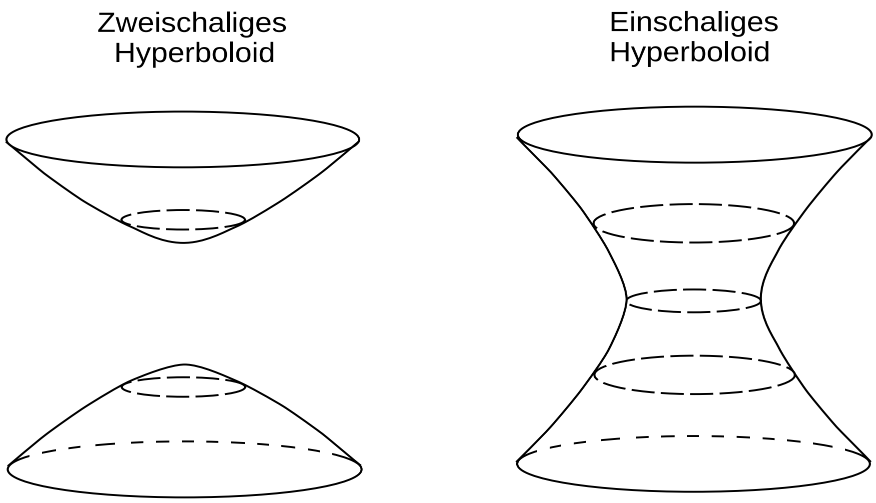
$x_1^2 - x_2^2 - x_3^2 = 1$ zweischaliges Hyperboloid;

$x_1^2 + x_2^2 - x_3^2 = 1$ einschaliges Hyperboloid;

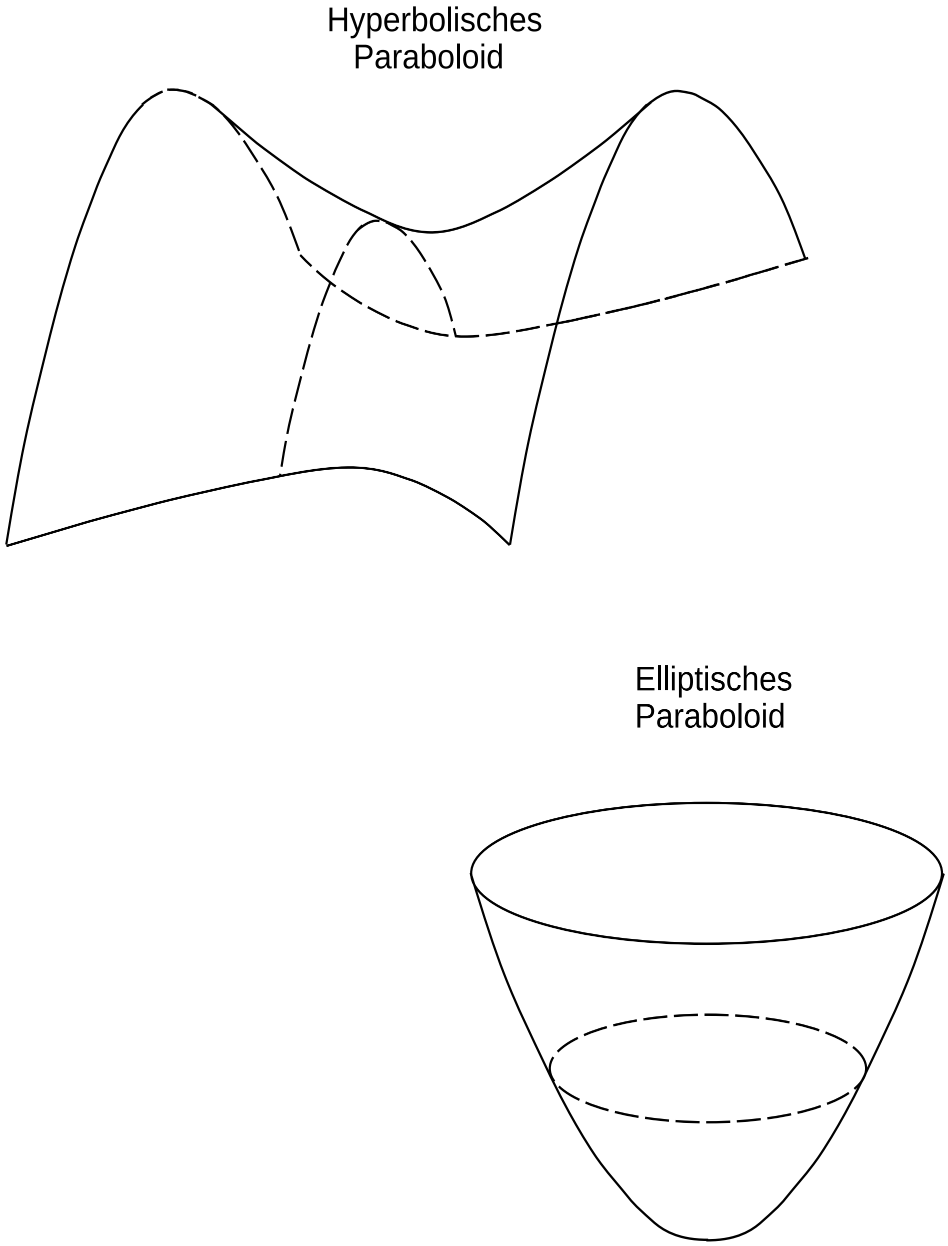
$x_1^2 + x_2^2 + x_3^2 = 1$ Kugel (Ellipse);

$x_1^2 - x_2^2 + 2\,x_3 = 0$ hyperbolisches Paraboloid;

$x_1^2 + x_2^2 - 2\,x_3 = 0$ elliptisches Paraboloid.



Figur 14.1



Figur 14.2

15 Die Methode der kleinsten Quadrate

Es sei $A \in M(m \times n, \mathbb{R})$, $b \in \mathbb{R}^m \setminus \{0\}$. Ist die Gleichung

$$(*) \qquad A\,x = b$$

lösbar, so wollen wir unter den evtl. vielen Lösungen eine *bestmögliche (im euklidischen Sinne)* auszeichnen. Ist $(*)$ dagegen nicht lösbar, so können wir immerhin versuchen, eine bestmögliche *Approximation* zu finden. Es wird hier darum gehen, die geometrischen und matrizentheoretischen Gesichtspunkte zur Lösung dieser Frage herauszuarbeiten, die in Verbindung stehen zu *Pseudoinversen*, dem *Schmidtschen Orthonormalisierungsverfahren*, und auch zu *normalen Matrizen* etc.

Bevor wir dieses Problem genauer angehen, wollen wir zunächst an den Begriff der *Pseudoinversen* aus der früheren Übungsaufgabe 44 erinnern: Eine Matrix $A^+ \in M(n \times m, \mathbb{R})$ heißt Pseudoinverse zu A, falls:

$$A\,A^+\,A = A\,, \quad A^+\,A\,A^+ = A^+\,.$$

Es wurde dort der folgende Satz behauptet, dessen einfachen Beweis wir hier anfügen:

Satz 15.1 *Die Gleichung $A\,x = b$ ist genau dann lösbar, wenn für eine Pseudoinverse A^+ von A die Beziehung $A\,A^+ b = b$ besteht. Es ist dann $x = A^+ b$ eine Lösung.*

Beweis. i) $A\,x = b$ ist höchstens dann lösbar, wenn $b = Ax = AA^+Ax = AA^+b$ gilt.
ii) Sei umgekehrt $b = AA^+b$, $x := A^+b$. Dann folgt $Ax = AA^+b = b$. □

Um dieses Ergebnis geometrisch zu deuten, mache man sich klar, daß die Existenz einer Pseudoinversen geeignete Zerlegungen der betroffenen Räume induziert (den einfachen Beweis überlassen wir dem Leser):

Satz 15.2 *Jede Pseudoinverse A^+ zu A liefert Zerlegungen*

$$V = \ker A \oplus \operatorname{im} A^+A\,, \quad W = \operatorname{im} A \oplus \ker A\,A^+\,.$$

Wegen der zweiten Zerlegung wird die kanonische Projektion von W parallel zu $\ker A\,A^+$ nach $\operatorname{im} A$ gerade durch die Komposition $A\,A^+$ gegeben, und die notwendige Bedingung $b = AA^+b$ besagt einfach, daß b mit seiner Projektion nach $\operatorname{im} A$ übereinstimmt. Nach unserem allgemeinen Strukturtheorem für lineare Abbildungen induziert A bezüglich der obigen Zerlegung eine Isomorphie

$$\operatorname{im} A^+A \longrightarrow \operatorname{im} A\,,$$

die konkret durch

$$A^+Ax \longmapsto AA^+Ax = Ax \in \operatorname{im} A$$

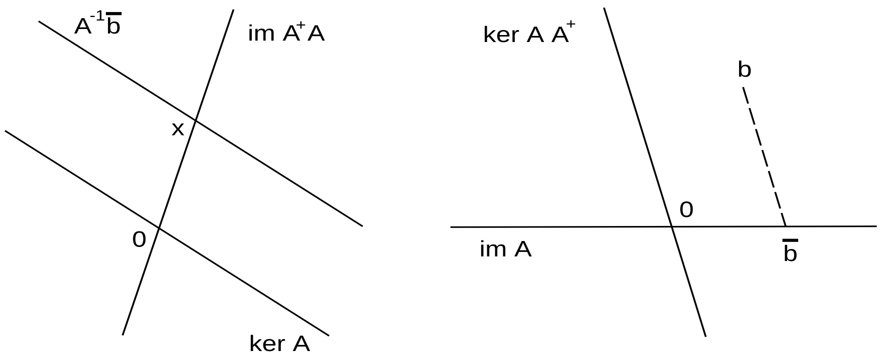
beschrieben wird. Ihre Umkehrabbildung ist dann nichts anderes als (die Einschränkung auf $\operatorname{im} A$) der Pseudoinversen A^+. Betrachten wir nun für beliebiges $b \in W$ die Abbildung A^+, so sehen wir, daß diese wegen

$$A^+b = A^+AA^+b = A^+\overline{b} \quad \text{mit} \quad A^+\overline{b} = A\,A^+b$$

über die Projektion von W nach $\operatorname{im} A$ faktorisiert. Infolgedessen ist $x = A^+b$ nichts anderes als die eindeutig bestimmte Lösung der Gleichung

$$A\,x = \overline{b} \quad \text{mit} \quad x \in \operatorname{im} A^+A\,,$$

wobei $\overline{b}$ die Projektion von b nach $\operatorname{im} A$ bezeichnet. Mit anderen Worten: Die Pseudoinverse ist gleich der Komposition der Projektion von W nach $\operatorname{im} A$, der Umkehrabbildung von A (eingeschränkt auf $\operatorname{im} A^+A$) und der Einbettung des letzten Raumes in V.



Figur 15.1

Hat man nun umgekehrt Zerlegungen

$$V = \ker A \oplus V' \,, \quad W = \operatorname{im} A \oplus W'$$

vorliegen, so induziert wieder A einen Isomorphismus von V' nach $\operatorname{im} A$. Bezeichnet man die entsprechende Umkehrabbildung mit A' und bildet man wie oben die Komposition von A' mit der entsprechenden Projektion bzw. Injektion

$$W \longrightarrow \operatorname{im} A \xrightarrow{A'} V' \hookrightarrow V \,,$$

so ist die resultierende Abbildung ersichtlich eine Pseudoinverse zu A. Somit sind Pseudoinverse und Zerlegungen von Urbild– und Bildraum bzgl. $\ker A$ bzw. $\operatorname{im} A$ ein und dasselbe.

