Lineare Algebra

Rainer Vogt

Wintersemester 2002/2003

Vorwort

Gemeinhin hält man Mathematik für eine Wissenschaft, mit deren Hilfe man etwas ausrechnen kann. Mit dieser Meinung wird man der Mathematik nicht gerecht. Überspitzt formuliert, könnte man sagen, Mathematik sei eine gewisse Denkstruktur. Das ist natürlich nicht korrekt, kommt aber der Sache wesentlich näher, wie Sie in den nächsten Wochen merken werden. Genauer gesagt, übt die Beschäftigung mit Mathematik eine analytische Denkstruktur, die in den nächsten Jahrzehnten zunehmend an Bedeutung gewinnen wird und die in vielen Bereichen der Forschung und des täglichen Lebens anwendbar ist.

Das Erlernen einer ungewohnten Denkstruktur ist mit viel Mühe verbunden. Das Ziel erreicht man nur dann, wenn man bereit ist, viel Zeit in die Beschäftigung mit und das Nachdenken über Mathematik zu investieren, sieht man von den wenigen Überfliegern einmal ab. Sechs bis acht Stunden über die Vorlesung und Übung hinaus dürfte dafür ein Minimum sein!

Sie werden sich fragen, was das mit Ihrem späteren Berufsziel zu tun hat. Schließlich müssten ein paar Rechenverfahren für das weitere Studium und die späteren beruflichen Anforderungen genügen. Dazu ein paar Bemerkungen. Für Diplommathematiker, Studenten im Bachelorstudiengang und Mathematiklehrer sollte es eine Selbstverständigkeit sein, die für die Mathematik nötige Denkstruktur zu beherrschen. Von Fachmathematikern wird analytisches Denken im Berufsleben gefordert und das Berufsfeld des Lehrers geht weit über das Abarbeiten von Rahmenrichtlinien hinaus: Ideal wäre, wenn Schülern in einem Alter mathematische Denkweisen vermittelt werden können, in dem sie für neue Denkprozesse noch besonders aufnahmefähig sind.

Kommen wir nun zu den Anwendern von Mathematik: Kognitionswissenschaftlern, Physikern und Systemwissenschaftlern. In diesen Bereichen gehen die Anforderungen an Abstraktionsvermögen und analytischem Denken sogar noch einen Schritt weiter: Man hat ein konkretes Problem vorliegen, für das man in einem

Abstraktionsschritt zunächst ein mathematisches Modell entwickelt, das man anschließend mit mathematischen Methoden analysiert, um Prognosen machen zu können. Ein ausgezeichneter Tummelplatz, solche Abstraktionsschritte zu üben, ist die Mathematik selbst. Ich werde in der Vorlesung wiederholt auf solche Abstraktionsschritte eingehen. Sie sollten sich mit diesen Teilen besonders intensiv beschäftigen, um Fertigkeiten in diesen Bereichen zu erlangen.

Für alle diejenigen, die noch immer nicht davon überzeugt sind, dass die doch hohen Anforderungen in der Mathematik berechtigt sind, sei ein Beispiel aus der Statistik angeführt: Der Inhalt eines Lagerhauses besteht zu 2/3 aus Würsten und 1/3 aus Eiern. Ein Hund geht in das Lagerhaus. Als er herauskommt, besteht der Inhalt zur Hälfte aus Würsten und zur Hälfte aus Eiern. Ein Biologe wendet ein gerade gelerntes statistisches Rezept an und folgert: “Ein Hund legt Eier!”

Fazit: Es genügt nicht, eine mathematische Formel anzuwenden. Man muss sie **verstanden** haben!

Inhaltsverzeichnis

I Vektorräume 5

1 Lineare Gleichungssysteme 5

2 Der Begriff des Körpers 21

3 Grundlagen: Mengen und Abbildungen 29

4 Vektorräume 37

5 Basis und Dimension 51

II Lineare Abbildungen 63

6 Grundbegriffe 63

7 Matrizen 74

8 Grundlagen: *Äquivalenzrelationen* 82

9 Quotienten- und Dualräume 84

10 Ergänzung zu linearen Gleichungssystemen 95

11 Determinanten 101

III Lineare Endomorphismen 118

12 Eigenwerte und Eigenvektoren 118

13 Polynomringe 131

14 Trigonalisierbare Endomorphismen 135

IV Vektorräume mit Skalarprodukt 139

15 Skalarprodukte 139

16 Bilinearformen und Matrizen 151

17 Orthogonale und unitäre Abbildungen 162

Teil I
Vektorräume

Die lineare Algebra entstand aus der Behandlung linearer Gleichungssysteme.

1 Lineare Gleichungssysteme

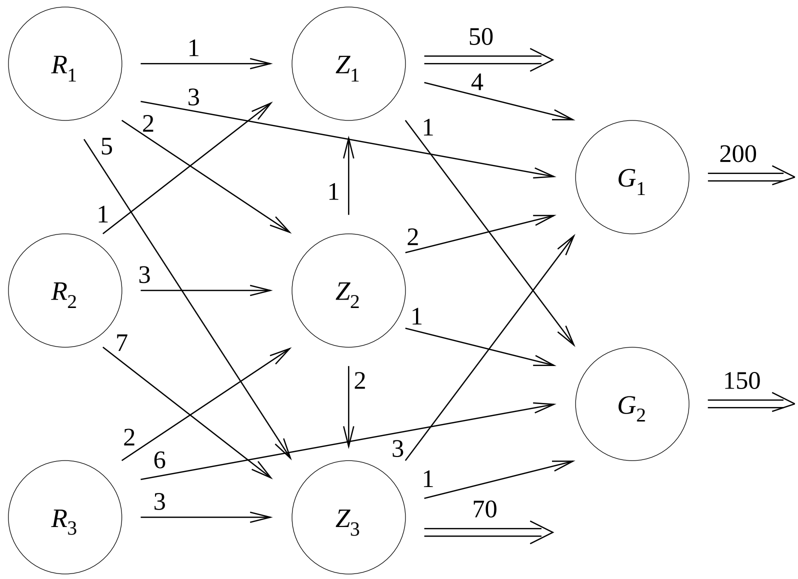
Eine Grundtechnik der linearen Algebra ist das Lösen linearer Gleichungssysteme. Die Lösungsmengen solcher Systeme und ihre Beziehungen zum Ausgangspunkt sind Grundbausteine der linearen Algebra. Wir beginnen mit Beispielen.

1.1 Teilebedarfsrechnung Aus Rohstoffen R_1, R_2, R_3 werden Zwischenprodukte Z_1, Z_2, Z_3 und Endprodukte G_1, G_2 gefertigt. Verkauft werden die Zwischenprodukte Z_1 und Z_3 und beide Endprodukte. Der Betrieb erhält den Auftrag, 200 Mengeneinheiten (*ME*) des Gutes G_1, 150 *ME* des Gutes G_2, 50 *ME* des Zwischenproduktes Z_1 und 70 *ME* des Zwischenproduktes Z_3 zu liefern.

Problem: Wie viel *ME* der Rohstoffe R_1, R_2, R_3 werden für den Auftrag benötigt?

Um dieses Problem zu lösen, stellt man am besten ein Bedarfsdiagramm auf, auch GOZINTOGRAPH genannt (s. nächste Seite). Unser Gozintograph ist folgendermaßen zu lesen: die einfachen Pfeile geben an, wie viele *ME* von einem Rohstoff bzw. Zwischenprodukt benötigt werden, um das Produkt an der Pfeilspitze zu produzieren. Die Doppelpfeile geben den gewünschten "Output" des Betriebs an. Bezeichnen wir mit $g_1, g_2, z_1, z_2, z_3, r_1, r_2, r_3$ die jeweils benötigten *ME* der Güter G_1, G_2, Zwischenprodukte Z_1, Z_2, Z_3 und Rohstoffe R_1, R_2, R_3 erhalten wir aus dem Gozintographen folgendes Gleichungssystem

$$
\begin{array}{ll}
g_1 = 200 & \text{bestimmt durch den Auftrag} \\
g_2 = 150 & \text{bestimmt durch den Auftrag} \\
z_1 = 4g_1 + 1g_2 + 50 & \text{(die 50 für den Auftrag)} \\
z_2 = 2g_1 + 1g_2 + 1z_1 + 2z_3 & \\
z_3 = 3g_1 + 1g_2 + 70 & \text{(die 70 für den Auftrag)} \\
r_1 = 3g_1 + 1z_2 + 2z_2 + 5z_3 & \\
r_2 = 1z_1 + 3z_2 + 7z_3 & \\
r_3 = 6g_2 + 2z_2 + 3z_3 &
\end{array}
$$



Für jeden “Knoten” unseres Diagramms erhalten wir also eine Gleichung.

Es ist üblich, die Unbekannnten auf die linke Seite und die “absoluten” Terme auf die rechte Seite des Gleichheitszeichen zu schreiben. Wir erhalten somit ein Gleichungssystem.

1.2

$$\begin{array}{ll}
g_1 = 200 & \mathrm{I} \\
g_2 = 150 & \mathrm{II} \\
z_1 - 4g_1 - g_2 = 50 & \mathrm{III} \\
z_2 - 2g_1 - g_2 - z_1 - 2z_3 = 0 & \mathrm{IV} \\
z_3 - 3g_1 - g_2 = 70 & \mathrm{V} \\
r_1 - 3g_1 - z_1 - 2z_2 - 5z_3 = 0 & \mathrm{VI} \\
r_2 - z_1 - 3z_2 - 7z_3 = 0 & \mathrm{VII} \\
r_3 - 6g_2 - 2z_2 - 3z_3 = 0 & \mathrm{VIII}
\end{array}$$

Der besseren Übersicht halber schreiben wir diese Gleichungen in ein Schema. Der senkrechte Strich nimmt die Stelle des Gleichheitszeichen ein, über dem waagerechten Strich stehen die Unbekannten. Außerdem ordnen wir Gleichungen und Unbekannte möglichst geschickt. Wir erhalten folgendes Schema.

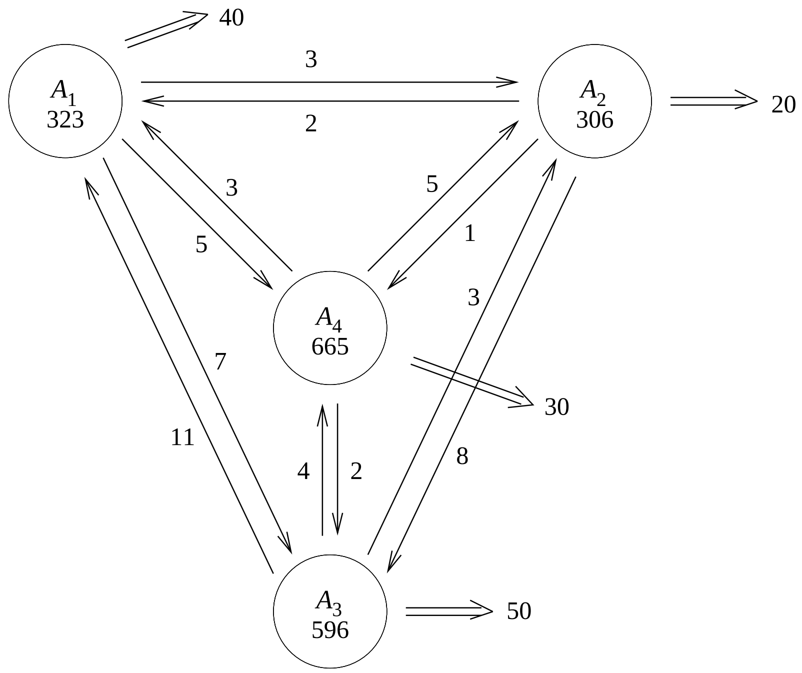
1.3

r_1	r_2	r_3	z_2	z_1	z_3	g_1	g_2		
1	0	0	-2	-1	-5	-3	0	0	VI
0	1	0	-3	-1	-7	0	0	0	VII
0	0	1	-2	0	-3	0	-6	0	VIII
0	0	0	1	-1	-2	-2	-1	0	IV
0	0	0	0	1	0	-4	-1	50	III
0	0	0	0	0	1	-3	-1	70	V
0	0	0	0	0	0	1	0	200	I
0	0	0	0	0	0	0	1	150	II

Dieses Schema ist in **OBERER DREIECKSFORM**, d.h. in der Diagonalen stehen Werte 1 und unterhalb der Diagonalen nur Werte 0. Ein solches Gleichungssystem kann man durch **SUKZESSIVES EINSETZEN** leicht lösen. Wir lesen, von unten nach oben arbeitend, sofort ab:

$$
\begin{aligned}
g_2 &= 150\\
g_1 &= 200\\
z_3 &= 3g_1 + g_2 + 70 = 820\\
z_1 &= 4g_1 + g_2 + 50 = 1000\\
z_2 &= z_1 + 2z_3 + 2g_1 + g_2 = 3190\\
r_3 &= 2z_2 + 3z_3 + 6g_2 = 9740\\
r_2 &= 3z_2 + z_1 + 7z_3 = 16310\\
r_1 &= 2z_2 + z_1 + 5z_3 + 3g_1 = 12080
\end{aligned}
$$

1.4 Leistungsverrechnung in einem System: In jedem Knoten eines Systems (etwa Abteilungen eines Betriebs) werden Leistungen erbracht. Der Wert einer Einheit aus dem Knoten A_i sei x_i Geldeinheiten (GE). Zur Aufrechterhaltung des Betriebs im Knoten sind Primärkosten nötig, wie in den Knoten angegeben. Daneben erhält jeder Knoten von anderen Knoten Leistungen (etwa vernetzte Computer). Die Doppelpfeile geben den Output an.



Hat ein System n Knoten A_i und sind die Primärkosten für A_i gerade b_i GE, so werden

$$\sum_{i=1}^{n} b_i := b_1 + b_2 + \ldots, b_n$$

GE in das System gesteckt. Ist der Output des Knotens A_i (Doppelpfeil) c_i ME im Wert von je x_i GE, so hat der Gesamtoutput den Wert

$$\sum_{i=1}^{n} c_i \cdot x_i.$$

Es gilt also:

$$\sum_{i=1}^{n} b_i = \sum_{i=1}^{n} c_i \cdot x_i.$$

Unser Problem: Da die Knoten Leistungen austauschen, kennen wir den Wert von x_i nicht!

Sei a_{ij} die Leistung (in ME), die Knoten A_j an Knoten A_i abgibt. In jedem Knoten ist die Leistungsaufnahme einschließlich Primärkosten gleich der Leistungsabgabe. Für jeden Knoten erhalten wir damit eine Gleichung. In unserem Beispiel

$$\begin{array}{lllll}
 & \text{Leistungsaufnahme} & & \text{Leistungsabgabe} & (\text{in } GE) \\
I & 323+2x_2+11x_3+3x_4 & = & (40+3+7+5)x_1 & \\
II & 306+3x_1+3x_3+5x_4 & = & (20+2+8+1)x_2 & \\
III & 596+7x_1+8x_2+2x_4 & = & (50+11+3+4)x_3 & \\
IV & 665+5x_1+1x_2+4x_3 & = & (30+3+5+2)x_4 &
\end{array}$$

Resultierendes Gleichungssystem

$$\begin{array}{llll}
I & 55x_1-2x_2-11x_3-3x_4 & = & 323 \\
II & -3x_1+31x_2-3x_3-5x_4 & = & 306 \\
III & -7x_1-8x_2+68x_3-2x_4 & = & 596 \\
IV & -5x_1-x_2-4x_3+40x_4 & = & 665
\end{array}$$

Man sieht sofort, dass Probleme dieser Art sehr schnell zu großen Gleichungssystemen mit recht unfreundlich großen Zahlen führen. Um mit diesen umgehen zu können, benötigen wir einen guten Rechenalgorithmus.

Ein solcher Algorithmus ist der Eliminationsalgorithmus, der in der modernen Mathematik Gauß zugeschrieben wird, der aber schon hunderte von Jahren früher in China benutzt wurde.

1.5 Definition: Mit $\mathbb{R}$ bezeichnen wir die Menge der reellen Zahlen. Ein **LINEARES GLEICHUNGSSYSTEM** über $\mathbb{R}$ ist ein System von Gleichungen

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m
\end{array}$$

mit a_{ij} und b_i aus $\mathbb{R}$ für $i = 1, \dots, m$ und $j = 1, \dots, n$. Man spricht von einem Gleichungssystem mit m Gleichungen und n Unbekannten. Die n-Tupel $(x_1, \dots, x_n)$, die diese Gleichungen erfüllen, heißen Lösungen des Gleichungssystems, die Gesamtheit der Lösungen **LÖSUNGSMENGE**.

Im Spezialfall $b_1 = b_2 = \dots = b_m = 0$ nennt man das Gleichungssystem **HOMOGEN** und sonst **INHOMOGEN**.

Ersetzen wir im System (1.5) alle b_i durch 0, erhalten wir das **ZUGEHÖRIGE HOMOGENE GLEICHUNGSSYSTEM**.

Ein Gleichungssystem zu **LÖSEN** bedeutet, die Lösungsmenge so darzustellen, dass man sofort ablesen kann, ob ein gegebenes n-Tupel $(x_1, \dots, x_n)$ reeller Zahlen zu ihr gehört oder nicht. Dazu manipulieren wir das Gleichungssystem: Wir formen es in ein übersichtlicheres Gleichungssystem mit derselben Lösungsmenge um, etwa in ein System mit oberen Dreiecksschema.

Folgende Manipulationen sind offensichtlich erlaubt:

1.6 Erlaubte Manipulationen:

(1) Umordnen der Gleichungen und Umordnen der Summation.

(2) Multiplizieren einer Gleichung mit einer Zahl $c \neq 0$

(3) Addition des c-fachen einer Gleichung zu einer anderen, c beliebig.

Erläuterungen zu (3): Folgende Aussagen sind äquivalent

(1) $(x_1, \ldots, x_n)$ löst die Gleichungen

$$\begin{array}{ll} \text{I} & a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \text{II} & a_{21}x_1 + \ldots + a_{2n}x_n = b_2 \end{array}$$

(2) $(x_1, \ldots, x_n)$ löst die Gleichungen

$$\begin{array}{lrl} \text{III} & (a_{11} + ca_{21})x_1 + \ldots + (a_{1n} + ca_{2n})x_n = & b_1 + cb_2 \\ \text{II} & a_{21}x_1 + \ldots + a_{2n}x_n = & b_2 \end{array}$$

Beweis:

(1) $\Rightarrow$ (2) I $+c \cdot$II gibt III

(2) $\Rightarrow$ (1) III $-c \cdot$II gibt I.

$\square$

1.7 Lösungsverfahren Der besseren Übersicht wegen schreiben wir die Gleichungen wieder in einem System auf:

Ausgangsschema:

$$\begin{array}{ccccc|cl} x_1 & x_2 & x_3 & \text{------------} & x_n & b & \\ \hline a_{11} & a_{12} & a_{13} & \text{------------} & a_{1n} & b_1 & \text{1. Gleichung} \\ a_{21} & a_{22} & a_{23} & \text{------------} & a_{2n} & b_2 & \text{2. Gleichung} \\ a_{31} & a_{32} & a_{33} & \text{------------} & a_{2n} & b_3 & \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \\ a_{m1} & a_{m2} & a_{m3} & \text{------------} & a_{mn} & b_m & m\text{-te Gleichung} \end{array}$$

Beispiel: Das Gleichungssystem

$$\begin{array}{rcrcrcr} 2x_1 & - & 3x_2 & + & 2x_3 & = & -1 \\ 3x_1 & + & 2x_2 & - & 3x_3 & = & 5 \\ 7x_1 & - & 4x_2 & + & x_3 & = & 3 \end{array}$$

hat das Ausgangsschema

$$\begin{array}{rrr|r} x_1 & x_2 & x_3 & b \\ \hline 2 & -3 & 2 & -1 \\ 3 & 2 & -3 & 5 \\ 7 & -4 & 1 & 3 \end{array}$$

1. Schritt: Wir suchen ein Element $a_{pq} \neq 0$.

Wir dividieren die p-te Gleichung durch a_{pq}.

Wir erhalten eine neue p-te Gleichung.

$$\overline{a}_{p1}x_1 + \overline{a}_{p2}x_2 + \ldots + \overline{a}_{pn}x_n = \overline{b}_p$$

mit

$$\overline{a}_{pj} = \frac{a_{pj}}{a_{pq}}, \text{ also } \overline{a}_{pq} = 1, \text{ und } \overline{b}_{pj} = \frac{b_p}{a_{pq}}$$

Im Beispiel: Wir können a_{11} nehmen, hier 2 in der linken oberen Ecke. Also ist $p = 1$ und $q = 1$. Die neue erste Gleichung ist

$$x_1 - \frac{3}{2}x_2 + x_3 = -\frac{1}{2}$$

2. Schritt: Wir schreiben die neue p-te Gleichung als erste Gleichung und die Spalte mit der q-ten Unbestimmten x_q als erste Spalte:

$$\begin{array}{ccccc|cl} x_q & x_1 & x_2 & \cdots & x_n & b & \\ \hline 1 & \overline{a}_{p1} & \overline{a}_{p2} & \cdots & \overline{a}_{pn} & \overline{b}_p & \text{1. neue Zeile} \\ a_{1q} & a_{11} & a_{12} & \cdots & a_{1n} & b_1 & \text{alte 1. Zeile} \\ a_{2q} & a_{21} & a_{22} & \cdots & a_{2n} & b_2 & \text{alte 2. Zeile} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \\ a_{mq} & a_{m1} & a_{m2} & \cdots & a_{mn} & b_m & \text{alte } m\text{-te Zeile} \end{array}$$

Im Beispiel hat das Schema nach dem 2. Schritt die Form

$$
\begin{array}{rrr|r}
x_1 & x_2 & x_3 & b \\
\hline
1 & -\frac{3}{2} & 1 & -\frac{1}{2} \\
3 & 2 & -3 & 5 \\
7 & -4 & 1 & 3
\end{array}
$$

3. Schritt: Wir subtrahieren das a_{iq}-fache der 1. Zeile des neuen Schemas von der alten i-ten Zeile. Diese Zeile geht über in

$$
\underbrace{(a_{iq} - a_{iq}1)}_{0} x_q + \underbrace{(a_{i1} - a_{iq}\overline{a}_{p1})}_{\overline{a}_{i1}} x_1 + \underbrace{(a_{i2} - a_{iq}\overline{a}_{p2})}_{\overline{a}_{i2}} x_2 + \dots
$$
$$
+ \underbrace{(a_{in} - a_{iq}\overline{a}_{pn})}_{\overline{a}_{in}} x_n = \underbrace{b_i - a_{iq}\overline{b}_p}_{\overline{b}_i}
$$

Wir erhalten ein Schema

$$
\begin{array}{cccccc|c}
x_q & x_1 & x_2 & \cdots & x_n & & b \\
\hline
1 & \overline{a}_{p1} & \overline{a}_{p2} & \cdots & \overline{a}_{pn} & & \overline{b}_p \\
0 & \overline{a}_{11} & \overline{a}_{12} & \cdots & \overline{a}_{1n} & & \overline{b}_1 \\
0 & \overline{a}_{21} & \overline{a}_{22} & \cdots & \overline{a}_{2n} & & \overline{b}_2 \\
\vdots & \vdots & \vdots & & \vdots & & \vdots \\
0 & \overline{a}_{m1} & \overline{a}_{m2} & \cdots & \overline{a}_{mn} & & \overline{b}_m
\end{array}
$$

Im Beispiel ziehen wir das 3-fache der ersten Zeile von der zweiten ab und das 7-fache von der ersten Zeile von der dritten. Wir erhalten

$$
\begin{array}{rrr|r}
x_1 & x_2 & x_3 & b \\
\hline
1 & -\frac{3}{2} & 1 & -\frac{1}{2} \\
0 & \frac{13}{2} & -6 & \frac{13}{2} \\
0 & \frac{13}{2} & -6 & \frac{13}{2}
\end{array}
$$

4. Schritt: Wir wiederholen das Verfahren auf dem Teilsystem ohne die erste Zeile und Spalte, wobei aber bei **Spaltungsvertauschungen die Spalten des Gesamtsystems vertauscht werden**.

5. Schritt: Das Verfahren bricht ab, wenn im letzten Teilsystem alle $\overline{a}_{ij} = 0$ sind.

Wir wiederholen das Verfahren im Beispiel

1. Schritt: $a_{22} = \frac{13}{2} \neq 0$. Wir dividieren die 2. Gleichung durch $\frac{13}{2}$ und erhalten im 2. Schritt das neue Schema

$$
\begin{array}{ccc|c}
x_1 & x_2 & x_3 & b \\
\hline
1 & -\frac{3}{2} & 1 & -\frac{1}{2} \\
0 & 1 & -\frac{12}{13} & 1 \\
0 & \frac{13}{2} & -6 & \frac{13}{2}
\end{array}
$$

Im 3. Schritt subtrahieren wir das $\frac{13}{2}$-fache der zweiten Zeile von der dritten und erhalten

$$
\begin{array}{ccc|c}
x_1 & x_2 & x_3 & b \\
\hline
1 & -\frac{3}{2} & 1 & -\frac{1}{2} \\
0 & 1 & -\frac{12}{13} & 1 \\
0 & 0 & 0 & 0
\end{array}
$$

Wir erhalten so ein Schema in **DREIECKSFORM**

1.8

$$
\begin{array}{ccccccc|c}
x_{q_1} & x_{q_2} & \cdots & x_{q_r} & x_{q_{r+1}} & \cdots & x_{q_n} & b \\
\hline
1 & a'_{12} & & a'_{1r} & a'_{1,r+1} & \cdots & a'_{1n} & b'_1 \\
0 & 1 & \ddots & & \vdots & & \vdots & \vdots \\
\vdots & 0 & & & & & & \\
0 & & & 1 & a'_{r,r+1} & & a'_{rn} & b'_r \\
0 & \cdots & \cdots & 0 & 0 & \cdots & 0 & b'_{r+1} \\
\vdots & & & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & \cdots & 0 & 0 & \cdots & 0 & b'_m
\end{array}
$$

Unser Gleichungssystem 1.5 hat somit dieselbe Lösungsmenge wie das zum Schema 1.8 gehörende Gleichungssystem.

1.9

$$\begin{array}{ccccccccc}
x_{q_1} & +a'_{12}x_{q_2}+ & \dots & +a'_{1r}x_{q_r} & +a'_{1q,r+1}x_{q_{r+1}}+ & \dots & +a'_{1n}x_{q_n} & = & b'_1 \\
 & x_{q_2}+ & \dots & +a'_{2r}x_{q_r} & +a'_{2q,r+1}x_{q_{r+1}}+ & \dots & +a'_{2n}x_{q_n} & = & b'_2 \\
 & & \ddots & \vdots & \vdots & & \vdots & & \\
 & & & x_{q_r} & +a'_{r,r+1}x_{q_{r+1}}+ & \dots & +a'_{rn}x_{q_n} & = & b'_r \\
 & & & & & & 0 & = & b'_{r+1} \\
 & & & & & & 0 & = & b'_m
\end{array}$$

Im Beispiel

$$\begin{array}{ccccccc}
x_1 & - & \frac{3}{2}\cdot x_2 & + & x_3 & = & -\frac{1}{2} \\
 & & x_2 & - & \frac{12}{13}x_3 & = & 1 \\
 & & & & 0 & = & 0
\end{array}$$

Man sieht sofort

1.10 Lösbarkeitskriterium: Das Gleichungssystem (1.5) hat genau dann eine Lösung, wenn in (1.8) gilt

$$b'_{r+1} = b'_{r+2} = \ldots = b'_m = 0.$$

In unserem Beispiel ist $r = 2$, $m = 3$, also $r+1 = m = 3$ und $b'_3 = 0$. Also ist unser Gleichungssystem lösbar.

Ist das Lösbarkeitskriterium erfüllt, erhalten wir die Gesamtlösung durch **SUKZESSIVES EINSETZEN**.

1.11

$$\begin{array}{rl}
x_{q_r} & = b'_r - a'_{r,r+1}, x_{q_{r+1}} - \ldots - a'_{r_n} x_{q_n} \\
x_{q_{r-1}} & = b'_{r-1} - a'_{r-1,r} x_{q_r} - a'_{r-1,r+1}, x_{q_{r+1}} - \ldots - a'_{r-1,n} x_{q_n} \ (x_{q_r} \text{ ist bekannt}) \\
\vdots &
\end{array}$$

Dabei sind $x_{q_{r+1}}, \ldots, x_{q_n}$ frei wählbar.
In unserem Beispiel erhalten wir

$$\begin{array}{rcrcl}
x_2 & = & 1 & + & \frac{12}{13}\cdot x_3 \\
x_1 & = & -\frac{1}{2} & + & \frac{3}{2}\cdot x_2 - x_3 \\
 & = & -\frac{1}{2} & + & \frac{3}{2} + \frac{18}{13}\cdot x_3 - x_3 \\
 & = & 1 & + & \frac{5}{13}\cdot x_3
\end{array}$$

mit frei wählbaren x_3, d.h. jedes Tripel

$$(x_1,x_2,x_3) = (1+\frac{5}{13}\cdot x_3, 1+\frac{12}{13}\cdot x_3, x_3),$$

wobei x_3 eine beliebige reelle Zahl ist, ist eine Lösung.

Wir sehen aber auch, dass das sukzessive Einsetzen mit Aufwand verbunden ist. Wir wollen es ebenfalls schematisieren.

1.12 Vereinfachung der Dreiecksform Liegt ein lösbares Gleichungssystem vor, hat das obere Dreiecksschema folgende Form (wir lassen die Nullzeilen weg!).

$$\begin{array}{cccccccc|c}
x_{q_1} & x_{q_2} & & x_{q_r} & x_{q_{r+1}} & \cdots & x_{q_n} & & b \\ \hline
1 & a'_{12} & \cdots & a'_{1r} & a'_{1,r+1} & \cdots & a'_{1n} & & b'_1 \\
0 & 1 & & a'_{2r} & a'_{2,r+1} & & a'_{2n} & & b'_2 \\
 & & \ddots & \vdots & \vdots & & \vdots & & \vdots \\
 & 0 & & & & & & & \\
 & & & 1 & a'_{r,r+1} & \cdots & a'_{r,n} & & b'_r
\end{array}$$

1. Schritt: Wir subtrahieren das a'_{ir}-fache der r-ten Gleichung von der i-ten Gleichung für $1 \le i \le r-1$. Wir erhalten eine neues Schema

$$\begin{array}{cccccccc|c}
x_{q_1} & x_{q_2} & & x_{q_{r-1}} & x_{q_r} & x_{q_{r-1}} & \cdots & x_{q_n} & b \\ \hline
1 & a'_{12} & & a'_{1,r-1} & 0 & a''_{1,r+1} & \cdots & a''_{1n} & b''_1 \\
0 & 1 & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
 & & \ddots & & & & & & \\
 & 0 & & & & & & & \\
 & & & 1 & 0 & a''_{r-1,r+1} & \cdots & a''_{r-1,n} & b''_{r-1} \\
 & & & & 1 & a'_{r,r+1} & \cdots & a'_{r,n} & b'_r
\end{array}$$

Im Beispiel ist $r = 2$, also kann i nur den Wert 1 annehmen. $a'_{12} = -3/2$. Wir müssen also das $(-3/2)$-fache der zweiten Zeile von der ersten abziehen, d.h. das $(+3/2)$-fache der zweiten Zeile zur ersten addieren. Wir erhalten

$$\begin{array}{ccc|c}
x_1 & x_2 & x_3 & b \\ \hline
1 & 0 & -\frac{5}{13} & 1 \\
0 & 1 & -\frac{12}{13} & 1
\end{array}$$

2. Schritt: Wir subtrahieren das $a'_{i,r-1}$-fache der neuen $(r-1)$-ten Gleichung von der i-ten Gleichung für $1 \le i \le r-2$ und fahren mit diesem Verfahren fort, bis wir folgendes **REDUZIERTE SYSTEM** erhalten.

1.13 Reduziertes System

$$
\begin{array}{ccccccc|c}
x_{q_1} & x_{q_2} & & x_{q_r} & x_{q_{r+1}} & \cdots & x_{q_n} & b \\
\hline
1 & 0 & \cdots & 0 & a^*_{1,r+1} & \cdots & a^*_{1,n} & b^*_1 \\
0 & 1 & & \vdots & \vdots & & \vdots & \vdots \\
 & & \ddots \quad 0 & & & & & \\
 & 0 & & & & & & \\
 & & & 1 & a^*_{r,r+1} & \cdots & a^*_{r,n} & b^*_r
\end{array}
$$

Als **Gesamtlösung** erhalten wir

1.14

$$
\begin{array}{rcrclcccl}
x_{q_1} & = & b^*_1 & - & a^*_{1,r+1}x_{q_{r+1}} & - & \cdots & - & a^*_{1n}x_{q_n} \\
x_{q_2} & = & b^*_2 & - & a^*_{2,r+1}x_{q_{r+1}} & - & \cdots & - & a^*_{2n}x_{q_n} \\
\vdots & & & & & & & & \\
x_{q_r} & = & b^*_r & - & a^*_{r,r+1}x_{q_{r+1}} & - & \cdots & - & a^*_{rn}x_{q_n}
\end{array}
$$

wobei $x_{q_{r-1}}, \dots, x_{q_n}$ beliebig gewählt werden können.
In unserem Beispiel haben wir schon nach dem ersten Schritt das reduzierte System erreicht. Wir können sofort die Lösung ablesen:

$$
\begin{aligned}
x_1 &= 1 + \tfrac{5}{13} \cdot x_3 \\
x_2 &= 1 + \tfrac{12}{13} \cdot x_3
\end{aligned}
$$

mit beliebigem x_3.

1.15 Rechenbeispiel: Zu lösen

$$
\begin{array}{rcrcrcrcrcr}
x_1 & - & 2x_2 & + & x_3 & + & x_4 & - & x_5 & = & -1 \\
2x_1 & + & x_2 & + & 2x_3 & + & x_4 & + & 2x_5 & = & 4 \\
-x_1 & - & 2x_2 & - & x_3 & - & x_4 & - & x_5 & = & -3 \\
3x_1 & + & x_2 & + & 3x_3 & + & 2x_4 & + & 2x_5 & = & 5
\end{array}
$$

Ausgangsschema: Da $a_{11} = 1$, entfällt der erste Schritt

x_1	x_2	x_3	x_4	x_5	b	
1	-2	1	1	-1	-1	
2	1	2	1	2	4	
-1	-2	-1	-1	-1	-3	
3	1	3	2	2	5	
1	-2	1	1	-1	-1	
0	5	0	-1	4	6	II $-$ 2 $\cdot$I
0	-4	0	0	-2	-4	III $+$ I
0	7	0	-1	5	8	IV $-$ 3 $\cdot$I

Um unangenehme Bruchrechnungen zu ersparen, nehmen wir das neue a_{24} für den ersten Schritt und vertauschen Spalten

x_1	x_4	x_2	x_3	x_5	b	
1	1	-2	1	-1	-1	
0	1	-5	0	-4	-6	$-1 \cdot$ II
0	0	-4	0	-2	-4	
0	-1	7	0	5	8	
1	1	-2	1	-1	-1	
0	1	-5	0	-4	-6	
0	0	-4	0	-2	-4	
0	0	2	0	1	2	IV + II

Für den nächsten Schritt wählen wir das neue a_{35} und vertauschen Spalten

x_1	x_4	x_5	x_2	x_3	b	
1	1	-1	-2	1	-1	
0	1	-4	-5	0	-6	
0	0	1	2	0	2	$-\frac{1}{2}\cdot$ III
0	0	1	2	0	2	
1	1	-1	-2	1	-1	
0	1	-4	-5	0	-6	
0	0	1	2	0	2	
0	0	0	0	0	0	IV - III

Wir haben die obere Dreiecksform erreicht und erkennen, dass das Gleichungssystem lösbar ist.