Im *euklidischen* Fall können wir nun statt der nicht kanonischen direkten Summenzerlegungen die eindeutig bestimmten *orthogonalen* Komplemente zu $\ker A$ und $\operatorname{im} A$ heranziehen. Es gibt dann bzgl. dieser Zerlegung eine eindeutig bestimmte Pseudoinverse, die wir kurz als *orthogonale* Pseudoinverse bezeichnen werden. Da eine orthogonale Projektion einen gegebenen Vektor auf denjenigen im Bildraum abbildet, der zu ihm minimalen Abstand hat (siehe Übungsaufgaben zu Kapitel 10), so löst die Gleichung $x = A^+ b$ das eingangs gestellte Approximationsproblem im folgenden Sinne:

i) Minimiere $\| A x - b \|$, d. h. finde (alle) $\widetilde{x} \in \mathbb{R}^n$ mit

$$\| A\widetilde{x} - b \| \le \| A x - b \| \quad \text{für alle} \quad x \in \mathbb{R}^n \,.$$

ii) Suche unter diesen $\widetilde{x}$ das Element kleinster Norm $\|\widetilde{x}\|$.

Es ist unmittelbar klar, daß es stets genau eine solche Lösung gibt. Wir nennen sie die *beste Approximation bzgl. kleinster Quadrate*.

Satz 15.3 *Bezeichnet* A^+ *die* orthogonale *Pseudoinverse zu einer linearen Abbildung* $A : V \to W$ *endlich–dimensionaler euklidischer Vektorräume, so ist* $\overline{b} = A\,A^+ b$ *die orthogonale Projektion eines beliebigen Elementes* $b \in W$ *nach* $\operatorname{im} A$ *und* $x = A^+ b = A^+ \overline{b}$ *ist die gesuchte beste Approximation der Gleichung* $(*)$.

Das *Fazit* aus diesen Überlegungen ist also: kann man die orthogonale Pseudoinverse A^+ von A rechnerisch in den Griff bekommen, so hat man ein effektives Verfahren, um x in Abhängigkeit von b zu bestimmen.

Nun wissen wir von früher, daß die *orthogonalen* Komplemente durch die *transponierte Matrix* zu A bestimmbar sind:

$$V = \ker A \perp \operatorname{im} {}^tA \,, \quad W = \operatorname{im} A \perp \ker {}^tA \,.$$

Unsere Aufgabe besteht also schlicht darin, die zu diesen Zerlegungen gehörige Pseudoinverse, also die Umkehrabbildung zu A eingeschränkt auf im ${}^tA$ zu berechnen. – Aber Vorsicht!

Warnung. Einfache Beispiele zeigen, daß die naheliegende Vermutung $A^+ = {}^tA$ nicht korrekt ist. M. a. W: Im allgemeinen ist $(A|\,\mathrm{im}\,{}^tA)^{-1} \neq {}^tA|\,\mathrm{im}\,A$!

Diese Beobachtung bringt uns immerhin auf die Idee, die Abweichung von ${}^tA\,A$ von der Identität auf im ${}^tA$ zu berechnen. Also untersuchen wir zunächst allgemein die Matrix ${}^tA\,A$. Sie ist eine symmetrische $n \times n$–Matrix:

$${}^t({}^tA\,A) = {}^tA\,{}^{tt}A = {}^tA\,A\,.$$

Nach dem Satz über die Hauptachsentransformation gibt es zu $S := {}^tA\,A$ eine orthogonale Matrix Q, so daß

$${}^tQ\,S\,Q = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n)$$

mit reellen Eigenwerten $\lambda_1, \ldots, \lambda_n$. Wir interpretieren diese Aussage um: Es seien $q_1, \ldots, q_n$ die Spalten von Q; dann bilden diese eine Orthonormalbasis, bestehend aus Eigenvektoren:

$$S\,q_j = \lambda_j\,q_j\,, \quad {}^tq_j\,q_k = 0\,, \quad j \neq k\,, \quad {}^tq_j\,q_j = 1\,.$$

In unserem Fall ergibt sich damit sofort

$$0 \leq \|A\,q_j\|^2 = {}^tq_j{}^tA\,A\,q_j = {}^tq_j\,S\,q_j = \lambda_j\,\|q_j\|^2 = \lambda_j\,.$$

Folgerung. Die Eigenwerte von ${}^tA\,A$ sind nicht–negativ.

Seien nun die Eigenwerte so geordnet, daß $\lambda_1, \ldots, \lambda_r > 0$, $\lambda_{r+1} = \cdots = \lambda_n = 0$. Wir setzen $\mu_j = \sqrt{\lambda_j}$, $j \leq r$. Dann ist $A\,q_j = 0$, $j > r$, und die Elemente

$$y_j := \frac{A\,q_j}{\mu_j}\,, \quad 1 \leq j \leq r\,,$$

bilden ein Orthonormalsystem in $\mathbb{R}^m$, da

$$\langle\, y_j,\, y_k\,\rangle = \mu_j^{-1}\,\mu_k^{-1}\,\langle\, A\,q_j,\, A\,q_k\,\rangle = \mu_j^{-1}\,\mu_k^{-1}\,{}^tq_j\,{}^tAA\,q_k = \mu_j^{-1}\,\mu_k^{-1}\,\lambda_j\delta_{jk}$$

für j, $k \leq r$. Man ergänzt nun dieses System nach Schmidt zu einer Orthonormalbasis $(y_1, \ldots, y_m)$ von $\mathbb{R}^m$ und definiert $Q_2 = Q$ und Q_1 als die orthonormale $m \times m$–Matrix, deren Spalten die y_k sind. Damit ist

$${}^tQ_1\,A\,Q_2 = \underbrace{\left(\begin{array}{ccc|cc} \mu_1 & & 0 & 0 & 0 \\ & \ddots & & & \\ 0 & & \mu_r & 0 & 0 \\ \hline 0 & & 0 & 0 & 0 \\ & & & & \\ 0 & & 0 & 0 & 0 \end{array}\right)}_{n} \left.\vphantom{\begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array}}\right\}m \quad =: \Sigma\,.$$

Wir haben damit unser eigentliches Ziel wesentlich genauer getroffen, indem wir mit den obigen Überlegungen kartesische Koordinaten auf im ${}^tA$ und im A gefunden haben, so daß A dort Diagonalgestalt mit positiven Diagonalelementen besitzt.

Wir fassen noch einmal zusammen:

Satz 15.4 *Für jede reelle* $m \times n$*–Matrix* A *gibt es orthogonale Matrizen* Q_1 , Q_2 *und* Σ *wie oben mit positiven Einträgen* $\mu_1, \ldots, \mu_r$ *, so daß* $A = Q_1 \, \Sigma \, {}^t Q_2$ *.*

Es ist offensichtlich, daß die orthogonale Pseudoinverse Σ^+ von Σ durch die Matrix

$$\Sigma^+ = \underbrace{\left(\begin{array}{ccc|cc} \mu_1^{-1} & & 0 & 0 & 0 \\ & \ddots & & & \\ 0 & & \mu_r^{-1} & 0 & 0 \\ \hline 0 & & 0 & 0 & 0 \\ & & & & \\ 0 & & 0 & 0 & 0 \end{array} \right)}_{m} \left. \vphantom{\begin{array}{c} a \\ a \\ a \\ a \\ a \\ a \end{array}} \right\} n$$

dargestellt wird. – Damit ist klar:

Satz 15.5 *Ist* $A = Q_1 \, \Sigma \, {}^t Q_2$ *eine Zerlegung wie oben, so ist* $A^+ = Q_2 \, \Sigma^+ \, {}^t Q_1$ *die orthogonale Pseudoinverse von* A *.*

Wir beweisen zum Schluß noch einmal, wie aus der obigen Gestalt der orthogonalen Pseudoinversen ihre charakteristische Eigenschaft folgt. Wir benötigen dazu nur die Definition für orthogonale Abbildungen:

$$\langle \, Qu, \, Qv \, \rangle = \langle \, u, \, , v \, \rangle \, , \quad \text{insbesondere} \quad \| \, Qv \, \| = \| \, v \, \| \, .$$

Nach der obigen Zerlegung ist

$$\| \, Ax - b \, \| = \| \, Q_1 \, \Sigma \, {}^t Q_2 \, x - b \, \| = \| \, \Sigma \, {}^t Q_2 \, x - {}^t Q_1 \, b \, \| \, ,$$

so daß sich für $y = {}^t Q_2 \, x$ ergibt:

$$\| \, y \, \| = \| \, {}^t Q_2 \, x \, \| = \| \, Q_2^{-1} \, x \, \| = \| \, x \, \| \, .$$

Also entspricht einem *minimalen* x mit minimalem Abstand $\| \, A \, x - b \, \|$ ein *minimaler* Vektor y , welcher den Abstand $\| \, \Sigma \, y - {}^t Q_1 \, b \, \|$ minimiert. Dieser ist aber, wie man bei der speziellen Form von Σ sofort sieht, notwendig gegeben durch

$$y = \Sigma^+ \, {}^t Q_1 \, b \, ,$$

und daraus folgt, wie wir schon wissen,

$$x = Q_2 \, y = Q_2 \, \Sigma^+ \, {}^t Q_1 \, b = A^+ \, b \, . \qquad \square$$

Bemerkung. Aus dem Vorstehenden entnimmt man leicht, daß die *orthogonale* unter allen Pseudoinversen eindeutig bestimmt ist durch die zusätzliche Symmetrieforderung

$${}^t(A^+ A) = A^+ A \, , \quad {}^t(A \, A^+) = A \, A^+ \, .$$

In dieser Form wurde sie von Moore und Penrose eingeführt.

Supplement: Sphärische Trigonometrie

Im dreidimensionalen Raum $\mathbb{R}^3$ berechnet sich die Determinante D (und damit das Volumen $V = |D|$ eines durch die Vektoren u, v, w bestimmten *Spates* oder *Parallelotopes*) zu

$$\begin{aligned} D(u, v, w) &= D\Big(\sum_i u_i e_i\,,\ \sum_j v_j e_j\,,\ \sum_k w_k e_k\Big) = \sum_{i,j,k} u_i v_j w_k D(e_i, e_j, e_k) \\ &= u_1 v_2 w_3 - u_1 v_3 w_2 + u_2 v_3 w_1 - u_2 v_1 w_3 + u_3 v_1 w_2 - u_3 v_2 w_1 \\ &=: \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} . \end{aligned}$$

Führt man das sogenannte *Vektor–* oder *Kreuzprodukt* in $\mathbb{R}^3$ ein durch

$$v \times w := {}^t(v_2 w_3 - v_3 w_2\,,\ v_3 w_1 - v_1 w_3\,,\ v_1 w_2 - v_2 w_1)\,,$$

so kann man formal schreiben

$$v \times w = \begin{vmatrix} e_1 & e_2 & e_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} ,$$

und es wird

$$D(u, v, w) = \langle u, v \times w \rangle\,, \quad D(w, u, v) = \langle w, u \times v \rangle\,, \quad \text{etc.}$$

Man sieht sofort, daß der Vektor $v \times w$ die folgenden Bedingungen erfüllt:

1. $$v \times w \perp v\,, \quad v \times w \perp w\,,$$

2. $$||\, v \times w \,|| = ||\, v \,|| \cdot ||\, w \,|| \cdot |\sin \varphi| = F(v, w)\,.$$

(Hierbei bezeichnet φ den Winkel zwischen den Vektoren v und w, und $F(v, w)$ ist der Flächeninhalt des von diesen Vektoren aufgespannten Parallelogramms). Die zweite Formel ergibt sich aus

$$\|\, v \times w \,\|^2 = \|\, v \,\|^2 \,\|\, w \,\|^2 - \langle v, w \rangle^2 = \|\, v \,\|^2 \,\|\, w \,\|^2 \,(1 - \cos^2 \varphi)\,.$$

Es gibt zu linear unabhängigen Vektoren v und w genau zwei Vektoren, die den Bedingungen 1. und 2. genügen. Sie unterscheiden sich nur durch ihr Vorzeichen; die Auswahl des Kreuzproduktes $v \times w$ ist bestimmt durch die dritte Forderung

3. $$D(v, w, v \times w) = D(v \times w, v, w) \ge 0\,,$$

d. h. das System $(v, w, v \times w)$ ist (bei linear unabhängigen v, w) *positiv orientiert (rechtshändig).*

Einfach einzusehende *Rechenregeln* für das Vektorprodukt sind:

$$\begin{aligned} (v_1 + v_2) \times w &= v_1 \times w + v_2 \times w\,, \\ (\lambda v) \times w &= \lambda\,(v \times w)\,, \\ w \times v &= -v \times w\,. \end{aligned}$$

Wir rechnen noch ein paar weitere nützliche Formeln aus:

(a) $$(u \times v) \times w = \langle u, w \rangle\, v - \langle v, w \rangle\, u\,.$$

(b) $$\langle u_1 \times u_2, v_1 \times v_2 \rangle = \begin{vmatrix} \langle u_1, v_1 \rangle & \langle u_1, v_2 \rangle \\ \langle u_2, v_1 \rangle & \langle u_2, v_2 \rangle \end{vmatrix} .$$

(c) $$(u_1 \times u_2) \times (v_1 \times v_2) \, = \, D\,(u_1,\, v_1,\, v_2)\, u_2 \, - \, D\,(u_2,\, v_1,\, v_2)\, u_1 \; .$$

Beweis. Zu (a):

$$\begin{aligned}(u \times v) \times w \, &= \, \Big((\textstyle\sum u_i\, e_i) \times (\sum v_j\, e_j) \Big) \times (\sum w_k\, e_k) \\ &= \textstyle\sum u_i v_j w_k\, ((e_i \times e_j) \times e_k) \\ &= \textstyle\sum u_i v_j w_k\, (\langle\, e_i,\, e_k \,\rangle\, e_j \, - \, \langle\, e_j,\, e_k \,\rangle\, e_i) \\ &= \langle \textstyle\sum u_i\, e_i,\, \sum w_k\, e_k \,\rangle \sum v_j\, e_j \, - \, \langle \sum v_j\, e_j,\, \sum w_k\, e_k \,\rangle \sum u_i\, e_i \\ &= \langle\, u,\, w \,\rangle\, v \, - \, \langle\, v,\, w \,\rangle\, u \; .\end{aligned}$$

Zu (b):

$$\begin{aligned}\langle\, u_1 \times u_2 \,,\; v_1 \times v_2 \,\rangle \, &= \, D\,(u_1 \times u_2,\, v_1,\, v_2) \, = \, D\,(v_2,\, u_1 \times u_2,\, v_1) \\ &= \langle\, v_2,\, (u_1 \times u_2) \times v_1 \,\rangle \\ &= \langle\, v_2,\, \langle\, u_1,\, v_1 \,\rangle\, u_2 \, - \, \langle\, u_2,\, v_1 \,\rangle\, u_1 \,\rangle \; .\end{aligned}$$

Zu (c): $(u_1 \times u_2) \times (v_1 \times v_2) \, = \, \langle\, u_1,\, v_1 \times v_2 \,\rangle\, u_2 \, - \, \langle\, u_2,\, v_1 \times v_2 \,\rangle\, u_1$.

Wir wenden diese Formeln an, um die Grundergebnisse der *Sphärischen Trigonometrie* abzuleiten. Es handelt sich dabei um geometrische Aussagen auf der Kugeloberfläche S^2: Kürzeste Verbindungen (Geraden) sind hier die *(Segmente von) Großkreise(n)*, d. h. Durchschnitte der Sphäre S^2 mit Ebenen durch ihren Mittelpunkt. Ohne Einschränkung sei S^2 als *Einheitskugel* angenommen:

$$S^2 \, = \, \{u \in \mathbb{R}^3 : \, \|\, u \,\| \, = \, 1\} \; .$$

Wir können damit *Dreiecke* auf der Kugel studieren; gegeben seien also drei Punkte $P_j \in S^2$, die wir uns so orientiert denken, daß für ihre Ortsvektoren $u_j = \overrightarrow{0P_j}$ die Beziehung

$$D\,(u_1,\, u_2,\, u_3) \, > \, 0$$

gilt. Wir definieren dann die Vektoren $v_1,\, v_2,\, v_3$ durch

$$\begin{aligned}u_1 \times u_2 \, &= \, ||\, u_1 \times u_2 \,||\, v_3 \; , \\ u_2 \times u_3 \, &= \, ||\, u_2 \times u_3 \,||\, v_1 \; , \\ u_3 \times u_1 \, &= \, ||\, u_3 \times u_1 \,||\, v_2 \; .\end{aligned}$$

Die v_j sind Einheitsvektoren, die wegen

$$\begin{aligned}D\,(u_2 \times u_3,\, u_3 \times u_1,\, u_1 \times u_2) \, &= \, \langle\, u_2 \times u_3,\, (u_3 \times u_1) \times (u_1 \times u_2) \,\rangle \\ &= \, \langle\, u_2 \times u_3,\, D\,(u_3,\, u_1,\, u_2)\, u_1 \,\rangle \\ &= \, D\,(u_1,\, u_2,\, u_3)^2 \, > \, 0\end{aligned}$$

ein *positiv orientiertes* Koordinatensystem bilden. Außerdem folgt aus

$$\|\, u_2 \times u_3 \,\|\, \langle\, u_1,\, v_1 \,\rangle \, = \, \langle\, u_1,\, u_2 \times u_3 \,\rangle \, = \, D\,(u_1,\, u_2,\, u_3) \, > \, 0$$

auch

$$\langle\, u_1,\, v_1 \,\rangle \, > \, 0$$

und entsprechend für die anderen Indizes.

Die Längen der Seiten bestimmen sich als Bogenlängen der Winkel zwischen den Ortsvektoren. Man setze also a_1 gleich dem Bogenmaß des Winkels zwischen $\overrightarrow{0P_2}$ und $\overrightarrow{0P_3}$ und entsprechend zyklisch vertauscht für die anderen Koordinaten. Dann gilt

$$\langle u_2, u_3 \rangle = \| u_2 \| \cdot \| u_3 \| \cos a_1 \quad \text{etc.}$$

und damit

$$\begin{cases} \cos a_1 = \langle u_2, u_3 \rangle \ , \\ \cos a_2 = \langle u_3, u_1 \rangle \ , \\ \cos a_3 = \langle u_1, u_2 \rangle \ . \end{cases}$$

Ebenso bestimmt man die Winkel α_j zwischen den Seiten als die Winkel zwischen den (Loten auf die) Ebenen, die diese Seiten ausschneiden. Wegen $\langle u_1, v_1 \rangle > 0$ weist das Lot v_1 in die Richtung der Halbkugel, auf der P_1 liegt. Somit ergibt sich

$$\begin{cases} \cos \alpha_3 = -\langle v_1, v_2 \rangle \ , \\ \cos \alpha_2 = -\langle v_3, v_1 \rangle \ , \\ \cos \alpha_1 = -\langle v_2, v_3 \rangle \ . \end{cases}$$

Aus der Definition des Vektorproduktes folgt weiter:

$$\sin a_1 = \| u_2 \times u_3 \| \ , \quad \text{etc.}$$
$$\sin \alpha_1 = \| v_2 \times v_3 \| \ , \quad \text{etc.}$$

Nun steht $v_1 \times v_2$ senkrecht auf v_1 und v_2, und es gilt $\langle v_1 \times v_2, v_3 \rangle = D(v_1, v_2, v_3) > 0$. u_3 steht ebenfalls senkrecht auf v_1 und v_2, und $\langle u_3, v_3 \rangle > 0$. Damit ergibt sich

$$v_1 \times v_2 = \| v_1 \times v_2 \| \, u_3 \ ,$$

und entsprechend für die anderen Indizes.

Wir sind jetzt in der Lage, die drei Hauptsätze der sphärischen Trigonometrie, die Beziehungen zwischen trigonometrischen Funktionen der Seiten und Winkel in einem sphärischen Dreieck zum Ausdruck bringen, im Schnellverfahren abzuleiten.

Satz 15.6 (Winkelcosinussatz)

$$\cos \alpha_2 \cdot \cos \alpha_3 + \cos \alpha_1 = \sin \alpha_2 \cdot \sin \alpha_3 \cdot \cos a_1 \ .$$

Beweis. Es gilt

$$\begin{aligned} \cos \alpha_2 \cdot \cos \alpha_3 + \cos \alpha_1 &= \begin{vmatrix} \langle v_3, v_1 \rangle & \langle v_3, v_2 \rangle \\ \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle \end{vmatrix} \\ &= \langle (v_3 \times v_1), (v_1 \times v_2) \rangle \\ &= \| v_3 \times v_1 \| \cdot \| v_1 \times v_2 \| \, \langle u_2, u_3 \rangle \\ &= \sin \alpha_2 \cdot \sin \alpha_3 \cdot \cos a_1 \ . \end{aligned}$$

Entsprechend erhält man durch Berechnung von $\langle u_3 \times u_1, u_1 \times u_2 \rangle$ den

Satz 15.7 (Seitencosinussatz)

$$\cos a_1 = \cos a_2 \cdot \cos a_3 + \sin a_2 \cdot \sin a_3 \cdot \cos \alpha_1 \ .$$

Weiter ist

$$\begin{aligned} D\,(u_3,\, u_1,\, u_2)\, u_1 &= (u_3 \times u_1) \times (u_1 \times u_2) \\ &= \|\, u_3 \times u_1 \,\| \cdot \|\, u_1 \times u_2 \,\| \,\cdot\, v_2 \times v_3 \\ &= \sin\, a_2 \cdot \sin\, a_3 \cdot \sin\, \alpha_1 \cdot u_1 \; . \end{aligned}$$

Wegen $u_1 \neq 0$ folgt hieraus

$$D\,(u_3,\, u_1,\, u_2) = \sin\, a_2 \cdot \sin\, a_3 \cdot \sin\, \alpha_1 \; .$$

Durch zyklische Vertauschung erhält man entsprechend

$$D\,(u_1,\, u_2,\, u_3) = \sin\, a_3 \cdot \sin\, a_1 \cdot \sin\, \alpha_2 \; , \quad D\,(u_2,\, u_3,\, u_1) = \sin\, a_1 \cdot \sin\, a_2 \cdot \sin\, \alpha_3 \; ,$$

und wegen der Gleichheit der linken Seiten folgt schließlich

Satz 15.8 (Sinussatz)

$$\frac{\sin\, \alpha_1}{\sin\, a_1} = \frac{\sin\, \alpha_2}{\sin\, a_2} = \frac{\sin\, \alpha_3}{\sin\, a_3} \; .$$

Aufgaben

Einleitung

Es bezeichne in den ersten beiden Aufgaben $\mathfrak{P}(M)$ stets die *Potenzmenge* einer Menge M, d.h. die Menge aller Teilmengen A von M (einschließlich der leeren Menge $\emptyset$).