Wir bringen jetzt die Dreiecksform auf reduzierte Form

x_1	x_4	x_5	x_2	x_3	b	
1	1	0	0	1	1	I + III
0	1	0	3	0	2	$\mathrm{II} + 4 \cdot \mathrm{III}$
0	0	1	2	0	2	
1	0	0	-3	1	-1	I - II
0	1	0	3	0	2	
0	0	1	2	0	2	

Wir haben die reduzierte Form erreicht. Die Lösungsmenge L besteht aus allen Tupeln $(x_1, x_2, x_3, x_4, x_5)$, für die gilt

1.16
$$\begin{aligned} x_1 &= -1 + 3x_2 - x_3 \\ x_4 &= 2 - 3x_2 \\ x_5 &= 2 - 2x_2 \end{aligned}$$

Die Lösungsmenge $\mathbb{L}$ ist also

$$\mathbb{L} = \{x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5; x_2 \in \mathbb{R}, x_3 \in \mathbb{R} \textbf{ und } x_1, x_4, x_5 \text{ erfüllen die Gleichung 1.16}\}$$

1.17 Rechenbeispiel: Zu lösen

$$\begin{array}{ccccccc} 0 \cdot x_1 & + & 2x_2 & + & 3x_3 & = & 5 \\ 0 \cdot x_1 & + & 3x_2 & - & 2x_3 & = & 1 \end{array}$$

Ausgangsschema:

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & b \\ \hline 0 & 2 & 3 & 5 \\ 0 & 3 & -2 & 1 \end{array}$$

Wir nehmen $a_{12} = 2$ für den ersten Schritt und ordnen um

$$\begin{array}{ccc|c} x_2 & x_3 & x_1 & b \\ \hline 2 & 3 & 0 & 5 \\ 3 & -2 & 0 & 1 \end{array}$$

Wir teilen die erste Gleichung durch 2

$$\begin{array}{ccc|cl} x_2 & x_3 & x_1 & b & \\ \hline 1 & \frac{3}{2} & 0 & \frac{5}{2} & \\ 3 & -2 & 0 & 1 & \\ \hline 1 & \frac{3}{2} & 0 & \frac{5}{2} & \\ 0 & -\frac{13}{2} & 0 & -\frac{13}{2} & \mathrm{II} - 3 \cdot \mathrm{I} \\ \hline 1 & \frac{3}{2} & 0 & \frac{5}{2} & \\ 0 & 1 & 0 & 1 & -\frac{2}{13} \cdot \mathrm{II} \\ \hline 1 & 0 & 0 & 1 & \mathrm{I} - \frac{3}{2} \cdot \mathrm{II} \\ 0 & 1 & 0 & 1 & \end{array}$$

Das letzte Schema ist die reduzierte Form, das vorletzte die Dreiecksform.
Wie können die Lösbarkeit direkt ablesen:

$$x_2 = 1 \qquad x_3 = 1 \qquad x_1 \text{ ist beliebig}$$

Also $\mathbb{L} = \{(x_1, 1, 1) \in \mathbb{R}^3; x_1 \in \mathbb{R} \text{ beliebig}\}$.

1.18 Rechenbeispiel: Zu lösen

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & + & x_3 & = & 1 \\ x_1 & - & x_2 & + & 2x_3 & = & 2 \\ 3x_1 & + & x_2 & + & 4x_3 & = & 5 \end{array}$$

x_1	x_2	x_3	b	
1	1	1	1	
1	-1	2	2	
3	1	4	5	
1	1	1	1	
0	-2	1	1	II – I
0	-2	1	2	III – $3 \cdot$ II
1	1	1	1	
0	1	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2} \cdot$ II
0	0	0	1	III – II

Damit haben wir unser Gleichungssystem auf folgende Form gebracht:

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 1 \\ & & x_2 & - & \frac{1}{2}x_3 & = & -\frac{1}{2} \\ & & & & 0 & = & 1 \end{array}$$

Die letzte Gleichung kann nie erfüllt werden. Also ist unser Gleichungssystem **NICHT LÖSBAR**!

Oft fragt man sich, ob Gleichungssysteme mit ganzzahligen Koeffizienten auch ganzzahlige Lösungen haben. Am einfachsten wäre die Antwort, wenn man den vorliegenden Algorithmus über der Menge $\mathbb{Z}$ der ganzen Zahlen durchführen könnte. Der erste Schritt im Lösungsverfahren 1.7 zeigt aber schon, dass das nicht immer geht: Wir können in $\mathbb{Z}$ nicht immer dividieren.

Um das Gaußsche Lösungsverfahren in möglichst großer Allgemeinheit zu haben, fragen wir uns, welche Eigenschaften der reellen Zahlen wir benutzt haben. Wir formalisieren das Verfahren, eine typisch mathematische Vorgehensweise.

2 Der Begriff des Körpers

Der Lösungsalgorithmus für Gleichungssysteme machte sich zu Nutze, dass man in $\mathbb{R}$ addieren, subtrahieren, multiplizieren und durch Elemente $a \neq 0$ dividieren kann. Offensichtlich hätte das gleiche Verfahren auch für Gleichungssysteme mit rationalen Koeffizienten funktioniert, also über $\mathbb{Q}$. Was sind nun die formalen Eigenschaften der Addition und Multiplikation, die wir benutzt haben?

Dazu betrachten wir anstelle der Mengen $\mathbb{Q}$ oder $\mathbb{R}$ eine beliebige Menge $\mathbb{K}$. Wir fordern, dass auf $\mathbb{K}$ zwei Rechenoperationen, genannt "Addition" und "Multiplikation" definiert sind. Formal ausgedrückt, bedeutet das, dass wir zwei Abbildungen

$$\begin{array}{lll} \text{Addition:} & \mathbb{K}\times\mathbb{K}\to\mathbb{K}, & (x,y)\mapsto x+y \\ \text{Multiplikation:} & \mathbb{K}\times\mathbb{K}\to\mathbb{K}, & (x,y)\mapsto x\cdot y \end{array}$$

haben (zum Begriff "Abbildung" und zur Definition von $\mathbb{K}\times\mathbb{K}$ siehe nächsten Paragraphen). Wir nennen $x+y$ die **Summe** und $x\cdot y$ das **Produkt** von x und y.

Addition und Multiplikation sollen gewissen Regeln genügen, die wir in möglichst wenigen Axiomen fordern. Aus diesen Axiomen wollen wir dann allgemeinere Regeln ableiten.

Diese Axiome sind dann Ausgangspunkt einer mathematischen Theorie, im vorliegenden Fall der **KÖRPERTHEORIE**. Bei ihrer Entwicklung dürfen nur die vorgegebenen Axiome benutzt werden. Daher sollte man ein solches Axiomensystem erst nach gründlicher Kenntnis geeigneter Beispiele aufstellen.

Wir lassen und von den Modellen $\mathbb{Q}$ und $\mathbb{R}$ und den Rechenregeln leiten, die wir in §1 stillschweigend benutzt haben. Nach Aufstellung der Axiome werden wir sehen, dass es unendlich viele wesentlich verschiedene Körper gibt. Allen ist gemeinsam, dass man in ihnen wie in $\mathbb{Q}$ oder $\mathbb{R}$ rechnen kann.

2.1 Definition: Ein **Körper** ist eine Menge $\mathbb{K}$ zusammen mit einer

$$\begin{array}{rll} \text{"Addition:"} & \mathbb{K}\times\mathbb{K}\to\mathbb{K}, & (x,y)\mapsto x\nabla y \\ \text{und "Multiplikation"} & \mathbb{K}\times\mathbb{K}\to\mathbb{K}, & (x,y)\mapsto x*y, \end{array}$$

so dass folgende Axiome erfüllt sind:

(A1) Assoziativität: $\forall\, x,y,z\in\mathbb{K}$ gilt $(x\nabla y)\nabla z = x\nabla(y\nabla z)$.

(A2) Kommutativität: $\forall\, x,y\in\mathbb{K}$ gilt $x\nabla y = y\nabla x$.

(A3) (Axiom vom additiven neutralen Element): Man kann (mindestens) ein Element $n_\nabla\in\mathbb{K}$ finden, so dass gilt: $\forall\, x\in\mathbb{K}$ ist $\quad x\nabla n_\nabla = x$.

(A4) (Axiom vom additiven Inversen): Zu jedem $x\in\mathbb{K}$ kann man (mindestens) ein Element $x'\in\mathbb{K}$ finden, so dass $x\nabla(x') = n_\nabla$.

(M1) Assoziativität: $\forall\, x,y,z \in \mathbb{K}$ gilt $(x*y)*z = x*(y*z)$.

(M2) Kommutativität: $\forall\, x,y \in \mathbb{K}$ gilt $x*y = y*x$.

(M3) (Axiom vom multiplikativen neutralen Element): Man kann (mindestens) ein Element $n_* \neq n_\nabla$ in $\mathbb{K}$ finden, so dass gilt: $\forall\, x \in \mathbb{K}$ ist $x*n_* = x$.

(M4) (Axiom vom multiplikativen Inversen): Zu jedem $x \neq n_\nabla$ aus $\mathbb{K}$ kann man (mindestens) ein $\overline{x} \in \mathbb{K}$ finden, so dass $x*\overline{x} = n_*$.

(D) Distributivität: $\forall\, x,y,z \in \mathbb{K}$ gilt $x*(y\nabla z) = x*y \;\nabla\; x*z$.

2.2 Bemerkung: Wir benutzen die übliche Konvention ‘Punkt vor Strich’, d.h. hier $*$ vor ∇. Genauer müsste (D) heißen: $\forall\, x,y,z \in \mathbb{K}$ gilt

$$x*(y\nabla z) = (x*y)\nabla(x*z).$$

2.3 Logische Zeichen: Wir haben in (2.1) und auch bereits in §1 logische Zeichen benutzt, die wir jetzt kurz erläutern:

$A \Rightarrow B$:	Aus (der Aussage) A folgt (die Aussage) B, d.h. A ist hinreichend für B.
$B \Leftarrow A$:	Aus (der Aussage) A folgt (die Aussage) B. Natürlich bedeutet das auch: Falls B falsch ist, ist A falsch, d.h. B ist notwendig für A.
$A \Leftrightarrow B$:	A ist äquivalent zu B, d.h. aus A folgt B und aus B folgt A.
$\forall\, x \in M$:	Für jedes (Element) x der Menge M (gilt) ...
$\exists\, x \in M$:	Es gibt ein (Element) x in der Menge M, (so dass gilt ...)
$\exists!\, x \in M$:	Es gibt genau ein (Element) x in M,...
$M := N$:	Die Menge M ist **definitionsgemäß** gleich der Menge N.

2.4 Warnung: Gehen Sie mit logischen Kürzeln sparsam und sorgfältig um. In verbindenden **Texten** haben sie nichts zu suchen.

2.5 Folgerungen aus den Axiomen 2.1 und Bezeichnungen:

(1) Es gibt genau ein Element $n_\nabla \in \mathbb{K}$, das (A3) erfüllt. Wir nennen dieses Element die **NULL** in $\mathbb{K}$ und bezeichnen es in Zukunft mit 0.

(2) Zu jedem $x \in \mathbb{K}$ gibt es genau ein Element $x' \in \mathbb{K}$, so dass $x\nabla x' = n_\nabla$, wir nennen dieses x' das **NEGATIVE ELEMENT** von x und bezeichnen es in Zukunft mit $-x$.

(3) Es gibt genau ein Element $n_* \in \mathbb{K}$, das (M3) erfüllt. Wir nennen dieses Element die **EINS** in $\mathbb{K}$ und bezeichnen es in Zukunft mit 1.

(4) Zu jedem $x \neq n_\nabla$ in $\mathbb{K}$ gibt es genau ein Element $\overline{x}$ mit $x * \overline{x} = n_*$. Wir nennen $\overline{x}$ das **INVERSE** von x und bezeichnen es in Zukunft mit x^{-1}.

Beweis:

(1) Sei m_∇ ein weiteres Element, das (A3) erfüllt, d.h.

$$\forall\, x \in \mathbb{K} \quad \text{gilt} \quad x \nabla m_\nabla = x$$

. Nehmen wir $x = n_\nabla$, erhalten wir unter Ausnutzung von (A3) und (A2)

$$n_\nabla = n_\nabla \;\nabla\; m_\nabla = m_\nabla \;\nabla\; n_\nabla = m_\nabla$$

(2) Sei $x \in \mathbb{K}$ und seien $\widetilde{x}, x'$ zwei Elemente, so dass $x \nabla \widetilde{x} = n_\nabla$ und $x \nabla x' = n_\nabla$.

$$\begin{aligned} \widetilde{x} &= \widetilde{x} \nabla n_\nabla = \widetilde{x} \nabla (x \nabla x') = (\widetilde{x} \nabla x) \nabla x' \\ &= (x \nabla \widetilde{x}) \nabla x' = n_\nabla \nabla x' = x' \nabla n_\nabla = x' \end{aligned}$$

Wie haben die Axiome (A1),...,(A4) ausgenutzt.

Die Beweise für (3) und (4) verlaufen genauso.

□

2.6 Bezeichnung: Es ist üblich, die Addition in Körpern stets mit $+$ und die Multiplikation mit $\cdot$ zu bezeichnen. Statt ∇ und $*$ schreiben wir daher fast immer $+$ und $\cdot$.

Für $x, y \in \mathbb{K}$ schreiben wir $x - y$ für $x + (-y)$ und auch

$$\frac{x}{y} \text{ für } x \cdot y^{-1} \text{ oder } y^{-1} \cdot x \qquad (y \neq 0)$$

2.7 Rechenregeln: In einem Körper $\mathbb{K}$ gilt

(1) $0 \cdot x = 0 \quad \forall\, x \in \mathbb{K}$

(2) $x \cdot y = 0 \iff x = 0$ oder $y = 0$

(3) $(-x)y = -(xy), \quad (-x)(-y) = xy, \quad -x = (-1) \cdot x$

(4) $\dfrac{x}{y} = \dfrac{u}{v} \iff x \cdot v = u \cdot y \quad (y \neq 0, v \neq 0)$

(5) $\dfrac{x}{y} + \dfrac{u}{v} = \dfrac{x \cdot v + u \cdot y}{y \cdot v} \quad (y \neq 0, v \neq 0)$

(6) $\dfrac{x}{y} \cdot \dfrac{u}{v} = \dfrac{x \cdot u}{y \cdot v} \quad (y \neq 0, v \neq 0)$

(7) $\left(\dfrac{x}{y}\right)^{-1} = \dfrac{y}{x}, \quad (x \cdot y)^{-1} = x^{-1} \cdot y^{-1} \quad (x \neq 0, y \neq 0)$

Der Beweis ist eine einfache Übung. Wir zeigen exemplarisch:

(1) $$\begin{aligned} 0\cdot x &= (0+0)\cdot x = 0\cdot x + 0\cdot x \qquad | -0\cdot x \\ 0 &= 0\cdot x \end{aligned}$$

(3) $x - x = 0 \Rightarrow x\cdot y + (-x)\cdot y = 0 \Rightarrow (-x)\cdot y$ ist negativ zu $x\cdot y$, also $(-x)\cdot y = -(x\cdot y)$

(5) $\frac{x}{y} + \frac{u}{v} = x\cdot y^{-1} + u\cdot v^{-1} = x\cdot y^{-1}\cdot v\cdot v^{-1} + u\cdot v^{-1}\cdot y\cdot y^{-1} = (xv)\cdot(y\cdot v)^{-1} + (u\cdot y)\cdot(y\cdot v)^{-1} = (x\cdot v + u\cdot y)\cdot(y\cdot v)^{-1} = \frac{xv+uy}{y\cdot v}$

□

2.8 Beispiel: Aus der Schule kennt man die Körper $\mathbb{Q}$ und $\mathbb{R}$.

2.9 Beispiel: Sei p eine fest gegebene Primzahl. Sei $\mathbb{F}_p$ eine Menge von p Elementen, die wir mit

$$[0], [1], \ldots, [p-1]$$

bezeichnen. Wir definieren eine Abbildung

$$\mathbb{Z} \to \mathbb{F}_p, \quad n \mapsto \text{“}n \mod p\text{''} := [r], \quad n \mapsto [r], \text{ genannt } n \mod p,$$

wobei r der positive Rest beim Teilen von n durch p ist.
Beispiel:

$$\begin{aligned} 25 \mod 13 &= [12] \qquad p = 13 \\ -5 \mod 7 &= [2] \qquad p = 7 \end{aligned}$$

Wir definieren auf $\mathbb{F}_p$ eine **Addition** $\oplus$ und eine **Multiplikation** $\odot$ durch

$$\begin{aligned} [r_1]\oplus[r_2] &= (r_1+r_2) \mod p \\ [r_1]\odot[r_2] &= (r_1\cdot r_2) \mod p \end{aligned}$$

Satz: $(\mathbb{F}_p, \oplus, \odot)$ ist ein Körper.
Beweis:

$$\begin{aligned} (k \mod p)\oplus(l \mod p) &= (k+l) \mod p \qquad (1) \\ (k \mod p)\odot(l \mod p) &= (k\cdot l) \mod p \end{aligned}$$

Denn sei $(k \mod p) = [r]$ und $(l \mod p) = [s]$, dann gibt es $a, b \in \mathbb{Z}$ mit $k = a\cdot p + r$, $l = b\cdot p + s$. Also

$$\begin{aligned} k + l &= (a+b)p + (r+s) \\ k\cdot l &= (a\cdot b\cdot p + a\cdot s + b\cdot r)\cdot p + r\cdot s \end{aligned}$$

Also haben $k+l$ und $r+s$ sowie $k \cdot l$ und $r \cdot s$ denselben Rest bei der Division durch p. Statt mit r und s können wir also mit k und l rechnen. Aus (1) folgen nun sofort die Körperaxiome aus den Rechenregeln für $\mathbb{Z}$ mit Ausnahme des Axioms (M4), wobei $[0]$ das Nullelement und $[1]$ das Einselement ist. Wir zeigen exemplarisch das Axiom (D)

$$\begin{aligned}
[r_1] \odot ([r_2] \oplus [r_3]) &= (r_1 \mod p) \odot ((r_2 + r_3) \mod p) && \text{nach Def.} \\
&= r_1 \cdot (r_2 + r_3) \mod p && \text{nach (1)} \\
&= (r_1 \cdot r_2 + r_1 \cdot r_3) \mod p \\
([r_1] \odot [r_2]) \oplus ([r_1] \odot [r_3]) &= (r_1 \cdot r_2 \mod p) \oplus (r_1 \cdot r_3) \mod p && \text{nach Def.} \\
&= (r_1 \cdot r_2 + r_1 \cdot r_3) \mod p && \text{nach (1)}
\end{aligned}$$

Für den Nachweis von (M4) zeigen wir zunächst: Ist

$$0 < r < p \text{ und } [r] \odot [s] = [0] \text{ so ist } s = 0. \tag{2}$$

Beweis: $[r] \odot [s] = [0] \iff r \cdot s \mod p = [0] \Rightarrow p$ teilt $r \cdot s$. Da $0 < r < p$, ist p kein Teilern von r. Also ist p Teiler von s. Da $0 \leq s < p$, folgt $s = 0$. Hier geht ein, dass p eine Primzahl ist.

Nachweis von (M4): Sei $0 < r < p$, also $[r] \neq [0]$. Die Abbildung

$$f : \mathbb{F}_p \to \mathbb{F}_p, \quad [x] \mapsto [r] \odot [x]$$

ist **INJEKTIV**, d.h. ist $[x] \neq [y]$, so ist $f([x]) \neq f([y])$:
Denn sei $f([x]) = f([y])$, also $[x] \odot [r] = [y] \odot [r]$. Dann folgt aus den bereits bewiesenen Axiomen, dass

$$[r] \odot ([x] \ominus [y]) = [0], \text{ also } [x] \ominus [y] = [0] \quad \text{nach (2)}$$

und somit $[x] = [y]$. Da $\mathbb{F}_p$ genau p Elemente hat, und verschiedene Elemente auf verschiedene Elemente abgebildet werden, muss es ein $[s]$ in $\mathbb{F}_p$ geben mit $f[s] = [1]$, denn jedes $[x] \in \mathbb{F}_p$ muss als Bild auftreten. Für dieses $[s]$ gilt

$$[r] \odot [s] = [1].$$

$\square$

Bemerkung: Der Körper $\mathbb{F}_p$ erscheint auf den ersten Blick als recht künstlich konstruiert. Er spielt aber sowohl in der Mathematik als auch in ihren Anwendungen eine große Rolle. Zu diesen Anwendungen gehört die Multiplikation großer Zahlen in Computern.

2.10 Beispiel: Der Körper $\mathbb{C}$ der komplexen Zahlen.

Für die Lösung quadratischer Gleichungen

$$x^2 + px + q = 0$$

gilt die bekannte Formel

$$x = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}.$$

Man sieht sofort, dass die Gleichung $x^2 - 2x + 3 = 0$ keine reelle Lösung hat; die Formel ergibt

$$x = 1 \pm \sqrt{-2}$$

In der modernen Mathematik würde man versuchen, den Zahlbereich künstlich zu erweitern, damit solche unlösbaren Gleichungen eine Lösung bekommen und man so eine einheitliche Theorie der quadratischen Gleichungen erhält.

In der frühen Gleichungslehre war man aber von solchen abstrakten Gedankengängen noch weit entfernt. Man erklärte die Gleichung $x^2 - 2x + 3 = 0$ als unlösbar. Deshalb wurden Mathematiker erst bei der Untersuchung kubischer Gleichungen im 16. Jahrhundert ernsthaft mit komplexen Zahlen konfrontiert.

1545 entwickelten Cardano, Ferro und Tartiglia eine Lösungsformel für kubische Gleichungen

$$x^3 + bx^2 + cx + d = 0.$$

Wendet man diese Formel auf die Gleichung

$$x^3 - 15x - 4 = 0$$

an, erhält man als Lösungen

$$\begin{aligned} x_1 = u + v, \quad x_2 &= (-\tfrac{1}{2} + \tfrac{\sqrt{-3}}{2})u - (\tfrac{1}{2} + \tfrac{\sqrt{-3}}{2})v \\ x_3 &= -(\tfrac{1}{2} + \tfrac{\sqrt{-3}}{2})u + (-\tfrac{1}{2} + \tfrac{\sqrt{-3}}{2})v \end{aligned}$$

mit $u = \sqrt[3]{2 + \sqrt{-121}}$ und $v = \sqrt[3]{2 - \sqrt{-121}}$.

Diese Ausdrücke sind im Reellen sinnlos, weil man aus negativen Zahlen keine Wurzel ziehen kann. Aber man sieht sofort, dass 4 eine Lösung der Gleichung ist. D.h. einer dieser Ausdrücke muss sich in einer geeigneten Zahlbereichserweiterung von $\mathbb{R}$ in 4 umrechnen lassen. Wir erhalten also aus einem “sinnlosen” Ausdruck ein sinnvolles Ergebnis. Bereits 1560 rechnete Bombelli so systematisch mit komplexen Zahlen, aber erst mit der expliziten Beschreibung von Hamilton 1837 wurden sie endgültig entmystifiziert.

Dabei ist das Rechnen in geeigneten Zahlbereichserweiterungen durchaus nahe liegend: Wollen wir ein ganzzahliges lineares Gleichungssystem nach ganzzahligen Lösungen untersuchen, lösen wir es erst über dem Körper $\mathbb{Q}$ der rationalen Zahlen und untersuchen dann, ob die Lösungsmenge ganzzahlige Lösungen enthält.

Konstruktion von $\mathbb{C}$: $\mathbb{C} = \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ als Menge. Addition und Multiplikation sind definiert durch

$$\begin{array}{rcl} (x_1,y_1)+(x_2,y_2) & = & (x_1+x_2,y_1+y_2) \\ (x_1,y_1)\cdot(x_2,y_2) & = & (x_1x_2-y_1y_2,x_1y_2+y_1x_2) \end{array}$$

Die Körperaxiome rechnet man leicht nach. Das Nullelement ist $(0,0)$, das Einselement ist $(1,0)$. Nur (M4) bereitet wieder etwas Mühe.

Sei $(x,y) \neq (0,0)$, d.h. x und y sind nicht gleichzeitig 0. Also ist $x^2+y^2 \neq 0$. Man rechnet leicht nach, dass

$$\left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right) = (x,y)^{-1}$$

ist.

2.11 Definition: Sei $\mathbb{K}$ ein Körper $\mathbb{F} \subset \mathbb{K}$ Teilmenge. $\mathbb{F}$ heißt **UNTERKÖRPER** von $\mathbb{K}$, wenn $\mathbb{F}$ unter der Addition und Multiplikation $\mathbb{K}$ wieder ein Körper ist. Man nennt dann auch $\mathbb{K}$ eine **KÖRPERERWEITERUNG** von $\mathbb{F}$.

2.12 Beispiel: $\mathbb{R}$ kann als Unterkörper von $\mathbb{C}$ aufgefasst werden

$$\mathbb{R} = \{(x,0) \in \mathbb{C}; x \in \mathbb{R}\}$$

Denn
$$\begin{array}{rcl} (x_1,0)+(x_2,0) & = & (x_1+x_2,0) \\ (x_1,0)\cdot(x_2,0) & = & (x_1\cdot x_2,0) \\ 0_{\mathbb{C}} & = & (0_{\mathbb{R}},0) \\ -(x,0) & = & (-x,0) \\ 1_{\mathbb{C}} & = & (1_{\mathbb{R}},0) \\ (x,0)^{-1} & = & (x^{-1},0) \end{array}$$

Setzen wir $i = (0,1_{\mathbb{R}})$, erhalten wir $i^2 = -1_{\mathbb{C}}$. Ist $x \in \mathbb{R}$, schreiben wir statt $(x,0)$ kürzer nur x. Mit dieser Konvention gilt

$$z = (x,y) = x + y \cdot i.$$

Diese Darstellung ist eindeutig.

2.13 Bezeichnung: Sei $z = (x,y) = x+iy \in \mathbb{C}$

(1) $x = Re(z)$ heißt **REALTEIL** und $y = Im(z)$ **IMAGINÄRTEIL** von z

(2) Durch $z = (x,y)$ wird $\mathbb{C}$ mit der Ebene $\mathbb{R}^2$ identifiziert. Die x-Achse heißt dann **REELLE ACHSE** $\mathbb{R}$, die y-Achse **IMAGINÄRE ACHSE** $i \cdot \mathbb{R}$.

(3) $\bar{z} = x - iy$ heißt zu z **KONJUGIERTE KOMPLEXE ZAHL**.

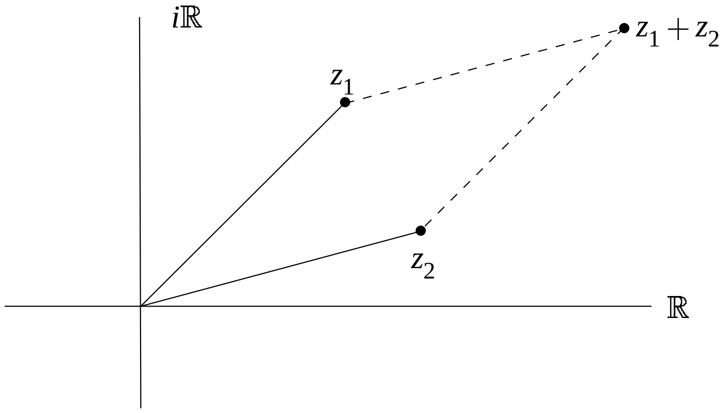
(4) $|z| := {}_{+}\sqrt{z \cdot \overline{z}} = \sqrt{x^2 + y^2}$ heißt **BETRAG** von z.

2.14 Für $z_1, z_2 \in \mathbb{C}$ gilt $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$ und $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$

Beweis: Nachrechnen $\square$

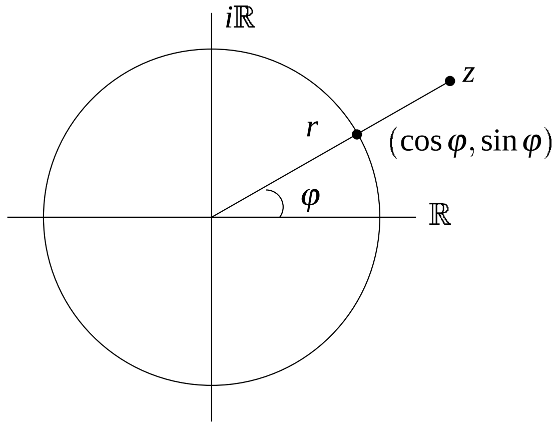
Geometrische Beschreibung

2.15 Addition:



Für die Multiplikation gehen wir zur Polarkoordinatendarstellung über

2.16 Polarkoordinatendarstellung: $z = (x, y) = x + iy \in \mathbb{C}$. Ist $z \neq 0$, so ist $r := |z| > 0$ und $|\frac{z}{r}| = 1$. Die Elemente $z' \in \mathbb{C}$ mit $|z'| = 1$ liegen nach (2.13.4) auf dem Einheitskreis. Insbesondere gibt es ein $\varphi \in \mathbb{R}$, so dass



$$\frac{z}{r} = \cos\varphi + i \sin\varphi$$

$$\boxed{z = r(\cos\varphi + i\sin\varphi)}$$

Ist nun $w = r'(\cos\vartheta + i\sin\vartheta)$ eine weitere komplexe Zahl in Polarkoordinaten, dann folgt aus den Additionstheoremen für Sinus und Cosinus:

$$\begin{aligned} \underline{\underline{z \cdot w}} &= r \cdot r'[\cos\varphi\cos\vartheta - \sin\varphi\sin\vartheta + i(\cos\varphi\sin\vartheta + \sin\varphi\cos\vartheta)] \\ &= \underline{\underline{r \cdot r'(\cos(\varphi + \vartheta) + i\sin(\varphi + \vartheta))}}. \end{aligned}$$

3 Grundlagen: Mengen und Abbildungen

Wir werden in dieser Vorlesung immer wieder auf Grundlagen zurückgreifen, die wir nicht erarbeitet haben. Diese werden als Ergänzungen nachträglich eingeschoben, damit nicht auf Vorrat gelernt werden muss. Für einen mehr systematischen Aufbau sei auf die Literatur hingewiesen.

Wir setzen den Begriff der **MENGE** als bekannt voraus. Eine Menge ist so definiert, dass es klar ist, wann ein Element zu ihr gehört und wann nicht.

3.1 Bezeichnung: $\mathbb{N} = \{0,1,2,\dots\}$, die Menge der natürlichen Zahlen.
$\mathbb{Z} = \{0,\pm1,\pm2,\pm3,\dots\}$, die Menge der ganzen Zahlen.
$\mathbb{Q} = \{\frac{a}{b}; a \in \mathbb{Z}, b \neq 0$ aus $\mathbb{N}\}$, die Menge der rationalen Zahlen.
$\mathbb{R}$, die Menge der reellen Zahlen.
$\mathbb{C}$, die Menge der komplexen Zahlen.
$\emptyset$, die leere Menge, charakterisiert durch: $x \notin \emptyset \ \forall x$.

3.2 Definition: Seien M,N Mengen. N ist **Teilmenge** von M, in Zeichen $N \subset M$, wenn gilt: $x \in N \Rightarrow x \in M$.

3.3 Für Mengen M,N,L gilt

(1) $\emptyset \subset M$

(2) $M \subset M$

(3) $M \subset N$ und $N \subset L \Rightarrow M \subset L$

(4) $M = N \iff M \subset N$ und $N \subset M$

Beweis: trivial $\square$

3.4 Definition: Sei J eine Menge (genannt **Indexmenge**). Für jedes $j \in J$ sei eine Menge M_j gegeben. Dann heißt

$$\bigcup_{j\in J} M_j = \{x; \exists\, j \in J \text{ mit } x \in M_j\} \quad \text{die } \textbf{VEREINIGUNG}$$
$$\bigcap_{j\in J} M_j = \{x; x \in M_j\ \forall\, j \in J\} \quad \text{der } \textbf{DURCHSCHNITT}$$

der M_j, $j \in J$. Ist $J = \{1,2,\dots,n\}$ bzw. $J = \mathbb{N}$ schreiben wir auch

$$\bigcup_{j\in J} M_j = \bigcup_{j=1}^{n} M_j = M_1, \cup \ldots \cup M_n \quad \text{bzw.} \quad \bigcup_{j\in \mathbb{N}} M_j = \bigcup_{j=1}^{\infty} M_j$$
$$\bigcap_{j\in J} M_j = \bigcap_{j=1}^{n} M_j = M_1, \cap \ldots \cap M_n \quad \text{bzw.} \quad \bigcap_{j\in \mathbb{N}} M_j = \bigcap_{j=1}^{\infty} M_j$$

Man beachte, dass wir zwei Darstellungsweisen für Mengen benutzen: Die **aufzählende** Darstellung, etwa bei $\mathbb{N}$ oder $\mathbb{Z}$, und die **charakterisierende** Darstellung, die durch eine Aussageform gegeben ist.

Sei M eine Menge und $A(x)$ eine Aussage über $x \in M$. Dann ist

$$N = \{x \in M; A(x)\}$$

die Teilmenge aller $x \in M$, für die die Aussage $A(x)$ wahr ist.