1. Man zeige: Die Abbildung

$$\begin{cases} \mathfrak{P}(M \cup N) \longrightarrow \mathfrak{P}(M) \times \mathfrak{P}(N) \\ \quad A \quad \longmapsto (\, A \cap M \,, \, A \cap N \,) \end{cases}$$

ist bijektiv, falls $M \cap N = \emptyset$. Bleibt diese Aussage richtig, wenn die Voraussetzung $M \cap N = \emptyset$ nicht erfüllt ist?

2. Eine Menge M heißt *endlich*, wenn sie bijektiv auf einen Abschnitt

$$\{j \in \mathbb{N} : 1 \le j \le n\}$$

der natürlichen Zahlen abgebildet werden kann. (Man zeigt leicht durch Induktion, daß dann die Zahl n durch M eindeutig festgelegt ist; wir nennen sie die *Anzahl* oder *Mächtigkeit* oder *Kardinalität* von M, in Zeichen: $n = \mathrm{card}\, M$). Man zeige durch vollständige Induktion: Ist $\mathrm{card}\, M = n$ endlich, so gilt

$$\mathrm{card}\, \mathfrak{P}(M) = 2^n .$$

Kapitel 1

3. Es sei $\mathbb{K}$ ein beliebiger Körper. Auf $\mathbb{K}^2$ erkläre man eine Addition $+$ und eine Multiplikation $\cdot$ vermöge

$$(*) \qquad (a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2) ,$$

$$(**) \qquad (a_1, b_1) \cdot (a_2, b_2) = (a_1 \cdot a_2, b_1 \cdot b_2) .$$

Welche der Körperaxiome sind für $(\mathbb{K}^2, +, \cdot)$ erfüllt, welche nicht ?

4. Erklärt man zur Addition $+$ wie in $(*)$ von Aufgabe 3 eine Multiplikation auf $\mathbb{K}^2$ vermöge

$$(***) \qquad (a_1, b_1) \cdot (a_2, b_2) = (a_1 a_2 - b_1 b_2 \,, \, a_1 b_2 + b_1 a_2) ,$$

so wird $(\mathbb{K}^2, +, \cdot)$ genau dann zu einem Körper, wenn $\mathbb{K}$ der folgenden Bedingung genügt:

$$\text{Für alle } (a, b) \in \mathbb{K}^2 \text{ gilt : } \quad (a^2 + b^2 = 0 \iff a = b = 0) .$$

5. Es bezeichne $\mathbb{K}_0$ die Menge der reellen Zahlen der Form

$$r = \alpha + \beta\sqrt{2} \,, \quad \alpha \,, \beta \in \mathbb{Q} .$$

Man zeige: $\mathbb{K}_0$ ist unter der Addition und Multiplikation in $\mathbb{R}$ abgeschlossen (d. h. $r_1 \,, r_2 \in \mathbb{K}_0 \Longrightarrow r_1 + r_2 \in \mathbb{K}_0$ und $r_1 r_2 \in \mathbb{K}_0$) und bildet mit diesen beiden Verknüpfungen einen Körper. (Man bezeichnet den Körper $\mathbb{K}_0$ wegen seines Bildungsgesetzes auch mit $\mathbb{K}_0 = \mathbb{Q}(\sqrt{2})$).

6. Es sei m eine positive ganze Zahl, $m \ge 2$. $\mathbb{Z}_m$ sei die Menge der nichtnegativen ganzen Zahlen, die kleiner als m sind:

$$\mathbb{Z}_m := \{0, 1, 2, \dots , m - 1\} .$$

Für $\alpha\,,\ \beta \in \mathbb{Z}_m$ sei $\gamma := \alpha \oplus \beta \in \mathbb{Z}_m$ die eindeutig bestimmte Zahl $\gamma \in \mathbb{Z}_m$ mit m teilt $\alpha + \beta - \gamma$ in $\mathbb{Z}$ (hier bezeichnet $+$ die übliche Addition in $\mathbb{Z}$). Entsprechend definiere man eine Multiplikation. Man zeige: $\mathbb{Z}_m$ ist ein Körper genau dann, wenn m eine Primzahl ist. Was ist -1 in $\mathbb{Z}_5$? Was ist $\frac{1}{3}$ in $\mathbb{Z}_7$?

Hinweis: Aus dem *euklidischen Algorithmus* folgt für zwei beliebige Zahlen $p\,,\ q \in \mathbb{N}^* = \mathbb{N} \setminus \{0\}$ die Existenz von Elementen $a\,,\ b \in \mathbb{Z}$, so daß $ap + bq$ der größte gemeinsame Teiler von p und q ist.

Kapitel 2

7. Man betrachte in $\mathbb{K}^3 = \{(x_1,\, x_2,\, x_3) : x_i \in \mathbb{K}\}$ die Mengen

$$\begin{aligned} V_1 &= \{(x_1,\, x_2,\, x_3) \in \mathbb{K}^3 : \quad x_1 = 0\}\,, \\ V_2 &= \{(x_1,\, x_2,\, x_3) \in \mathbb{K}^3 : \quad x_1 = 0 \quad \text{oder} \quad x_2 = 0\}\,, \\ V_3 &= \{(x_1,\, x_2,\, x_3) \in \mathbb{K}^3 : \quad x_1 + x_2 + x_3 = 0\}\,, \\ V_4 &= \{(x_1,\, x_2,\, x_3) \in \mathbb{K}^3 : \quad x_1 + x_2 = 1\}\,. \end{aligned}$$

Welche dieser Mengen sind $\mathbb{K}$–Untervektorräume von $\mathbb{K}^3$?

8. Es sei $\mathbb{R}\,[\,t\,]$ der Vektorraum aller Polynome

$$P = a_0 + a_1\,t + \cdots + a_n\,t^n\,, \quad a_0, \ldots, a_n \in \mathbb{R}\,, \quad n \in \mathbb{N}$$

mit reellen Koeffizienten. Welche der folgenden Mengen sind $\mathbb{R}$–Untervektorräume von $\mathbb{R}\,[\,t\,]$?

$$\begin{aligned} W_1 &= \{P \in \mathbb{R}\,[\,t\,] : \deg P \le 3\}\,, \\ W_2 &= \{P \in \mathbb{R}\,[\,t\,] : \deg P = 3\}\,, \\ W_3 &= \{P \in \mathbb{R}\,[\,t\,] : 2\,P\,(0) = P\,(1)\}\,, \\ W_4 &= \{P \in \mathbb{R}\,[\,t\,] : P\,(t) \ge 0 \quad \text{für} \quad 0 \le t \le 1\}\,, \\ W_5 &= \{P \in \mathbb{R}\,[\,t\,] : P\,(-t) = -P\,(t) \quad \text{für alle} \quad t \in \mathbb{R}\}\,. \end{aligned}$$

9. Man beweise, daß die vier Vektoren

$$(1,\, 0,\, 0)\,, \quad (0,\, 1,\, 0)\,, \quad (0,\, 0,\, 1) \quad \text{und} \quad (1,\, 1,\, 1)$$

in $\mathbb{R}^3$ linear abhängig sind, je drei von ihnen aber stets linear unabhängig.

10. Man gebe eine notwendige und hinreichende Bedingung an die drei Skalare $x_1,\, x_2,\, x_3 \in \mathbb{K}$ dafür an, daß die drei Vektoren $(1,\, x_i,\, x_i^2)\,,\ i = 1, 2, 3\,,$ linear unabhängig in $\mathbb{K}^3$ sind.

11. a) Man finde zwei Basen von $\mathbb{R}^4$, denen nur die Vektoren $(1,\, 1,\, 0,\, 0)$ und $(0,\, 0,\, 1,\, 1)$ gemeinsam sind.

b) Man finde zwei disjunkte Basen von $\mathbb{R}^4$, von denen die erste die Vektoren $(1,\, 0,\, 0,\, 0)$ und $(1,\, 1,\, 0,\, 0)$, die zweite die Vektoren $(1,\, 1,\, 1,\, 0)$ und $(1,\, 1,\, 1,\, 1)$ enthält.

12. a) Unter welcher Bedingung an $c \in \mathbb{R}$ sind die drei Vektoren $(c,\, 1,\, 0)\,,\ (1,\, c,\, 1)$ und $(0,\, 1,\, c)$ linear abhängig in $\mathbb{R}^3$?

b) Wie lautet die Antwort zu $a)$, wenn man den reellen Zahlkörper $\mathbb{R}$ durch $\mathbb{Q}$ ersetzt ?

c) Sei $\mathbb{K}$ ein beliebiger Körper. Unter welcher Bedingung an $c \in \mathbb{K}$ sind die Vektoren $(1 + c,\, 1 - c)$

und $(1 - c, 1 + c)$ in $\mathbb{K}^2$ linear abhängig ?

13. Man zeige: Wird der $\mathbb{K}$–Vektorraum V von m Elementen erzeugt, so sind n Vektoren $v_1, \ldots, v_n \in V$ stets linear abhängig, wenn $n > m$. Man benutze hierzu (ohne Beweis) die folgende Aussage: Jedes homogene Gleichungssystem

$$\sum_{k=1}^{n} a_{jk}\, x_k = 0\,, \quad j = 1, \ldots, m\,, \quad n > m$$

besitzt mindestens eine nichttriviale Lösung $(c_1, \ldots, c_n) \neq (0, \ldots, 0)$.

14. Eine reelle Zahl x heißt *algebraisch*, wenn es (von x abhängige) Zahlen $n \in \mathbb{N}^*$ und $a_1, \ldots, a_n \in \mathbb{Q}$ gibt, so daß

$$(*) \qquad x^n + a_1\, x^{n-1} + \cdots + a_n = 0\,.$$

(Beispiele: jede rationale Zahl, oder $x = \sqrt{2}$). Man zeige:

a) Ist x algebraisch, so ist der von allen Potenzen $x^0 = 1$, $x^1 = x$, $x^2, \cdots$ erzeugte $\mathbb{Q}$–Untervektorraum $\mathbb{Q}\,[\,x\,]$ von $\mathbb{R}$ endlich erzeugt;

b) ist A eine Teilmenge von $\mathbb{R}$, die $\mathbb{Q}$ umfaßt (also $\mathbb{Q} \subset A \subset \mathbb{R}$) und abgeschlossen ist bzgl. der Addition und Multiplikation in $\mathbb{R}$ (d. h. $a_1, a_2 \in A \Longrightarrow a_1 + a_2 \in A$, $a_1\, a_2 \in A$), so trägt A eine natürliche $\mathbb{Q}$–Vektorraumstruktur. Ist dieser Vektorraum endlich erzeugt, so ist jedes Element $x \in A$ algebraisch.

15. Man zeige, daß für je zwei algebraische Zahlen $x, y \in \mathbb{R}$ der $\mathbb{Q}$–Vektorraum $\mathbb{Q}\,[x, y]$ aller „polynomialen Ausdrücke“

$$\sum_{j,k=0}^{n} a_{jk}\, x^j y^k\,, \quad a_{jk} \in \mathbb{Q}$$

in x und y endlich–dimensional ist, und schließe daraus, daß die Gesamtheit $\overline{\mathbb{Q}}$ *aller* algebraischen Zahlen in $\mathbb{R}$ einen Körper bildet.

Kapitel 3

16. Es sei $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ die durch

$$(x_1, x_2, x_3) \longmapsto (x_2, x_2 - x_3, x_2 - x_1)$$

gegebene lineare Abbildung.

a) Man berechne F^4.

b) Man bestimme die Matrix zu F bzgl. der Standardbasis e_1, e_2, e_3.

c) Man bestimme die Matrix zu F bzgl. der Basis $v_1 = e_2 + e_3$, $v_2 = e_3 + e_1$, $v_3 = e_1 + e_2$.

17. Sei V ein endlich erzeugter Vektorraum über $\mathbb{R}$ und $f : V \to V$ ein Endomorphismus. Man beweise: Wird f bezüglich jeder Basis $\mathcal{B}$ von V durch dieselbe Matrix $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ dargestellt, so existiert ein Element $\lambda \in \mathbb{R}$, so daß $f = \lambda\, \mathrm{id}_V$.

18. Das folgende Diagramm von Vektorräumen und Homomorphismen sei „kommutativ“, d. h. $\varphi_{i-1} \circ f_i = g_i \circ \varphi_i$, $i = 1, 2, 3, 4$, und die beiden Zeilen seien „exakt“, d. h. $\ker f_i = \operatorname{im} f_{i+1}$, $\ker g_i = \operatorname{im} g_{i+1}$ für $i = 1, 2, 3$.

$$\begin{tikzcd} V_4 \arrow[r, "f_4"] \arrow[d, "\varphi_4"] & V_3 \arrow[r, "f_3"] \arrow[d, "\varphi_3"] & V_2 \arrow[r, "f_2"] \arrow[d, "\varphi_2"] & V_1 \arrow[r, "f_1"] \arrow[d, "\varphi_1"] & V_0 \arrow[d, "\varphi_0"] \\ W_4 \arrow[r, "g_4"'] & W_3 \arrow[r, "g_3"'] & W_2 \arrow[r, "g_2"'] & W_1 \arrow[r, "g_1"'] & W_0 \end{tikzcd}$$

Für die „vertikalen“ Homomorphismen mögen die folgenden Bedingungen erfüllt sein: φ_4 ist surjektiv, φ_3 und φ_1 sind Isomorphismen, φ_0 ist injektiv. Man zeige, daß unter diesen Umständen φ_2 ein Isomorphismus ist.

19. Es sei V ein endlich–dimensionaler $\mathbb{K}$–Vektorraum mit Basis $(v_1, \ldots, v_n)$, und V^* bezeichne den Vektorraum $\mathrm{Hom}(V, \mathbb{K})$ aller $\mathbb{K}$–linearen Abbildungen $\lambda : V \longrightarrow \mathbb{K}$. (Man nennt V^* den *Dualraum* von V). Dann existieren (eindeutig bestimmte) Elemente $v_j^* \in V^*$, $j = 1, \ldots, n$, mit

$$v_j^*(v_k) = \delta_{jk} = \begin{cases} 1 \quad , & j = k \, , \\ 0 \quad , & j \neq k \, . \end{cases}$$

Man zeige: $(v_1^*, \ldots, v_n^*)$ ist eine Basis von V^*. [Siehe auch Kapitel 9].

20. Es seien V und W endlich–dimensionale $\mathbb{K}$–Vektorräume mit Basen $(v_1, \ldots, v_n)$ bzw. $(w_1, \ldots, w_m)$, und $F : V \longrightarrow W$ sei eine lineare Abbildung. Man zeige:

a) Ist $\lambda \in W^* = \mathrm{Hom}\,(W, \mathbb{K})$, so ist die Komposition $\lambda \circ F \in V^* = \mathrm{Hom}\,(V, \mathbb{K})$;

b) die Zuordnung $\lambda \mapsto \lambda \circ F$ induziert eine lineare Abbildung $F^* : W^* \longrightarrow V^*$;

c) wird F bzgl. der obigen Basen durch die Matrix A beschrieben, so wird F^* bzgl. der dualen Basen $(w_1^*, \ldots, w_m^*)$ und $(v_1^*, \ldots, v_n^*)$ (siehe Aufgabe 19) durch die transponierte Matrix ${}^tA$ beschrieben. [Siehe auch Kapitel 9].

21. Es sei $F \in \mathrm{End}\, V$ für einen n–dimensionalen $\mathbb{K}$–Vektorraum V mit $F^{n-1} \neq 0$ und $F^n = 0$. Es sei $v_0 \in V$ ein Vektor mit $F^{n-1}(v_0) \neq 0$.

a) Man zeige, daß die n Vektoren $v_0, v_1 = F(v_0), v_2 = F^2(v_0), \ldots, v_{n-1} = F^{n-1}(v_0)$ eine Basis von V bilden.

b) Man bestimme die Matrix von F bzgl. dieser Basis.

22. Es sei $P : \mathbb{R}^2 \to \mathbb{R}^2$ eine lineare Abbildung mit $P^2 = P$. Man zeige:

a) Ist $\mathrm{rang}\, P = 2$, so ist $P = \mathrm{id}$;

b) ist $\mathrm{rang}\, P = 1$, so gibt es eine Basis v_1, v_2 von $\mathbb{R}^2$ mit $P(v_1) = v_1$, $P(v_2) = 0$. Man interpretiere dieses Ergebnis geometrisch.

23. Es sei $A \in M\,(2 \times 2, \mathbb{K})$ geschrieben in der Form

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} .$$

Man zeige:

$$A^2 - (a + d)\,A + (a\,d - b\,c)\,E = 0 \, .$$

24. Es seien A, $B \in M\,(n \times n, \mathbb{K})$ quadratische Matrizen, so daß A mit $A\,B - B\,A$ vertauscht. Man zeige:

$$A^k B - B\,A^k = k\,A^{k-1}\,(A\,B - B\,A) \quad \text{für alle} \quad k \geq 1 \, .$$

Kapitel 4

25. Es sei V ein $\mathbb{K}$–Vektorraum und $P \in \operatorname{End} V$ eine *Projektion*, d. h. eine lineare Abbildung $P : V \longrightarrow V$ mit $P^2 = P$. Man zeige: $v \in \operatorname{im} P$ genau dann, wenn $P(v) = v$, und schließe hieraus:

$$V = \ker P \oplus \operatorname{im} P .$$

(Dies verallgemeinert Aufgabe 22). Wie sieht P bzgl. dieser Zerlegung von V aus ?

26. Es seien u_1, u_2, v_1, v_2 Vektoren in $\mathbb{R}^4$. Setze $U = \operatorname{span}(u_1, u_2)$ und $V = \operatorname{span}(v_1, v_2)$. Wann ist in den folgenden Fällen $\mathbb{R}^4 = U \oplus V$?

a) $u_1 = (1, 1, 0, 0)$, $u_2 = (1, 0, 1, 0)$, $v_1 = (0, 1, 0, 1)$, $v_2 = (0, 0, 1, 1)$;

b) $u_1 = (-1, 1, 1, 0)$, $u_2 = (0, 1, -1, 1)$, $v_1 = (1, 0, 0, 0)$, $v_2 = (0, 0, 0, 1)$;

c) $u_1 = (1, 0, 0, 1)$, $u_2 = (0, 1, 1, 0)$, $v_1 = (1, 0, 1, 0)$, $v_2 = (0, 1, 0, 1)$.

27. Es sei $(v_\iota)_{\iota \in I}$ eine nicht leere Familie von Vektoren $v_\iota \neq 0$ in V. Man zeige: $(v_\iota)_{\iota \in I}$ ist genau dann eine Basis von V, wenn

$$V = \operatorname{span}(v_\iota)_{\iota \in J} \oplus \operatorname{span}(v_\kappa)_{\kappa \in K}$$

für alle Zerlegungen $I = J \cup K$, $J \cap K = \emptyset$, von I.

28. Es seien $W_1, \ldots, W_k$ Untervektorräume von V, die zusammen V erzeugen, d. h. $V = W_1 + \cdots + W_k$. Man zeige die Äquivalenz der folgenden Aussagen:

a) Die Darstellung $v = w_1 + \cdots + w_k$ der Vektoren $v \in V$ durch Vektoren $w_j \in W_j$, $j = 1, \ldots, k$, ist eindeutig;

b) $\displaystyle\sum_{j=1}^{k} w_j = 0 , \quad w_j \in W_j$ impliziert $w_j = 0$ für alle $j = 1, \ldots, k$;

c) $W_j \cap (W_1 + \cdots + W_{j-1}) = \{0\}$, $\quad j = 2, \ldots, k$;

d) $W_j \cap (W_1 + \cdots + W_{j-1} + W_{j+1} + \cdots + W_k) = \{0\}$, $\quad j = 1, \ldots, k$.

Ist überdies V endlich–dimensional, so sind diese Aussagen äquivalent zu

e)
$$\sum_{j=1}^{k} \dim W_j = \dim V .$$

Kapitel 5

29. Man definiere auf $\mathbb{N} \times \mathbb{N}$, wobei $\mathbb{N} = \{0, 1, 2, \ldots\}$, eine Relation $\sim$ durch

$$(a_1, b_1) \sim (a_2, b_2) \Longleftrightarrow a_1 + b_2 = a_2 + b_1$$

und zeige, daß $\sim$ eine Äquivalenzrelation ist. Man erkläre auf der Quotientenmenge

$$Z = \mathbb{N} \times \mathbb{N} / \sim$$

eine Addition und Multiplikation und eine injektive Abbildung

$$j : Z \hookrightarrow \mathbb{R} ,$$

s. d. $j(Z) = \mathbb{Z}$ und die algebraischen Operationen auf Z und $\mathbb{Z}$ vermöge j „übereinstimmen“.

30. Man betrachte die Menge $\mathcal{C}$ aller Folgen (q_j) mit rationalen Gliedern $q_j \in \mathbb{Q}$, die (als Folgen reeller Zahlen) die Cauchy–Eigenschaft besitzen. Entsprechend sei $\mathcal{N} \subset \mathcal{C}$ die Menge der Nullfolgen. Man nenne zwei Folgen $(q_j)_j$, $(q'_j)_j$ in $\mathcal{C}$ äquivalent, falls $(q_j - q'_j)_j \in \mathcal{N}$, und zeige, daß hierdurch tatsächlich eine Äquivalenzrelation $\sim$ auf $\mathcal{C}$ definiert wird. Man erkläre auf der Quotientenmenge $\mathcal{R} = \mathcal{C}/\sim$ eine Addition und eine Multiplikation und eine Bijektion $\ell : \mathcal{R} \to \mathbb{R}$, die mit den Ringstrukturen verträglich ist.

Kapitel 6

31. Man löse das folgende lineare Gleichungssystem mit Koeffizienten in $\mathbb{R}$:

$$\begin{aligned} x_1 &+ 2x_2 + 3x_3 = 1 , \\ 4x_1 &+ 5x_2 + 6x_3 = 2 , \\ 7x_1 &+ 8x_2 + 9x_3 = 3 , \\ 5x_1 &+ 7x_2 + 9x_3 = 4 . \end{aligned}$$

32. Man beweise die folgende Aussage: Ist $U \subset \mathbb{K}^n$ ein Untervektorraum und $x_0 \in \mathbb{K}^n$, so gibt es ein Gleichungssystem mit Koeffizienten in $\mathbb{K}$ mit n Gleichungen für n Unbekannte, dessen Lösungsraum genau die affine Menge $x_0 + U$ ist.

33. Es sei g eine Gerade im zweidimensionalen reellen affinen Raum $\mathbb{A}^2$, und a, b, p seien paarweise verschiedene Punkte von g. Dann gibt es eine eindeutig bestimmte reelle Zahl $\lambda \neq 0$ mit

$$\overrightarrow{a\,p} = \lambda \ \overrightarrow{p\,b} \ .$$

Wir sagen, daß *die Strecke von a nach b durch den Punkt p im Verhältnis $\lambda =: \lambda(a, b; p)$ geteilt wird*. Es sei $c \notin g$ ein weiterer Punkt, q liege auf der Geraden durch b und c, der Punkt r auf der Geraden durch c und a, wobei $q \neq b, c$ und $r \neq c, a$. Man zeige:

a) p, q, r liegen genau dann auf einer Geraden, wenn

$$\lambda(a, b; p)\,\lambda(b, c; q)\,\lambda(c, a; r) = -1 .$$

b) Was bedeutet es geometrisch, wenn dieses Produkt der Teilungsverhältnisse gleich $+1$ ist ? (Beweis !)