Beispiel: $\{x \in \mathbb{Z}; \exists\, y \in \mathbb{Z} \text{ mit } x \cdot y = 1\} = \{\pm 1\}$

3.5 Distributivgesetz: Sei $J \neq \emptyset$ eine Indexmenge, M_j, $j \in J$, und N seien Mengen. Dann gilt:

(1) $N \cap \bigcup\limits_{j \in J} M_j = \bigcup\limits_{j \in J} (N \cap M_j)$

(2) $N \cup \bigcap\limits_{j \in J} M_j = \bigcap\limits_{j \in J} (N \cup M_j)$

Beweis: $x \in N \cap \bigcup\limits_{j \in J} M_j \iff x \in N$ **und** $\exists\, j \in J$ mit $x \in M_j \iff \exists\, j \in J$ mit $x \in N \cap M_j \iff x \in \bigcup\limits_{j \in J} (N \cap M_J)$

Das beweist (1). Zum Beweis von (2)

$x \in N \cup \bigcap\limits_{j \in J} M_j \iff x \in N$ oder $x \in M_j\ \forall\, j \in J \iff \forall j \in J$ gilt: $x \in N$ oder $x \in M_j \iff \forall\, j \in J$ gilt $x \in N \cap M_j \iff x \in \bigcap\limits_{j \in J} N \cup M_j)$ $\square$

3.6 Definition: Für Mengen M und N definieren wir die **DIFFERENZMENGE** $M \backslash N$ durch

$$M \backslash N = \{x \in M; x \notin N\}.$$

3.7 de Morgan'sche Regeln:

(1) $M \backslash \bigcup\limits_{j \in J} N_j = \bigcap\limits_{j \in J} (M \backslash N_j)$

(2) $M \backslash \bigcap\limits_{j \in J} N_j = \bigcup\limits_{j \in J} (M \backslash N_j)$

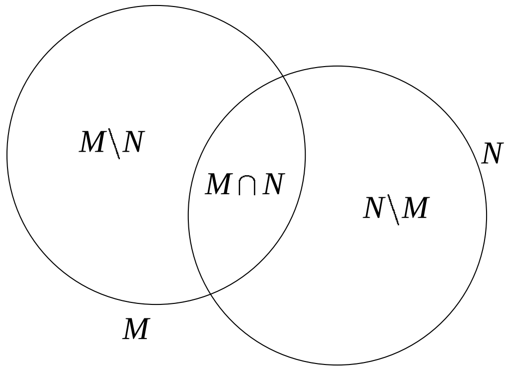
Beweis: Übung $\square$

3.8 Definition: Zwei Mengen M und N heißen **DISJUNKT**, wenn $M \cap N = \emptyset$. Sind M und N disjunkt, nennen wir $M \cup N$ die **DISJUNKTIVE VEREINIGUNG** und kennzeichnen das durch das Symbol $M \sqcup N$.

3.9 Sind M und N beliebige Mengen, so gilt

$$\begin{aligned} M \cup N &= (M \backslash N) \sqcup (N \backslash M) \sqcup (M \cap N) \\ M &= (M \backslash N) \sqcup (M \cap N) \\ M \cap N &= M \backslash (M \backslash N) \end{aligned}$$

Beweis: Übung



□

Warnung: Ein solches Diagramm dient der Anschauung, ist aber kein Beweis!

3.10 Definition: Das **KARTESISCHE PRODUKT** $M \times N$ zweier Mengen M und N ist

$$M \times N = \{(x,y); x \in M, y \in N\}.$$

Allgemeiner definieren wir für Mengen $M_1, M_2, \ldots, M_n$ das **KARTESISCHE PRODUKT**

$$M_1 \times M_2 \times \ldots \times M_n = \{(x_1, x_2, \ldots, x_n);\ x_i \in M_i,\ \text{für } i = 1, \ldots, n\}$$

Das Produkt haben wir bereits in §2 verwendet.

3.11 Definition: Eine **ABBILDUNG** f von der Menge M in eine Menge N, in Zeichen $f: M \to N$, ist eine Vorschrift, die jedem $x \in M$ genau ein $y = f(x) \in N$ zuordnet. $f(x)$ heißt **BILD** von x unter f, wir schreiben auch

$$x \mapsto f(x)$$

M heißt **DEFINITIONSBEREICH** von f und N heißt **WERTEBEREICH** von f.

Damit sind zwei Abbildungen $f: M \to N$ und $g: K \to L$ genau dann gleich, wenn $M = K$, $N = L$ und $f(x) = g(x)\ \forall x \in M = K$.

3.12 Definition und Satz: Ist $f : M \to N$ eine Abbildung, dann heißt

$$G_f := \{(x,y) \in M \times N; \quad y = f(x)\}$$

der **GRAPH** von f. Eine Teilmenge $G \subset M \times N$ ist genau dann Graph einer Abbildung, wenn gilt

$$x \in M \Rightarrow \exists!\, y \in N \text{ mit } (x,y) \in G$$

$\square$

Auf Grund von 3.12 definiert man eine Abbildung auch oft über ihre Graphen.

3.13 Definition: Sei $f : M \to N$ eine Abbildung und $A \subset M$. Dann heißt

$$f(A) := \{f(x); x \in A\} = \{y \in N; \exists\, x \in A \text{ mit } y = f(x)\} \subset N$$

das **BILD** von A unter f. Wir nennen $f(M) =: Bild\ f$ das **BILD** von f. Für $B \subset N$ heißt

$$f^{-1}(B) := \{x \in M; f(x) \in B\}$$

die **URBILDMENGE** von B unter f.

3.14 Beispiel: Sei $f : \mathbb{R}^2 \to \mathbb{R}$ die Multiplikation, $f(x,y) = x \cdot y$. Sei $A_x = \{(x,y) \in \mathbb{R}^2, y \in \mathbb{R}\}$. Dann gilt

$$f(A_0) = \{0\}, \quad f(A_x) = \mathbb{R} \quad \forall\, x \neq 0$$
$$f^{-1}(\{0\}) \text{ ist das Achsenkreuz}$$
$$f^{-1}(\{c\}) \text{ ist die Hyperbel } y = \tfrac{c}{x} \quad c \neq 0$$

Wir kommen nun zu einem wichtigen Begriff, den wir auch schon benutzt haben.

3.15 Definition: Eine Abbildung $f : M \to N$ heißt

(1) **INJEKTIV**, wenn gilt: $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$ (anders formuliert: $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$)

(2) **SURJEKTIV**, wenn es zu jedem $y \in N$ ein $x \in M$ gibt, so dass $f(x) = y$.

(3) **BIJEKTIV**, wenn f injektiv und surjektiv ist.

Ist $f : M \to N$ bijektiv, dann gibt es zu $y \in N$ ein $x \in M$ mit $f(x) = y$, da f surjektiv ist. Da f injektiv ist, gibt es kein weiteres $x' \in M$ mit $f(x') = y$. Also gibt es genau ein x mit $f(x) = y$. Damit ist

$$f^{-1} : N \to M, \quad y = f(x) \mapsto x$$

eine Abbildung.

3.16 Definition: Ist f bijektiv, dann heißt die Abbildung $f^{-1}: N \to M$ die **UMKEHRABBILDUNG** von f

Beachte: $f^{-1}(y) = x \iff f(x) = y$

Injektivität und Surjektivität von Abbildungen haben oft interessante Konsequenzen. Daher wollen wir einige Kriterien für diese Eigenschaften herleiten. Dazu benötigen wir den Begriff der Komposition.

3.17 Definition: Die **KOMPOSITION** zweier Abbildungen $f: K \to M$, $g: M \to N$ ist die Abbildung

$$g \circ f : K \to N, \quad (g \circ f)(x) = g(f(x)).$$

Die Abbildung

$$id_M : M \to M, \quad x \mapsto x$$

heißt die **IDENTITÄT** von M.

3.18 Die Komposition von Abbildungen ist assoziativ und hat die Identitäten als "neutrale Elemente". D.h. für Abbildungen $f: K \to L$, $g: L \to M$ und $h: M \to N$ gilt:

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

(2) $f \circ id_K = id_L \circ f = f$

Beweis: (1)

$$\begin{array}{lclcl} (h \circ (g \circ f))(x) & = & (h((g \circ f)(x))) & = & h(g(f(x))) \\ ((h \circ g) \circ f)(x) & = & (h \circ g)(f(x)) & = & h(g(f(x))) \end{array}$$

Da $h \circ (g \circ f)$ und $(h \circ g) \circ f$ beide den Definitionsbereich K und den Wertebereich N haben, sind sie somit gleich.

Analog zeigt man die 2. Aussage. □

3.19 Gegeben seien Abbildungen $f: K \to L$ und $g: L \to M$.

(1) Sind f und g injektiv, so ist auch $g \circ f$ injektiv.

(2) Ist $g \circ f$ injektiv, so ist f injektiv.

(3) Sind f und g surjektiv, so ist auch $g \circ f$ surjektiv.

(4) Ist $g \circ f$ surjektiv, so ist g surjektiv.

Beweis: (1): Aus $(g \circ f)(x_1) = (g \circ f)(x_2)$, also $g(f(x_1)) = g(f(x_2))$, folgt $f(x_1) = f(x_2)$, weil g injektiv ist, und somit $x_1 = x_2$, weil f injektiv ist. Damit ist $g \circ f$ injektiv.

(2): Für $x_1, x_2 \in K$ gilt

$$f(x_1) = f(x_2) \Rightarrow g(f(x_1)) = g(f(x_2)) \Rightarrow (g \circ f)(x_1) = (g \circ f)(x_2)$$

Da $g \circ f$ injektiv ist, folgt $x_1 = x_2$. Also ist f injektiv.

(3): Sei $z \in M$. Da g surjektiv ist, gibt es ein $y \in L$ mit $g(y) = z$. Da f surjektiv ist, gibt es ein $x \in K$ mit $f(x) = y$. Für dieses x gilt $(g \circ f)(x) = g(f(x)) = g(y) = z$. Also ist $g \circ f$ surjektiv.

(4): Sei $z \in M$. Da $g \circ f$ surjektiv ist, gibt es ein $x \in K$ mit $g(f(x)) = z$. Für $y = f(x) \in L$ gilt dann $g(y) = z$. Also ist g surjektiv. □

3.20 Satz: Sei $f : M \to N$ eine Abbildung von **nicht-leeren** Mengen. Dann gilt

(1) f ist injektiv $\iff$ $\exists$ Abbildung $g : N \to M$, so dass $g \circ f = id_M$

(2) f ist surjektiv $\iff$ $\exists$ Abbildung $h : N \to M$, so dass $f \circ h = id_N$

(3) f ist bijektiv $\iff$ $\exists$ Abbildung $g : N \to M$, so dass $g \circ f = id_M$ und $f \circ g = id_N$.

3.21 Bezeichnung: Die Abbildung g aus (1) nennt man **LINKSINVERS** zu f, die Abbildung h aus (2) nennt man **RECHTSINVERS** zu f und die Abbildung g aus (3) **INVERS** zu f.

Beweis: Beachte zunächst, dass die Identität id_M immer bijektiv ist. Damit folgt die Rückrichtung "$\Leftarrow$" von (1), (2) und (3) aus (3.19).

(1) "$\Rightarrow$" Wähle ein $x_0 \in M$ fest. Definiere $g : N \to M$ wie folgt

$$g(y) = \begin{cases} x, & \text{falls } f(x) = y \\ x_0, & \text{falls } y \notin f(M) \end{cases}$$

Dann ist g "wohldefiniert", weil es höchstens ein $x \in M$ mit $f(x) = y$ gibt, und es gilt $(g \circ f)(x) = g(y) = x$ mit $y = f(x)$, also $g \circ f = id_M$.

(2) "$\Rightarrow$" Da f surjektiv ist, gibt es zu jedem ein $y \in N$ mindestens ein $x \in M$ mit $f(x) = y$. Wir **wählen** für jedes y ein solches x und nennen es $g(y)$. Dadurch erhalten wir eine Abbildung $g : N \to M$. Da $f(x) = y$ für dieses $x = g(y)$, folgt $f(x) = f(g(y)) = y$. Also $f \circ g = id_N$.

(3) "$\Rightarrow$" nach (1) und (2) gibt es Abbildungen $g,h : N \to M$, so dass $f \circ h = id_N$ und $g \circ f = id_M$. Damit folgt die Behauptung aus

3.22 Ist $f : M \to N$ eine Abbildung und gibt es Abbildungen $g,h : N \to M$ mit $f \circ h = id_N$ und $g \circ f = id_M$, dann folgt $g = h$. Insbesondere hat f höchstens ein Inverses, die Umkehrabbildung f^{-1}.

Beweis: $h = id_M \circ h = (g \circ f) \circ h = g \circ (f \circ h) = g \circ id_N = g$. □

3.23 Folgerung: Ist $f : M \to N$ bijektiv, dann auch $f^{-1} : N \to M$ und es gilt $(f^{-1})^{-1} = f$.
Sind $g : L \to M$ und $f : M \to N$ bijektiv, dann auch $f \circ g$, und es gilt

$$(f \circ g)^{-1} = g^{-1} \circ f^{-1}.$$

Beweis: Es gilt $f^{-1} \circ f = id_M$ und $f \circ f^{-1} = id_N$. Nach (3.20) ist f^{-1} bijektiv mit Umkehrabbildung f. Damit ist der erste Teil gezeigt.
Weiter gilt $g^{-1} \circ g = id_L$ und $g \circ g^{-1} = id_M$. Es folgt

$$\begin{aligned}(g^{-1} \circ f^{-1}) \circ (f \circ g) &= g^{-1} \circ (f^{-1} \circ f) \circ g = g^{-1} \circ id_M \circ g = g^{-1} \circ g = id_L\\ (f \circ g) \circ (g^{-1} \circ f^{-1}) &= f \circ (g \circ g^{-1}) \circ f^{-1} = f \circ f^{-1} = id_N\end{aligned}$$

Also ist $f \circ g$ nach (3.20) bijektiv mit Umkehrabbildung $g^{-1} \circ f^{-1}$. □
Wir erhalten nun eine merkwürdige Charakterisierung endlicher Mengen. Einen Teil des Satzes haben wir bereits im Beispiel 2.9 benutzt.

3.24 Satz: Für eine Menge sind folgende Aussagen äquivalent

(1) M ist endlich.

(2) Jede injektive Abbildung $f : M \to M$ ist bijektiv.

(3) Jede surjektive Abbildung $g : M \to M$ ist bijektiv.

Beweis: (1) $\Rightarrow$ (3): M habe n Elemente. Gilt $g(x_1) = g(x_2)$ für $x_1 \neq x_2$, so hat $\mathrm{Bild}\, f = f(M)$ höchstens $n-1$ Elemente. Da g surjektiv ist, gilt aber $\mathrm{Bild}\, f = M$.
(3) $\Rightarrow$ (2): Ist $f : M \to M$ injektiv, gibt es nach (3.20) ein $g : M \to M$ mit $g \circ f = id_M$. Nach (3.19) ist g surjektiv und damit nach Voraussetzung bijektiv. Es folgt

$$g^{-1} = g^{-1} \circ id_M = g^{-1} \circ (g \circ f) = (g^{-1} \circ g) \circ f = f.$$

Also ist f die Umkehrabbildung von g und damit ebenfalls bijektiv.

(2) $\Rightarrow$ (1): Angenommen M ist unendlich. Dann gibt es eine injektive Abbildung $h : \mathbb{N} \to M$. Definiere

$$f : M \to M, \quad x \mapsto \begin{cases} x & \text{, falls } x \notin h(\mathbb{N}) \\ h(2n) & \text{, falls } x = h(n) \end{cases}$$

f ist injektiv, aber nicht surjektiv, da $h(1)$ nicht im Bild von f liegt. $\square$

4 Vektorräume

Aus §1 und §2 wissen wir, dass wir mit dem Gauß-Schema lineare Gleichungssysteme über beliebigen Körpern $\mathbb{K}$ lösen können. Nach einem Einschub über mathematische Grundlagen – solche Einschübe wird es häufiger geben – kehren wir zur Betrachtung linearer Gleichungssysteme zurück. Gegeben sei also ein lineares Gleichungssystem über $\mathbb{K}$

4.1

$$\begin{array}{cccccc} a_{11}x_1 & +\cdots+ & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +\cdots+ & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & +\cdots+ & a_{mn}x_n & = & b_m \end{array} \qquad a_{ij} \in \mathbb{K},\ b_i \in \mathbb{K}$$

Dann ist

4.2

$$\begin{array}{ccccc} a_{11}x_1 & +\cdots+ & a_{1n}x_n & = & 0 \\ \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & +\cdots+ & a_{mn}x_n & = & 0 \end{array}$$

das zugehörige homogene System.

Wir wissen, wie man solche Gleichungssysteme löst, haben aber keine Vorstellung darüber, was Lösbarkeit bedeutet, noch wie die Lösungsmenge aussieht. Vorstellung bedeutet hier das geometrische Bild der Lösungsmenge. Wir haben also weder den Begriff der Lösbarkeit noch die Lösungsmenge richtig verstanden. Für die Entwicklung einer Theorie der linearen Gleichungssysteme ist dieses Verständnis aber Voraussetzung.

Um hier weiterzukommen, untersuchen wir die **Struktur** der Lösungsmenge, eine für die Mathematik typische Vorgehensweise.

Sei L die Lösungsmenge von (4.1).

4.3 Satz: Sei $(z_1,\dots,z_n)$ eine (spezielle) Lösung des Systems (4.1). Dann gilt:

$(x_1,\dots,x_n) \in L \Longleftrightarrow$ es gibt eine Lösung $(y_1,\dots,y_n)$ des zugehörigen homogenen Systems, so dass $x_i = y_i + z_i,\ i = 1,\dots,n$

4.4 Um die Lösungsmenge zu bestimmen, müssen wir also **eine spezielle** Lösung des inhomogenen und **alle Lösungen** des zugehörigen homogenen Systems bestimmen.

Zur Vereinfachung der Notation benutzen wir (kleiner Einschub):

4.5 Das Summenzeichen

$\sum_{i=k}^{n} a_i$ bedeutet $a_k + a_{k+1} + \cdots + a_n, \quad k \leq n$

$\sum_{j=1}^{n} a_{ij}x_j$ bedeutet $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$

4.6 Rechenregeln für das Summenzeichen

(1) $\sum_{i=k}^{n} c \cdot a_i = c \cdot \sum_{i=k}^{n} a_i$ (Ausklammern)

(2) $\sum_{i=k}^{n} (a_i + b_i) = \sum_{i=k}^{n} a_i + \sum_{i=k}^{n} b_i$ (Umsummieren)

(3) $\sum_{i=k}^{n} a_i = \sum_{i=k}^{m} a_i + \sum_{i=m+1}^{n} a_i \quad k \leq m < n$ (Zerlegung der Summe)

(4) $\sum_{i=k}^{m} a_{i+\ell} = \sum_{j=k+\ell}^{m+\ell} a_j$ (Umindizieren)

(5) $\sum_{i=k}^{m} \left(\sum_{j=\ell}^{n} a_{ij} \right) = \sum_{j=\ell}^{n} \left(\sum_{i=k}^{m} a_{ij} \right)$ (Doppelsummenregel)

Beweis: 4.3: „$\Longrightarrow$“: $(x_1, \dots, x_n) \in L \Longrightarrow \sum_{j=1}^{n} a_{ij}x_j = b_i, \; i = 1, \dots, m$

Nach Voraussetzung gilt: $\sum_{j=1}^{n} a_{ij}z_j = b_i, \; i = 1, \dots, m$. Mit $y_i = x_i - z_i$ folgt

$$\sum_{j=1}^{n} a_{ij}y_j = \sum_{j=1}^{n} a_{ij}(x_j - z_j) = \sum_{j=1}^{n} a_{ij}x_j - \sum_{j=1}^{n} a_{ij}z_j = b_i - b_i = 0$$

für $i = 1, \dots, m$. Also löst $(y_1, \dots, y_n)$ das homogene System.
„$\Longleftarrow$“ Sei $(y_1, \dots, y_n)$ eine Lösung des homogenen Systems, d.h.

$$\sum_{j=1}^{n} a_{ij}y_j = 0 \qquad i = 1, \dots, m.$$

Dann gilt für $x_i = y_i + z_i$

$$\sum_{j=1}^{n} a_{ij}x_j = \sum_{j=1}^{n} a_{ij}(y_j + z_j) = \sum_{j=1}^{n} a_{ij}y_j + \sum_{j=1}^{n} a_{ij}z_j = 0 + b_i = b_i.$$

Also ist $(x_1, \ldots, x_n) \in L$. $\square$

Das Gauß-Verfahren liefert uns sofort eine spezielle Lösung. Aus (1.13) folgern wir

4.7 Mit der Bezeichnungsweise aus dem reduzierten Schema (1.13) erhalten wir eine spezielle Lösung $(z_1, \ldots, z_n)$ von (4.1), nämlich

$$\begin{aligned} z_{q_i} &= b_i^* \quad \text{für } i = 1, \ldots, r \\ z_{q_i} &= 0 \quad \text{für } i = r+1, \ldots, n. \end{aligned}$$

Nach (4.3) kommt der Lösungsmenge L^0 des homogenen Gleichungssystems (4.2) besondere Bedeutung zu.

4.8 Satz: Sind $(x_1, \ldots, x_n)$ und $(y_1, \ldots, y_n)$ Lösungen des homogenen Systems (4.2), dann sind auch $(x_1 + y_1, \ldots, x_n + y_n)$ und $(cx_1, \ldots, cx_n)$ mit beliebigem $c \in \mathbb{K}$ Lösungen von (4.2).

Beweis: Nach Voraussetzung gilt für $i = 1, \ldots, m$

$$\sum_{j=1}^{n} a_{ij}x_j = 0 \qquad \sum_{j=1}^{n} a_{ij}y_j = 0.$$

Also

$$\begin{aligned} \sum_{j=1}^{n} a_{ij}(x_j + y_j) &= \sum_{j=1}^{n} a_{ij}x_j + \sum_{j=1}^{n} a_{ij}y_j = 0 + 0 = 0 \\ \sum_{j=1}^{n} a_{ij}(c \cdot x_j) &= \sum_{j=1}^{n} c \cdot a_{ij}x_j = c \cdot \sum_{j=1}^{n} a_{ij}x_j = c \cdot 0 = 0. \end{aligned}$$

$\square$

(4.3) und (4.4) legen nun folgende Bezeichnungen und Rechenoperationen auf $\mathbb{K}^n$ nahe:

4.9 Bezeichnung: Für ein n-Tupel $(x_1, \ldots, x_n) \in \mathbb{K}^n$ schreiben wir kürzer x, also $x = (x_1, \ldots, x_n)$. Die Zahl $x_j \in \mathbb{K}$ heißt j-te **KOORDINATE** von x.

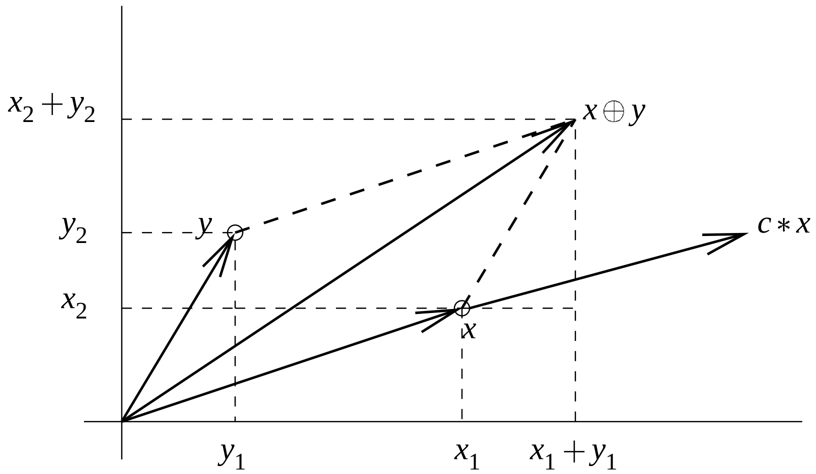
4.10 Definition: Auf $\mathbb{K}^n$ definieren wir eine **ADDITION**

$$\mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K}^n, \quad (x,y) \longmapsto x \oplus y$$

und eine **SKALARE MULTIPLIKATION**

$$\mathbb{K} \times \mathbb{K}^n \longrightarrow \mathbb{K}^n, \quad (c,x) \longmapsto c * x$$

durch $x \oplus y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$ und $c * x = (c \cdot x_1, c \cdot x_2, \ldots, c \cdot x_n)$ für $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$.
Diese Rechenoperationen kann man sich im $\mathbb{R}^1, \mathbb{R}^2, \mathbb{R}^3$ zeichnerisch veranschaulichen. Wir machen das für den $\mathbb{R}^2$.



$x + y$ erhält man als Diagonale des durch die Punkte x, y und 0 definierten Parallelogramms, cx liegt auf der Geraden durch 0 und x.
Jetzt können wir (4.3) und (4.8) eleganter ausdrücken.

4.11 Sei L die Lösungsmenge von (4.1) und L^0 die Lösungsmenge des homogenen Systems (4.2). Sei $z \in L$ fest gewählte spezielle Lösung. Dann gilt

$$L = \{z \oplus x;\ x \in L^0\}.$$

Weiter gilt: Sind $x, y \in L^0$ und $c \in \mathbb{K}$ beliebig, so ist $x \oplus y \in L^0$ und $c * x \in L^0$.

4.12 Rechenregeln:
Für die Addition:

(1) $x \oplus (y \oplus z) = (x \oplus y) \oplus z \quad \forall\, x, y, z \in \mathbb{K}^n$

(2) $x \oplus y = y \oplus x \quad \forall\, x, y \in \mathbb{K}^n$

(3) Man kann ein Element $\underline{0} \in \mathbb{K}^n$ finden, so dass für alle $x \in \mathbb{K}^n$ gilt:

$$x \oplus \underline{0} = x,$$

nämlich $\underline{0} = (0, 0, \dots, 0)$

(4) Zu $x \in \mathbb{K}^n$ gibt es ein $\ominus x$, so dass

$$x \oplus (\ominus x) = \underline{0},$$

nämlich $\ominus x = (-x_1, -x_2, \dots, -x_n)$ für $x = (x_1, x_2, \dots, x_n)$.

Für die skalare Multiplikation: $\quad \forall\, a, b \in \mathbb{K}$ und $\forall\, x, y \in \mathbb{K}^n$ gilt

(5) $(a \cdot b) * x = a * (b * x)$

(6) $a * (x \oplus y) = a * x \oplus a * y$

(7) $(a + b) * x = a * x \oplus b * x$

(8) $1 * x = x$

Beweis: Man rechnet diese Regeln einfach nach. Wir machen das exemplarisch für die Regeln (3) und (7):

$$\begin{aligned} x \oplus \underline{0} &= (x_1 + 0, x_2 + 0, \dots, x_n + 0) = (x_1, x_2, \dots, x_n) \\ (a+b) * x &= ((a+b) \cdot x_1, (a+b) \cdot x_2, \dots, (a+b) \cdot x_n) \\ &= (ax_1 + bx_1, ax_2 + bx_2, \dots, ax_n + bx_n) \\ &= (ax_1, ax_2, \dots, ax_n) \oplus (bx_1, bx_2, \dots, bx_n) \\ &= a * x \oplus b * x. \end{aligned}$$

$\square$

Die Rechenregeln für die Addition sind dieselben wie die eines Körpers. Bei Körpern haben wir bereits gesehen, dass eine axiomatische Behandlung sinnvoll ist, weil man dann allgemein gültige Resultate für **alle** Körper erhält. Es liegt daher nahe, bei der Behandlung der Addition und skalaren Multiplikation sich von dem speziellen Fall $\mathbb{K}^n$ zu lösen und ebenfalls axiomatisch vorzugehen, um dann für die Resultate ein größeres Anwendungsspektrum zu haben. Auch das ist eine in der Mathematik typische Vorgehensweise: Für einige gegebene Beispiele entwickelt man ein mathematisches Modell, das man untersucht, wobei die gegebenen Beispiele Leitlinien für die Untersuchung bilden. Man hofft, auf diese

Weise Einsichten über gut untersuchte Beispiele auf weniger bekannte Objekte mit gleicher Struktur übertragen zu können.

Unsere zentrale Anwendung sind nach wie vor Gleichungssysteme, und für deren Lösungen gelten die Rechenregeln 4.12. Es liegt daher nahe, diese Rechenregeln für die axiomatische Behandlung heranzuziehen. Wir werden im Anschluss daran erläutern, warum wir ausgerechnet die Rechenregeln 4.12 und nicht andere Rechenregeln, wie etwa

$$0 * x = \underline{0}$$

gewählt haben.

4.13 Definition: Sei $(\mathbb{K}, +, \cdot)$ ein Körper. Ein $\mathbb{K}$-**VEKTORRAUM** ist eine Menge V zusammen mit einer **ADDITION**

$$V \times V \longrightarrow V, \qquad (x,y) \longmapsto x \oplus y$$

und einer **SKALAREN MULTIPLIKATION**

$$\mathbb{K} \times V \longrightarrow V, \qquad (a,x) \longmapsto a * x,$$

so dass folgende Axiome erfüllt sind:

Für die Addition:

(1) $x \oplus (y \oplus z) = (x \oplus y) \oplus z \quad \forall\, x,y,z \in V$

(2) $x \oplus y = y \oplus x \quad \forall\, x,y \in V$

(3) Man kann ein Element $\underline{0} \in V$ finden, so dass für alle $x \in V$ gilt: $x \oplus \underline{0} = x$

(4) Zu $x \in V$ gibt es ein $\ominus x$, so dass

$$x \oplus (\ominus x) = \underline{0}$$

Für die skalare Multiplikation: $\quad \forall\, a,b \in \mathbb{K}$ und $\forall x,y \in V$ gilt

(5) $(a \cdot b) * x = a * (b * x)$

(6) $a * (x \oplus y) = a * x \oplus a * y$

(7) $(a + b) * x = a * x \oplus b * x$

(8) $1 * x = x$

Die Elemente von V nennt man **VEKTOREN**.

4.14 Bemerkung: Die Axiome der Addition sind für unser Beispiel $\mathbb{K}^n$ erfüllt. Sie sind identisch mit den Axiomen der Addition für Körper, die wir bereits als sinnvoll eingesehen haben. Die Axiome der skalaren Multiplikation gelten ebenfalls in $\mathbb{K}^n$. Wichtig ist hier aber ein weiterer mathematischer Sachverhalt: In $\mathbb{K}$ haben wir eine Addition und eine Multiplikation. Die Axiome (5),...,(8) geben nun an, wie sich diese Strukturen zur Struktur auf V, d.h. zur skalaren Multiplikation und Addition auf V verhalten.

Bevor wir erste Folgerungen aus den Axiomen ziehen, stellen wir zwei weitere Beispiele vor:

4.15 Beispiel: Sei $\mathbb{K}$ ein Körper und M eine beliebige Menge. Dann ist die Menge

$$\mathrm{Abb}(M,\mathbb{K})$$

aller Abbildungen von M nach $\mathbb{K}$ mit folgenden Verknüpfungen ein $\mathbb{K}$-Vektorraum: Sei $f,g \in \mathrm{Abb}(M,\mathbb{K}),\ a \in \mathbb{K}$; wir definieren

$$\begin{aligned} (f \oplus g)(x) &:= f(x) + g(x) && \forall\, x \in M \\ (a * f)(x) &:= a \cdot f(x) && \forall\, x \in M \end{aligned}$$

Der Nachweis der Axiome ist dem Leser überlassen.

4.16 Beispiel: Wir erinnern daran, dass $\mathbb{R} \subset \mathbb{C}$. Dann ist $\mathbb{C}$ mit der üblichen Addition und der üblichen Multiplikation

$$\mathbb{R} \times \mathbb{C} \longrightarrow \mathbb{C} \qquad (a,x) \longmapsto a \cdot x$$

ein $\mathbb{R}$-Vektorraum. Auf die gleiche Art kann man $\mathbb{R}$ und $\mathbb{C}$ mit der üblichen Addition und Multiplikation als $\mathbb{Q}$-Vektorräume auffassen.

Vektorräume diese Art spielen in der Algebra und der Zahlentheorie eine zentrale Rolle.

4.17 Folgerungen aus den Axiomen:

(1) Da die Axiome der Addition mit denen der Addition in Körpern übereinstimmen, erhalten wir aus (2.5): Es gibt genau ein Element $\underline{0} \in V$, das das Axiom 3 erfüllt. Wir nennen es den **Nullvektor**.

(2) $\forall\, x \in V$ gilt: $0 * x = \underline{0}$.

(3) $\forall\, a \in \mathbb{K}$ gilt $a * \underline{0} = \underline{0}$.

(4) $\forall\, a \in \mathbb{K},\ \forall\, x \in V$ gilt: $(-a) * x = \ominus(a * x) = a * (\ominus x)$.

(5) $\forall\, x \in V$ gilt: $(-1) * x = \ominus x$.

(6) $a * x = \underline{0} \Longrightarrow a = 0$ oder $x = \underline{0}$.

(7) $\forall\, a \in \mathbb{K}, \quad \forall\, x, y \in V$ gilt:

$$a * (x \ominus y) = a * x \ \ominus\ a * y$$

Hier schreiben wir kurz $x \ominus y$ für $x \oplus (\ominus y)$.

Beweis: (2) $0 * x = (0+0) * x = 0 * x \oplus 0 * x$. Addieren wir $\ominus(0 * x)$ ab, erhalten wir $0 * x = \underline{0}$.

(3) wird genauso gezeigt.

(4) $a * x \oplus (-a) * x = (a-a) * x = 0 * x = \underline{0}$. Addieren von $\ominus(a * x)$ gibt $(-a) * x = \ominus(a * x)$. Genauso zeigt man, dass $a * (\ominus x) = \ominus(a * x)$.

(5) folgt aus (4): $(-1) * x = \ominus(1 * x) = \ominus x$

(6) Angenommen $a \neq 0$. Dann gilt $a^{-1} * (a * x) = a^{-1} * \underline{0} = \underline{0}$. Aber $a^{-1} * (a * x) = (a^{-1} \cdot a) * x = 1 * x = x$. Also $x = \underline{0}$.

(7) $a * (x \ominus y) = a * x \oplus a * (\ominus y) = a * x \oplus (\ominus(a * y)) = a * x \ominus a * y$. $\square$

4.18 Bezeichnung: Wie in den meisten Lehrbüchern üblich, schreiben wir $+$ statt $\oplus$ und $\cdot$ statt $*$. Dem Leser sollte es aus dem jeweiligen Zusammenhang klar sein, ob unter $+$ die Addition von Körperelementen oder Vektoren zu verstehen ist und unter $\cdot$ die Multiplikation im Körper oder die skalare Multiplikation.

4.19 Definition: Sei V ein $\mathbb{K}$-Vektorraum. Eine Teilmenge $W \subset V$ heißt **UNTERVEKTORRAUM**, wenn W mit der Addition in V und der skalaren Multiplikation in V ein $\mathbb{K}$-Vektorraum ist.

Achtung: Wir haben jetzt Unterkörper und Untervektorräume kennen gelernt. Hier wird ein typischer Grundsatz der Mathematik angewandt. Wir haben eine Menge M mit einer (hier algebraischen) Struktur gegeben. Eine Teilmenge $N \subset M$ heißt dann Unterobjekt von M, wenn die Struktur auf M dieselbe Struktur auf N definiert.

Als erstes machen wir uns klar, welche Bedingungen $W \subset V$ erfüllen muss, damit W ein Untervektorraum ist.

4.20 Sei V ein $\mathbb{K}$-Vektorraum und $W \subset V$ Teilmenge. Dann gilt:

W ist Untervektorraum $\Longleftrightarrow$
(i) $W \neq \emptyset$
(ii) $x, y \in W \Longrightarrow x + y \in W$
(iii) $x \in W,\ a \in \mathbb{K} \Longrightarrow a \cdot x \in W$

Beweis: Die Bedingungen (ii) und (iii) sind dafür notwendig, dass die Addition und skalare Multiplikation auf V auch Verknüpfungen auf W sind. Damit ist „$\Longrightarrow$“ klar, weil auch $0 \in W$ sein muss.

„$\Longleftarrow$“ Wie schon erwähnt, definiert die Struktur auf V wegen (ii) und (iii) eine Struktur auf W. Nach (iii) ist mit x auch $(-1) \cdot x = -x$ aus W und damit auch $0 = x - x$. Die übrigen Axiome gelten für W, weil sie bereits auf der größeren Menge V gelten. $\square$

4.21 Beispiel: Die Lösungsmenge L^0 des homogenen Gleichungssystems (4.2) ist ein Untervektorraum von $\mathbb{K}^n$.

Wenn wir also etwas über die Struktur von L^0 erfahren wollen, müssen wir Untervektorräume studieren.