Kapitel 7

34. a) Man bestimme (durch Modifikation der Methode zur Bestimmung des *Zeilenranges* von Matrizen durch elementare Zeilenumformung) den *Spaltenrang* der Koeffizienten–Matrix und der erweiterten Koeffizienten–Matrix des linearen Gleichungssystems in Aufgabe 31.

b) Nach der gleichen Methode wie in Teil a) untersuche man, ob das Gleichungssystem

$$\begin{array}{rcrcrcrcr} x_1 & - & x_2 & + & 2x_3 & - & 3x_4 & = & 7 \\ 4x_1 & & & + & 3x_3 & + & x_4 & = & 9 \\ 2x_1 & - & 5x_2 & + & x_3 & & & = & -2 \\ 3x_1 & - & x_2 & - & x_3 & + & 2x_4 & = & -2 \end{array}$$

lösbar ist, und bestimme gegebenenfalls die Lösungsmenge.

35. Es sei $\mathbb{R}[x]$ der $\mathbb{R}$–Vektorraum der Polynome mit reellen Koeffizienten. Für

$$P = \sum_{j=0}^{n} a_j\, x^j$$

definiere man

$$D(P) := \sum_{j=1}^{n} j a_j \, x^{j-1} \,, \quad I(P) = \sum_{j=0}^{n} \frac{a_j}{j+1} \, x^{j+1} \,.$$

Man zeige: D, $I \in \operatorname{End}(\mathbb{R}[x])$, D ist surjektiv, aber nicht injektiv, I ist injektiv, aber nicht surjektiv, D ist Linksinverses zu I (d. h. $D \circ I = \mathrm{id}_{\mathbb{R}[x]}$).

36. Es sei θ eine reelle Zahl. Man beweise, daß die beiden Matrizen

$$A_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{und} \quad B_\theta = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

über dem Körper $\mathbb{C}$ der komplexen Zahlen ähnlich sind, d. h. daß es eine invertierbare Matrix C_θ gibt mit $B_\theta = C_\theta^{-1} A_\theta C_\theta$.

37. Es sei $F \in \operatorname{End}\mathbb{C}^2$ der Endomorphismus ${}^t(x_1, x_2) \mapsto {}^t(x_1, 0)$. Ferner sei $\mathcal{K}$ die kanonische geordnete Basis $({}^t(1,0), {}^t(0,1))$ und $\mathcal{A}$ die geordnete Basis mit den Vektoren $({}^t(1, i), {}^t(-i, 2))$. Man bestimme die Matrizen $M_{\mathcal{K}}^{\mathcal{K}}(F)$, $M_{\mathcal{A}}^{\mathcal{K}}(F)$, $M_{\mathcal{K}}^{\mathcal{A}}(F)$, $M_{\mathcal{A}}^{\mathcal{A}}(F)$ und $M_{\mathcal{A}'}^{\mathcal{A}'}(F)$, wobei $\mathcal{A}'$ aus $\mathcal{A}$ durch Vertauschung der beiden Basisvektoren entsteht.

38. Es sei

$$A = \begin{pmatrix} 3 & 21 & 0 & 9 & 0 \\ 1 & 7 & -1 & -2 & -1 \\ 2 & 14 & 0 & 6 & 1 \\ 6 & 42 & -1 & 13 & 0 \end{pmatrix} \in M(4 \times 5, \mathbb{R}) \,.$$

Man bestimme eine Basis des Spaltenraumes von A.

39. Es sei $A \in M(n \times n, \mathbb{K})$. Man zeige:

a) Ist A invertierbar, so folgt aus $AB = 0$, $B \in M(n \times n, \mathbb{K})$ das Verschwinden der Matrix B, d. h. $B = 0$.

b) Ist A nicht invertierbar, so gibt es ein $B \in M(n \times n, \mathbb{K})$, $B \neq 0$, mit $AB = 0$.

c) Gilt $A^2 - A + E = 0$, so ist A invertierbar.

40. Es sei V ein endlich–dimensionaler $\mathbb{K}$–Vektorraum mit geordneter Basis $\mathcal{A} = (v_1, \ldots, v_n)$. Man zeige: Zu jeder invertierbaren Matrix $A \in M(n \times n, \mathbb{K})$ und jedem Automorphismus $F \in \operatorname{Aut} V$ gibt es eine Basis $\mathcal{B} = (w_1, \ldots, w_n)$ von V mit $M_{\mathcal{B}}^{\mathcal{A}}(F) = A$.

41. Man löse das lineare Gleichungssystem $Ax = b$ für

$$A = \begin{pmatrix} 1 & 2 & 0 & 3 & 0 \\ 1 & 2 & -1 & -1 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 2 & 4 & 1 & 10 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \in M(5 \times 5, \mathbb{R}) \quad , \quad b = \begin{pmatrix} 2 \\ 6 \\ -4 \\ 2 \\ 2 \end{pmatrix} \in \mathbb{R}^5 \,.$$

42. Man bestimme zu den folgenden Matrizen die Inversen (falls existent):

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix} , \quad \begin{pmatrix} 2 & 5 & -1 \\ 4 & -1 & 2 \\ 6 & 4 & 1 \end{pmatrix} .$$

43. Man beweise die folgenden Aussagen:

a) Ist $A \in M(m \times n, \mathbb{K})$, $B \in M(n \times m, \mathbb{K})$, $n < m$, so ist AB nicht invertierbar.

b) Sind F, G Endomorphismen eines (nicht notwendig endlich–dimensionalen) Vektorraums V, so sind F und G genau dann Automorphismen, wenn $F \circ G$ und $G \circ F$ Automorphismen sind.

44. Es sei $A \in M(m \times n, \mathbb{K})$. Eine Matrix $B \in M(n \times m, \mathbb{K})$ heißt ein *Pseudoinverses* zu A, wenn gilt: $ABA = A$, $BAB = B$. Man zeige:

i) Ist $m = n$ und A invertierbar, so ist $B = A^{-1}$ das einzige Pseudoinverse von A;

ii) sind $S \in \mathrm{GL}(m, \mathbb{K})$, $T \in \mathrm{GL}(n, \mathbb{K})$, und ist B Pseudoinverses zu A, so ist $T^{-1}BS^{-1}$ Pseudoinverses zu SAT;

iii) jede Matrix A besitzt mindestens ein Pseudoinverses.

(Hinweis: Man bringe A in „Normalform“).

Kapitel 8

45. Es sei $V = M(n \times n, \mathbb{K})$ der $\mathbb{K}$–Vektorraum der quadratischen n–reihigen Matrizen über dem Körper $\mathbb{K}$. Für $C \in V$, $C = (c_{jk})_{1 \le j,k \le n}$, bezeichne $\mathrm{spur}\, C$ die Summe der Hauptdiagonalelemente: $\sum_{j=1}^{n} c_{jj}$ („Spur von C“).

a) Ist $B \in M(n \times n, \mathbb{K})$ eine fest gewählte Matrix, so zeige man, daß durch

$$f_B(A) := \mathrm{spur}\,(B \cdot {}^t A)$$

eine lineare Abbildung $f_B : V \to \mathbb{K}$ definiert wird.

b) Die Zuordnung $B \mapsto f_B$ induziert einen Isomorphismus $V \to V^*$. (Insbesondere ist jede Linearform auf V von dieser Gestalt, d. h. : für alle $f \in V^*$ existiert ein $B \in V$ mit $f = f_B$.)

46. a) Man berechne die Determinanten der folgenden Matrizen (mit Einträgen in $\mathbb{R}$):

$$\begin{pmatrix} 1 & -1 & 2 & 3 \\ 2 & 2 & 0 & 2 \\ 4 & 1 & -1 & -1 \\ 1 & 2 & 3 & 0 \end{pmatrix} , \quad \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} .$$

b) Man benutze die Cramersche Regel zur Lösung der Gleichungssysteme

$$\left\{ \begin{array}{rcrcrcr} x & + & y & + & z & = & 11 \\ 2x & - & 6y & - & z & = & 0 \\ 3x & + & 4y & + & 2z & = & 0 \end{array} \right. \quad \text{bzw.} \quad \left\{ \begin{array}{rcrcr} 3x & - & 2y & = & 7 \\ 3y & - & 2z & = & 6 \\ 3z & - & 2x & = & -1 \end{array} \right. .$$

47. Es sei $A \in M(2 \times 3, \mathbb{K})$, $\quad A = (a_{jk})_{\substack{j=1,2 \\ k=1,2,3}}$ und

$$c_1 = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} , \quad c_2 = \begin{vmatrix} a_{13} & a_{11} \\ a_{23} & a_{21} \end{vmatrix} , \quad c_3 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} .$$

Man zeige: Ist $\operatorname{rang} A = 2$, so ist der Vektor ${}^t(c_1, c_2, c_3)$ eine Basis des Lösungsraumes des homogenen Gleichungssystems $A\,x = 0$.

48. Es sei $D : M\,(2 \times 2, \mathbb{R}) \longrightarrow \mathbb{R}$ eine Funktion mit den folgenden Eigenschaften:

$$D\,(A\,B) = D\,(A) \cdot D\,(B)\,, \quad D\,(S) \neq D\,(E_2)\,,$$

wobei $A, B \in M\,(2 \times 2, \mathbb{R})$ beliebig sind und S die Matrix

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

bezeichnet. Man beweise die folgenden Aussagen:

a) $D\,(E_2) = 1\,,\; D\,(S) = -1\,,\; D\,(0) = 0\,;$

b) $D\,(B) = -D\,(A)$, wenn B aus A durch Spalten– oder Zeilenvertauschung entsteht;

c) $D\,(A) = 0$ genau dann, wenn $\det A = 0$.

Man gebe eine solche Funktion D an, die von $\det$ verschieden ist.

49. Es seien $A, B, C, D \in M\,(n \times n, \mathbb{K})$, s. d. C und D vertauschen und D invertierbar ist. Man zeige:

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det\,(A\,D - B\,C)\,.$$

Bleibt diese Beziehung auch bestehen, wenn C und D nicht vertauschbar oder D nicht invertierbar ist ?

Kapitel 9

50. In $\mathbb{R}^3$ seien $\alpha_1 = (1, 0, 1)$, $\alpha_2 = (0, 1, -2)$ und $\alpha_3 = (-1, -1, 0)$ gegeben.

a) Ist λ eine Linearform auf $\mathbb{R}^3$ mit

$$\lambda(\alpha_1) = 1\,, \quad \lambda(\alpha_2) = -1\,, \quad \lambda(\alpha_3) = 3\,,$$

so bestimme man $\lambda(\alpha)$ für $\alpha = (a, b, c) \in \mathbb{R}^3$.

b) Man bestimme alle Linearformen $\mu \in (\mathbb{R}^3)^*$ mit $\mu(\alpha_1) = \mu(\alpha_2) = 0, \mu(\alpha_3) \neq 0$ und zeige für diese $\mu((2, 3, -1)) \neq 0$.

51. Es sei V der Vektorraum aller Polynome über dem reellen Grundkörper $\mathbb{R}$ vom Grad ≤ 2, und t_1, t_2, t_3 seien drei paarweise verschiedene reelle Zahlen. Man definiere $L_i : V \to \mathbb{R}$, $i = 1, 2, 3$, durch $L_i(P) = P\,(t_i)$, $i = 1, 2, 3$, und zeige:

(i) $L_i \in V^*$, $i = 1, 2, 3$;

(ii) $\{L_1, L_2, L_3\}$ ist eine Basis von V^*.

(iii) Man bestimme eine Basis $\{P_1, P_2, P_3\}$ von V mit $L_i(P_j) = \delta_{ij}$ für alle $1 \leq i, j \leq 3$ und schreibe jedes beliebige Polynom $P \in V$ als Linearkombination der P_j.

52. Es sei V wie in der vorherstehenden Aufgabe. Man definiere für $P \in V$ die drei Funktionale

$$\lambda_1(P) = \int_0^1 P\,(x)\,dx\,, \quad \lambda_2(P) = \int_0^2 P\,(x)\,dx\,, \quad \lambda_3(P) = -\int_{-1}^0 P\,(x)\,dx$$

und zeige, daß $\Lambda = \{\lambda_1, \lambda_2, \lambda_3\}$ eine Basis von V^* bildet, indem man z. B. Polynome $\{P_1, P_2, P_3\}$ in V bestimmt, für die $\lambda_j(P_k) = \delta_{jk}$, $1 \leq j, k \leq 3$ gilt.

Kapitel 10

53. Es sei V ein euklidischer Vektorraum mit Skalarprodukt $\langle \,.\,, \,.\, \rangle$. Man beweise die folgenden Aussagen:

a) Zwei Vektoren v, w in V sind genau dann orthogonal, wenn $\| v + w \|^2 = \| v \|^2 + \| w \|^2$.

b) Für zwei Vektoren v, w in V mit $\| v \| = \| w \|$ sind stets $v - w$ und $v + w$ orthogonal. (Zeichnung ?)

Für einen unitären Raum V ist a) falsch (Beispiel ?). Man zeige in diesem Fall, daß zumindest die folgende Aussage gilt:

c) Zwei Vektoren v, w in V sind genau dann orthogonal, wenn $\| a v + b w \|^2 = \| a v \|^2 + \| b w \|^2$ für alle $a, b \in \mathbb{C}$ gilt.

54. Man zeige, daß durch

$$\langle \sum_j a_j x^j \, , \, \sum_k b_k x^k \rangle = \sum_{j,k} \frac{a_j b_k}{j + k + 1}$$

ein Skalarprodukt auf $\mathbb{R}[x]$ definiert wird. Für kleine n finde man zu den Polynomen $Q_j(x) = x^j$, $j = 0, \ldots, n - 1$, ein nichttriviales Polynom P_n n–ten Grades, das auf diesen senkrecht steht.

55. Es sei W ein Untervektorraum des euklidischen Vektorraums V, $v_0 \in V$ sei fest vorgegeben. Nach Definition heißt ein Vektor $w_0 \in W$ eine *beste Approximation von* v_0 *bzgl.* W , falls

$$\| v_0 - w_0 \| \leq \| v_0 - w \| \quad \text{für alle} \quad w \in W \, .$$

Man zeige:

i) w_0 ist eine beste Approximation von v_0 bzgl. W genau dann, wenn
$w_0 - v_0 \perp w$ für alle $w \in W$;

ii) wenn eine solche beste Approximation w_0 existiert, so ist sie eindeutig bestimmt;

iii) ist W endlich–dimensional und $\{w_1, \ldots, w_n\}$ eine Basis von W mit der Eigenschaft $\langle w_j, w_k \rangle = \delta_{jk}$ (*orthonormale Basis*), so ist w_0 gegeben durch

$$w_0 = \sum_{j=1}^{n} \langle v_0, w_j \rangle \, w_j \; .$$

56. Man wende das Schmidtsche Orthonormalisierungsverfahren auf die Vektoren

$$w_1 = (3, 0, 4), \; w_2 = (-1, 0, 7), \; w_3 = (2, 9, 11)$$

im $\mathbb{R}^3$ an.

57. Es sei $A x = b$ ein (nicht notwendig lösbares) reelles lineares Gleichungssystem mit m Gleichungen und n Unbestimmten. Man versehe $\mathbb{R}^m$ und $\mathbb{R}^n$ mit dem kanonischen euklidischen Skalarprodukt $\langle \,.\,, \,.\, \rangle$ und beweise die folgende Aussage:

Es gibt genau einen Vektor $x_0 \in \mathbb{R}^n$ minimaler Länge, der die Abstandsfunktion $x \mapsto \| A x - b \|$, $x \in \mathbb{R}^n$, minimiert.

Man finde den Vektor x_0 in dem folgenden Fall:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} .$$

Man stelle fest, daß x_0 in dem Zeilenraum von A enthalten ist, und begründe, warum dies immer der Fall sein muß.

58. Es sei $V = \mathcal{C}(I)$ der reelle Vektorraum aller reell–wertigen stetigen Funktionen auf dem Intervall $I = [-1, 1]$ mit dem euklidischen Skalarprodukt

$$\langle f, g \rangle = \int_{-1}^{1} f(x)\, g(x)\, dx .$$

Man finde das orthogonale Komplement zu dem Untervektorraum $U \subset V$ der *ungeraden* Funktionen h ($h \in U \iff h(-x) = -h(x)$ für alle $x \in I$).

59. Es sei $\| \, . \, \|$ eine Norm auf einem reellen Vektorraum V, welche der Parallelogramm–Gleichung genügt. Man zeige, daß es ein Skalarprodukt $\langle \, . \, , \, . \, \rangle$ auf V gibt mit $\langle v, v \rangle = \| v \|^2$ für alle Vektoren $v \in V$.

(Hinweis: Man betrachte V zunächst nur als $\mathbb{Q}$–Vektorraum. Siehe auch: M. Koecher, Lineare Algebra etc., p. 155).

60. a) Es sei $F : V \longrightarrow W$ eine lineare Abbildung von endlich–dimensionalen $\mathbb{K}$–Vektorräumen, und F^* bezeichne die duale Abbildung $W^* \to V^*$. Man zeige:

F ist injektiv $\iff F^*$ ist surjektiv;

F ist surjektiv $\iff F^*$ ist injektiv.

(Hinweis: $\ker F = (\operatorname{im} F^*)^\perp$, $\operatorname{im} F = (\ker F^*)^\perp$).

Bleiben diese Aussagen auch für beliebige $\mathbb{K}$–Vektorräume richtig?

b) Es seien U_1, U_2 Untervektorräume des endlich–dimensionalen euklidischen (oder unitären) Vektorraumes V. Man zeige:

$$(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp , \quad (U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp .$$

61. Es sei ℓ^2 die Menge der reellen Folgen $a = (a_j)_{j\in\mathbb{N}}$ mit $\displaystyle\sum_{j=0}^{\infty} a_j^2 < \infty$.

a) Man folgere aus der Cauchy–Schwarzschen Ungleichung für $\mathbb{R}^{n+1}$, daß mit $a = (a_j)_{j\in\mathbb{N}}$, $b = (b_j)_{j\in\mathbb{N}}$ in ℓ^2 die Reihe

$$\langle a, b \rangle := \sum_{j=0}^{\infty} a_j\, b_j$$

absolut konvergiert.

b) Aus a) schließe man, daß ℓ^2 bzgl. der üblichen Addition und Multiplikation mit reellen Zahlen einen $\mathbb{R}$–Vektorraum bildet, auf dem $\langle \, . \, , \, . \, \rangle$ ein euklidisches Skalarprodukt definiert.

62. Es sei $\{v_1, \ldots, v_n\}$ ein orthogonales System von Vektoren $v_j \neq 0$ in dem euklidischen Vektorraum V. Man beweise die *Besselsche Ungleichung*:

$$\sum_{j=1}^{n} \frac{\langle v, v_j \rangle^2}{\| v_j \|^2} \leq \| v \|^2 , \quad v \in V .$$

Wann gilt Gleichheit ?

(Hinweis: Betrachte die orthogonale Projektion $\widetilde{v}$ von v in den von $v_1, \ldots, v_n$ erzeugten Unterraum).

63. Es sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^n$. Man zeige: $\|\cdot\|$ kommt genau dann her von einem Skalarprodukt, wenn $\{x \in \mathbb{R}^n : \|x\| = 1\}$ in einem Ellipsoid (mit dem Ursprung als Mittelpunkt) enthalten ist. (Ein solches Ellipsoid ist das Bild der euklidischen Einheitskugel unter einem linearen Koordinatenwechsel).

64. Es sei ein nicht ausgeartetes Dreieck in der euklidischen Ebene gegeben mit dem Schnittpunkt s der Seitenhalbierenden und dem Mittelpunkt m des Umkreises (d. h. dem Schnittpunkt der Mittelsenkrechten). Ferner sei f der Mittelpunkt des Kreises durch die drei Seitenmitten. Man zeige: f liegt auf der Geraden durch m und s (bzw. $f = m = s$, falls $m = s$) und s teilt die Strecke von m nach f im Verhältnis $2:1$.

Kapitel 11

65. Man finde alle (komplexen) Eigenwerte der folgenden Matrizen und entscheide, ob diese diagonalisierbar sind:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & i \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

66. a) Es sei $N \in M(2\times 2, \mathbb{C})$ mit $N^2 = 0$ gegeben. Man zeige, daß entweder N gleich der Nullmatrix oder ähnlich zu der Matrix

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

ist. (Hinweis: Aufg. 21).

b) Man schließe aus a): Jede komplexe 2×2–Matrix ist ähnlich zu einer der beiden folgenden:

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \quad \begin{pmatrix} a & 0 \\ 1 & a \end{pmatrix}, \quad a,\, b \in \mathbb{C}.$$

(Hinweis: Aufg. 23).

67. Es sei $\sigma \in \mathfrak{S}_n$ eine Permutation der Zahlen $1, \ldots, n$. A_σ sei der Automorphismus ${}^t(x_1, \ldots, x_n) \mapsto {}^t(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ von $\mathbb{C}^n$. Man bestimme die Eigenwerte von A_σ. (Hinweis: Zerlege die Permutation σ in elementfremde Zykel).

68. Es seien $A, B \in M(n \times n, \mathbb{K})$. Man zeige: $\chi_{AB} = \chi_{BA}$.

(Hinweis: Man kann folgendermaßen vorgehen:

i) Ist $A = P$ eine Projektion ($P^2 = P$), so kann man P in eine geeignete Normalform bringen, aus der leicht $\chi_{PB} = \chi_{BP}$ folgt.

ii) Ist A beliebig, so gibt es ein invertierbares Q, s. d. $QA = P$ eine Projektion ist. Man wende dann i) auf $P = QA$ und BQ^{-1} an.
Eine andere Möglichkeit wird in F. Lorenz, *Lineare Algebra II*, Aufgabe 71, p.184 angedeutet.)

69. Es sei $V = \mathcal{C}^0(\mathbb{R}, \mathbb{R})$ der Raum der stetigen Funktionen $f : \mathbb{R} \longrightarrow \mathbb{R}$, und $T \in \operatorname{End} V$ sei definiert durch

$$(Tf)(x) = \int_0^x f(t)\, dt.$$

Man zeige: T besitzt keine Eigenwerte.