4.22 Satz: Sei V ein $\mathbb{K}$-Vektorraum und $J \neq \emptyset$ eine Indexmenge. Sei W_i für $i \in J$ ein Untervektorraum von V. Dann ist

$$W = \bigcap_{i \in J} W_i$$

ein Untervektorraum von V.

Beweis: $0 \in W_i$ für alle $i \in J$. Also $0 \in W$ und damit $W \neq \emptyset$ (die Bedingung $W \neq \emptyset$ testet man stets am einfachsten daran, ob $0 \in W$).

Sind nun $x, y \in W$ und $a \in \mathbb{K}$. Dann gilt $x, y \in W_i \quad \forall\, i \in J$. Da W_i ein Untervektorraum von V ist, folgt $x + y \in W_i$ und $a \cdot x \in W_i \quad \forall\, i \in J$. Also $x + y \in W$ und $ax \in W$. $\square$

Wir brauchen (4.22), damit folgende Definition sinnvoll ist:

4.23 Definition: Sei V ein $\mathbb{K}$-Vektorraum und $M \subset V$ eine beliebige Teilmenge. Wir definieren $\text{Span}(M)$ als den **kleinsten Untervektorraum** von V, der M enthält. D.h.

(1) $M \subset \text{Span}(M)$

(2) Ist W ein Untervektorraum von V, so dass $M \subset W$, dann gilt

$$\text{Span}(M) \subset W.$$

4.24 $\text{Span}(M)$ existiert. Genauer gilt: $\text{Span}(M)$ ist der Durchschnitt aller Untervektorräume W von V mit $M \subset W$. $\square$

4.25 Beispiel: (1) $\text{Span}(\emptyset) = \{0\}$

(2) $\text{Span}(V) = V$

(3) Sei $v \in V$. Dann ist $\text{Span}(v) = \mathbb{K} \cdot v$

4.26 Definition: Eine Teilmenge $M \subset V$, für die $\mathrm{Span}(M) = V$ gilt, heißt **ERZEUGENDENSYSTEM** von V. Besitzt V ein endliches Erzeugendensystem, heißt V **ENDLICH ERZEUGT**.

Die Beschreibung (4.24) von $\mathrm{Span}(M)$ ist ziemlich abstrakt. Wir werden gleich eine erheblich konkretere geben. Aber (4.23) und (4.24) sind Konzepte, die man sofort auf andere Strukturen übertragen kann. Um nur ein Beispiel zu nennen, könnte man genauso den kleinsten Unterkörper von $\mathbb{K}$, der eine vorgegebene Teilmenge $M \subset \mathbb{K}$ enthält, definieren und angeben. Wir haben wieder ein weiteres Standardrezept der Mathematik kennen gelernt.

4.27 Satz: Sei V ein $\mathbb{K}$-Vektorraum und $M \subset V$ eine nicht-leere Teilmenge. Dann gilt

$$\begin{aligned}\mathrm{Span}(M) = &\{x \in V;\ \exists y_1, \dots, y_k \in M \text{ und } a_1, \dots, a_k \in \mathbb{K}, \text{ so dass} \\ &x = a_1 y_1 + \cdots + a_k \cdot y_k\}\end{aligned}$$

4.28 Bezeichnung: Ein Ausdruck der Form

$$x = \sum_{i=1}^{k} a_i y_i = a_1 y_1 + \cdots + a_k y_k \qquad y_i \in V,\ a_i \in \mathbb{K}$$

heißt **LINEARKOMBINATION** von $y_1, \dots, y_k$. $\mathrm{Span}(M)$ ist also die Menge aller Linearkombinationen von Vektoren aus M.

Beweis: 4.26: Sei W die Menge auf der rechten Seite des Gleichheitszeichens. W ist Untervektorraum von V, denn:

(i) $W \neq \emptyset$, weil $M \neq \emptyset$.

(ii) Sind $x = a_1 y_1 + \cdots + a_k y_k$ und $x' = b_1 z_1 + \cdots + b_\ell z_\ell$ aus W, d.h. $y_1, \dots, y_k$, $z_1, \dots, z_\ell \in M$ und $a_1, \dots, a_k, b_1, \dots, b_\ell \in \mathbb{K}$. Dann ist

$$x + x' = a_1 y_1 + \cdots + a_k y_k + b_1 z_1 + \cdots + b_\ell z_\ell$$

ebenfalls eine Linearkombination von Vektoren aus M. Also $x + x' \in W$.

(iii) Ist $a \in \mathbb{K}$, so ist $a \cdot x = (a \cdot a_1) \cdot y_1 + \cdots + (a \cdot a_k) \cdot y_k$ wieder eine Linearkombination von Vektoren aus M, also $a \cdot x \in W$.

Nun gilt $M \subset W$, denn für $y \in M$ ist $y = 1 \cdot y \in W$. Es folgt

$$\mathrm{Span}(M) \subset W.$$

Da aber $M \subset \mathrm{Span}(M)$ und $\mathrm{Span}(M)$ ein Untervektorraum ist, sind alle Linearkombinationen von Vektoren aus M in $\mathrm{Span}(M)$. D.h. $W \subset \mathrm{Span}(M)$.

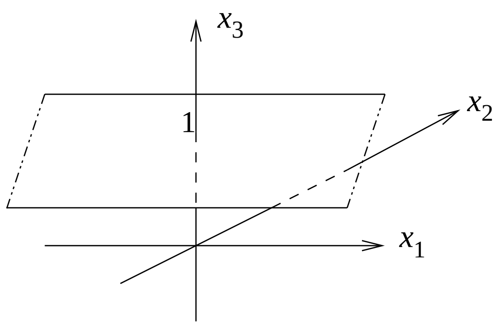
□

4.29 Beispiel: (1) Sei $e^1 = (1,0,0),\ e^2 = (0,1,0),\ e^3 = (0,0,1)$ aus $\mathbb{R}^3$. Dann gilt

$$\mathrm{Span}(e^1, e^2, e^3) = \mathbb{R}^3.$$

(2) In $\mathbb{R}^3$ betrachten wir die „Ebene“

$$M = \{(x_1, x_2, x_3) \in \mathbb{R}^3;\ x_3 = 1\}$$



Dann ist $\mathrm{Span}(M) = \mathbb{R}^3$.

Denn sei $(x_1, x_2, x_3) \in \mathbb{R}^3$ ein beliebiger Vektor, dann ist

$$(x_1, x_2, x_3) = (x_3 - 1) \cdot (0,0,1) + 1 \cdot (x_1, x_2, 1) \in \mathrm{Span}(M).$$

Durchschnitte von Untervektorräumen sind wieder Untervektorräume. Für Vereinigungen ist das in den seltensten Fällen richtig. Als „Ersatz“ nehmen wir

4.30 Definition: Sind $W_1, \dots, W_n$ Untervektorräume eines $\mathbb{K}$-Vektorraumes V, dann heißt der Untervektorraum

$$W_1 + W_2 + \cdots + W_n := \mathrm{Span}(W_1 \cup W_2 \cup \cdots \cup W_n) = \{x_1 + \cdots + x_n; x_i \in W_i \quad \forall\, i\}$$

die **SUMME** von $W_1, \dots, W_n$.
Gilt außerdem $W_i \cap (W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_n) = \{0\} \quad \forall\ i = 1, \dots, n$, heißt $W_1 + \cdots + W_n$ **DIREKTE SUMME** und wird auch mit $W_1 \oplus \cdots \oplus W_n$ bezeichnet.

Bemerkung: Da die W_i Untervektorräume sind, kann man Linearkombinationen von Vektoren in W_i zu einem Vektor zusammenfassen. Daher gilt

$$\mathrm{Span}(W_1 \cup \cdots \cup W_n) = \{x_1 + \cdots + x_n; x_i \in W_i,\ i = 1, \dots, n\}.$$

Diese Gleichung legt auch die Bezeichnung $W_1 + \cdots + W_n$ nahe.
Die direkte Summe besitzt folgende wichtige Eigenschaft:

4.31 Satz: Es seien $W_1, \ldots, W_n$ Untervektorräume eines $\mathbb{K}$-Vektorraumes V. Dann sind für $W = W_1 + \cdots + W_n \subset V$ folgende Aussagen äquivalent:

(1) $W = W_1 \oplus \cdots \oplus W_n$

(2) Jedes $x \in W$ besitzt genau eine Darstellung $x = x_1 + x_2 + \cdots + x_n$ mit $x_i \in W_i$ für $i = 1, \ldots, n$.

Beweis: (1) $\Longrightarrow$ (2): Da $x \in W = W_1 + \cdots + W_n$, gibt es $x_1 \in W_1, \ldots, x_n \in W_n$ mit

$$x = x_1 + \cdots + x_n \qquad (*)$$

Sei $x = y_1 + \cdots + y_n$ eine weitere Darstellung mit $y_i \in W_i$ für $i = 1, \ldots, n$. Dann folgt

$$0 = x - x = (x_1 - y_1) + \cdots + (x_n - y_n).$$

Also gilt

$$-(x_i - y_i) = (x_1 - y_1) + \cdots + (x_{i-1} - y_{i-1}) + (x_{i+1} - y_{i+1}) + \cdots + (x_n - y_n).$$

Die rechte Seite liegt in $W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_n$ und die linke in W_i. Da aber $W_i \cap (W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_n) = \{0\}$, folgt

$$x_i - y_i = 0.$$

Damit ist die Darstellung $(*)$ eindeutig.
(2) $\Longrightarrow$ (1) Sei $x \in W_i \cap (W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_n)$. Dann besitzt x eine Darstellung

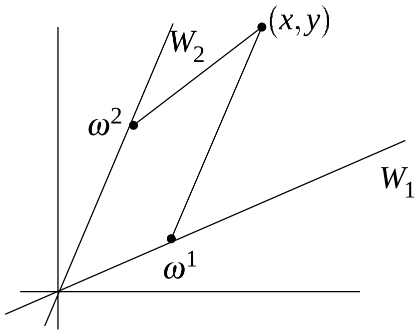
$$x = x_1 + \cdots + x_{i-1} + x_{i+1} + \cdots + x_n$$

mit $x_k \in W_k$ für alle k. Dann sind aber

$$x = 0 + \cdots 0 + \underset{i}{x} + 0 + \cdots + 0$$
$$x = x_1 + \cdots + x_{i-1} + 0 + x_{i+1} + \cdots + x_n$$

zwei Darstellungen von x gemäß (2). Wegen der Eindeutigkeit folgt $x = 0$, also $W_1 \cap (W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_n) = \{0\}$. $\square$

4.32 Beispiel:



In $\mathbb{R}^2$ sei $W_1 = \{(2x, x); x \in \mathbb{R}\}$ und $W_2 = \{(x, 2x); x \in \mathbb{R}\}$
W_1 und W_2 sind Untervektorräume und $W_1 \cap W_2 = \{0\}$. Es gilt $\mathbb{R}^2 = W_1 \oplus W_2$, denn

$$(x, y) = \underbrace{\left(2 \cdot \frac{2x-y}{3}, \frac{2x-y}{3}\right)}_{\in W_1} + \underbrace{\left(\frac{2y-x}{3}, 2 \cdot \frac{2y-x}{3}\right)}_{\in W_2}$$

Da sich jeder Vektor aus $\mathbb{R}^2$ **eindeutig** als Summe von Vektoren aus W_1 und W_2 darstellen lässt und W_1, W_2 eindimensional sind, können wir $\{W_1, W_2\}$ als **schiefwinkliges Koordinatensystem** auffassen.

Auf geeignete Wahlen von Koordinatensystemen werden wir später noch eingehen.

Zunächst wollen wir klären, was wir unter der Dimension eines Vektorraumes verstehen, wobei wir uns natürlich von der Anschauung leiten lassen. Weiterhin müssen wir untersuchen, wie wir direkte Summenzerlegungen von V in 1-dimensionale Untervektorräume konstruieren können.

Wir wollen den Begriff der direkten Summe noch etwas erweitern.

4.33 Definition und Satz: Seien V und W $\mathbb{K}$-Vektorräume. Dann ist $V \times W$ mit den Verknüpfungen

Addition: $(V \times W) \times (V \times W) \longrightarrow V \times W,\ (v_1, w_1) + (v_2, w_2) := (v_1 + v_2, w_1 + w_2)$
Skalarmultiplikation: $\mathbb{K} \times (V \times W) \longrightarrow V \times W,\ a \cdot (v, w) := (a \cdot v, a \cdot w)$

ein $\mathbb{K}$-Vektorraum, genannt das **PRODUKT** von V und W.

Der Nachweis der Vektorraumaxiome ist einfach. Wir beweisen exemplarisch die Axiome (4.13 (3) und (6)):

$(0, 0)$ ist der Nullvektor in $V \times W$, denn

$$\begin{aligned}
(v, w) + (0, 0) = (v + 0, w + 0) &= (v, w) \\
a \cdot ((v_1, w_1) + (v_2, w_2)) &= a \cdot (v_1 + v_2, w_1 + w_2) \\
&= (a \cdot (v_1 + v_2), a \cdot (w_1 + w_2)) \\
&= (a \cdot v_1 + a \cdot v_2), a \cdot w_1 + a \cdot w_2) \\
&= (a \cdot v_1, a \cdot w_1) + (a \cdot v_2, a \cdot w_2) \\
&= a \cdot (v_1, w_1) + a \cdot (v_2, w_2).
\end{aligned}$$

$\square$

Man sieht sofort, dass

$$V \times 0 := \{(v,0);\, v \in V\} \text{ und } 0 \times W := \{(0,w);\, w \in W\}$$

Untervektorräume von $V \times W$ sind. Da

$$(v,w) = (v,0) + (0,w),$$

gilt $V \times W = V \times 0 + 0 \times W$. Da außerdem $V \times 0 \cap 0 \times W = \{(0,0)\}$, folgt

4.34 $V \times W = V \times 0 \,\oplus\, 0 \times W$

Aus diesem Grunde nennt man $V \times W$ oft auch die (äußere) **DIREKTE SUMME** von V und W, bezeichnet mit $V \oplus W$.

5 Basis und Dimension

Sei $\mathbb{K}$ ein Körper und V ein $\mathbb{K}$-Vektorraum.

5.1 Definition: Vektoren $x_1, \dots, x_n$ aus V heißen **LINEAR UNABHÄNGIG** (über K) und $\{x_1, \dots, x_n\}$ heißt **LINEAR UNABHÄNGIGES SYSTEM**, wenn aus

$$0 = a_1 x_1 + \cdots + a_n x_n \quad \text{mit } a_i \in \mathbb{K} \text{ für } i = 1, \dots, n$$

folgt, dass $a_1 = a_2 = \ldots = a_n = 0$. Gibt es $a_1, \dots, a_n \in \mathbb{K}$, so dass $(a_1, \dots, a_n) \neq (0, \dots, 0)$ und $a_1 x_1 + \cdots + a_n x_n = 0$, heißen $x_1, \dots, x_n$ **LINEAR ABHÄNGIG**.

5.2 $x_1, \dots, x_n$ sind linear abhängig $\Longleftrightarrow$ mindestens ein x_i ist Linearkombination der übrigen Vektoren $x_j,\ j \neq i$.

Beweis: $x_1, \dots, x_n$ linear abhängig $\Longleftrightarrow$ $\exists\ (a_1, \dots, a_n) \neq (0, \dots, 0)$ in $\mathbb{K}^n$ mit $a_1 x_1 + \cdots + a_n x_n = 0$.
Sei etwa $a_i \neq 0$. Dann folgt aus $a_i x_i = -a_1 x_1 - \cdots - a_{i-1} x_{i-1} - a_{i+1} x_{i+1} - \cdots - a_i x_n$

$$x_i = -\frac{a_1}{a_i} x_1 - \cdots - \frac{a_{i-1}}{a_i} x_{i-1} - \frac{a_{i+1}}{a_i} x_{i+1} - \cdots - \frac{a_n}{a_i} x_n.$$

Sei umgekehrt $x_i = b_1 x_1 + \cdots + b_{i-1} x_{i-1} + b_{i+1} x_{i+1} + \cdots + b_n x_n$, dann gilt

$$0 = b_1 x_1 + \cdots + b_{i-1} x_{i-1} - x_i + b_{i+1} x_{i+1} + \cdots + b_n x_n.$$

Da der Koeffizient von x_i von 0 verschieden ist, sind $x_1, \dots, x_n$ linear abhängig. $\square$

5.3 Definition: Eine Menge $\mathscr{B} = \{x_1, \dots, x_n\}$ von Vektoren aus V heißt **BASIS** von V (genauer $\mathbb{K}$-Basis von V), wenn $V = \operatorname{Span}(\mathscr{B})$, d.h. $\mathscr{B}$ ist ein Erzeugendensystem von V, und $\mathscr{B}$ ist ein linear unabhängiges System.

Eine Basis von V kann man als „schiefwinkliges" Koordinatensystem auffassen, denn es gilt

5.4 Satz: Ist $\{x_1, \dots, x_n\}$ eine Basis von V, dann besitzt jedes $x \in V$ genau eine Darstellung als Linearkombination

$$x = a_1 x_1 + \cdots + a_n x_n \qquad a_i \in \mathbb{K},\ i = 1, \dots, n$$

der x_i. Es folgt (beachte: $\mathbb{K} \cdot v = \operatorname{Span}\{v\}$)

$$V = \mathbb{K} \cdot x_1 \oplus \cdots \oplus \mathbb{K} \cdot x_n.$$

Beweis: Da $V = \text{Span}\{x_1, \dots, x_n\}$, hat jedes x eine Darstellung

$$x = a_1 x_1 + \cdots + a_n x_n$$

mit $a_i \in \mathbb{K},\ i = 1, \dots, n$. Ist $x = b_1 x_1 + \cdots + b_n x_n$ eine zweite solche Darstellung, so folgt

$$0 = x - x = (a_1 - b_1)x_1 + \cdots + (a_n - b_n)x_n.$$

Da $\{x_1, \dots, x_n\}$ linear unabhängig ist, folgt $a_i - b_i = 0 \quad \forall\, i$, also $a_i = b_i \quad \forall\, i$. Da für jeden Vektor x die Menge $\mathbb{K} \cdot x = \{a \cdot x;\, a \in \mathbb{K}\}$ ein Untervektorraum von V ist, folgt der zweite Teil aus (4.31). $\square$

5.5 Definition: Sei V ein $\mathbb{K}$-Vektorraum. Die **DIMENSION** von $V, \dim V$, ist die größte Zahl $n \in \mathbb{N}$ mit der Eigenschaft, dass V n linear unabhängige Vektoren besitzt. Gibt es kein solches größtes n, setzen wir $\dim V = \infty$.

5.6 Satz: Ist V endlich dimensional, $\dim V = n$, und sind $x_1, \dots, x_n \in V$ linear unabhängig, dann ist $\{x_1, \dots, x_n\}$ eine Basis von V. Insbesondere ist jeder endlich dimensionale Vektorraum auch endlich erzeugt.

Beweis: Wir brauchen nur zu zeigen, dass $V \subset \text{Span}\{x_1, \dots, x_n\}$. Sei $x \in V$. Da $\dim V = n$, ist $\{x_1, \dots, x_n, x\}$ linear abhängig (da sonst $\dim V \geq n + 1$). Also hat man eine Gleichung

$$a_1 x_1 + \cdots + a_n x_n + a_{n+1} x = 0 \qquad a_i \in \mathbb{K},$$

in der nicht alle $a_i = 0$ sind. Es muss $a_{n+1} \neq 0$ gelten, da sonst $x_1, \dots, x_n$ linear abhängig wären. Es folgt

$$x = -\frac{1}{a_{n+1}}(a_1 x_1 + \cdots + a_n x_n) \in \text{Span}\{x_1, \dots, x_n\}.$$

$\square$

5.7 Basisergängzungssatz: Seien $x_1, \dots, x_r$ linear unabhängige Vektoren in V und $y_1, \dots, y_s$ weitere Vektoren in V, so dass

$$V = \text{Span}\{x_1, \dots, x_r, y_1, \dots, y_s\}.$$

Dann gibt es eine Teilmenge $\{y_{i_1}, \dots, y_{i_k}\} \subset \{y_1, \dots, y_s\}$, so dass

$$\{x_1, \dots, x_r, y_{i_1}, \dots, y_{i_k}\}$$

Basis von V ist.

Beweis: Unter allen Teilmengen $\{y_{i_1},\ldots,y_{i_k}\}$ von $\{y_1,\ldots,y_s\}$, für die gilt

$$V = \operatorname{Span}\{x_1,\ldots,x_r,y_{i_1},\ldots,y_{i_k}\}$$

wählen wir eine mit minimalem k. Wir nummerieren die y_i so um, dass $\{y_1,\ldots,y_k\}$ diese gewählte Teilmenge ist. Es gilt also

$$V = \operatorname{Span}\{x_1,\ldots,x_r,y_1,\ldots,y_k\},$$

aber lässt man ein y_i weg, wird V vom Rest nicht mehr erzeugt.

Behauptung: $\{x_1,\ldots,x_r,y_1,\ldots,y_k\}$ ist Basis von V.
Beweis: Wir müssen nur die lineare Unabhängigkeit nachweisen. Sei also eine Gleichung

$$a_1x_1+\cdots+a_rx_r+b_1y_1+\cdots+b_ky_k=0 \qquad a_i,b_j\in\mathbb{K}$$

gegeben. Angenommen $b_i\neq 0$, dann folgt wie im Beweis von (5.6)

$$y_i\in\operatorname{Span}\{x_1,\ldots,x_r,y_1,\ldots,y_{i-1},y_{i+1},\ldots,y_k\}$$

und damit

$$V=\operatorname{Span}\{x_1,\ldots,x_r,y_1,\ldots,y_{i-1},y_{i+1},\ldots,y_k\}.$$

Wegen der Minimalität von k ist das unmöglich. Also sind alle $b_i=0$. Da $x_1,\ldots,x_r$ linear unabhängig sind, folgt dann aber, dass alle $a_i=0$. □

5.8 Auswahlsatz: Aus jedem endlichen Erzeugendensystem $\{y_1,\ldots,y_s\}$ eines Vektorraumes $V\neq\{0\}$ kann man eine Basis auswählen.

Beweis: Wende den Basisergänzungssatz mit $r=0$ an. □

5.9 Austauschsatz: Sind $x_1,\ldots,x_r$ linear unabhängige Vektoren in V und ist M ein Erzeugendensystem von V, dann kann man r Vektoren aus M so auswählen und gegen die $x_1,\ldots,x_r$ austauschen, dass man wieder ein Erzeugendensystem von V erhält.

Beweis: Wir nehmen an, wir hätten schon $x_1,\ldots,x_{k-1}$ gegen Vektoren $y_1,\ldots,y_{k-1}$ aus M so ausgetauscht, dass die Menge

$$M_{k-1}=\{x_1,\ldots,x_{k-1}\}\cup(M\backslash\{y_1,\ldots,y_{k-1}\})$$

ebenfalls V erzeugt. Für den Fall $k=1$ setzen wir hier $M_0=M$. Da M_{k-1} ein Erzeugendensystem ist, hat x_k eine Darstellung

$$x_k=a_1x_1+\cdots+a_{k-1}x_{k-1}+a_ky_k+\cdots+a_sy_s$$

mit $a_i \in \mathbb{K}$ für $i = 1, \dots, s$ und geeignete Vektoren $y_k, \dots, y_s$ aus $M \backslash \{y_1, \dots, y_{k-1}\}$. Da $\{x_1, \dots, x_k\}$ linear unabhängig ist, können nicht alle $a_k, a_{k+1}, \dots, a_s$ Null sein. Durch Umnummerierung dürfen wir annehmen, dass $a_k \neq 0$. Es folgt durch Auflösen nach y_k und Division durch a_k, dass

$$y_k \in \operatorname{Span}\{x_1, \dots, x_{k-1}, x_k, y_{k+1}, \dots, y_s\}.$$

Also ist $M_k = \{x_1, \dots, x_{k-1}, x_k\} \cup (M \backslash \{y_1, \dots, y_k\})$ ein Erzeugendensystem von V.

Beginnend mit $M_0 = M$ setzen wir das Verfahren solange fort, bis alle x_i gegen geeignete y_i getauscht sind. $\square$

Wir fassen unsere Ergebnisse zusammen:

5.10 Satz: Sei V endlich dimensionaler Vektorraum

(1) Ist $\{x_1, \dots, x_r\}$ eine Basis von V und $\{y_1, \dots, y_s\}$ ein Erzeugendensystem, dann gilt $r \leq s$.

(2) Ist $\{x_1, \dots, x_r\}$ eine Basis von V, so gilt $\dim V = r$. Also haben alle Basen von V dieselbe Anzahl von Elementen.

(3) $\dim V = \operatorname{Min}\{n \in \mathbb{N};\ \exists\, x_1, \dots, x_n \in V \text{ mit } V = \operatorname{Span}\{x_1, \dots, x_n\}\}$

Beweis: (1) folgt aus dem Austauschsatz.

Da eine Basis auch ein Erzeugendensystem ist, haben nach (1) zwei Basen gleichviel Elemente. Nach (5.6) gibt es aber eine Basis mit $\dim V$ Elementen. Also folgt (2). Aus (1) und (2) folgt wiederum (3). $\square$

5.11 Beispiel: $\dim \mathbb{K}^n = n$: Sei $e^i = (0, 0, \dots, 0, \overset{i}{1}, 0, \dots, 0)$. Dann ist $\{e^1, \dots, e^n\}$ eine Basis von $\mathbb{K}^n$, genannt die **STANDARDBASIS**. Der Beweis dafür ist trivial.

Anwendungen auf $\mathbb{K}^n$ und lineare Gleichungssysteme

Seien $v^1 := (a_{11}, \dots, a_{1n}), v^2 = (a_{21}, \dots, a_{2n}), \dots, v^m = (a_{m1}, \dots, a_{mn})$ Vektoren im $\mathbb{K}$-Vektorraum $\mathbb{K}^n$.

5.12 Wie testet man die lineare Unabhängigkeit von $v^1, \dots, v^m$?

$v^1, \dots, v^m$ sind linear unabhängig $\Longleftrightarrow$ die vektorielle Gleichung

$$x_1 \cdot v^1 + \dots + x_m \cdot v^m = 0 \qquad x_i \in \mathbb{K}$$

hat nur die **triviale Lösung** $(x_1, \dots, x_m) = (0, \dots, 0)$. Schreiben wir diese Gleichung koordinatenweise aus, erhalten wir das Gleichungssystem

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{21}x_2 & + \cdots + & a_{m1}x_m & = & 0 \\ \vdots & & \vdots & & \vdots & & \\ a_{1n}x_1 & + & a_{2n}x_2 & + \cdots + & a_{mn}x_m & = & 0 \end{array}$$

von n Gleichungen mit m Unbekannten. Man beachte, dass die Indices der a_{ij} anders geordnet sind als in (4.2).

Da $\mathbb{K}^n$ nur Systeme mit höchstens n linear unabhängigen Vektoren besitzt (s. 5.11), brauchen wir nur den Fall $m \leq n$ zu erörtern. Wir sehen sofort: $v^1, \dots, v^m$ sind genau dann linear unabhängig, wenn die Dreiecksform (1.8) des Gleichungssystems die Gestalt

$$\begin{array}{ccccc|c} x_{q_1} & x_{q_2} & \dots & \dots & x_{q_m} & b \\ \hline 1 & a'_{21} & \dots & \dots & a'_{m1} & 0 \\ 0 & 1 & & & \vdots & \vdots \\ & & \ddots & & \vdots & \vdots \\ & 0 & & \ddots & \vdots & \vdots \\ & & & & 1 & 0 \\ 0 & \dots & \dots & \dots & 0 & \vdots \\ \vdots & & & & \vdots & \vdots \\ 0 & \dots & \dots & \dots & 0 & 0 \end{array}$$

hat.

Beispiel: Im $\mathbb{R}^3$ testen wir die Vektoren

$$\begin{aligned} v^1 &= (1, 1, 1) \\ v^2 &= (1, 0, 3) \\ v^3 &= (3, 5, -1) \end{aligned}$$

auf lineare Unabhängigkeit. Wie in (5.12) angegeben, schreiben wir die Vektoren als Spalten eines Gleichungssystem-Schemas auf und formen nach dem

Gauß'schen Algorithmus um:

x_1	x_2	x_3	b	
1	1	3	0	
1	0	5	0	
1	3	-1	0	
1	1	3	0	
0	-1	2	0	$II-I$
0	2	-4	0	$III-I$
1	1	3	0	
0	1	-2	0	$-II$
0	0	0	0	$III+2\cdot II$

Die Einsen auf der Diagonalen reichen nicht an die vertikale Trennungslinie. Also sind die Vektoren linear abhängig. In der Tat ist $(-5,2,1)$ Lösung der vektoriellen Gleichung

$$x_1 \cdot v^1 + x_2 \cdot v^2 + x_3 \cdot v^3 = 0,$$

so dass

$$-5 \cdot v^1 + 2 \cdot v^2 + v^3 = 0.$$

Beispiel: Sind v^1 und v^2 wie im letzten Beisiel, dagegen aber $v^3 = (3,5,0)$, so ist die obere Dreiecksform des entsprechenden Gleichungssystems

x_1	x_2	x_3	b
1	1	3	0
0	1	-2	0
0	0	1	0

Folglich sind die drei Vektoren linear unabhängig in $\mathbb{R}^3$.

5.13 Geometrische Interpretation der Lösungsmenge L eines linearen Gleichungssystems

Wir betrachten das **reduzierte System** (1.13) eines lösbaren linearen Gleichungssystems

$$\begin{array}{cccccccc|c}
x_{q_1} & x_{q_2} & \dots & \dots & x_{q_r} & x_{q_{r+1}} & \dots & x_{q_n} & b \\
\hline
1 & 0 & \dots & \dots & 0 & a^*_{1,r+1} & \dots & a^*_{1n} & b^*_1 \\
0 & 1 & \dots & \dots & 0 & \vdots & & \vdots & \vdots \\
 & & \ddots & 0 & \vdots & \vdots & & \vdots & \vdots \\
 & & & \ddots & \vdots & \vdots & & \vdots & \vdots \\
 & 0 & & & 1 & a^*_{r,r+1} & \dots & a^*_{r,n} & b^*_r
\end{array}$$

Wir lesen sofort die spezielle Lösung ab:

$$x^0 = (x_1^0, \dots, x_n^0) \text{ mit } x_{q_i}^0 = \begin{cases} b_i^* & 1 \leq i \leq r \\ 0 & r+1 \leq i \leq n \end{cases}$$

Ist L die Gesamtlösungsmenge und L^0 die Lösungsmenge des zugehörigen homogenen Gleichungssystems, so wissen wir

$$L = x^0 + L^0 := \{x^0 + x;\ x \in L^0\},$$

und L^0 ist ein Untervektorraum von $\mathbb{K}^n$. Das reduzierte Schema für das homogene Gleichungssystem sieht genauso aus, nur dass $b_i^* = 0$ ist für alle i. Damit können wir aus dem Schema sofort eine Basis für L^0 ablesen:

5.14 Das reduzierte Schema liefert uns eine Basis von $n-r$ Vektoren

$$v^1, \dots, v^{n-r}$$

von L^0. Für die Koordinaten $v_1^i, \dots, v_n^i$ von v^i gilt

$$v_{q_j}^i = \begin{cases} -a_{j,r+i}^* & 1 \leq j \leq r \\ 1 & j = r+i \\ 0 & \text{sonst} \end{cases}$$

Um die Bezeichnungsweise nicht zu überfrachten, erläutern wir das für den Fall, dass die x_i in der üblichen Reihenfolge aufgelistet sind:

$$\begin{array}{cccccccc|c} x_1 & x_2 & \dots & \dots & x_r & x_{r+1} & \dots & x_n & b \\ \hline 1 & 0 & \dots & \dots & 0 & a_{1,r+1}^* & \dots & a_{1,n}^* & 0 \\ 0 & 1 & \dots & \dots & 0 & a_{2,r+1}^* & \dots & a_{2,n}^* & 0 \\ & & & & 0 & \vdots & & \vdots & \vdots \\ & & \ddots & 0 & \vdots & \vdots & & \vdots & \vdots \\ & & & \ddots & \vdots & \vdots & & \vdots & \vdots \\ & 0 & & & 1 & a_{r,r+1}^* & \dots & a_{r,n}^* & 0 \end{array} \qquad (*)$$

In der b-Spalte haben wir Nullen, weil wir das homogene System betrachten. Die Idee ist nun, die Gleichungsteile mit $x_{r+1}, \dots, x_n$ auf die rechte Seite zu bringen, daher gehen die $a_{i,j}^*$ nach $-a_{i,j}^*$ über. Dann setzen wir ein $x_j = 1$ für $j \geq r+1$ und die übrigen 0 und lesen dann die Werte für $x_1, \dots, x_r$ ab. Wir erhalten die Vektoren

$$\begin{aligned} v^1 &= (-a_{1,r+1}^*, -a_{2,r+1}^*, \dots, -a_{r,r+1}^*, 1, 0, \dots, 0) \\ v^2 &= (-a_{1,r+2}^*, -a_{2,r+2}^*, \dots, -a_{r,r+2}^*, 0, 1, 0, \dots, 0) \\ &\ \vdots \\ v^{n-r} &= (-a_{1,n}^*, -a_{2,n}^*, \dots, -a_{r,n}^*, 0, \dots, 0, 1) \end{aligned}$$

Die Vektoren sind linear unabhängig. Aus

$$c_1 v^1 + \cdots + c_{n-r} v^{n-r} = 0 \qquad c_i \in \mathbb{K}$$

folgt $c_1 = c_2 = \ldots = c_{n-r} = 0$, denn die letzten $n-r$ Koordinaten des Summenvektors auf der linken Seite sind

$$c_1, c_2, \ldots, c_{n-r}.$$

Nach (1.14) ist $x = (x_1, \ldots, x_n)$ genau dann Lösung unseres homogenen Systems, wenn

$$\begin{array}{rcl} x_1 & = & -a^*_{1,r+1} x_{r+1} - \ldots - a^*_{1,n} x_n \\ x_2 & = & -a^*_{2,r+1} x_{r+1} - \ldots - a^*_{2,n} x_n \\ \vdots & & \\ x_r & = & -a^*_{r,r+1} x_{r+1} - \ldots - a^*_{r,n} x_n \end{array}$$

Für dieses x gilt aber

$$x = x_{r+1} \cdot v^1 + \cdots + x_n \cdot v^{n-r}.$$

Also ist $\{v^1, \ldots, v^{n-r}\}$ auch ein Erzeugendensystem von L^0. Wir erhalten:

5.15 L^0 ist ein $(n-r)$-dimensionaler Untervektorraum von $\mathbb{K}^n$ mit Basis $v^1, \ldots, v^{n-r}$.

5.16 Definition: Sei $W \subset \mathbb{K}^n$ ein Untervektorraum und $x^0 \in \mathbb{K}^n$ fest gewählt. Dann heißt

$$x^0 + W = \{x^0 + y;\ y \in W\}$$

AFFINER UNTERRAUM von $\mathbb{K}^n$ der **DIMENSION** $\dim W$.