70. Es sei A eine Diagonalmatrix in $M(n \times n, \mathbb{K})$ mit charakteristischem Polynom

$$(x - c_1)^{d_1} \cdot \ldots \cdot (x - c_k)^{d_k} ,$$

$c_1, \ldots, c_k$ paarweise verschieden. Man betrachte den Vektorraum

$$V = \{B \in M(n \times n, \mathbb{K}) : BA = AB\}$$

und zeige:

$$\dim_{\mathbb{K}} V = d_1^2 + \cdots + d_k^2 .$$

71. Es sei $W = M(n \times n, \mathbb{K})$ und $W_0 = \{C \in W : \operatorname{Spur} C = 0\}$, wobei $\operatorname{Spur}((c_{jk})_{1 \le j,k \le n}) = \sum_{j=1}^{n} c_{jj}$. Man zeige (evtl. durch Dimensionsüberlegungen), daß $W_0 = \operatorname{span} W_1$, wobei

$$W_1 = \{C \in W : \exists\, A, B \in W, \quad \text{s. d.} \quad C = AB - BA\} .$$

72. Es sei V ein endlich–dimensionaler $\mathbb{K}$–Vektorraum, $F, G \subset \operatorname{End} V$ seien beide trigonalisierbar, und F und G mögen miteinander vertauschen: $F \circ G = G \circ F$. Man zeige: F und G sind *simultan* trigonalisierbar (d. h. es existiert *eine* Basis von V bzgl. der F *und* G obere Dreiecksgestalt besitzen).

(Hinweis: Man kopiere den Beweis von Theorem 5. 4. 4 in Fischer, *Lineare Algebra*, p. 175; zum Induktionsanfang betrachte man G auf $\operatorname{Eig}(f; \lambda)$).

Man überlege sich, ob die entsprechende Aussage für *diagonalisierbare* Endomorphismen richtig bleibt.

Kapitel 12

73. a) Man finde die Jordansche Normalform für die Matrizen

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} , \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & -1/2 & 0 \end{pmatrix} .$$

b) Bestimme die Matrix des Ableitungsoperators D auf dem Vektorraum der komplexen Polynome vom Grade ≤ 3 bzgl. der natürlichen Basis $1, x, x^2, x^3$ und finde die Jordansche Normalform dieser Matrix.

74. Wieviele verschiedene Jordansche Normalformen können die komplexen 6×6–Matrizen annehmen, deren charakteristisches Polynom von der Gestalt

$$\chi_A(t) = (t + 2)^4 (t - 1)^2$$

ist ?

75. Es sei N eine $n \times n$–Matrix über $\mathbb{C}$ mit $N^{n-1} \ne 0$, $N^n = 0$. Man zeige, daß N zu ${}^tN$ ähnlich ist. Man folgere hieraus mit Hilfe der Jordanschen Normalform, daß *jede* komplexe $n \times n$–Matrix ähnlich zu ihrer Transponierten ist.

76. Es sei $\mathbb{K}$ ein Körper der Charakteristik 0 (d. h. $n \cdot 1 \ne 0$ für alle $n \in \mathbb{N} \setminus \{0\}$). Man zeige: Ist A eine $n \times n$–Matrix mit $\operatorname{Spur} A = 0$, so ist A ähnlich zu einer Matrix $\widetilde{A} = (\widetilde{a}_{jk})$ mit

$\widetilde{a}_{jj} = 0\,,\; j = 1, \ldots, n\,.$

(Hinweis: Ist $A \neq \lambda E$, so gibt es ein $x \in \mathbb{K}^n$, s. d. x und $A\,x$ linear unabhängig sind. Danach führe man Induktion nach n).

Warum verschärft diese Aussage Aufgabe 71?

77. a) Ist N eine nilpotente komplexe 3×3–Matrix, so zeige man

$$A := E_3 + N = B^2\,, \quad \text{wobei} \quad B = E_3 + \frac{1}{2}\,N - \frac{1}{8}\,N^2\,.$$

b) Mit Hilfe der Binomialreihe für $(1 + t)^{1/2}$ zeige man, daß auch $A = E_n + N$, N eine nilpotente komplexe $n \times n$–Matrix, eine Quadratwurzel $B \in M\,(n \times n,\, \mathbb{C})$ besitzt.

c) Wann besitzt eine *nichtsinguläre* komplexe $n \times n$–Matrix eine Quadratwurzel ?

Kapitel 13

78. Es sei $V = \mathbb{C}^2$ mit der unitären Standardstruktur versehen, und $F \in \mathrm{End}(V)$ sei bzgl. der kanonischen Basis repräsentiert durch die Matrix

$$A = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}.$$

Man zeige, daß F normal ist, und finde eine orthonormale Basis von V, die aus Eigenvektoren von F besteht.

79. Es sei $F \in \mathrm{End}\, V$, V ein endlich–dimensionaler unitärer Vektorraum. Man zeige:

a) Aus $\|\, F\,(v)\,\| = \|\, F^{ad}(v)\,\|$ für alle $v \in V$ folgt, daß F normal ist, und umgekehrt. (*Hinweis*: Polarisierungsgleichung).

b) Ist F normal und nilpotent, so ist $F = 0$.

c) Ist F normal und idempotent ($F^2 = F$), so ist F selbstadjungiert (d. h. $F^{ad} = F$).

80. Es sei F wie in Aufgabe 79; $\lambda_1, \ldots, \lambda_n$ seien die (nicht notwendig paarweise verschiedenen) Eigenwerte von F, $n = \dim V$. Man zeige:

$$\sum_{j=1}^{n} |\lambda_j|^2 \leq \mathrm{Spur}\,(F^{ad}F)\,.$$

Gleichheit gilt genau dann, wenn F normal ist.

81. Es sei V der Körper der komplexen Zahlen, aufgefaßt als $\mathbb{R}$–Vektorraum.

a) Man zeige, daß durch $\langle\, z,\, w\,\rangle := \mathrm{Re}\,(\overline{z}w)$ ein euklidisches Skalarprodukt auf V erklärt ist.

b) Man finde eine Isometrie von V auf den $\mathbb{R}^2$ mit dem euklidischen Standard–Skalarprodukt.

c) Für jedes $\gamma \in V$ sei A_γ der Endomorphismus $z \mapsto \gamma z$. Man zeige: $A_\gamma^{ad} = A_{\overline{\gamma}}$.

d) Für welche γ ist A_γ selbstadjungiert, für welche orthogonal, für welche positiv definit (d. h. wann gilt $\langle\, z,\, A_\gamma z\,\rangle > 0$ für alle $z \neq 0$)?

e) Man finde einen orthogonalen Endomorphismus von V, der nicht von der Gestalt A_γ ist.

82. Es sei V_n der $\mathbb{C}$–Vektorraum der Polynome mit komplexen Koeffizienten vom Grad $\leq n$, versehen mit dem Skalarprodukt

$$\langle\, P,\, Q\,\rangle = \int_0^1 \overline{P\,(t)}\, Q\,(t)\, dt\,.$$

Ist der Differentiationsoperator D selbstadjungiert, ist er unitär ?

83. Man zeige: Ein Endomorphismus F des endlich–dimensionalen unitären Vektorraumes V ist genau dann normal, wenn für jeden F–invarianten Unterraum $U \subset V$ gilt: $F(U^\perp) \subset U^\perp$.

84. Der kleine Hamilton pflegte seinen (später so berühmten) Vater oft zu fragen: "Daddy, can you now multiply triples ?" In der Tat hat sich Hamilton fast ein ganzes Gelehrtenleben lang mit dieser Frage beschäftigt und für die Quadrupel schließlich die Antwort in Form der Quaternionen gefunden. Sie können jedenfalls mit Hilfe der Eigenwerttheorie schon heute beweisen: Für $n \geq 3$, n ungerade, kann man auf $\mathbb{R}^n$ mit der üblichen Addition *keine* Multiplikation erklären, die $\mathbb{R}^n$ zu einem Körper macht, in den $\mathbb{R}$ vermöge $r \mapsto r \cdot 1$, 1 das Einselement des Körpers, so eingebettet ist, daß die Einschränkung der Multiplikation auf die reellen Zahlen mit der $\mathbb{R}$–Vektorraum–Multiplikation übereinstimmt.

85. Es sei F ein *normaler* Endomorphismus eines unitären Vektorraumes V der Dimension $\dim V = n < \infty$. Ferner seien $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von F . Man zeige:

a) Ist $|\lambda_j| = 1$, $j = 1, \ldots, n$, so ist F unitär.

b) Ist $\lambda_j \in \mathbb{R}$, $j = 1, \ldots, n$, so ist F selbstadjungiert.

c) Ist $\lambda_j \in \mathbb{R}$ und positiv für alle $j = 1, \ldots, n$, so ist F positiv–definit, d. h. $\langle F(v), v \rangle > 0$ für alle $v \neq 0$.

Bleiben diese Folgerungen richtig, wenn F nicht als normal vorausgesetzt wird ?

86. Man interpretiere und beweise Korollar 6. 4. 14 in Fischer, *Lineare Algebra*.

87. Man zeige: Zu jeder Matrix $S \in \mathrm{O}(2)$ gibt es eine Matrix $U \in \mathrm{O}(2)$, so daß $S^{-1} = U^{-1} S U$. Man interpretiere diesen Sachverhalt geometrisch und verallgemeinere ihn auf beliebige Dimension n .

88. Man zeige, daß die beiden Matrizen

$$\frac{1}{4} \begin{pmatrix} 2+\sqrt{3} & -2+\sqrt{3} & -\sqrt{2} \\ -2+\sqrt{3} & 2+\sqrt{3} & -\sqrt{2} \\ \sqrt{2} & \sqrt{2} & 2\sqrt{3} \end{pmatrix} , \quad \frac{1}{2} \begin{pmatrix} 1+r & -1 & r \\ 1 & r & -1-r \\ r & 1+r & 1 \end{pmatrix} ,$$

wobei $r = \dfrac{1}{2} (\sqrt{5} - 1)$, in $\mathrm{SO}(3)$ liegen. Man finde jeweils die Drehachse und den (absoluten) Drehwinkel.

Hinweis: Man zeige, daß der Drehwinkel α einer Matrix $Q \in \mathrm{SO}(3)$ die folgende Bedingung erfüllt:

$$\cos \alpha = \frac{1}{2} (\mathrm{spur}\, Q - 1) .$$

(Siehe auch Aufgabe 90).

89. Man zeige:

a) Permutationsmatrizen A_σ sind orthogonal, und es gilt $A_\sigma \in \mathrm{SO}(3)$ genau dann, wenn die Permutation σ gerade ist.

b) Es gilt $\det(A_\sigma + A_\tau) = 0$, wenn σ gerade und τ ungerade ist.

90. Man zeige: Für jede Matrix $Q = (q_{jk}) \in \mathrm{SO}(3)$ mit Drehwinkel α gilt:

a) Q ist genau dann symmetrisch, wenn $\sin \alpha = 0$;

b) ist Q nicht symmetrisch, so bestimmt der Vektor

$$q = {}^t(q_{32} - q_{23},\, q_{13} - q_{31},\, q_{21} - q_{12})$$

die Drehachse.

Hinweis: Es sei v ein normierter Fixvektor von Q. Man ergänze v zu einer Orthonormalbasis (v_1, v_2, v_3), wobei v_2 mindestens eine verschwindende Koordinate besitzt, und zeige $q = (2 \sin \alpha)\, v$.

Kapitel 14

91. Man bringe die Quadrik

$$41\, x_1^2 - 24\, x_1 x_2 + 34\, x_2^2 - 30\, x_1 - 40\, x_2 = 0$$

auf (metrische) Normalform.

92. Man zeige: Jedes Element in $\mathrm{SO}(3)$ ist Produkt von zwei Spiegelungen. (Folgerung: Jedes Element in $\mathrm{O}(3)$ ist Produkt von höchstens drei Spiegelungen. Wie lautet vermutlich die entsprechende Verallgemeinerung für die Gruppe $\mathrm{O}(n)$?)