5.17 Die Lösungsmenge eines lösbaren inhomogenen linearen Gleichungssystems mit n Unbekannten ist ein affiner Unterraum von $\mathbb{K}^n$.

Beispiel: Die reduzierte Form des Beispiels 1.15 ist

$$\begin{array}{ccccc|c} x_1 & x_4 & x_5 & x_2 & x_3 & b \\ \hline 1 & 0 & 0 & -3 & 1 & -1 \\ 0 & 1 & 0 & 3 & 0 & 2 \\ 0 & 0 & 1 & 2 & 0 & 2 \end{array}$$

Für die spezielle Lösung x^0 setzen wir $x_2 = 0$ und $x_3 = 0$ und erhalten $x_1 = -1$, $x_4 = 2$, $x_5 = 2$, so dass

$$x^0 = (-1, 0, 0, 2, 2).$$

Die Lösungsmenge L^0 des zugehörigen homogenen Gleichungssystems ist ein $(5-3)$-dimensionaler, also 2-dimensionaler Untervektorraum von $\mathbb{R}^5$. Eine Basis erhalten wir nach dem Verfahren 5.14, indem wir (in den beiden letzten Spalten)

(1) $x_2 = 1$ und $x_3 = 0$ setzen und dann an der x_2-Spalte ablesen

$$x_1 = 3,\ x_4 = -3,\ x_5 = -2,$$

$$v^1 = (3, 1, 0, -3, -2)$$

(2) $x_2 = 0$ und $x_3 = 1$ setzen. Wir erhalten aus der x_3-Spalte

$$x_1 = -1,\ x_4 = 0,\ x_5 = 0,$$

also

$$v^2 = (-1, 0, 1, 0, 0).$$

5.18 Auffinden einer Basis eines Untervektorraumes $W \subset \mathbb{K}^n$.

Wir wollen annehmen, dass wir einen Untervektorraum $W \subset \mathbb{K}^n$ durch ein Erzeugendensystem gegeben haben und suchen eine Basis von W. Sei also $\{y_1, \dots, y_m\}$ ein Erzeugendensystem von W, d.h.

$$W = \operatorname{Span}\{y_1, \dots, y_m\}.$$

5.19 Lemma:

(1) $\operatorname{Span}\{y_1, \dots, y_i, \dots, y_j, \dots, y_m\} = \operatorname{Span}\{y_1, \dots, y_j, \dots, y_i, \dots, y_m\}$

(2) Für $c \in \mathbb{K}\backslash\{0\}$ ist

$$\operatorname{Span}\{y_1, \dots, y_i, \dots, y_m\} = \operatorname{Span}\{y_1, \dots, c \cdot y_i, \dots, y_m\}$$

(3) Für $i \neq j$ und $c \in \mathbb{K}$ ist

$$\operatorname{Span}\{y_1, \dots, y_m\} = \operatorname{Span}\{y_1, \dots, y_i + cy_j, \dots, y_j, \dots, y_m\}$$

Beweis: (1) und (2) sind trivial. Für (3) beachte:
$y_i + cy_j \in \operatorname{Span}\{y_1, \dots, y_m\}$, und wegen $y_i = y_i + cy_j - cy_j$

$$y_i \in \operatorname{Span}\{y_1, \dots, y_i + cy_j, \dots, y_j, \dots, y_m\}.$$

$\square$

Sei nun $y_i = (a_{i1}, \dots, a_{in}),\ i = 1, \dots, m$. Wir schreiben die y_i als Zeilen eines Schemas

$$\begin{array}{cccc|}
x_1 & x_2 & & x_n \\
\hline
a_{11} & a_{12} & \cdots\cdots & a_{1n} \\
\vdots & \vdots & & \vdots \\
a_{m1} & a_{m2} & \cdots\cdots & a_{mn}
\end{array}$$

Nach (5.19) spannen die Zeilenvektoren bei Umformungen nach dem Gauß'schen Verfahren wieder W auf. Aber **Achtung:** Spaltenvertauschung bedeutet Koordinatenvertauschung; d.h. die x_i-Spalte gibt die i-ten Koordinaten an.

Wir erläutern das Ergebnis am Endschema (5.14($*$)). (Im allgemeinen Fall muss man noch die Koordinaten zurück tauschen.) Danach hat W die Vektoren

$$\begin{array}{rcl}
w^1 & = & (1, 0, \dots, 0, a^*_{1,r+1}, \dots, a^*_{1,n}) \\
\vdots & & \\
w^r & = & (0, \dots, 0, 1, a^*_{r,r+1}, \dots, a^*_{r,n})
\end{array}$$

als Basis: Nach (5.19) spannen sie W auf. Sie sind linear unabhängig, wie man an den ersten r Koordinaten sofort sieht.

Beispiel: Wir suchen eine Basis von

$$W = \mathrm{Span}\{v^1, v^2, v^3, v^4\} \subset \mathbb{R}^5$$

mit

$$\begin{aligned}
v^1 &= (1, -2, 1, 1, -1) \\
v^2 &= (2, 1, 2, 1, 2) \\
v^3 &= (-1, -2, -1, -1, -1) \\
v^4 &= (3, 1, 3, 2, 2).
\end{aligned}$$

Das Schema mit den Zeilen $v^1, \dots, v^4$ ist das des Beispiels (1.15). Aus dem reduzierten Schema (siehe vorausgegangenes Beispiel) entnehmen wir als Basis für W die drei Vektoren

$$\begin{aligned}
w^1 &= (1, -3, 1, 0, 0) \\
w^2 &= (0, 3, 0, 1, 0) \\
w^3 &= (0, 2, 0, 0, 1).
\end{aligned}$$

Die i-te Koordinate entnimmt man der x_i-Spalte.

Wir schließen den Abschnitt mit zwei Dimensionsformeln.

5.20 Satz: Seien W_1 und W_2 endlich dimensionale Untervektorräume eines $\mathbb{K}$-Vektorraumes V. Dann gilt

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2).$$

Beweis: Sei $\{x_1, \dots, x_n\}$ Basis von $W_1 \cap W_2$. Wir ergänzen diese Basis zu Basen $\{x_1, \dots, x_n, y_1, \dots, y_k\}$ von W_1 und $\{x_1, \dots, x_n, z_1, \dots, z_\ell\}$ von W_2.

Behauptung: $\{x_1, \dots, x_n, y_1, \dots, y_k, z_1, \dots, z_\ell\}$ ist Basis von $W_1 + W_2$. (Damit folgt der Satz, denn $\dim W_1 + \dim W_2 = n+k+n+\ell$ und $\dim(W_1+W_2)+\dim(W_1 \cap W_2) = n+k+\ell+n$).

Sei $x \in W_1 + W_2$, d.h. $x = u + v$ mit $u \in W_1$ und $v \in W_2$. Dann ist u $\mathbb{K}$-Linearkombination von $\{x_1, \dots, x_n, y_1, \dots, y_k\}$ und v ist $\mathbb{K}$-Linearkombination von $\{x_1, \dots, x_n, z_1, \dots, z_\ell\}$. Also ist x Linearkombination von $\{x_1, \dots, x_n, y_1, \dots, y_k, z_1, \dots, z_\ell\}$, d.h.

$$W_1 + W_2 \subset \mathrm{Span}\{x_1, \dots, x_n, y_1, \dots, y_k, z_1, \dots, z_\ell\}.$$

Da alle $x_i, y_i, z_i \in W_1 + W_2$, ist $\{x_1, \dots, x_n, y_1, \dots, y_k, z_1, \dots, z_\ell\}$ ein Erzeugendensystem von $W_1 + W_2$. Es ist auch linear unabhängig: Sei

$$a_1x_1 + \cdots + a_nx_n + b_1y_1 + \cdots + b_ky_k + c_1z_1 + \cdots + c_\ell z_\ell = 0$$

$a_i, b_i, c_i \in \mathbb{K}$. Sei nun

$$\begin{aligned} x &= a_1x_1 + \cdots + a_nx_n \\ y &= b_1y_1 + \cdots + b_ny_k \\ z &= c_1z_1 + \cdots + c_\ell z_\ell \end{aligned}$$

Dann gilt $x + y = -z \in W_1 \cap W_2$. Da aber $\{x_1, \dots, x_n\}$ Basis von $W_1 \cap W_2$ und $\{x_1, \dots, x_n, y_1, \dots, y_k\}$ Basis von W_1 ist, folgt, dass alle $b_i = 0$ sein müssen, also $x = -z$. Dann folgt analog, dass alle c_i und damit auch alle a_i Null sein müssen. $\square$

5.21 Satz: Sei V ein $\mathbb{K}$-Vektorraum und $W \subset V$ ein Untervektorraum, der die direkte Summe von k endlich dimensionalen Untervektorräumen W_i ist, also $W = W_1 \oplus \cdots \oplus W_k$. Dann gilt

$$\dim W = \dim W_1 + \cdots + \dim W_k.$$

Für den Beweis benutzen wir mal wieder eine Ergänzung aus den mathematischen Grundlagen.

5.22 Ergänzung: Das **INDUKTIONSPRINZIP**: Sei $A(n)$ eine Behauptung über natürliche Zahlen n für $n \geq n_0$. Gilt

(1) $A(n_0)$ ist eine wahre Aussage.

(2) $A(n) \Longrightarrow A(n+1)$ für $n \geq n_0$.

Dann ist $A(n)$ eine wahre Behauptung für alle $n \geq n_0$.
Denn $A(n_0)$ ist wahr, nach (2) also auch $A(n_0+1)$, also auch $A(n_0+2)$ usw. Das Induktionsprinzip folgt sofort aus den sog. Peano-Axiomen für die natürlichen Zahlen.

Beweis: 5.21: Sei $A(k)$ die Aussage des Satzes.
$A(1)$ ist wahr. Klar!
$A(k) \Longrightarrow A(k+1)$: Wir dürfen voraussetzen, dass $A(k)$ wahr ist. Sei nun $W = W_1 \oplus \cdots \oplus W_k \oplus W_{k+1}$ und $W' = W_1 \oplus \cdots \oplus W_k$. Dann gilt nach Def. 4.30: $W = W' + W_{k+1}$ und $W' \cap W_{k+1} = \{0\}$. Aus (5.20) folgt

$$\dim W = \dim W' + \dim W_{k+1} - \dim(W' \cap W_{k+1}) = \dim W' + \dim W_{k+1}.$$

Da $A(k)$ gilt, haben wir

$$\dim W' = \dim W_1 + \cdots + \dim W_k.$$

□

5.23 Beispiel: $A(n)$ sei die Aussage: $n^2 - 4n + 3 \geq 0$

Behauptung: $A(n)$ ist wahr für $n \geq 3$.
Beweis: $A(3)$ ist wahr, denn $3^2 - 4 \cdot 3 + 3 = 0$.
$A(n) \Longrightarrow A(n+1)$: Wir untersuchen

$$\begin{aligned}(n+1)^2 - 4(n+1) + 3 &= n^2 + 2n + 1 - 4n - 4 + 3 \\ &= \underbrace{n^2 - 4n + 3}_{\geq 0 \text{ wegen } A(n)} + \underbrace{2n - 3}_{\geq 0 \text{ wegen } n \geq 3} \geq 0\end{aligned}$$

□

Teil II

Lineare Abbildungen

6 Grundbegriffe

6.1 **Achtung! Mathematisches Grundprinzip:** Will man Mengen mit einer mathematischen Struktur miteinander vergleichen, betrachtet man Abbildungen zwischen ihnen, die diese Struktur erhalten!

6.2 Definition: Eine Abbildung $f : V \longrightarrow W$ zwischen zwei $\mathbb{K}$-Vektorräumen heißt **LINEAR**, genauer $\mathbb{K}$-linear, wenn $\forall\, x,y \in V$ und $\forall\, a \in \mathbb{K}$ gilt

(1) $f(x+y) = f(x) + f(y)$

(2) $f(a \cdot x) = a \cdot f(x)$

D.h. eine lineare Abbildung erhält die Addition und die skalare Multiplikation auf einem Vektorraum.

6.3 **Einfache Eigenschaften:** Für eine lineare Abbildung $f : V \longrightarrow W$ gilt:

(1) $f(0) = 0$

(2) $f(-x) = -f(x) \quad \forall\, x \in V$

(3) $f(\sum_{i=1}^n a_i x_i) = \sum_{i=1}^n a_i \cdot f(x_i) \quad \forall\, x_1, \dots, x_n \in V, \quad \forall\, a_1, \dots, a_n \in \mathbb{K}$

(4) $\{x_1, \dots, x_n\}$ linear abhängig in $V \Longrightarrow \{f(x_1), \dots, f(x_n)\}$ linear abhängig in W.

(5) Die Abbildung $f : V \longrightarrow W$ ist genau dann linear, wenn

$$f(a \cdot x + y) = a \cdot f(x) + f(y) \quad \forall\, x,y \in V, \quad \forall\, a \in \mathbb{K}.$$

Beweis: (1) $f(0) = f(0+0) = f(0) + f(0)$. Addiere $-f(0)$.
(2) $f(x) + f(-x) = f(x-x) = f(0) = 0$. Addiere $-f(x)$.
(3) Beweis durch Induktion: Sei $A(n)$ die Aussage (3). Dann ist $A(1)$ wahr.

$A(n) \Longrightarrow A(n+1)$: Setze $y = \sum_{i=1}^{n} a_i \cdot x_i$. Dann gilt unter Benutzung von $A(n)$

$$
\begin{aligned}
&f\left(\sum_{i=1}^{n+1} a_i x_i\right) \\
&= f(y + a_{n+1} \cdot x_{n+1}) \overset{6.2.1}{=} f(y) + f(a_{n+1} x_{n+1}) \overset{6.2.2}{=} f(y) + a_{n+1} f(x_{n+1}) \\
&= \sum_{i=1}^{n} a_i f(x_i) + a_{n+1} f(x_{n+1}) = \sum_{i=1}^{n+1} a_i \cdot f(x_i)
\end{aligned}
$$

(5) „$\Longrightarrow$“ ist (6.3.3). „$\Longleftarrow$“: Mit $a = 1$ folgt Axiom 6.2.1. Wie in (1) folgert man daraus, dass $f(0) = 0$. Mit $y = 0$ folgt Axiom 6.2.2.

(4) $x_1, \ldots, x_n$ linear abhängig $\Longrightarrow$ $\exists\ (a_1, \ldots, a_n) \neq (0, \ldots, 0)$ in $\mathbb{K}^n$ mit $\sum\limits_{i=1}^{n} a_i x_i = 0$. Mit diesen a_i gilt

$$\sum_{i=1}^{n} a_i f(x_i) \overset{(3)}{=} f\left(\sum_{i=1}^{n} a_i x_i\right) = f(0) = 0.$$

Also sind $f(x_1), \ldots, f(x_n)$ linear abhängig. $\square$

6.4 Lemma: Es seien $f : V_1 \longrightarrow V_2$ und $g : V_2 \longrightarrow V_3$ lineare Abbildungen von $\mathbb{K}$-Vektorräumen. Dann gilt

(1) id_{V_1} ist linear.

(2) $g \circ f$ ist linear.

Beweis: (1) ist trivial.

(2) $(g \circ f)(ax + y) = g(f(ax + y)) = g(a \cdot f(x) + f(y)) = a \cdot g(f(x)) + g(f(y))$

$$= a \cdot (g \circ f)(x) + (g \circ f)(y)$$

$\square$

6.5 Definition: Eine lineare Abbildung $f : V \longrightarrow W$ von $\mathbb{K}$-Vektorräumen heißt **ISOMORPHISMUS**, wenn es eine lineare Abbildung $g : W \longrightarrow V$ gibt, so dass $g \circ f = id_V$ und $f \circ g = id_W$. Wir schreiben $V \cong W$, V und W heißen **ISOMORPH**. g heißt **INVERS** zu f. (Nach 3.22 ist $g = f^{-1}$.)

6.6 <u>Achtung!</u> Mathematisches Grundprinzip: Ist $V \cong W$, so sind V und W aus mathematischer Sicht im wesentlichen gleich: Aus der Struktur von V erhält man über f und g die Struktur von W. Z.B. gilt für $w_1, w_2 \in W$

$$w_1 + w_2 = f \circ g(w_1 + w_2) = f(g(w_1) + g(w_2)).$$

Links haben wir die Addition in W. Sie ist durch die Addition $g(w_1)+g(w_2)$ in V eindeutig bestimmt.

Aus demselben Grund haben V und W als Vektorräume dieselben Eigenschaften. Kennt man V, so kennt man auch W als $\mathbb{K}$-Vektorraum. Definition 6.5 lässt sich natürlich auch auf andere Strukturen übertragen, etwa auf Körper.

In vielen Lehrbüchern werden Isomorphismen anders eingeführt. Dabei vergisst man aber leicht den allgemeinen mathematischen Grundsatz. Der folgende Satz enthält die häufiger benutzte Beschreibung eines linearen Isomorphismus!

6.7 Satz: Es seien U, V, W $\mathbb{K}$-Vektorräume. Dann gilt

(1) $f : U \longrightarrow V$ ist ein Isomorphismus $\Longleftrightarrow$ f ist linear und bijektiv.

(2) id_V ist ein Isomorphismus.

(3) Sind $f : U \longrightarrow V$ und $g : V \longrightarrow W$ Isomorphismen, dann ist $g \circ f$ ein Isomorphismus.

Beweis: (1) „$\Longrightarrow$" folgt aus (3.20).

„$\Longleftarrow$" Sei $g : V \longrightarrow U$ die Umkehrabbildung von f. Wir müssen nur zeigen, dass g linear ist. Seien $v_1, v_2 \in V$ und $u_1 = g(v_1), u_2 = g(v_2)$. Dann gilt $f(u_1) = v_1$ und $f(u_2) = v_2$. Es folgt

$$\begin{aligned} g(a \cdot v_1 + v_2) &= g(a \cdot f(u_1) + f(u_2)) = (g \circ f)(a \cdot u_1 + u_2) \\ &= a \cdot u_1 + u_2 = a \cdot g(v_1) + g(v_2). \end{aligned}$$

(2) und (3) folgen sofort aus (1) und (6.4). □

6.8 Satz und Definition: Sei $f : V \longrightarrow W$ eine lineare Abbildung von $\mathbb{K}$-Vektorräumen. Dann sind

$$\begin{aligned} \operatorname{Kern} f &= \{x \in V;\ f(x) = 0\} = f^{-1}(0) \\ \operatorname{Bild} f &= f(V) = \{y \in W;\ \exists\, x \in V \text{ mit } f(x) = y\} \end{aligned}$$

Untervektorräume von V bzw. W, genannt der **KERN** und das **BILD** von f.

Wir müssen zeigen, dass Kern f und Bild f Untervektorräume sind. Da $\{0\} \subset W$ und $V \subset V$ Untervektorräume sind (genannt die **TRIVIALEN** Untervektorräume), folgt das aus folgendem allgemeineren Resultat.

6.9 Satz: Sei $f : V \longrightarrow V'$ eine lineare Abbildung. Dann gilt

(1) Ist $U \subset V$ Untervektorraum, dann ist $f(U) \subset V'$ Untervektorraum.

(2) Ist $U' \subset V'$ Untervektorraum, dann ist $f^{-1}(U') \subset V$ Untervektorraum.

Beweis: (1) Da $U \neq \emptyset$, ist $f(U) \neq \emptyset$. Seien nun $y, y' \in f(U)$ und sei $a \in \mathbb{K}$. Dann existiert $x, x' \in U$ mit $f(x) = y,\ f(x') = y'$. Es folgt

$$\begin{array}{rcll} y+y' & = & f(x)+f(x') = f(x+x') \in f(U) & , \quad \text{da } x+x' \in U \\ a\cdot y & = & a\cdot f(x) = f(a\cdot x) \in f(U) & , \quad \text{da } a\cdot x \in U. \end{array}$$

(2) Da $f(0) = 0 \in U'$, ist $0 \in f^{-1}(U')$, also $f^{-1}(U') \neq \emptyset$. Seien x, x' aus $f^{-1}(U')$, d.h. $f(x), f(x') \in U'$, und sei $a \in \mathbb{K}$. Es gilt

$$\begin{array}{rcll} f(x+x') & = & f(x)+f(x') \in U' & , \quad \text{also } x+x' \in f^{-1}(U') \\ f(a\cdot x) & = & a\cdot f(x) \in U' & , \quad \text{also } a\cdot x \in f^{-1}(U'). \end{array}$$

□

Ein oft benutztes Kriterium für die Injektivität einer linearen Abbildung ist

6.10 Satz: Eine lineare Abbildung $f: V \longrightarrow W$ ist genau dann injektiv, wenn $\operatorname{Kern} f = \{0\}$ (wir schreiben in Zukunft kürzer $\operatorname{Kern} f = 0$).

Beweis: Sei f injektiv. Für $x \in \operatorname{Kern} f$ gilt $f(x) = 0 = f(0)$. Da f injektiv ist, folgt $x = 0$.

Sei $\operatorname{Kern} f = 0$ und $f(x) = f(y)$. Dann gilt

$$0 = f(x) - f(y) = f(x-y).$$

Also $x - y \in \operatorname{Kern} f$, d.h. $x - y = 0$. Also $x = y$. □

6.11 Satz: Sei $f: V \longrightarrow W$ eine lineare Abbildung und V endlich dimensional. Dann gilt:

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

Wir vermerken zunächst

6.12 Ist V ein endlichdimensionaler Vektorraum und $U \subset V$ ein Untervektorraum, dann gilt $\dim U \leq \dim V$, denn U kann nicht mehr linear unabhängige Vektoren haben als V.

Beweis des Satzes: Als Untervektorraum eines endlich dimensionalen Vektorraumes ist Kern f endlich dimensional. Sei $\{x_1, \ldots, x_k\}$ Basis von $\operatorname{Kern} f$ und $\{x_1, \ldots, x_k, y_1, \ldots, y_r\}$ die Ergänzung zu einer Basis von V (benutze (5.7)).

Behauptung: $\{f(y_1),\ldots,f(y_r)\}$ ist Basis von Bild f (damit folgt der Satz).
Beweis: Sei $x \in V$ beliebig, $x = a_1 \cdot x_1 + \cdots + a_k x_k + b_1 y_1 + \cdots + b_r y_r$. Dann gilt

$$f(x) = b_1 \cdot f(y_1) + \cdots + b_r \cdot f(y_r)$$

da $f(x_i) = 0$ für $i = 1, \cdots, k$. Also ist $\{f(y_1),\ldots,f(y_r)\}$ ein Erzeugendensystem von Bild f. Das System ist linear unabhängig, denn aus

$$a_1 \cdot f(y_1) + \cdots + a_r \cdot f(y_r) = 0$$

folgt

$$f(a_1 y_1 + \cdots + a_r y_r) = 0, \text{ also } a_1 y_1 + \cdots + a_r y_r \in \text{Kern } f.$$

Das ist aber nur möglich, wenn $a_1 = \ldots = a_r = 0$. $\square$

6.13 Definition: Die Zahl $\operatorname{rg} f = \dim(\text{Bild } f)$ heißt **RANG** von f.

$\operatorname{crg} f = \dim(\text{Kern } f)$ heißt **CORANG** von f.

Lineare Abbildungen haben die schöne Eigenschaft, dass sie durch ihre Werte auf einigen wenigen Vektoren schon festgelegt sind.

6.14 Satz: Seien V, W $\mathbb{K}$-Vektorräume, $\mathscr{B} = \{x_1,\ldots,x_n\}$ eine **Basis** von V und $y_1,\ldots,y_n$ beliebige Vektoren aus W. Dann gibt es genau eine lineare Abbildung

$$f : V \longrightarrow W$$

mit $f(x_i) = y_i$ für $i = 1,\ldots,n$.

Beweis: Unsere Vorgangsweise ist typisch für viele Existenz- und Eindeutigkeitsbeweise: Man nimmt zunächst an, dass f existiert und untersucht, wie f aussehen muss. Dieses Ergebnis verwendet man dann für den Nachweis der Existenz.
Sei also $f : V \longrightarrow W$ eine lineare Abbildung, so dass $f(x_i) = y_i$. Da $\mathscr{B}$ eine Basis ist, hat jedes $x \in V$ eine eindeutige Darstellung der Form

$$x = a_1 \cdot x_1 + \cdots + a_n x_n.$$

Weil f linear ist, muss gelten

$$f(x) = \sum_{i=1}^{n} a_i \cdot f(x_i) = \sum_{i=1}^{n} a_i \cdot y_i. \qquad (*)$$

Damit kann es höchstens eine solche lineare Abbildung geben. Das Bild $f(x)$ von x ist durch $(*)$ für jedes $x \in V$ vorgegeben.

Für die Existenz von f brauchen wir jetzt nur zu zeigen, dass durch die Vorschrift $(*)$ eine lineare Abbildung definiert wird. Sei

$$y = b_1 x_1 + \cdots + b_n x_n$$

ein weiterer Vektor aus V. Dann ist $a \cdot x = (a \cdot a_1)x_1 + \cdots + (a \cdot a_n)x_n$, und

$$x + y = (a_1 + b_1)x_1 + \cdots + (a_n + b_n)x_n.$$

Nach Vorschrift $(*)$ gilt

$$f(x+y) = \sum_{i=1}^{n} (a_i + b_i)y_i, \quad f(y) = \sum_{i=1}^{n} b_i y_i, \quad f(ax) = \sum_{i=1}^{n} aa_i \cdot y_i.$$

Es folgt: $f(x+y) = f(x) + f(y), \; f(ax) = a \cdot f(x)$. $\square$

6.15 Folgerung: Sind $f, g : V \longrightarrow W$ zwei lineare Abbildungen, die auf einem Erzeugendensystem M von V übereinstimmen, dann gilt $f = g$.

Denn jedes Erzeugendensystem enthält eine Basis. $\square$

Dieser einfache Satz hat eine Reihe weitreichender Konsequenzen:

6.16 Satz: Sei V ein endlich dimensionaler $\mathbb{K}$-Vektorraum, $\dim V = n$. Dann gilt

$$V \cong \mathbb{K}^n$$

(dabei wird $\mathbb{K}^0$ als $\{0\}$ definiert).

Beweis: Sei $\{x_1, \ldots, x_n\}$ eine Basis von V und $\{e^1, \ldots, e^n\}$ die Standardbasis von $\mathbb{K}^n$. Dann gibt es genau eine lineare Abbildung $f : V \longrightarrow \mathbb{K}^n$ mit $f(x_i) = e^i$ für $i = 1, \ldots, n$ und genau eine lineare Abbildung $g : \mathbb{K}^n \longrightarrow V$ mit $g(e^i) = x_i$ für $i = 1, \ldots, n$. Für die linearen Abbildungen $f \circ g$ und $g \circ f$ gilt $f \circ g(e^i) = e^i$, $g \circ f(x_i) = x_i$ für $i = 1, \ldots, n$. Die linearen Abbildungen $id_{\mathbb{K}^n}$ und id_V erfüllen dieselben Bedingungen. Aus dem Eindeutigkeitsteil von (6.14) folgt $f \circ g = id_{\mathbb{K}^n}$ und $g \circ f = id_V$. $\square$

Isomorphe Vektorräume haben die gleichen Eigenschaften. Also muss insbesondere gelten:

6.17 Ist $f : V \longrightarrow W$ ein Isomorphismus von Vektorräumen und sind $x_1, \ldots, x_n$ aus V linear unabhängig, so sind auch $f(x_1), \ldots, f(x_n)$ aus W linear unabhängig.

Wir kontrollieren das nach: Sei $g : W \longrightarrow V$ der inverse Isomorphismus und

$$a_1 f(x_1) + \cdots + a_n f(x_n) = 0.$$

Dann folgt

$$\begin{aligned} 0 = g(0) &= g(a_1 f(x_1) + \cdots + a_n f(x_n)) \\ &= a_1 \cdot (g \circ f)(x_1) + \cdots + a_n \cdot (g \circ f)(x_n) \\ &= a_1 x_1 + \cdots + a_n x_n. \end{aligned}$$

Da $x_1, \ldots, x_n$ linear unabhängig sind, folgt $a_1 = a_2 = \ldots = a_n = 0$. $\square$

Als Folgerung erhalten wir

6.18 Satz: (1) Sind V, W beliebige $\mathbb{K}$-Vektorräume und $V \cong W$, dann folgt

$$\dim V = \dim W.$$

(2) Sind V und W endlich dimensional, so gilt: $V \cong W \Longleftrightarrow \dim V = \dim W$.

Beweis: Ist $f : V \longrightarrow W$ ein Isomorphismus mit Inversem $g : W \longrightarrow V$. Da f und g linear unabhängige Systeme auf linear unabhängige Systeme abbilden, gilt $\dim V \leq \dim W$ wegen f und $\dim W \leq \dim V$ wegen g. Sind nun V und W endlich dimensional und $\dim V = \dim W = n$, folgt $V \cong \mathbb{K}^n$, $W \cong \mathbb{K}^n$, also auch $V \cong W$. $\square$

Aus (6.18 (2)) und (6.11) folgt sofort

6.19 Satz: Ist $f : V \longrightarrow W$ linear und V endlich dimensional, dann gilt

$$V \cong \operatorname{Kern} f \oplus \operatorname{Bild} f \qquad (= \operatorname{Kern} f \times \operatorname{Bild} f).$$

Der Isomorphismus ist von einer Basiswahl abhängig. Sei also $\{x_1, \ldots, x_k\}$ Basis von $\operatorname{Kern} f$ und $\{x_1, \ldots, x_k, y_1, \ldots, y_r\}$ eine Ergänzung zu einer Basis von V. Aus dem Beweis von (6.11) wissen wir, dass $\{f(y_1), \ldots, f(y_r)\}$ Basis von $\operatorname{Bild} f$ ist. Dann ist $\{(x_1, 0), \ldots, (x_k, 0), (0, f(y_1)), \ldots, (0, f(y_r))\}$ Basis von $\operatorname{Kern} f \times \operatorname{Bild} f$ und $x_i \longmapsto (x_i, 0), i = 1, \ldots, k$, $y_j \longmapsto (0, f(y_j))$ bildet die Basis von V bijektiv auf diese Basis von $\operatorname{Kern} f \times \operatorname{Bild} f$ ab. Nach (6.14) wird dadurch ein linearer Isomorphismus $V \longrightarrow \operatorname{Kern} f \times \operatorname{Bild} f$ definiert.

6.20 Bezeichnung: Sind V und W $\mathbb{K}$-Vektorräume, dann bezeichnen wir die Menge der $\mathbb{K}$-linearen Abbildungen $V \longrightarrow W$ mit

$$\operatorname{Hom}_{\mathbb{K}}(V, W).$$

6.21 Satz: $\operatorname{Hom}_{\mathbb{K}}(V, W)$ ist mit den Verknüpfungen

$$\begin{aligned} (f + g)(v) &:= f(v) + g(v) \qquad && f, g \in \operatorname{Hom}_{\mathbb{K}}(V, W) \\ (a \cdot f)(v) &:= a \cdot f(v) && a \in \mathbb{K}, f \in \operatorname{Hom}_{\mathbb{K}}(V, W) \end{aligned}$$

ein $\mathbb{K}$-Untervektorraum des $\mathbb{K}$-Vektorraumes $\operatorname{Abb}(V, W)$ (vergl. (4.15)).

Beweis: Nach (4.20) müssen wir zeigen: Sind $f, g : V \longrightarrow W$ $\mathbb{K}$-linear, dann sind $f+g$ und $a \cdot f$ mit $a \in \mathbb{K}$ auch $\mathbb{K}$-linear. Seien $v, w \in V$ und sei $r \in \mathbb{K}$.

$$\begin{aligned}
(f+g)(r \cdot v + w) &= f(r \cdot v + w) + g(r \cdot v + w) \\
&= r \cdot f(v) + f(w) + r \cdot g(v) + g(w) \\
&= r \cdot (f(v) + g(v)) + f(w) + g(w) \\
&= r \cdot (f+g)(v) + (f+g)(w) \\
(a \cdot f)(r \cdot v + w) &= a \cdot f(r \cdot v + w) \\
&= a \cdot r \cdot f(v) + a \cdot f(w) \\
&= r \cdot (a \cdot f)(v) + (a \cdot f)(w)
\end{aligned}$$

$\mathrm{Hom}_{\mathbb{K}}(V, W) \neq \emptyset$, denn die **NULLABBILDUNG**

$$\underline{0}^{W,V} : V \longrightarrow W, \quad v \longmapsto 0 \quad \forall\, v \in V$$

ist linear. □

Wenden wir uns wieder der Komposition zu. Man prüft leicht nach

6.22 Für $\mathbb{K}$-lineare Abbildungen $f : U \longrightarrow V, g, g_1, g_2 : V \longrightarrow W$ und $h : W \longrightarrow X$ gilt:

(1) $\underline{0}^{W,V} \circ f = \underline{0}^{W,U} = g \circ \underline{0}^{V,U}$

(2) $id_V \circ f = f = f \circ id_U$

(3) $(g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f$

(4) $h \circ (g_1 + g_2) = h \circ g_1 + h \circ g_2$

(5) $(a \cdot g) \circ f = g \circ (a \cdot f) = a \cdot (g \circ f) \quad \forall\, a \in \mathbb{K}$

6.23 **WARNUNG:** Die Komposition

$$\mathrm{Hom}_{\mathbb{K}}(V, W) \times \mathrm{Hom}_{\mathbb{K}}(U, V) \xrightarrow{c} \mathrm{Hom}_{\mathbb{K}}(U, W)$$

ist **NICHT** linear. Hier geben wir der linken Seite die Produktstruktur.

Wäre die Komposition linear, müsste für $f_1, f_2 \in \mathrm{Hom}_{\mathbb{K}}(U, V)$ und $g_1, g_2 \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ gelten

$$c((g_1, f_1) + (g_2, f_2)) = c(g_1, f_1) + c(g_2, f_2) = g_1 \circ f_1 + g_2 \circ f_2.$$

Aber $c((g_1,f_1)+(g_2,f_2)) = c(g_1+g_2,f_1+f_2) = (g_1+g_2)\circ(f_1+f_2)$

$$\underset{(6.21)}{=} g_1\circ(f_1+f_2)+g_2\circ(f_1+f_2) = g_1\circ f_1+g_1\circ f_2+g_2\circ f_1+g_2\circ f_2$$

Die Komposition definiert auf $\mathrm{Hom}_{\mathbb{K}}(V,V)$ eine Verknüpfung, für die nach (6.22) die Distributivgesetze gelten. Wir erhalten fast so etwas wie einen Körper.

6.24 Definition: Eine lineare Abbildung $f : V \longrightarrow V$ heißt **LINEARER ENDOMORPHISMUS**. Wir schreiben

$$\mathrm{End}_{\mathbb{K}}(V) \text{ für } \mathrm{Hom}_{\mathbb{K}}(V,V).$$

6.25 Satz: $\mathrm{End}_{\mathbb{K}}(V)$ ist mit der Komposition als „Multiplikation“ und der Addition aus (6.21) einen RING, genauer eine $\mathbb{K}$-ALGEBRA.

6.26 Definition: Ein **RING** ist eine Menge R mit zwei Verknüpfungen

Addition: $R\times R \longrightarrow R, \quad (a,b)\longmapsto a+b$

Multiplikation: $R\times R \longrightarrow R, \quad (a,b)\longmapsto a\cdot b$

Es gelten die Axiome (vergl. die Körperxiome 2.1)

(A1) $(a+b)+c = a+(b+c) \quad \forall\, a,b,c\in R$

(A2) $a+b = b+a \quad \forall\, a,b\in R$

(A3) Es gibt ein Nullelement $0\in R$, so dass für alle $a\in R$ gilt: $\quad a+0=a$

(A4) Zu $a\in R$ existiert $-a\in R$, so dass $a+(-a)=0$

(M1) $(a\cdot b)\cdot c = a\cdot(b\cdot c) \quad \forall\, a,b,c\in R$

(M3) Es gibt ein Einselement $1\in R$, so dass für alle $a\in R$ gilt: $\quad 1\cdot a = a = a\cdot 1$

(D) Es gelten die Distributivgesetze

$$\left.\begin{array}{rcl} a\cdot(b+c) &=& a\cdot b+a\cdot c \\ (b+c)\cdot a &=& b\cdot a+c\cdot a \end{array}\right. \qquad \forall\, a,b,c\in R$$

Gilt außerdem (M2): $a\cdot b = b\cdot a \quad \forall\, a,b\in R$, nennt man R einen **KOMMUTATIVEN RING**.

Beispiele: Der Ring $\mathbb{Z}$ der ganzen Zahlen und die “Restklassenringe” $\mathbb{Z}/n$.

6.27 Definition: Eine $\mathbb{K}$-**ALGEBRA** ist ein $\mathbb{K}$-Vektorraum $(A,+,\cdot)$ mit i einer Multiplikation

$$A\times A \longrightarrow A, \quad (a,b)\longmapsto a*b,$$

so dass gilt

(1) $(A,+,*)$ ist ein Ring.

(2) $\forall\, r \in \mathbb{K}, \quad \forall\, a,b \in A$ gilt

$$r \cdot (a * b) = (r \cdot a) * b = a * (r \cdot b)$$

6.28 Definition: Sei $g : V \longrightarrow W$ eine lineare Abbildung. Eine Abbildung $f : W \longrightarrow V$ heißt **RECHTSINVERS** zu g und $h : W \longrightarrow V$ **LINKSINVERS** zu g, falls $g \circ f = id_W$ bzw. $h \circ g = id_V$.

6.29 Lemma: Sei $g : V \longrightarrow W$ eine lineare Abbildung.

(1) Zwei linksinverse $h_1, h_2 : W \longrightarrow V$ und zwei rechtsinverse $f_1, f_2 : W \longrightarrow V$ von g können verschieden sein.

(2) Besitzt g sowohl links als auch rechtsinverse, so sind diese gleich. g ist linearer Isomorphismus und die links- und rechtsinversen sind die Umkehrabbildung.

Beweis: (2): Sei f ein beliebiges rechtsinverses und h ein beliebiges linksinverses, also $g \circ f = id_W,\ h \circ g = id_V$. Dann gilt

$$f = id_V \circ f = (h \circ g) \circ f = h \circ (g \circ f) = h \circ id_W = h.$$

Damit ist g linearer Isomorphismus mit Inversen f. $\square$

Beispiel für (6.28 (1)): Die Abbildungen $g : \mathbb{R} \longrightarrow \mathbb{R}^2,\ x \longmapsto (x,0)$ und

$$h : \mathbb{R} \longrightarrow \mathbb{R}^2, \quad x \longmapsto (x,x)$$

sind $\mathbb{R}$-linear. Die Abbildungen

$$p_1 : \mathbb{R}^2 \longrightarrow \mathbb{R},\ (x_1,x_2) \longmapsto x_1$$
$$p_2 : \mathbb{R}^2 \longrightarrow \mathbb{R},\ (x_1,x_2) \longmapsto x_2$$

sind ebenfalls $\mathbb{R}$-linear. Es gilt $p_1 \circ h = p_2 \circ h = id_{\mathbb{R}}$ und $p_1 \circ g = p_1 \circ h = id_{\mathbb{R}}$. $\square$

Zum Abschluss zeigen wir folgendes einfaches und doch zunächst überraschendes Resultat.

6.30 Satz: Seien V und W endlich dimensionale Vektorräume **gleicher** Dimension und $f : V \longrightarrow W$ sei lineare Abbildung. Dann gilt

(1) f injektiv $\Longleftrightarrow$ f bijektiv

(2) f surjektiv $\Longleftrightarrow$ f bijektiv.

Beweis: Die Rückrichtung ist klar. Für die Hinrichtung nutzen wir (6.11):

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

Ist f injektiv, dann ist $\operatorname{Kern} f = 0$, so dass $\dim \operatorname{Bild} f = \dim V = \dim W$, also $\operatorname{Bild} f = W$. Folglich ist f auch surjektiv.

Ist f surjektiv, also $\operatorname{Bild} f = W$, dann ist $\dim \operatorname{Bild} f = \dim W = \dim V$, also dim Kern $f = 0$, und somit $\operatorname{Kern} f = \{0\}$. Folglich ist f auch injektiv. $\square$

7 Matrizen

Satz 6.14 eröffnet uns eine Möglichkeit, lineare Abbildungen rechnerisch zu erfassen.

Sei $f : V \longrightarrow W$ eine lineare Abbildung von $\mathbb{K}$-Vektorräumen, $\dim V = n$ und $\dim W = m$. Wir **WÄHLEN** feste Basen $\mathscr{B} = \{v^1, \dots, v^n\}$ von V und $\mathscr{B}' = \{w^1, \dots, w^m\}$ von W. Nach (6.14) ist f eindeutig durch die Bilder $f(v^j)$ der Basisvektoren $v^1, \dots, v^n$ gegeben. Diese drücken wir als Linearkombination der Basisvektoren von W aus:

7.1 $$f(v^j) = \sum_{i=1}^{m} a_{ij} w^i \qquad j = 1, \dots, n;\ a_{ij} \in \mathbb{K}.$$

Die a_{ij} ordnen wir zu einem Schema

7.2
$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Ein solches Schema nennt man eine $(m \times n)$-**MATRIX** über $\mathbb{K}$. Die a_{ij} nennt man **ELEMENTE** oder **EINTRÄGE** der Matrix A. Den Vektor $z^i = (a_{i1}, a_{i2}, \dots, a_{in}) \in \mathbb{K}^n$ nennt man den i-ten **ZEILENVEKTOR**, den „Vektor“

$$s^j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in \mathbb{K}^m$$

den j-ten **SPALTENVEKTOR** von A. Ist $m = n$, nennt man A auch n-**QUADRATISCHE** Matrix. $\mathbb{K}^{m,n}$ bezeichnet die Menge aller $(m \times n)$-Matrizen. Die **EINHEITSMATRIX** E_n und die **NULLMATRIX** $0_{m,n}$ sind

$$E_n = \underbrace{\begin{pmatrix} 1 & & & \\ & \ddots & & 0 \\ 0 & & \ddots & \\ & & & 1 \end{pmatrix}}_{n} \left.\vphantom{\begin{pmatrix} 1\\ \\ \\ 1 \end{pmatrix}}\right\} n \qquad 0_{m,n} = \underbrace{\begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}}_{n} \left.\vphantom{\begin{pmatrix} 0\\ \vdots \\ 0 \end{pmatrix}}\right\} m$$

7.3 WARNUNG: A hängt von f und der Wahl der Basen $\mathscr{B}$ von V und $\mathscr{B}'$ von W ab. Eine andere Wahl der Basis liefert eine andere Matrix.

7.4 Bezeichnung: Da A von f, $\mathscr{B}$ und $\mathscr{B}'$ abhängt, schreiben wir

$$A = M_{\mathscr{B}}^{\mathscr{B}'}(f).$$

Der j-te Spaltenvektor von $M_{\mathscr{B}}^{\mathscr{B}'}(f)$ besteht aus den Koordinaten von $f(v^j)$ bzgl. der Basis $\mathscr{B}'$.

7.5 Nach (6.14) ist die Abbildung

$$\begin{aligned} \varphi_{\mathscr{B}}^{\mathscr{B}'} : \mathrm{Hom}_{\mathbb{K}}(V,W) &\longrightarrow \mathbb{K}^{m,n} \qquad m = \dim W,\ n = \dim V \\ f &\longmapsto M_{\mathscr{B}}^{\mathscr{B}'}(f) \end{aligned}$$

bijektiv.
Damit überträgt sich die Struktur von $\mathrm{Hom}_{\mathbb{K}}(V,W)$ auf $\mathbb{K}^{m,n}$. Seien

$$V \underset{g}{\overset{f}{\rightrightarrows}} W \xrightarrow{h} X$$

lineare Abbildungen von $\mathbb{K}$-Vektorräumen V mit Basis $\mathscr{B} = \{v^1,\dots,v^n\}$, W mit Basis $\mathscr{B}' = \{w^1,\dots,w^m\}$ und X mit Basis $\mathscr{B}'' = \{x^1,\dots,x^\ell\}$. Sei

$$A = (a_{ij}) = M_{\mathscr{B}}^{\mathscr{B}'}(f),\ B = (b_{ij}) = M_{\mathscr{B}}^{\mathscr{B}'}(g),\ C = (c_{ij}) = M_{\mathscr{B}'}^{\mathscr{B}''}(h)$$

d.h.

$$f(v^j) = \sum_{i=1}^m a_{ij}w^i,\ g(v^j) = \sum_{i=1}^m b_{ij}w^i,\ h(w^i) = \sum_{k=1}^\ell c_{ki}x^k.$$

7.6 Satz: Mit den eben eingeführten Bezeichnungen gilt

(1) $M_{\mathscr{B}}^{\mathscr{B}'}(r\cdot f) = (e_{ij}) \in \mathbb{K}^{m,n}$ mit $e_{ij} = r\cdot a_{ij} \quad \forall\, r \in \mathbb{K}$

(2) $M_{\mathscr{B}}^{\mathscr{B}'}(f+g) = (r_{ij}) \in \mathbb{K}^{m,n}$ mit $r_{ij} = a_{ij} + b_{ij}$

(3) $M_{\mathscr{B}}^{\mathscr{B}''}(h\circ f) = (u_{kj}) \in \mathbb{K}^{\ell,n}$ mit $u_{kj} = \sum\limits_{i=1}^m c_{ki}\cdot a_{ij}$

Beweis: (1) Nach (7.1) gilt $(r\cdot f)(v^j) = \sum\limits_{i=1}^m e_{ij}w^i$. Es gilt aber

$$(r\cdot f)(v^j) = r\cdot f(v^j) = r\cdot \sum_{i=1}^m a_{ij}w^i = \sum_{i=1}^m (ra_{ij})w^i.$$

Nach (5.4) gilt $e_{ij} = ra_{ij}$.
(2) Analog gilt

$$\sum_{i=1}^{m} r_{ij} w^i = (f+g)(v^j) = f(v^j) + g(v^j)$$
$$= \sum_{i=1}^{n} a_{ij} w^i + \sum_{i=1}^{n} b_{ij} w^i = \sum_{i=1}^{n} (a_{ij} + b_{ij}) \cdot w^i.$$

Es folgt nach (5.4)

$$r_{ij} = a_{ij} + b_{ij}$$

(3) $\sum\limits_{k=1}^{\ell} u_{kj} x^k \overset{(7.1)}{=} (h \circ f)(v^j) = h(f(v^j)) = h\left(\sum\limits_{i=1}^{m} a_{ij} w^i \right)$

$$= \sum_{i=1}^{m} h(a_{ij} w^i) = \sum_{i=1}^{m} a_{ij} h(w^i) = \sum_{i=1}^{m} a_{ij} \cdot \sum_{k=1}^{\ell} c_{ki} x^k$$
$$= \sum_{k=1}^{\ell} \left(\sum_{i=1}^{m} c_{ki} a_{ij} \right) \cdot x^k.$$

Es folgt: $u_{kj} = \sum\limits_{i=1}^{m} c_{ki} \cdot a_{ij}$. □

Das motiviert folgende Strukturen auf den Mengen von Matrizen:

7.7 Definition: (1) Skalarmultiplikation von Matrizen

$$\mathbb{K} \times \mathbb{K}^{m,n} \longrightarrow \mathbb{K}^{m,n}, \quad (r, (a_{ij})) \longmapsto (r \cdot a_{ij})$$

(2) Addition (beachte: gleiche Indices)

$$\mathbb{K}^{m,n} \times \mathbb{K}^{m,n} \longrightarrow \mathbb{K}^{m,n}, \quad ((a_{ij}), (b_{ij})) \longmapsto (a_{ij} + b_{ij})$$

(3) Multiplikation (beachte die Indices!!)

$$\mathbb{K}^{\ell,m} \times \mathbb{K}^{m,n} \longrightarrow \mathbb{K}^{\ell,n}, \quad ((c_{ki}), (a_{ij})) \longmapsto (r_{kj})$$

mit $r_{kj} = \sum\limits_{i=1}^{m} c_{ki} \cdot a_{ij}$, $k = 1, \ldots, \ell$; $j = 1, \ldots, n$.

ACHTUNG: Das Produkt $C \cdot A$ von Matrizen ist nur dann definiert, wenn Spaltenzahl (C) = Zeilenzahl (A) ist!

Mit Def(7.7) liest sich Satz 7.6

7.8 Mit den Bezeichnungen von Satz 7.6 gilt

(1) $M_{\mathscr{B}}^{\mathscr{B}'}(r \cdot f) = r \cdot M_{\mathscr{B}}^{\mathscr{B}'}(f)$

(2) $M_{\mathscr{B}}^{\mathscr{B}'}(f+g) = M_{\mathscr{B}}^{\mathscr{B}'}(f) + M_{\mathscr{B}}^{\mathscr{B}'}(g)$

(3) $M_{\mathscr{B}}^{\mathscr{B}''}(h \circ f) = M_{\mathscr{B}'}^{\mathscr{B}''}(h) \cdot M_{\mathscr{B}}^{\mathscr{B}'}(f)$

(4) $\varphi_{\mathscr{B}}^{\mathscr{B}} : \mathrm{End}_{\mathbb{K}}(V) = \mathrm{Hom}_{\mathbb{K}}(V,V) \longrightarrow \mathbb{K}^{n,n}$ erfüllt

$$\varphi_{\mathscr{B}}^{\mathscr{B}}(h \circ f) = M_{\mathscr{B}}^{\mathscr{B}}(h \circ f) = M_{\mathscr{B}}^{\mathscr{B}}(h) \cdot M_{\mathscr{B}}^{\mathscr{B}}(f) = \varphi_{\mathscr{B}}^{\mathscr{B}}(h) \circ \varphi_{\mathscr{B}}^{\mathscr{B}}(f)$$

Insbesondere ist $\mathbb{K}^{n,n}$ eine zu $\mathrm{End}_{\mathbb{K}}(V)$ isomorphe $\mathbb{K}$-Algebra.

7.9 Beispiel:

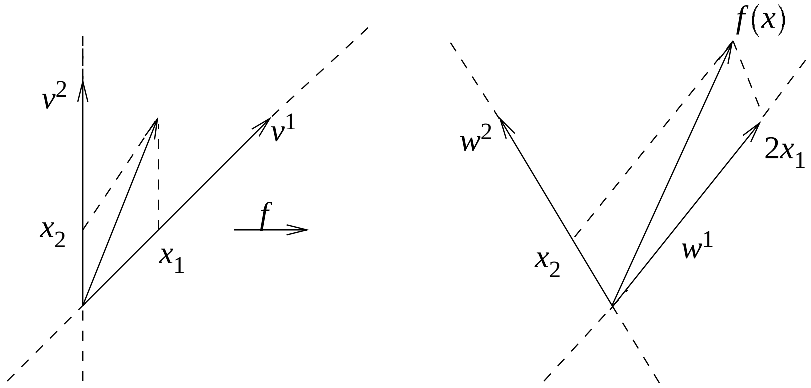
$$A = \begin{pmatrix} 2 & 1 \\ -1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Spaltenzahl $(B) = 3 =$Zeilenzahl (A). Also ist $B \cdot A$ definiert:

$$B \cdot A = \begin{pmatrix} 3 & 4 \\ 1 & 2 \\ -1 & 2 \end{pmatrix}$$

Kehren wir nun zur Ausgangsfrage zurück, der rechnerischen Beschreibung linearer Abbildungen. Wir beginnen mit einem Beispiel, das auch zeigen soll, dass beliebige Basen selbst für die Vektorräume $\mathbb{K}^n$ durchaus sinnvoll sein können.

7.10 Beispiel: $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2,\ (x,y) \longmapsto (3x-y, x+y)$
ist eine lineare Abbildung. Als Basen wählen wir $\mathscr{B} = \{v^1, v^2\},\ \mathscr{B}' = \{w^1, w^2\}$ mit $v^1 = w^1 = (1,1)$ und $v^2 = (0,1),\ w^2 = (-1,1)$. Wir erhalten schiefwinklige Koordinatensysteme



$$\begin{array}{ll} (*) & f(v^1) = (3-1, 1+1) = (2,2) = 2 \cdot w^1 \\ (**) & f(v^2) = (-1,1) = w^2 \end{array} \quad . \quad \text{Also } M_{\mathscr{B}}^{\mathscr{B}'}(f) = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Sei nun $x \in \mathbb{R}^2$ beliebig. Dann gilt bzgl. der schiefwinkligen Koordinaten

$$x = x_1 \cdot v^1 + x_2 \cdot v^2 \qquad (\text{ d.h. } x = (x_1, x_1 + x_2) \in \mathbb{R}^2)$$

und damit

$$f(x) = x_1 \cdot f(v^1) + x_2 \cdot f(v^2) = 2x_1 \cdot w^1 + x_2 \cdot w^2.$$

Damit können wir das Bild sofort ablesen: Nach $(*)$ gilt, die w^1-Komponente ist $2x_1$; nach $(**)$ ist die w^2-Komponente gleich x_2.

7.11 Aufgabe: Ist $f : V \longrightarrow W$ eine lineare Abbildung, $\dim V = n, \dim W = m$, dann gibt es Basen $\mathscr{B}$ von V und $\mathscr{B}'$ von W, so dass

$$M_{\mathscr{B}}^{\mathscr{B}''}(f) = \begin{array}{c} \begin{array}{c} \uparrow \\ r \\ \downarrow \end{array} \\ \\ \\ \end{array} \begin{pmatrix} \begin{matrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{matrix} & 0 \\ 0 & 0 \end{pmatrix} \begin{array}{c} \uparrow \\ m \\ \downarrow \end{array}$$
$$\longleftarrow n \longrightarrow$$

7.12 Allgemeiner Fall: Sei $f : V \longrightarrow W$ wie in (7.1), $A = M_{\mathscr{B}}^{\mathscr{B}'}(f) = (a_{ij})$. Sei $x \in V$ beliebig und $(x_1, \ldots, x_n)$ der **KOORDINATENVEKTOR** von x bzgl. $\mathscr{B}$, d.h.

$$x = x_1 \cdot v^1 + x_2 \cdot v^2 + \cdots + x_n \cdot v^n = \sum_{j=1}^{n} x_j \cdot v^j.$$

Dann folgt

$$f(x) = \sum_{j=1}^{n} x_j \cdot f(v^j) = \sum_{j=1}^{n} x_j \cdot \sum_{i=1}^{m} a_{ij} w^i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \cdot x_j \right) \cdot w^i.$$

D.h. der Koordinatenvektor $(y_1, \ldots, y_m)$ von $f(x)$ bzgl. $\mathscr{B}'$ ist gegeben durch

$$y_i = \sum_{j=1}^{n} a_{ij} x_j$$

oder in **Matrixschreibweise**

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

7.13 Konvention: Wegen dieses Sachverhalts schreiben wir Vektoren in $\mathbb{K}^n$ als Spaltenvektoren!

7.14 Spezialfall: $f : \mathbb{K}^n \longrightarrow \mathbb{K}^m$

Ist $f : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ linear, wird (7.12) noch schöner, wenn wir die **STANDARDBASEN** $\mathscr{S}_n$ von K^n und $\mathscr{S}_m$ von K^m nehmen, denn dann stimmen Vektor und Koordinatenvektor überein. Ist A die Matrix von f bzgl. der Standardbasen, dann gilt

$$f(x) = A \cdot x \quad \forall\, x \in \mathbb{K}^n.$$

Damit ist die Abbildung $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m,\ f_A(x) = A \cdot x$ linear und $A = M^{\mathscr{S}_m}_{\mathscr{S}_n}(f_A)$.

7.15 Die Abbildung

$$\begin{aligned} \mathbb{K}^{m,n} &\longrightarrow \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m) \\ A &\longmapsto f_A \end{aligned}$$

ist nach (6.14) bijektiv, genauer ist sie die Umkehrabbildung von $\varphi^{\mathscr{S}_m}_{\mathscr{S}_n}$. Nach (7.8) geht dabei die Addition, Skalarmultiplikation und Matrizenmultiplikation in die Addition, Skalarmultiplikation und Komposition von linearen Abbildungen über. Damit gelten für $\mathbb{K}^{m,n}$ die Rechenregeln für lineare Abbildungen mit der Komposition durch die Multiplikation ersetzt. Z.B. zeigt man die Assoziativität wie folgt

$$\begin{aligned} A \cdot (B \cdot C) &= M^{\mathscr{S}_k}_{\mathscr{S}_l}(f_A) \cdot (M^{\mathscr{S}_l}_{\mathscr{S}_m}(f_B) \cdot M^{\mathscr{S}_m}_{\mathscr{S}_n}(f_C)) \\ &= M^{\mathscr{S}_k}_{\mathscr{S}_l}(f_A) \cdot M^{\mathscr{S}_l}_{\mathscr{S}_n}(f_B \circ f_C) \\ &= M^{\mathscr{S}_k}_{\mathscr{S}_n}(f_A \circ (f_B \circ f_C)) \\ &= M^{\mathscr{S}_k}_{\mathscr{S}_n}((f_A \circ f_B) \circ f_C) \\ &= M^{\mathscr{S}_k}_{\mathscr{S}_m}(f_A \circ f_B) \cdot M^{\mathscr{S}_m}_{\mathscr{S}_n}(f_C) \\ &= (M^{\mathscr{S}_k}_{\mathscr{S}_l}(f_A) \cdot M^{\mathscr{S}_l}_{\mathscr{S}_m}(f_B)) \cdot M^{\mathscr{S}_m}_{\mathscr{S}_n}(f_C) \\ &= (A \cdot B) \cdot C \end{aligned}$$

Natürlich kann man das auch direkt aus der Definition der Matrizenmultiplikation herleiten.

Es folgt nach (6.25)

7.16 Satz: $\mathbb{K}^{n,n}$ ist eine $\mathbb{K}$-Algebra.

7.17 WARNUNG: Die Matrizenmultiplikation ist i.a. nicht kommutativ: Für $A \in \mathbb{K}^{\ell,m}, B \in \mathbb{K}^{m,n}$ ist $A \cdot B \in \mathbb{K}^{\ell,n}$ definiert, aber $B \cdot A$ ist nur dann definiert, wenn $\ell = n$. In diesem Fall gilt

$$A \cdot B \in \mathbb{K}^{n,n} \qquad B \cdot A \in \mathbb{K}^{m,m}.$$

Also gilt $A \cdot B = B \cdot A$ höchstens dann, wenn $\ell = m = n$. Aber selbst in diesem Fall ist die Multiplikation i.a. nicht kommutativ:

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

Als Leser-Service seien die Rechenregeln für Matrizen nochmals aufgelistet.

7.18 Rechenregeln für Matrizen:

(1) $(A+B)+C = A+(B+C) \quad \forall\, A,B,C \in \mathbb{K}^{m,n}$

(2) $A+B = B+A \quad \forall\, A,B \in \mathbb{K}^{m,n}$

(3) Ist $0_{m,n} \in \mathbb{K}^{m,n}$ die Nullmatrix, so gilt für $A \in \mathbb{K}^{m,n}$

$$A + 0_{m,n} = 0_{m,n} + A = A$$

(4) $A+(-A) = 0_{m,n} \quad \forall A \in \mathbb{K}^{m,n}$

(5) $(a \cdot b) \cdot A = a \cdot (b \cdot A) \quad \forall\, a,b \in \mathbb{K}, \quad \forall\, A \in \mathbb{K}^{m,n}$

(6) $a \cdot (A+B) = a \cdot A + a \cdot B \quad \forall\, a \in \mathbb{K}, \quad \forall\, A,B \in \mathbb{K}^{m,n}$

(7) $(a+b) \cdot A = a \cdot A + b \cdot A \quad \forall\, a,b \in \mathbb{K}, \quad \forall\, A \in \mathbb{K}^{m,n}$

(8) $1 \cdot A = A \quad \forall\, A \in \mathbb{K}^{m,n}$

(9) $A \cdot (B \cdot C) = (A \cdot B) \cdot C \quad \forall\, A \in \mathbb{K}^{k,\ell}, \forall\, B \in \mathbb{K}^{\ell,m}, \forall\, C = \mathbb{K}^{m,n}$

(10) $E_m \cdot A = A \cdot E_n = A \quad \forall\, A \in \mathbb{K}^{m,n}$

(11) $\forall\, A \in \mathbb{K}^{k,\ell}, \forall\, B_1, B_2 \in \mathbb{K}^{\ell,m}, \forall\, C = \mathbb{K}^{m,n}$ gilt

$$A \cdot (B_1 + B_2) = A \cdot B_1 + A \cdot B_2 \quad \text{und} \quad (B_1 + B_2) \cdot C = B_1 \cdot C + B_2 \cdot C$$

(12) $(a \cdot A) \cdot B = A \cdot (a \cdot B) = a \cdot (A \cdot B) \quad \forall\, a \in \mathbb{K}, \forall\, A \in \mathbb{K}^{\ell,m}, \quad \forall\, B \in \mathbb{K}^{m,n}$

Der Vollständigkeit halber formulieren wir noch (6.28) in Matrizensprache.

7.19 Definition: Sei $A \in \mathbb{K}^{m,n}$. Ein **LINKSINVERSES (RECHTSINVERSES)** von A ist eine Matrix $B \in \mathbb{K}^{n,m}$ (bzw. $C \in \mathbb{K}^{n,m}$), so dass $B \cdot A = E_n$ (bzw. $A \cdot C = E_m$).

7.20 Satz: Besitzt $A \in \mathbb{K}^{m,n}$ ein linksinverses B und ein rechtsinverses C, dann gilt $m = n$ und $B = C$. Diese eindeutig gegebene Matrix wird mit A^{-1} bezeichnet. Es gilt also

$$A \cdot A^{-1} = A^{-1} \cdot A = E_n. \qquad \square$$

Wir übertragen nun den Rangbegriff 6.13 auf Matrizen:

7.21 Definition: Sei $A \in \mathbb{K}^{m,n}$. Wir definieren den RANG $\mathrm{rg} A$ durch

$$\mathrm{rg} A = \dim \mathrm{Bild}\, f_A$$

mit f_A aus (7.15).

Da der j-te Spaltenvektor von A gerade $f_A(e^j)$ ist, folgt

7.22 $\mathrm{rg}(A) =$ Maximalanzahl linear unabhängiger Spalten in A.

Sei nun $f : V \longrightarrow W$ lineare Abbildung, $\mathscr{B} = \{v^1, \ldots, v^n\}$ Basis von V und $\mathscr{B}' = \{w^1, \ldots, w^m\}$ Basis von W. Sei $A = M_{\mathscr{B}}^{\mathscr{B}'}(f)$. Der Übergang zum Koordinatenvektor bzgl. $\mathscr{B}$ definiert einen linearen Isomorphismus

$$\varphi_{\mathscr{B}} : V \longrightarrow \mathbb{K}^n.$$

Analog haben wir den linearen Isomorphismus $\varphi_{\mathscr{B}'} : W \longrightarrow \mathbb{K}^m$. Aus (7.12) folgt, dass

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\varphi_{\mathscr{B}}"', "\cong"] & W \arrow[d, "\cong"', "\varphi_{\mathscr{B}'}"] \\ \mathbb{K}^n \arrow[r, "f_A"] & \mathbb{K}^m \end{tikzcd} \qquad \varphi_{\mathscr{B}'} \circ f = f_A \circ \varphi_{\mathscr{B}}.$$

Damit ist $\mathrm{Bild}\, f \cong \mathrm{Bild}\, f_A$ vermöge des Isomorphismus $\varphi_{\mathscr{B}'}$. Wir erhalten also

7.23 Satz: Für jede Wahl von Basen $\mathscr{B}$ von V und $\mathscr{B}'$ von W gilt

$$\mathrm{rg}\, f = \mathrm{rg}\, M_{\mathscr{B}}^{\mathscr{B}'}(f). \qquad \square$$

8 Grundlagen: *Äquivalenzrelationen*

Für unsere weiteren Untersuchungen brauchen wir weitere Grundlagen.

8.1 Definition: Sei X eine Menge. Eine **RELATION** auf X ist eine Teilmenge R von $X \times X$. Wir schreiben

$$xRy \overset{def}{\Longleftrightarrow} (x,y) \in R$$

8.2 Beispiel: (1) Gleichheitsrelation „=“ auf X ist $R = \{(x,x);\, x \in X\}$. Es gilt also: $xRy \Longleftrightarrow x = y$.

(2) Kleiner-Relation auf $\mathbb{Z}$: $R = \{(x,y) \in \mathbb{Z} \times \mathbb{Z}; x < y\}$.

Für uns besonders wichtig sind Äquivalenzrelationen.

8.3 Definition: Eine Relation $\sim$ ($x \sim y$ ist Kurzform für xRy) auf einer Menge X heißt **ÄQUIVALENZRELATION**, wenn gilt:

(1) Reflexivität: $x \sim x \quad \forall\, x \in X$

(2) Symmetrie: $x \sim y \Longrightarrow y \sim x \quad \forall\, x,y \in X$

(3) Transitivität: $x \sim y$ **und** $y \sim z \Longrightarrow x \sim z \quad \forall\, x,y,z \in X$.

Die Menge

$$[x] := \{y \in X; x \sim y\} \subset X$$

heißt **ÄQUIVALENZKLASSE** von x, ein $y \in [x]$ **REPRÄSENTANT** von $[x]$.

8.4 Beispiel: Die Gleichheitsrelation ist eine Äquivalenzrelation, die Kleiner-Relation dagegen nicht.

8.5 Satz: Für eine Äquivalenzrelation $\sim$ auf der Menge X gilt:

(0) $x \in [x]\ \forall x \in X$

(1) $\forall\, x,y \in X$ gilt $[x] = [y]$ oder $[x] \cap [y] = \emptyset$

(2) $\forall\, x,y \in X$ gilt $[x] = [y] \Longleftrightarrow x \sim y$

(3) X zerfällt in disjunkte Äquivalenzklassen.

Beweis: (1) Angenommen, $\exists z \in [x] \cap [y]$. Dann gilt $x \sim z$ und $y \sim z$. Es folgt $x \sim z, z \sim y$, also $x \sim y$. Ist nun $w \in [y]$, so gilt $y \sim w$. Es folgt $x \sim w$, so dass $w \in [x]$. Also gilt $[y] \subset [x]$. Analog zeigt man $[y] \supset [x]$.

(2) „$\Longrightarrow$" haben wir gerade gezeigt.
„$\Longleftarrow$" $x \sim y$, also $y \in [x]$. Da ausserdem $y \in [y]$, folgt $y \in [x] \cap [y]$, also $[x] = [y]$ nach (1).

(3) folgt aus (1). $\square$

9 Quotienten- und Dualräume

Sei $\mathbb{K}$ ein Körper und V ein $\mathbb{K}$-Vektorraum. Folgende Schreibweise haben wir schon früher benutzt.

9.1 Definition: Seien A, B Teilmengen von V und sei $r \in \mathbb{K}$. Wir definieren

$$\begin{aligned} A + B &= \{a + b; a \in A, b \in \mathscr{B}\} \\ r \cdot A &= \{r \cdot a; a \in A\} \\ -A &= (-1) \cdot A \\ A - B &= \{a - b; a \in A, b \in B\} = A + (-B). \end{aligned}$$

9.2 Ist $U \subset V$ ein Untervektorraum, dann gilt

$$U + U = U, \quad r \cdot U \subset U, ist r \neq 0 gilt sogar r \cdot U = U, \quad -U = U.$$

Beweis: Für $u_1, u_2 \in U$ gilt $u_1 + u_2 \in U$. Es folgt $U + U \subset U$. Da $0 \in U$, gilt für $u \in U; u = 0 + u \in U + U$, also $U \subset U + U$. Mit $u \in U$ ist auch $r \cdot u \in U$, also $r \cdot U \subset U$.

Für $u \in U$ ist $-u \in U$, also $-U \subset U$. Aber $u = -(-u) \in -U$, also $U \subset -U$. $\square$

9.3 Definition: Sei U Untervektorraum von V. Wir definieren eine Relation $\sim_U$ auf V durch $v_1 \sim_U v_2 \iff v_1 - v_2 \in U$.

9.4 $\sim_U$ ist eine Äquivalenzrelation auf V. Die Äquivalenzklasse $[v]$ von $v \in V$ ist die Menge

$$[v] = v + U$$

(wir schreiben kürzer $v + U$ für $\{v\} + U$).

Beweis: Reflexivität: Da $0 \in U$ ist $v - v \in U$, also $v \sim_U v$.

Symmetrie: $v_1 \sim_U v_2 \Longrightarrow v_1 - v_2 =: u \in U \Longrightarrow v_2 - v_1 = -u \in U \Longrightarrow v_2 \sim_U v_1$.

Transitivität: $v_1 \sim_U v_2,\ v_2 \sim_U v_3 \Longrightarrow v_1 - v_2 =: u_1 \in U,\ v_2 - v_3 =: u_2 \in U \Longrightarrow v_1 - v_3 = v_1 - v_2 + v_2 - v_3 = u_1 + u_2 \in U \Longrightarrow v_1 \sim_U v_3$

$v_1 \in [v] \iff v \sim_U v_1 \iff v_1 \sim_U v \iff v_1 - v =: u \in U \iff \exists u \in U$ mit $v_1 = v + u \iff v_1 \in v + U$. $\square$

9.5 Satz und Definition: Sei U Untervektorraum von V und V/U die Menge der Äquivalenzklassen der Relation $\sim_U$. Dann wird durch die Verknüpfungen

$$\begin{aligned} [v] + [w] &:= [v + w] \\ r \cdot [v] &:= [r \cdot v] \qquad r \in \mathbb{K} \end{aligned}$$

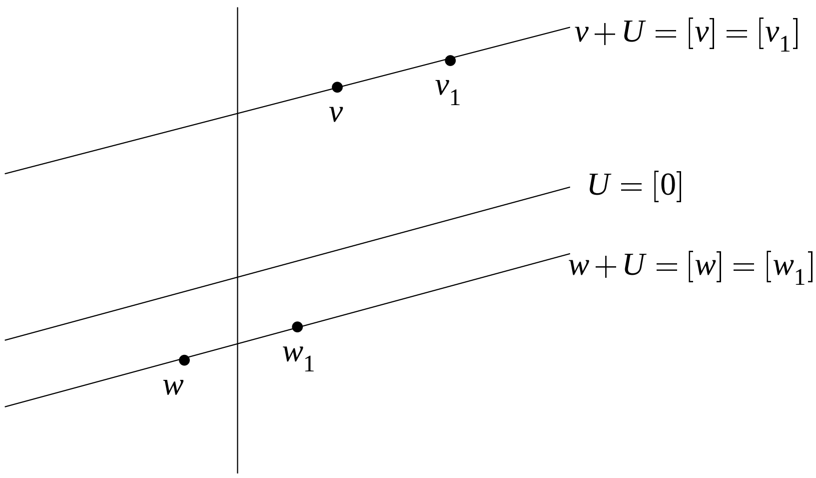
eine $\mathbb{K}$-Vektorraumstruktur auf V/U definiert. Wir nennen V/U mit dieser Struktur den **QUOTIENTENRAUM** von V modulo U. Die Abbildung

$$p : V \longrightarrow V/U, \qquad v \longmapsto [v]$$

ist linear und wird **KANONISCHE PROJEKTION** genannt.

Beweis: Wir müssen zunächst nachkontrollieren, dass unsere Verknüpfungen überhaupt vernünftig definiert sind, der Mathematiker nennt das **WOHLDEFINIERT**: Man muss beachten, dass $[v]$ und $[w]$ **Teilmengen** von V sind und v, w einzelne Elemente dieser Mengen, die willkürlich herausgegriffen werden.

Beispiel: $V = \mathbb{R}^2$



Wir ordnen $[v]$ und $[w]$ die Äquivalenzklasse von $v+w$ zu. Das wäre wenig sinnvoll, wenn bei der Addition anderer Repräsentanten $v_1 \in [v]$ und $w_1 \in [w]$ die Summe $v_1 + w_1$ in einer anderen Äquivalenzklasse als $v+w$ liegt. Wir müssen also zeigen:

$$\begin{aligned} v \sim_U v_1,\ w \sim_U w_1 \quad &\Longrightarrow \quad v+w \sim_U v_1 + w_1 \\ v \sim_U v_1,\ r \in \mathbb{K} \quad &\Longrightarrow \quad r \cdot v \sim_U r \cdot v_1. \end{aligned}$$

Nun gilt: $v - v_1 \in U$ und $w - w_1 \in U$; es folgt

$$\begin{aligned} (v+w) - (v_1 + w_1) = v + w - v_1 - w_1 &= (v - v_1) + (w - w_1) \in U \\ r \cdot v - r \cdot v_1 &= r \cdot (v - v_1) \in U. \end{aligned}$$

Also $v + w \sim_U v_1 + w_1$ und $r \cdot v \sim_U r \cdot v_1$.

Die Vektorraumaxiome folgen nun leicht. Wir beweisen exemplarisch die Axiome (4.13 (3)) und (4.13 (4)). Der Nullvektor in V/W ist die Klasse $[0]$, denn

$$[0] + [v] = [0 + v] = [v],$$

und $-[x]$ ist die Klasse $[-x]$, denn

$$[x]+[-x]=[x-x]=[0]=0.$$

Die Linearität von p prüft man ebenso leicht:

$$\begin{aligned} p(v+w) &= [v+w] = [v]+[w] = p(v)+p(w) \qquad v,w \in V \\ p(r\cdot v) &= [r\cdot v] = r\cdot [v] = r\cdot p(v) \qquad r\in \mathbb{K},\ v\in V \end{aligned}$$

$\square$

9.6 Ergänzung: $\operatorname{Kern} p = U$.

Beweis: $v \in \operatorname{Kern} p \Longleftrightarrow p(v)=[0] \Longleftrightarrow [v]=[0] \Longleftrightarrow v\in 0+U=U$ $\square$

Die folgenden beiden Ergebnisse machen die Bedeutung der Quotientenraumkonstruktion deutlich.

9.7 Satz: Sei $f: V \longrightarrow W$ eine lineare Abbildung und $U\subset V$ ein Untervektorraum, so dass $U\subseteq \operatorname{Kern} f$. Dann gibt es genau eine lineare Abbildung $\overline{f}: V/U \longrightarrow W$, so dass $f=\overline{f}\circ p$

$$\begin{tikzcd} V \arrow[rr, "f"] \arrow[dr, "p"'] & & W \\ & V/U \arrow[ur, "\overline{f}"'] & \end{tikzcd}$$

Beweis: Wir gehen wie beim Beweis von (6.14) vor. Der Eindeutigkeitsbeweis zeigt uns, wie wir $\overline{f}$ definieren müssen.
Eindeutigkeit: Es **muss** gelten

$$f(v)=\overline{f}\circ p(v)=\overline{f}([v]) \quad \forall\, v\in V.$$

Existenz: Es bleibt uns also nichts übrig, als zu definieren

$$\overline{f}: V/U \longrightarrow W, \quad \overline{f}([v]) = f(v) \qquad (*)$$

$\overline{f}$ ist wohldefiniert, denn für $v_1\in[v]$ gilt $v_1\in v+U$. Also

$$f(v_1)\in f(v+U)=f(v)+f(U)=f(v)+0=f(v).$$

Damit ist die Zuordnung $(*)$ unabhängig von der Wahl von $v\in[v]=v+U$.

$\overline{f}$ ist linear, denn

$$\begin{aligned}\overline{f}([v]+[w]) &= \overline{f}([v+w]) = f(v+w) = f(v)+f(w)\\ &= \overline{f}([v])+\overline{f}([w])\\ \overline{f}(r\cdot[v]) &= \overline{f}([rv]) = f(r\cdot v) = r\cdot f(v) = r\cdot\overline{f}([v]).\end{aligned}$$

□

Für den Spezialfall $U = \operatorname{Kern} f$ erhalten wir

9.8 Satz: Ist $f : V \longrightarrow W$ linear, dann ist die „**INDUZIERTE ABBILDUNG**" $\overline{f} : V/\operatorname{Kern} f \longrightarrow W$ aus (9.7) injektiv. Insbesondere definiert $\overline{f}$ einen Isomorphismus

$$V/\operatorname{Kern} f \cong \operatorname{Bild} f.$$

Beweis: Sei $\overline{f}([v]) = 0$, also $f(v) = 0$. Dann ist $v \in \operatorname{Kern} f = 0 + \operatorname{Kern} f = [0]$, d.h. $v \in [0]$, also $[v] = [0] = 0$. Damit ist $\operatorname{Kern}\overline{f} = \{[0]\}$ und $\overline{f}$ nach (6.10) injektiv. Da f als Abbildung

$$V \longrightarrow \operatorname{Bild} f$$

surjektiv ist, und $f = \overline{f} \circ p$, ist $\overline{f}$ als Abbildung nach $\operatorname{Bild} f$ surjektiv (3.24) (2), also ist $\overline{f}$ als Abbildung nach $\operatorname{Bild} f$ bijektiv. □

Mit (6.19) erhalten wir

9.9 Folgerung: Ist $f : V \longrightarrow W$ linear und V endlich dimensional, dann gilt

$$V \cong \operatorname{Kern} f \oplus V/\operatorname{Kern} f.$$

Insbesondere gilt $\dim V = \dim \operatorname{Kern} f + \dim V/\operatorname{Kern} f$. □

Die zweite wichtige Konstruktion dieses Abschnitts ist der Dualraum.

9.10 Definition: Sei V ein $\mathbb{K}$-Vektorraum. Der $\mathbb{K}$-Vektorraum

$$V^* = \operatorname{Hom}_{\mathbb{K}}(V, \mathbb{K})$$

heißt **DUALRAUM** von V. Die Elemente $\varphi \in V^*$ heißen **LINEARFORMEN** auf V.

Ist $f : V \longrightarrow W$ eine lineare Abbildung. Dann heißt die (nach 6.22 (3)) lineare Abbildung

$$f^* : W^* \longrightarrow V^*, \qquad \varphi \longmapsto \varphi \circ f$$

die **DUALE ABBILDUNG** von f.

9.11 Beispiel: Sei $\mathscr{P}$ der $\mathbb{R}$-Vektorraum aller Polynomfunktionen $q : \mathbb{R} \longrightarrow \mathbb{R}$ (wir überlassen den Nachweis der Vektorraumaxiome dem Leser).
Wir definieren $\varphi_1, \varphi_2, \varphi_3 \in \mathscr{P}^*$ durch

$$\varphi_1(q) = \int_0^1 q dx, \quad \varphi_2(q) = q(0), \quad \varphi_3(q) = q'(0)$$

wobei q' die Ableitung von q ist. Alle drei sind Linearformen, denn

$$\int_0^1 (q_1 + q_2) dx = \int_0^1 q_1 dx + \int_0^1 q_2 dx \text{ und } \int_0^1 rq dx = r \int_0^1 q dx, \; r \in \mathbb{K}$$
$$(q_1 + q_2)(0) = q_1(0) + q_2(0) \text{ und } (rq)(0) = r \cdot q(0) \text{ nach (4.15)}$$
$$(q_1 + q_2)' = q_1' + q_2' \text{ und } (rq)' = r \cdot q'$$

Behauptung: $\varphi_1, \varphi_2, \varphi_3$ sind linear unabhängig in $\mathscr{P}^*$.
Beweis: Seien $a, b, c \in \mathbb{R}$, so dass

$$a \cdot \varphi_1 + b\varphi_2 + c \cdot \varphi_3 = 0$$

d.h. $a \cdot \varphi_1(q) + b \cdot \varphi_2(q) + c \cdot \varphi_3(q) = 0 \quad \forall\, q \in \mathscr{P}$.
Für $q = x^2 - \frac{1}{3}$ erhalten wir

$$\varphi_1(q) = \left[\frac{x^3}{3} - \frac{1}{3}x\right]_0^1 = 0, \quad \varphi_2(q) = -\frac{1}{3}, \quad \varphi_3(q) = 0.$$

Es folgt $b = 0$. Für $q = 1$ erhalten wir

$$\varphi_1(q) = 1, \quad \varphi_3(q) = 0.$$

Es folgt $a = 0$, da ja $b = 0$. Für $q = x$ ist $\varphi_3(q) = 1$, also auch $c = 0$. □

9.12 Funktorielle Eigenschaften:

(1) $(id_V)^* = id_{V^*}$

(2) Sind $f : U \longrightarrow V$ und $g : V \longrightarrow W$ $\mathbb{K}$-linear, so gilt

$$(g \circ f)^* = f^* \circ g^* : W^* \longrightarrow U^*.$$

Beweis: (1) ist klar. Für (2) sei $\varphi \in W^* = \mathrm{Hom}_{\mathbb{K}}(W, \mathbb{K})$. Dann gilt

$$(g \circ f)^*(\varphi) = \varphi \circ (g \circ f) = (\varphi \circ g) \circ f = f^*(g^*(\varphi)) = (f^* \circ g^*)(\varphi).$$

$\square$

Sei nun V ein $\mathbb{K}$-Vektorraum mit Basis $\mathscr{B} = \{v^1, \dots, v^n\}$. Nehmen wir $1 \in \mathbb{K}$ als Basis von $\mathbb{K}$, so ist $\varphi \in V^*$ eindeutig durch den Zeilenvektor

$$M_{\mathscr{B}}^{\{1\}}(\varphi) = (y_1, \dots, y_n) \text{ mit } y_i = \varphi(v^i)$$

gegeben. Wir erhalten damit eine Zuordnung

$$V^* \longrightarrow \mathbb{K}^{1,n} = \mathbb{K}^n$$

die nach (6.14) und § 7 ein linearer Isomorphismus ist. Insbesondere ist $\dim V^* = n = \dim V$. Die Abbildung bildet die Linearform

$$\varphi^i : V \longrightarrow \mathbb{K},$$

die durch

9.13 $$\varphi^i(v^j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

eindeutig gegeben ist, auf den i-ten Einheitsvektor e^i ab. Wir erhalten

9.14 Satz und Definition: Die in (9.13) definierten Linearformen $\varphi^1, \dots, \varphi^n$ bilden eine Basis $\mathscr{B}^*$ von V^*, genannt die zu $\mathscr{B}$ **DUALE BASIS**.

9.15 Bezeichnung: Das Symbol $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ nennt man **KRONECKER**-Symbol.

9.16 Bemerkung: In der Physik schreibt man oft $\langle \varphi | v \rangle$ für $\varphi(v)$. Die **AUSWERTUNGSABBILDUNG**

$$V^* \times V \longrightarrow \mathbb{K}, \quad (\varphi, v) \longmapsto \varphi(v) = \langle \varphi | v \rangle$$

nennt man auch **DUALE PAARUNG**.
Ist $\mathscr{B} = \{v^1, \dots, v^n\}$ Basis von V, dann heißt

$$\langle \varphi | v^i \rangle = \varphi(v^i) \in \mathbb{K}$$

i-ter **FOURIER**-Koeffizient von φ bzgl. $\mathscr{B}$. Es gilt

$$\varphi = \langle \varphi | v^1 \rangle \cdot \varphi^1 + \dots + \langle \varphi | v^n \rangle \cdot \varphi^n = \sum_{i=1}^{n} \langle \varphi | v^i \rangle \varphi^i$$

wobei $\{\varphi^1, \dots, \varphi^n\}$ die duale Basis $\mathscr{B}^*$ ist. Nach (6.14) genügt es, diese Gleichung für die Basis $\mathscr{B}$ zu testen:

$$\begin{aligned}(\langle\varphi|v^1\rangle \cdot \varphi^1 + \dots + \langle\varphi|v^n\rangle\varphi^n)(v^j) &= \langle\varphi|v^1\rangle \cdot \varphi^1(v^j) + \dots + \langle\varphi|v^n\rangle\varphi^n(v^j)\\ &= \sum_{i=1}^{n}\langle\varphi|v^i\rangle\delta_{ij} = \langle\varphi|v^j\rangle = \varphi(v^j)\end{aligned}$$

Sei nun $f: V \longrightarrow W$ $\mathbb{K}$-lineare Abbildung, $\mathscr{B} = \{v^1, \dots, v^n\}$ Basis von $\mathscr{B}$, $\mathscr{B}' = \{w^1, \dots, w^m\}$ Basis von W, $\mathscr{B}^* = \{\varphi^1, \dots, \varphi^n\}$ und $\mathscr{B}'^* = \{\psi^1, \dots, \psi^m\}$ die dazu dualen Basen. Sei $A = M_{\mathscr{B}}^{\mathscr{B}'}(f) = (a_{ij}) \in \mathbb{K}^{m,n}$. Wir wollen

$$B = M_{\mathscr{B}'^*}^{\mathscr{B}^*}(f^*) = (b_{ij}) \in \mathbb{K}^{n,m}$$

bestimmen. Nach Definition gilt

$$f^*(\psi^i) = \sum_{j=1}^{n} b_{ji}\varphi^j \quad \text{und} \quad f(v^j) = \sum_{i=1}^{m} a_{ij}w^i$$

$f^*(\psi^i) = \psi^i \circ f = \sum_{j=1}^{n} \langle\psi^i \circ f|v^j\rangle \cdot \varphi^j$ nach (9.16).

Es folgt

$$\begin{aligned}b_{ji} &= (\psi^i \circ f)(v^j) = \psi^i\left(\sum_{k=1}^{m} a_{kj}w^k\right) = \sum_{k=1}^{n} a_{kj} \cdot \psi^i(w^k) = \sum_{k=1}^{m} a_{kj}\delta_{ik}\\ &= a_{ij}.\end{aligned}$$

9.17 Definition: Sei $A = (a_{ij}) \in \mathbb{K}^{m,n}$ ein Matrix. Dann heißt die Matrix $B = (b_{ji}) \in \mathbb{K}^{n,m}$ mit $b_{ji} = a_{ij}$ die zu A **TRANSPONIERTE** Matrix und wird mit A^t bezeichnet.

Wir haben damit gezeigt

9.18 Satz: Sei $f: V \longrightarrow W$ $\mathbb{K}$-linear, $\mathscr{B}$ Basis von V, $\mathscr{B}'$ Basis von W und $\mathscr{B}^*, \mathscr{B}'^*$ die dazu dualen Basen. Dann gilt für $f^*: W^* \longrightarrow V^*$

$$M_{\mathscr{B}'^*}^{\mathscr{B}^*}(f^*) = (M_{\mathscr{B}}^{\mathscr{B}'}(f))^t. \qquad \Box$$

9.19 Beispiel: $A = \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 3 & 1 \end{pmatrix} \qquad A^t = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 0 & 1 \end{pmatrix}$

9.20 Lemma: Sei $A \in \mathbb{K}^{m,n}$ und $B \in \mathbb{K}^{n,\ell}$. Dann gilt

$$(A \cdot B)^t = B^t \cdot A^t.$$

Beweis: Sei $A = (a_{ij}),\ B = (b_{jk})$ und $A \cdot B = (c_{ik})$, also

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

Sei $A^t = (a^t_{ji}),\ B^t = (b^t_{kj})$ und $(A \cdot B)^t = (c^t_{ki})$, also $a^t_{ji} = a_{ij}, b^t_{kj} = b_{jk}$ und $c^t_{ki} = c_{ik}$. Wir erhalten

$$c^t_{ki} = c_{ik} = \sum_{j=1}^{n} a_{ij} \cdot b_{jk} = \sum_{j=1}^{n} b^t_{kj} \cdot a^t_{ji},$$

was zu zeigen war. □

9.21 Aufgabe: Beweisen Sie (9.20) mit Hilfe von (9.12 (2)) und (9.18).

Wenden wir uns wieder linearen Abbildungen zu.

9.22 Satz: Sei $f : V \longrightarrow W$ lineare Abbildung endlich dimensionaler $\mathbb{K}$-Vektorräume. Dann gilt

(1) f injektiv $\Longrightarrow f^*$ surjektiv

(2) f surjektiv $\Longrightarrow f^*$ injektiv

Der Satz folgt direkt aus (9.12) und folgendem Ergebnis

9.23 Satz: Sei $f : V \longrightarrow W$ lineare Abbildung endlich dimensionaler $\mathbb{K}$-Vektorräume. Dann gilt

(1) f injektiv $\Longleftrightarrow$ es gibt eine lineare Abbildung $g : W \to V$, so dass $g \circ f = id_V$.

(2) f surjektiv $\Longleftrightarrow$ es gibt eine lineare Abbildung $g : W \to V$, so dass $f \circ g = id_W$.

Beweis: Die Rückrichtungen folgen aus (3.19). Wir beweisen die Hinrichtungen.
(1) Sei $\{v^1, \dots, v^n\}$ Basis von V. Da f injektiv ist, sind $f(v^1), \dots, f(v^n)$ linear unabhängig. Wir ergänzen diese Vektoren zu einer Basis

$$\mathscr{B} = \{f(v^1), \dots, f(v^n), z^{n+1}, \dots, z^m\}$$

von W und definieren g durch $g(f(v^i)) = v^i$ und $g(z^j) = 0$. Dann gilt $g \circ f = id_V$, weil $g \circ f$ die Basis von V identisch abbildet.

(2) Sei $\{w^1, \ldots, w^m\}$ Basis von W. Da f surjektiv ist, gibt es zu jedem w^j ein $u^j \in V$, so dass $f(u^j) = w^j$ ist. Wir definieren nun g durch $g(w^j) = u^j$. Dann gilt $f(g(w^j)) = f(u^j) = w^j$, so dass $f \circ g$ die Basis identisch abbildet. Es folgt $f \circ g = id_W$. □

9.24 Folgerung: Sei $f : V \longrightarrow W$ lineare Abbildung endlichdimensionaler $\mathbb{K}$-Vektorräume. Dann gilt

$$\operatorname{rg} f = \operatorname{rg} f^*.$$

Beweis: Sei $\overline{f} : V \longrightarrow \operatorname{Bild} f$ die lineare Abbildung $v \longmapsto f(v)$ und $g : \operatorname{Bild} f \longrightarrow W$ die **Einbettung** $w \longmapsto w$, d.h. es gilt $f = g \circ \overline{f}$, also $f^* = \overline{f}^* \circ g^*$. Da $\overline{f}$ surjektiv und g injektiv ist, ist g^* surjektiv und $\overline{f}^*$ injektiv. Es folgt

$$\operatorname{rg} f = \dim \operatorname{Bild} f = \dim(\operatorname{Bild} f)^* = \dim \operatorname{Bild} \overline{f}^* = \dim \operatorname{Bild}(f^*) = \operatorname{rg} f^*.$$

□

9.25 Folgerung: Sei $A \in \mathbb{K}^{m,n}$ eine Matrix. Dann gilt

$$\begin{aligned} \operatorname{rg} A &= \text{Maximalanzahl linear unabhängiger Spalten} \\ &= \text{Maximalanzahl linear unabhängiger Zeilen} \\ &= \text{Anzahl der von Null verschiedenen Zeilen in der zugehörigen oberen} \\ &\quad\ \text{Dreiecksform (die Zahl } r \text{ in (1.8))} \end{aligned}$$

Beweis: Die erste Gleichung ist (7.22). Sei $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ die lineare Abbildung $x \longmapsto A \cdot x$. Nach (9.18) ist $f_{A^t} : \mathbb{K}^m \longmapsto \mathbb{K}^n$ die zu f_A duale Abbildung f_A^*, so dass $\operatorname{rg} A^t = \operatorname{rg} A$ nach (9.24). Damit folgt die zweite Gleichung. Aus (5.18) folgt dann die dritte Gleichung. □

9.26 Aufgabe: Sei $f : V \longrightarrow W$ eine lineare Abbildung endlich dimensionaler $\mathbb{K}$-Vektorräume mit Matrix $A \in \mathbb{K}^{m,n}$ bzgl. beliebig vorgegebener Basen $\mathscr{B}$ von V und $\mathscr{B}'$ von W. Zeigen Sie

(1) f injektiv $\Longleftrightarrow \operatorname{rg} A = n \leq m$

(2) f surjektiv $\Longleftrightarrow \operatorname{rg} A = m \leq n$

(3) f bijektiv $\Longleftrightarrow \operatorname{rg} A = m = n \Longleftrightarrow A$ ist invertierbar.

Wir behandeln zum Abschluss einen für Mathematik und Physik gleichermaßen wichtigen Begriff.

9.27 Definition: Sei V ein $\mathbb{K}$-Vektorraum. Dann heißt

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

BIDUALRAUM von V.

9.28 Satz: Die Abbildung $\tau_V : V \longrightarrow V^{**},\ v \longmapsto \hat{v}$ mit

$$\hat{v} : V^* \longrightarrow \mathbb{K}, \quad \varphi \longmapsto \varphi(v)$$

ist $\mathbb{K}$-linear, und für jede $\mathbb{K}$-lineare Abbildung $V \xrightarrow{f} W$ ist

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\tau_V"'] & W \arrow[d, "\tau_W"] \\ V^{**} \arrow[r, "f^{**}"] & W^{**} \end{tikzcd}$$

kommutativ.

Beweis: (1) Wir müssen zeigen, dass $\hat{v}$ eine lineare Abbildung ist:

$$\begin{aligned} \hat{v}(\varphi + \psi) &= (\varphi + \psi)(v) = \varphi(v) + \psi(v) = \hat{v}(\varphi) + \hat{v}(\psi), \quad \varphi, \psi \in V^* \\ \hat{v}(r \cdot \varphi) &= (r \cdot \varphi)(v) = r \cdot \varphi(v) = r \cdot \hat{v}(\varphi), \quad \varphi \in V^*,\ r \in \mathbb{K} \end{aligned}$$

(2) τ_V ist $\mathbb{K}$-linear: Für $\varphi \in V^*,\ v, w \in V$ und $r \in \mathbb{K}$ gilt

$$\begin{aligned} \widehat{v+w}(\varphi) &= \varphi(v+w) = \varphi(v) + \varphi(w) = \hat{v}(\varphi) + \hat{w}(\varphi) = (\hat{v} + \hat{w})(\varphi) \\ \widehat{r \cdot v}(\varphi) &= \varphi(r \cdot v) = r \cdot \varphi(v) = r \cdot \hat{v}(\varphi) = (r \cdot \hat{v})(\varphi). \end{aligned}$$

Also $\tau_V(v+w) = \widehat{v+w} = \hat{v} + \hat{w} = \tau_V(v) + \tau_V(w)$ und analog $\tau_V(r \cdot v) = r \cdot \tau_V(v)$.
(3) Sei nun $f : V \longrightarrow W$ $\mathbb{K}$-linear und $v \in V$. Dann ist

$$(f^{**} \circ \tau_V)(v) = f^{**}(\hat{v}) = \hat{v} \circ f^* : W^* \longrightarrow \mathbb{K}$$

die lineare Abbildung $\varphi \longmapsto \hat{v} \circ f^*(\varphi) = f^*(\varphi)(v) = \varphi(f(v))$, und

$$(\tau_W \circ f)(v) = \tau_W(f(v)) = \widehat{f(v)} : W^* \longrightarrow \mathbb{K}$$

die lineare Abbildung $\varphi \longmapsto \widehat{f(v)}(\varphi) = \varphi(f(v))$. $\square$

9.29 Ergänzung: Ist V endlichdimensional, dann ist τ_V ein Isomorphismus.

Beweis: Nach (9.14) ist $\dim V = \dim V^* = \dim V^{**}$. Daher genügt es nach (9.22) zu zeigen, dass τ_V injektiv ist. Sei $v \neq 0$. Wir ergänzen v zu einer Basis $\mathscr{B} = \{v, v^2, v^3, \ldots, v^n\}$ von V und definieren eine Linearform $\varphi \in V^*$ durch $\varphi(v) = 1$ und $\varphi(v^i) = 0$ für $i \geq 2$. Es folgt

$$\hat{v}(\varphi) = \varphi(v) = 1.$$

Also ist $\hat{v} \neq 0$. Damit ist $\operatorname{Kern} \tau_V = \{0\}$, denn kein $v \neq 0$ liegt im Kern. Also ist τ_V injektiv. $\square$

9.30 Aufgabe: Sei $A \in \mathbb{K}^{m,n}$ vom Rang r. Zeigen Sie: Durch Wegstreichen von geeigneten $m - r$ Zeilen und $n - r$ Spalten kann man eine r-quadratische Untermatrix vom Rang r finden.

10 Ergänzung zu linearen Gleichungssystemen

10.1 Gegeben sei ein lineares Gleichungssystem

$$\begin{array}{ccccc} a_{11}x_1 & + \cdots + & a_{1n}x_n & = & b_1 \\ \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + \cdots + & a_{mn}x_n & = & b_m \end{array}$$

Sei

10.2

$$\begin{array}{cccccccc|c} x_{q_1} & \ldots & \ldots & x_{q_r} & x_{q_{r+1}} & \ldots & \ldots & x_{q_n} & b \\ \hline 1 & a'_{12} & & a'_{1r} & a'_{1r+1} & \ldots & \ldots & a'_{1n} & b'_1 \\ & \ddots & & & \vdots & & & \vdots & \\ & 0 & \ddots & & \vdots & & & \vdots & \\ & & & 1 & a'_{rr+1} & \ldots & \ldots & a'_{rn} & b'_r \\ 0 & \ldots & \ldots & 0 & 0 & \ldots & \ldots & 0 & b'_{r+1} \\ \vdots & & & \vdots & \vdots & & & \vdots & \vdots \\ 0 & \ldots & \ldots & 0 & 0 & \ldots & \ldots & 0 & b'_m \end{array}$$

das zugehörige obere Dreiecksschema und

10.3

$$\begin{array}{cccccccc|c} x_{q_1} & \ldots & \ldots & x_{q_r} & x_{q_{r+1}} & \ldots & \ldots & x_{q_n} & b \\ \hline 1 & & & & a^*_{1r+1} & \ldots & \ldots & a^*_{1n} & b^*_1 \\ & \ddots & 0 & & \vdots & & & \vdots & \\ & 0 & \ddots & & \vdots & & & \vdots & \\ & & & 1 & a^*_{rr+1} & \ldots & \ldots & a^*_{rn} & b^*_r \end{array}$$

das zugehörige reduzierte Schema. Weiter sei $A = (a_{ij})$ die Koeffizientenmatrix, und $x \in \mathbb{K}^n$, $b \in \mathbb{K}^m$ die Spaltenvektoren

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Sei $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ die lineare Abbildung $x \longmapsto A \cdot x$. In der Sprache der linearen Abbildungen suchen wir also $x \in \mathbb{K}^n$, so dass

$$f_A(x) = A \cdot x = b.$$

Die Lösbarkeit von (10.1) ist also gleichbedeutend mit der Bedingung, dass $b \in \operatorname{Bild} f_A$.

10.4 Definition: Sind $A \in \mathbb{K}^{m,n}$ und $B \in \mathbb{K}^{m,\ell}$, so bezeichnet $(A|B) \in \mathbb{K}^{m,n+\ell}$ die Matrix, die aus A durch Anhängen von B entsteht. Die Matrix $(A|b)$ nennt man die **ERWEITERTE MATRIX** des Gleichungssystems.

Aus (9.25) wissen wir, dass r der Rang der Matrix A ist. Die Lösbarkeitsbedingung (10.1) erhält somit die Form

(10.1) ist genau dann lösbar, wenn $\operatorname{rg} A = \operatorname{rg}(A|b)$.

Die Lösungsmenge $\mathbb{L}^0$ des zugehörigen homogenen Systems ist

$$\mathbb{L}^0 = \{x \in \mathbb{K}^n;\ f_A(x) = 0\} = \operatorname{Kern} f_A.$$

Da $\dim(\operatorname{Bild} f_A) = r$ und $\dim \operatorname{Bild} f_A + \dim \operatorname{Kern} f_A = \dim \mathbb{K}^n = n$, ist

$$\dim \mathbb{L}^0 = \dim \operatorname{Kern} f_A = n - r.$$

Wenn (10.1) also eine Lösung besitzt, ist die Lösungsmenge ein $(n-r)$-dimensionaler affiner Unterraum von $\mathbb{K}^n$.

Wir fassen zusammen:

10.5 Satz: (1) Für das Gleichungssystem (10.1) sind äquivalent

(a) $A \cdot x = b$ ist lösbar.

(b) $b \in \operatorname{Bild} f_A = \operatorname{Span}\{s^1, \dots, s^n\},\ s^j = j$-ter Spaltenvektor von A.

(c) $\operatorname{rg} A = \operatorname{rg}(A|b)$

(2) $\operatorname{Kern} f_A$ hat die Dimension $n - \operatorname{rg} A = n - r$ und ist die Lösungsmenge des zugehörigen homogenen Systems.

(3) Ist x^0 Lösung von $A \cdot x = b$, dann ist die Lösungsmenge

$$\mathbb{L} = x^0 + \operatorname{Kern} f_A$$

ein $(n-r)$-dimensionaler affiner Unterraum von $\mathbb{K}^n$.

Wir wenden uns jetzt den bisher etwas stiefmütterlich behandelten affinen Unterräumen zu.

10.6 Definition: Ist $L \subset \mathbb{K}^n$ ein k-dimensionaler affiner Unterraum, dann heißt

$$\operatorname{codim} L = n - k$$

die **KODIMENSION** von L. Ist $\operatorname{codim} L = 1$ (also $\dim L = n-1$), heißt L **HYPEREBENE** in $\mathbb{K}^n$.

10.7 Bemerkung: Die leere Menge $\emptyset$ wird auch als affiner Unterraum von $\mathbb{K}^n$ angesehen, doch legt man sich bei ihrer Dimension nicht fest. Man vereinbart nur $\dim\emptyset < 0$ und $\operatorname{codim}\emptyset > n$.

10.8 Satz: Für eine Teilmenge $L \subset \mathbb{K}^n$ sind äquivalent

(1) L ist affiner Unterraum von $\mathbb{K}^n$ der Dimension k.

(2) $\exists$ Matrix $A \in \mathbb{K}^{n-k,n}$ vom Rang $n-k$ und ein $b \in \mathbb{K}^{n-k}$, so dass L die Lösungsmenge des Gleichungssystems $A \cdot x = b$ ist.

Beweis: (2) $\Longrightarrow$ (1) ist (10.5 (3))

(1) $\Longrightarrow$ (2) Sei $L = x^0 + V$ mit einem k-dimensionalen Untervektorraum, $V \subset \mathbb{K}^n$. Sei $\mathscr{B} = \{v^1, \ldots, v^k\}$ Basis von V. Wir ergänzen $\mathscr{B}$ zu einer Basis $\{v^1, \ldots, v^k, w^1, \ldots, w^{n-k}\}$ von $\mathbb{K}^n$ und definieren eine lineare Abbildung $f : \mathbb{K}^n \to \mathbb{K}^m$ durch $f(v^i) = 0$ für $i = 1.\ldots, k$ und $f(w^j) = e^j$ für $j = 1.\ldots, n-k$. Sei A die Matrix von f bzgl. der Standardbasen, d.h. $f = f_A$. Dann gilt $V = \{x \in \mathbb{K}^n;\ A \cdot x = 0\}$. Setzen wir $b = f(x^0) = A \cdot x^0$, so ist L Lösungsmenge von $A \cdot x = b$. $\square$

10.9 Beispiel: $L = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \operatorname{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} \right\}$.

Also $x^0 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v^1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v^2 = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}$

Dann ist $\mathscr{B} = \{v^1, v^2, v^3\}$ mit $v^3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ Basis von $\mathbb{K}^3$. Das f aus dem Beweis von (10.8) bildet v^1 und v^2 auf $0 \in \mathbb{K}$ und v^3 auf den Einheitsvektor $1 \in \mathbb{K}$ ab. Es folgt $M^{\mathscr{S}_1}_{\mathscr{B}}(f) = (0,0,1)$. Wir bestimmen die gesuchte Matrix $A = M^{\mathscr{S}_1}_{\mathscr{S}_3}(f)$ mit der Formel

$$M^{\mathscr{S}_1}_{\mathscr{B}}(f) \cdot M^{\mathscr{B}}_{\mathscr{S}_3}(id) = M^{\mathscr{S}_1}_{\mathscr{S}_3}(f \circ id) = M^{\mathscr{S}_1}_{\mathscr{S}_3}(f)$$

Die Spaltenvektoren von $M^{\mathscr{B}}_{\mathscr{S}_3}(id)$ erhalten wir, indem wir die Einheitsvektoren als Linearkombinationen von v^1, v^2, v^3 darstellen.

$$\begin{aligned} e^1 &= v^1 + v^2 + v^3 \\ e^2 &= -v^2 - v^3 \\ e^3 &= v^3 \end{aligned}$$

Es folgt

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

Weiter gilt $b = A \cdot x^0 = 2$. Also ist L die Lösungsmenge des Gleichungssystems

$$x_1 - x_2 + x_3 = 2.$$

10.10 Satz: Sind $L_1, L_2 \subset \mathbb{K}^n$ affine Unterräume, dann ist auch $L_1 \cap L_2$ ein affiner Unterraum, und es gilt

$$L_1 \cap L_2 = \emptyset \text{ oder } \operatorname{codim}(L_1 \cap L_2) \leq \operatorname{codim} L_1 + \operatorname{codim} L_2.$$

Beweis: Sei $\dim L_1 = k, \dim L_2 = \ell$. Nach (10.8) gibt es $A \in \mathbb{K}^{n-k,n}, B \in \mathbb{K}^{n-\ell,n}$, $b \in \mathbb{K}^{n-k}$ und $c \in \mathbb{K}^{n-\ell}$, so dass L_1 und L_2 die Lösungsmengen von

$$A \cdot x = b \quad \text{bzw.} \quad B \cdot x = c$$

sind. $L_1 \cap L_2$ ist damit die Lösungsmenge beider Gleichungssysteme, d.h. aller der $x \in \mathbb{K}^n$, für die

$$\begin{pmatrix} A \\ B \end{pmatrix} \cdot x = \begin{pmatrix} b \\ c \end{pmatrix} \qquad (*)$$

Da $\operatorname{rg} \begin{pmatrix} A \\ B \end{pmatrix} \leq \operatorname{rg} A + \operatorname{rg} B$, gilt im Falle $L_1 \cap L_2 \neq \emptyset$

$$\operatorname{codim}(L_1 \cap L_2) = \operatorname{rg} \begin{pmatrix} A \\ B \end{pmatrix} \leq \operatorname{rg} A + \operatorname{rg} B = \operatorname{codim} L_1 + \operatorname{codim} L_2.$$

□

10.11 Definition: Man sagt, zwei affine Unterräume $L_1, L_2 \subset \mathbb{K}^n$ sind in **ALLGEMEINER LAGE**, wenn

$$\operatorname{codim}(L_1 \cap L_2) = \operatorname{codim} L_1 + \operatorname{codim} L_2.$$

10.12 Bemerkung: (1) Ist $\operatorname{codim} L_1 + \operatorname{codim} L_2 > n$, so sind L_1 und L_2 genau dann in allgemeiner Lage, wenn $L_1 \cap L_2 = \emptyset$.

(2) Ist $\operatorname{codim} L_1 + \operatorname{codim} L_2 \leq n$ und sind L_1 und L_2 in allgemeiner Lage, so ist $L_1 \cap L_2 \neq \emptyset$ und $L_1 \cap L_2$ hat die kleinstmögliche Dimension, nämlich

$$\dim(L_1 \cap L_2) = \dim L_1 + \dim L_2 - n.$$

10.13 Beispiel: (1) Zwei Geraden im $\mathbb{R}^3$ sind in allgemeiner Lage, wenn sie sich nicht schneiden. Ihre Kodimension ist 2.

(2) Eine Gerade und eine Ebene im $\mathbb{R}^3$ sind in allgemeiner Lage, wenn sie sich in einem Punkt schneiden.

(3) Zwei Ebenen im $\mathbb{R}^3$ sind in allgemeiner Lage, wenn sie nicht parallel sind, d.h. in einer Geraden schneiden.

An den Beispielen sehen wir: Sind zwei affine Unterräume in allgemeiner Lage und „wackelt" man an ihnen nur wenig, bleiben sie in allgemeiner Lage. Sind sie dagegen nicht in allgemeiner Lage, kann man sie durch eine geschickte kleine Änderung in allgemeine Lage bringen.

10.14 Lineare Gleichungssysteme mit genau einer Lösung

Gegeben sei ein lineares Gleichungssystem wie in (10.1), und wir wollen annehmen, dass es für jedes $b \in \mathbb{K}^m$ genau eine Lösung hat. Das ist nur möglich, wenn $m = n$. Dann ist $r = n = m$ im reduzierten Schema (10.3), und die Koeffizientenmatrix A ist nach (9.26) (3) invertierbar. Es folgt

$$A \cdot x = b \Longleftrightarrow x = A^{-1} \cdot b.$$

Kennt man A^{-1}, kann man die Lösung sofort angeben. Dies ist besonders dann von Bedeutung, wenn man das Gleichungssystem für verschieden b lösen will. Wir wollen jetzt ein Verfahren vorstellen, wie man A^{-1} effektiv berechnen kann. Sei also

$$A^{-1} = (x_{ij}).$$

Aus $A \cdot A^{-1} = E_n = (\delta_{ij})$ und der Definition des Matrizenproduktes folgt:

$$\sum_{j=1}^{n} a_{ij} x_{jk} = \delta_{ik}.$$

Für jedes fest gewählte k müssen wir also ein Gleichungssystem

$$(*) \qquad A \cdot y^k = e^k \qquad k = 1, \dots, n$$

mit $y^k = \begin{pmatrix} x_{1k} \\ \vdots \\ x_{nk} \end{pmatrix}$ und e^k, dem k-ten Einheitsvektor, lösen.

Das machen wir simultan. Das Ausgangsschema wäre damit

10.15

$$\begin{array}{ccccc|ccccc}
x_1 & x_2 & & & x_n & e^1 & e^2 & & & e^n \\
\hline
a_{11} & a_{12} & \dots & \dots & a_{1n} & 1 & 0 & \dots & \dots & 0 \\
a_{21} & a_{22} & \dots & \dots & a_{2n} & 0 & 1 & \dots & \dots & 0 \\
\vdots & \vdots & & & \vdots & \vdots & & \ddots & 0 & \\
\vdots & \vdots & & & \vdots & \vdots & 0 & & \ddots & \\
a_{n1} & a_{n2} & \dots & \dots & a_{nn} & 0 & & & & 1
\end{array}$$

Da $r = n$, kann man den Gauß-Algorithmus ohne Spaltenvertauschung durchführen, so dass man folgendes reduzierte Schema erhält:

10.16

$$
\begin{array}{ccccc|ccccc}
x_1 & x_2 & & & x_n & e^1 & e^2 & & & e^n \\
\hline
1 & 0 & \dots & \dots & 0 & b^*_{11} & b^*_{12} & \dots & \dots & b^*_{1n} \\
0 & 1 & \dots & \dots & 0 & b^*_{21} & b^*_{22} & \dots & \dots & b^*_{2n} \\
\vdots & & \ddots & 0 & & \vdots & \vdots & \ddots & & \vdots \\
\vdots & 0 & & \ddots & & \vdots & \vdots & & \ddots & \vdots \\
0 & 0 & \dots & \dots & 1 & b^*_{n1} & b^*_{n2} & \dots & \dots & b^*_{nn}
\end{array}
$$

Wir können sofort ablesen, dass

$$y^k = \begin{pmatrix} x_{1k} \\ \vdots \\ x_{nk} \end{pmatrix} = \begin{pmatrix} b^*_{1k} \\ \vdots \\ b^*_{nk} \end{pmatrix}$$

die eindeutige Lösung ist, d.h.

10.17 Satz: $A^{-1} = (b^*_{ij})$.

10.18 Beispiel: $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & -1 \\ 0 & 1 & 2 \end{pmatrix}$

$$
\begin{array}{rrr|rrrl}
1 & 1 & 0 & 1 & 0 & 0 & \\
0 & -1 & -1 & 0 & 1 & 0 & \\
0 & 1 & 2 & 0 & 0 & 1 & \\
\hline
1 & 0 & -1 & 1 & 1 & 0 & \quad I+II \\
0 & 1 & 1 & 0 & -1 & 0 & \quad (-1) \\
0 & 0 & 1 & 0 & 1 & 1 & \quad III-II \\
\hline
1 & 0 & 0 & 1 & 2 & 1 & \quad I+III \\
0 & 1 & 0 & 0 & -2 & -1 & \quad II-III \\
0 & 0 & 1 & 0 & 1 & 1 &
\end{array}
$$

Also

$$A^{-1} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -2 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

11 Determinanten

In vielen Anwendungen hat man Gleichungssysteme

$$A \cdot x = b,$$

die für jedes b genau eine Lösung besitzen. Wir wissen bereits, dass das gleichbedeutend damit ist, dass die zugehörige lineare Abbildung

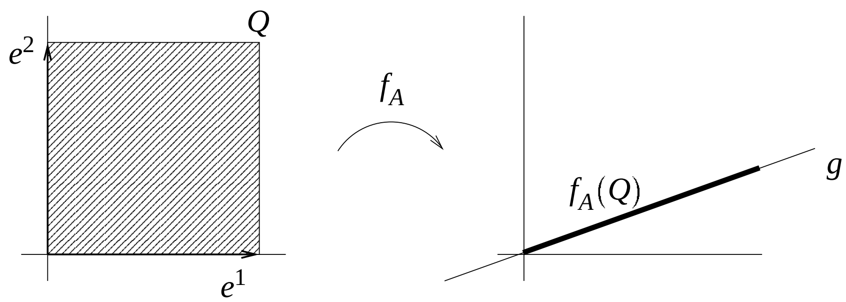
$$f_A : \mathbb{R}^n \to \mathbb{R}^m$$

bijektiv ist. Das wiederum ist gleichbedeutend damit, dass $m = n$, also A eine n-quadratische Matrix ist, und dass A invertierbar ist. Die Invertierbarkeit von A können wir mit Hilfe des Gauß-Algorithmus testen: A ist genau dann invertierbar, wenn $rg(A) = n$. Beim Gauß-Algorithmus müssen wir in der Regel dividieren, ein lästiges Unterfangen, das selbst beim Einsatz von Computern wegen eventueller kritischer Rundungsfehler nicht ganz unproblematisch ist. Schon ein kleiner Fehler kann eine Null zerstören. Wir hätten daher gern ein anderes Kriterium, das auch für den Einsatz von Computern besser geeignet ist.

Dazu eine Vorüberlegung: Sei A eine n-quadratische Matrix und $f_A : \mathbb{R}^n \to \mathbb{R}^n$ die zugehörige lineare Abbildung. Ist A nicht invertierbar, dann ist f_A nicht injektiv (6.29) und damit $Kern f_A \neq \{0\}$. Es gibt also ein $x \neq 0$ in $\mathbb{R}^n$ mit $f_A(x) = 0$. Schauen wir uns das im $\mathbb{R}^2$ an: Ist $\{e^1, e^2\}$ die Standardbasis, gilt für $x = (x_1, x_2)$

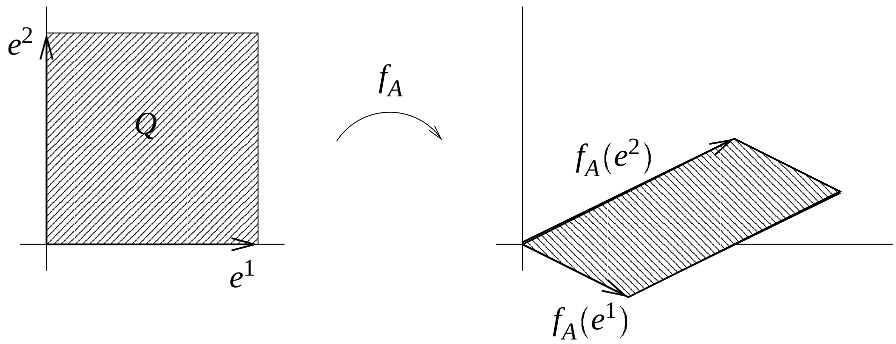
$$x = x_1 \cdot e^1 + x_2 \cdot e^2.$$

Es folgt: $0 = f_A(x) = x_1 \cdot f_A(e^1) + x_2 \cdot f_A(e^2)$, also $-x_1 \cdot f_A(e^1) = x_2 \cdot f_A(e^2)$. Damit liegen $f_A(e^1)$ und $f_A(e^2)$ auf einer Geraden g durch den Nullpunkt. Das Einheitsquadrat Q wird also auf einer Strecke auf dieser Geraden abgebildet.



Das bedeutet, dass die Fläche von $f_A(Q)$ Null ist.

Ist f_A dagegen bijektiv, so bildet $\{f_A(e^1), f_A(e^2)\}$ eine Basis des $\mathbb{R}^2$ und das Einheitsquadrat wird auf das von $f_A(e^1)$ und $f_A(e^2)$ definierte Parallelogramm abgebildet, dessen Flächeninhalt von 0 verschieden ist.



11.1 Vermutung: f_A ist genau dann bijektiv, wenn es den durch die Vektoren $e^1, \dots, e^n$ der Standardbasis des $\mathbb{R}^n$ gegebenen **Einheitswürfel** auf ein **Parallelotop** mit Volumen $\neq 0$ abbildet.

11.2 Aufgabe: Überlegen Sie sich präzise Definitionen für den Einheitswürfel in $\mathbb{R}^n$ und das durch Vektoren $v^1, \dots, v^n$ bestimmte Parallelotop.

Während diese Aufgabe recht einfach zu lösen ist, wenn man sich von der Anschauung im $\mathbb{R}^2$ und $\mathbb{R}^3$ leiten lässt, ist das Problem, wie man sinnvoll Volumina im $\mathbb{R}^n$ definiert, schon schwieriger. An diesem Beispiel lässt sich mathematische Formalisierung besonders schön demonstrieren.

11.3 Bezeichnung: Sind $v^1, \dots, v^n$ beliebige Vektoren in $\mathbb{R}^n$, dann bezeichne

$$D(v^1, \dots, v^n)$$

das noch zu definierende Volumen des von $v^1, \dots, v^n$ beschriebenen Parallelotops.

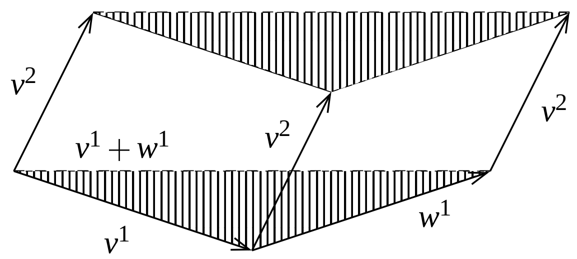
Wir verlangen nun von $D(v^1, \dots, v^n)$ ein paar nahe liegende Eigenschaften und wollen dann untersuchen, ob daraus eine vernünftige Definition wird.

11.4 Gewünschte Eigenschaften von $D(v^1, \dots, v^n)$:

(D1) Der n-dimensionale Einheitswürfel sollte das Volumen 1 haben, d.h.

$$D(e^1, e^2, \dots, e^n) = 1.$$

(D2) **Addition zweier Parallelotope:** Beispiel im $\mathbb{R}^2$



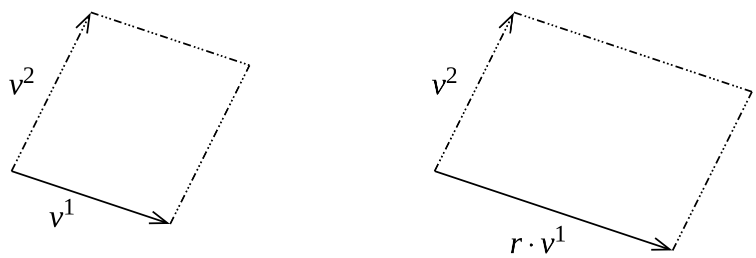
Die gestreiften Dreiecke haben denselben Flächeninhalt, so dass im $\mathbb{R}^2$ gilt

$$D(v^1, v^2) + D(w^1, v^2) = D(v^1 + w^1, v^2).$$

Daher fordern wir im $\mathbb{R}^n$: Sind $v^1, \dots, v^n \in \mathbb{R}^n$ und $w^i \in \mathbb{R}^n$, so gilt

$$D(v^1, \dots, v^i, \dots, v^n) + D((v^1, \dots, w^i, \dots, v^n) \\ = D(v^1, \dots, v^i + w^i, \dots, v^n)$$

Multiplikation eines $\mathbf{v^i} \in \mathbb{R}^\mathbf{n}$ mit einer $\mathbf{r} \in \mathbb{R}$



Im $\mathbb{R}^2$ gilt offensichtlich $D(r \cdot v^1, v^2) = r \cdot D(v^1, v^2)$. Allgemein:

$$D(v^1, v^2, \dots, r \cdot v^i, \dots, v^n) = r \cdot D(v^1, \dots, v^i, \dots, v^n)$$

(D3) Sind zwei der Vektoren $v^1, \dots, v^n \in \mathbb{R}^n$ gleich, liegt das durch sie definierte Parallelotop im Untervektorraum, der von diesen Vektoren aufgespannt wird, also in einem Teilraum von niederer Dimension. Das n-dimensionale Volumen dieses Parallelotops ist damit 0, so wie ein Parallelogramm in einer Ebene im $\mathbb{R}^3$ das 3-dimensionale Volumen Null hat. Also

$$D(v^1, \dots, v^n) = 0 \quad \text{falls } v^i = v^j \text{ für ein } i \neq j.$$

Das sind sicherlich Minimalforderungen, die wir an einen Volumenbegriff stellen können. Über eine Forderung sollten wir aber noch nachdenken: Ist r im zweiten Teil von (D2) negativ, erhalten wir ein negatives Volumen. Ist das sinnvoll? Wir könnten das durch die Forderung umgehen, dass

$$D(v^1, \dots, r \cdot v^i, \dots, v^n) = |r| \cdot D(v^1, \dots, v^i, \dots, v^n).$$

Eine solche Festlegung hätte aber viele Nachteile: (D2) bedeutet, dass für jedes $i = 1, \dots, n$ und fest gewählte Vektoren $v^1, \dots, v^{i-1}, v^{i+1}, \dots, v^n$ die Abbildung

$$\mathbb{R}^n \to \mathbb{R},\ x \mapsto D(v^1, v^{i-1}, x, v^{i+1}, \dots, v^n)$$

$\mathbb{R}$-linear ist, d.h.

$$D : \underbrace{\mathbb{R}^n \times \ldots \times \mathbb{R}^n}_{n-\text{mal}} \to \mathbb{R} \quad (v^1, \dots, v^n) \mapsto D(v^1, \dots, v^n)$$

ist in jeder der n Variablen (bei Festhalten der anderen) $\mathbb{R}$-linear. Eine solche Abbildung nennt man daher **MULTILINEAR**, und lineare Abbildungen haben wir hinlänglich studiert.

Zum anderen wissen zumindest einige Hörer, dass es im Zusammenhang der Integralrechnung durchaus sinnvoll ist, mit negativen Flächeninhalten zu argumentieren. Man spricht dann von orientierten Flächen. Eine Umkehrung der Orientierung führt eine Vorzeichenänderung beim Flächeninhalt mit sich. Um den geometrischen Flächeninhalt zu erhalten, muss man zum Absolutbetrag übergehen.

Nach diesen Vorbetrachtungen definieren wir (salopp)

11.5 Definition: Sei $\mathbb{K}$ ein Körper. Eine **ORIENTIERTE VOLUMENFUNKTION** für Parallelotope im $\mathbb{K}^n$, gewöhnlich **DETERMINANTE** genannt, ist eine Abbildung

$$D : \mathbb{K}^n \times \ldots \times \mathbb{K}^n \to \mathbb{K}, \quad (v^1, \dots, v^n) \mapsto D(v^1, \dots, v^n),$$

so dass folgende Axiome gelten

(D1) $D(e^1, \dots, e^n) = 1$

(D2) $D(v^1, \dots, v^i, \dots, v^n) + D(v^1, \dots, w^i, \dots, v^n) =$
$D(v^1, \dots, v^i + w^i, \dots, v^n)$

$D(v^1, \dots, r \cdot v^i, \dots, v^n) = r \cdot D(v^1, \dots, v^i, \dots, v^n)$

$\forall v^1, \dots, v^n, w^i \in \mathbb{K}^n,\ \forall r \in \mathbb{K},\ \forall i = 1, \dots, n$

(D3) $D(v^1,\dots,v^n)=0$, falls $v^i=v^{i+1}$ für ein $i\in\{1,2,\dots,n-1\}$

Wie bereits erwähnt, ist Axiom (D2) gleichbedeutend damit, dass D multilinear ist.

11.6 Definition: Seien $V_1,V_2,\dots,V_n$ und W $\mathbb{K}$-Vektorräume. Eine Abbildung

$$f: V_1\times V_2\times\dots\times V_n\to W$$

heißt **MULTILINEAR**, wenn für jedes $i=1,\dots,n$ und für jede Familie fest gewählter Vektoren $\{v^j\in V_j;\, j\neq i\}$ die Abbildung

$$f_i: V_i\to W, x\mapsto f(v^1,\dots,v^{i-1},x,v^{i+1},\dots,v^n)$$

$\mathbb{K}$-linear ist. Ist $W=\mathbb{K}$, spricht man auch von einer **MULTILINEARFORM**.

Noch wissen wir nicht, ob Determinanten existieren, und wenn ja, wie viele verschiedene es gibt. Antwort gibt folgender zentraler Satz, dessen Beweis aber einiger Vorbereitungen bedarf:

11.7 Satz: Für jedes $n\geq 1$ gibt es genau eine Determinante

$$D=D_n:\mathbb{K}^n\times\dots\times\mathbb{K}^n\to\mathbb{K}.$$

Wie im Beweis von 6.14 erläutert, empfiehlt es sich, zunächst die Eindeutigkeit von D zu zeigen in der Hoffnung, dass diese Untersuchung Hinweise auf die Konstruktion liefert. Wir nehmen daher zunächst an, dass es ein solches D gibt.

Zunächst beweisen wir die multilineare Version von Satz 6.14.

11.8 Satz: Seien $V_1,\dots,V_n$ endlich dimensionale $\mathbb{K}$-Vektorräume, $\dim V_i=r_i$ und $\mathscr{B}_i=\{v^{i1},\dots,v^{ir_i}\}$ sei Basis von V_i, $i=1,\dots,n$. Sei W ein beliebiger $\mathbb{K}$-Vektorraum. Zu jedem “Multiindex” $(j_1,\dots,j_n)$ mit $1\leq j_k\leq r_k$, $k=1,\dots,n$ sei ein Vektor $y^{j_1,\cdots,j_n}\in W$ gegeben. Dann gibt es genau eine multilineare Abbildung

$$f: V_1\times\dots\times V_n\to W,$$

so dass $f(v^{1j_1},\dots,v^{nj_n})\overset{(*)}{=}y^{j_1,\cdots,j_n}$ für jeden Multiindex $(j_1,\dots,j_n)$.

Bevor wir den Satz beweisen. wollen wir seine Aussage an einem Beispiel erläutern.

Beispiel: Seien V_1, V_2, V_3 $\mathbb{K}$-Vektorräume der Dimension $r_1=2$, $r_2=3$ und $r_3=2$. Seien $\mathscr{B}_1=\{v^{1,1},v^{1,2}\}$, $\mathscr{B}_2=\{v^{2,1},v^{2,2},v^{2,3}\}$ und $\mathscr{B}_3=\{v^{3,1},v^{3,2}\}$ Basen von V_1, V_2 bzw. V_3. Wir haben 12 verschiedene Multiindices, nämlich

$$\begin{array}{cccccc}(1,1,1) & (1,1,2) & (1,2,1) & (1,2,2) & (1,3,1) & (1,3,2)\\ (2,1,1) & (2,1,2) & (2,2,1) & (2,2,2) & (2,3,1) & (2,3,2)\end{array}$$

Ist uns zu jedem Multiindex ein Vektor in einem $\mathbb{K}$-Vektorraum W gegeben, also 12 nicht notwendig verschiedene Vektoren $y^{1,1,1}$, $y^{1,1,2}, \ldots,$ $y^{2,3,2}$ aus W, dann gibt es genau eine multilineare Abbildung

$$f : V_1 \times V_2 \times V_3 \to W,$$

so dass $f(v^{1,1}, v^{2,1}, v^{3,1}) = y^{1,1,1},\ f(v^{1,1}, v^{2,1}, v^{3,2}) = y^{1,1,2}, \ldots, f(v^{1,2}, v^{2,3}, v^{3,2}) = y^{2,3,2}$. Die Indices der y sind die zweiten Indices der Basisvektoren.

Beweis: Wie auch schon bei der Formulierung des Satzes unterschlagen wir zur Förderung der Lesbarkeit oft das Komma zwischen den Indices.

Es hat sich bisher als gute Strategie erwiesen, zunächst die Eindeutigkeit einer solchen multilinearen Abbildung anzugehen. Dieser Strategie bleiben wir treu:

Eindeutigkeit: Sei als f eine multilineare Abbildung wie in der Aussage des Satzes und $(x^1, \ldots, x^n) \in V_1 \times \ldots \times V_n$ ein beliebiges Element. Da $\mathscr{B}_i$ Basis von V_i ist, hat x^i eine eindeutige Darstellung der Form

$$x^i = a_{i1} v^{i1} + \ldots + a_{ir_i} v^{ir_i}, \quad a_{ij} \in \mathbb{K}.$$

Es folgt

$$
\begin{aligned}
(**) \quad f(x^1, \ldots, x^n) &= \sum_{j_1=1}^{r_1} a_{1j_1} f(v^{1j_1}, x^2, \ldots, x^n) \\
&\qquad\qquad \text{Linearität in 1. Variablen} \\
&= \sum_{j_1=1}^{r_1} a_{1j_1} \sum_{j_2=1}^{r_2} a_{2j_2} f(v^{1j_1}, v^{2j_2} x^3, \ldots, x^n) \\
&\qquad\qquad \text{Linearität in 2. Variablen} \\
&= \sum_{j_1=1}^{r_1} \sum_{j_2=1}^{r_2} \sum_{j_3=1}^{r_3} a_{1j_1} \cdot a_{2j_2} \cdot a_{3j_3} \cdot f(v^{1j_1}, v^{2j_2}, x^3, \ldots, x^n) \\
&\qquad\qquad \text{Linearität in 3. bis n. Variablen} \\
&\ \ \vdots \\
&= \sum_{j_1=1}^{r_1} \sum_{j_2=1}^{r_2} \cdots \sum_{j_n=1}^{r_n} a_{1j_1} \cdot a_{2j_2} \cdot \ldots \cdot a_{nj_n} \cdot f(v^{1j_1}, \ldots, v^{kj_n}) \\
&= \sum_{j_1=1}^{r_1} \sum_{j_2=1}^{r_2} \cdots \sum_{j_n=1}^{r_n} a_{1j_1} \cdot a_{2j_2} \cdot \ldots \cdot a_{nj_n} \cdot y^{j_1, \ldots, j_n}
\end{aligned}
$$

Da für jede multilineare Abbildung, die die Bedingung $(*)$ des Satzes erfüllt, die Gleichung $(**)$ gilt, sind diese Abbildungen alle gleich.

Existenz: Wir definieren f durch die Gleichungen $(**)$.

Ist $x^i = v^{ik_i}$, so gilt $a_{ik_i} = \delta_{ik_i}$. In der Gleichung $(**)$ für $f(v^{1k_1}, \ldots, v^{nk_n})$ sind alle Summanden $a_{1j_1} \cdot \ldots \cdot a_{nj_n}$ Null außer $a_{1k_1} \cdot \ldots \cdot a_{nk_n} = 1$. Also erfüllt f die Bedingung $(*)$ des Satzes.

f ist multilinear: Sei $z^i = \sum_{j_i=1}^{r_i} b_{ij_i} v^{ij_i}$, also $x^i + z^i = \sum_{j_i=1}^{r_i} (a_{ij_i} + b_{ij_i}) v^{ij_i}$. Nach Definition gilt:

$$\begin{aligned}
&f(x^1, \dots, x^i + z^i, \dots, x^n) \\
&= \sum_{j_1=1}^{r_1} \cdots \sum_{j_n=1}^{r_n} a_{1j_1} \cdot \dots \cdot (a_{ij_i} + b_{ij_i}) \cdot \dots \cdot a_{nj_n} y^{j_1, \dots, j_n} \\
&= \sum_{j_1=1}^{r_1} \cdots \sum_{j_n=1}^{r_n} a_{1j_1} \cdot \dots \cdot a_{nj_n} y^{j_1, \dots, j_n} \\
&\qquad + \sum_{j_1=1}^{r_1} \cdots \sum_{j_n=1}^{r_n} a_{1j_1} \cdot \dots \cdot b_{ij_i} \cdot \dots \cdot a_{nj_n} \cdot y^{j_1, \dots, j_n} \\
&= f(x^1, \dots, x^i, \dots, x^n) + f(x^1, \dots, z^i, \dots, x^n)
\end{aligned}$$

Genauso zeigt man, dass $f(x^1, \dots, rx^i, \dots, x^n) = r \cdot f(x^1, \dots, x^i, \dots, x^n)$. □

11.9 Definition: Eine multilineare Abbildung $f : V \times \ldots \times V \to W$ heißt **ALTERNIEREND**, wenn $f(x^1, \dots, x^n) = 0$, falls $x^i = x^{i+1}$ für ein $i \in \{1, \dots, n-1\}$.

11.10 Lemma: Für eine alternierende multilineare Abbildung $f : V \times \ldots \times V \to W$ gilt

(1) $f(x^1, \dots, x^i, \dots, x^j, \dots, x^n) = -f(x^1, \dots, \underbrace{x^j}_{i\text{-te Stelle}}, \dots, \underbrace{x^i}_{j\text{-te Stelle}}, \dots, x^n)$

(2) $f(x^1, \dots, , x^n) = 0$, falls $x^i = x^j$ für ein $i \neq j$.

Beweis: 1. Fall: $j = i+1$. Wir konzentrieren uns nur auf die Position i und $i+1$ und unterdrücken die übrigen Positionen beim Aufschreiben.

$$\begin{aligned}
0 &= f(x^i + x^{i+1}, x^i + x^{i+1}) \\
&= f(x^i, x^i) + f(x^i, x^{i+1}) + f(x^{i+1}, x^i) + f(x^{i+1}, x^{i+1}) \\
&= f(x^i, x^{i+1}) + f(x^{i+1}, x^i)
\end{aligned}$$

Also $f(x^i, x^{i+1}) = -f(x^{i+1}, x^i)$.

2. Fall: $j = i + k > i$: Um i an die j-te Stelle zu bringen, vertauschen wir es nacheinander mit $i+1, i+2, \dots, i+k = j$. Um danach $j = i+k$ an die i-te Stelle zu bringen, vertauschen wir es nacheinander mit $i+k-1, i+k-2, \dots, i+1$. Insgesamt machen wir dabei $k + (k-1) = 2k-1$ Vertauschungen benachbarter Elemente. Nach Fall 1 ändert f dabei das Vorzeichen um $(-1)^{2k-1} = -1$.

(2) Wir vertauschen j mit $i+1$ und erhalten

$$f(x^1, \dots, x^i, \dots, x^j, \dots, x^n) = -f(x^1, \dots, x^i, x^j, \dots, x^{i+1}, \dots, x^n) = 0,$$

weil $x^i = x^j$. □

11.11 Ist $f : \mathbb{K}^n \times \ldots \times \mathbb{K}^n \to \mathbb{K}$ eine alternierende n-**Form** und $\{e^1, \ldots, e^n\}$ die Standardbasis, dann folgt aus (11.8) und seinem Beweis

$$f(x^1, \ldots, x^n) = \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} x_{1j_1} \cdot \ldots \cdot x_{nj_n} \cdot f(e^{j_1}, \ldots, e^{j_n}),$$

falls $x^i = (x_{i1} \ldots, x_{in})$. Die Werte $f(e^{j_1}, \ldots, e^{j_n})$ sind nach (11.10) aber Null, falls $j_k = j_l$ für ein $k \neq l$, d.h. $f(e^{j_1}, \ldots, e^{j_n})$ ist höchstens dann von Null verschieden, falls alle Indices $j_1, \ldots, j_n$ verschieden sind, d.h. $(j_1, \ldots, j_n)$ besteht genau aus den Zahlen $1, \ldots, n$, aber in anderer Anordnung.

11.12 Definition: Eine **PERMUTATION** von $\{1, \ldots, n\}$ ist eine Umordnung dieser Menge, d.h. ein Tupel $(j_1, \ldots, j_n)$ von Zahlen aus $\{1, \ldots, n\}$, so dass $j_k \neq j_l$ für $k \neq l$.

Es folgt: $f(x^1, \ldots, x^n) = \sum\limits_{(j_1, \cdots, j_n)} x_{1j_1} \cdot \ldots \cdot x_{nj_n} \cdot f(e^{j_1}, \ldots, e^{j_n})$, wobei $(j_1, \ldots, j_n)$ alle Permutationen durchläuft.

Nach (11.10) sind die Werte $f(e^{j_1}, \ldots, e^{j_n})$ aber bereits durch $f(e^1, \ldots, e^n)$ festgelegt: Sei $(j_1, \ldots, j_n)$ eine Permutation. Wir vertauschen j_1 mit $1 \in (j_1, \ldots, j_n)$, falls $j_1 \neq 1$. Wir erhalten einen Ausdruck $(1, j_2, \ldots)$. Dann vertauschen wir j_2 mit 2, falls $j_2 \neq 2$ usw.

11.13 Bezeichnung: Sei k die Anzahl der Vertauschungen, die wir nach diesem speziellen Verfahren insgesamt gemacht haben, um $(j_1, \ldots, j_n)$ nach $(1, \ldots, n)$ zu überführen. Dann heißt $\operatorname{sign}(j_1, \ldots, j_n) := (-1)^k$ das **VORZEICHEN** der Permutation $(j_1, \ldots, j_n)$.

Wir erhalten

11.14 Satz: Für eine alternierende Form $f : \mathbb{K}^n \times \ldots \times \mathbb{K}^n \to \mathbb{K}$ gilt

$$f(x^1, \ldots, x^n) = \sum_{(j_1, \cdots, j_n)} \operatorname{sign}(j_1, \ldots, j_n) \cdot x_{1j_1} \cdot \ldots \cdot x_{nj_n} \cdot f(e^1, \ldots, e^n)$$

wobei $(j_1, \ldots, j_n)$ alle Permutationen von $\{1, \ldots, n\}$ durchläuft und $x^i = (x_{i1}, \ldots, x_{in})$.

Da eine Determinante eine alternierende n-Form ist, folgt mit (D1).

11.15 Satz: $D(x^1, \ldots, x^n) = \sum\limits_{(j_1, \cdots, j_n)} \operatorname{sign}(j_1, \ldots, j_n) \cdot x_{1j_1} \cdot \ldots \cdot x_{nj_n}$, wobei $(j_1, \ldots, j_n)$ die Menge aller Permutationen von $\{1, \ldots, n\}$ durchläuft.

Damit ist die Eindeutigkeit der Determinante gezeigt.

11.16 Bezeichnung: Um die Schreibarbeit zu minimieren, fasst man x^i als i-ten Zeilenvektor einer n-quadratischen Matrix $A = (x_{ij})$ auf und benutzt Symbole

$$D(A) := D(x^1, \dots, x^n) =: \det A =: |A|.$$

Nach bewährtem Rezept könnten wir nun die Formel (11.15) zur Definition einer Determinanten hernehmen und zeigen, dass die so definierte Abbildung die Axiome (D1), (D2), (D3) erfüllt. Das ist für (D1) und (D2) einfach, bei (D3) geraten wir aber in Schwierigkeiten. Deshalb beschreiten wir ausnahmsweise einen anderen Weg, der aber den zusätzlichen Vorteil hat, eine von (11.15) abweichende Berechnungsmöglichkeit der Determinante zu liefern. Mit der Bezeichnung von (11.16) ist die Determinante eine Abbildung

$$D = D_n : \mathbb{K}^{n,n} \to \mathbb{K}.$$

Wir konstruieren D_n durch Induktion nach n:

11.17 $D_1 : \mathbb{K}^{1,1} = \mathbb{K} \to \mathbb{K}$, $\qquad D_1(x) = x$ ist eine Determinante.

11.18 Bezeichnung: Für eine n-quadratische Matrix A und $1 \le i, j \le n$ bezeichne A_{ij} die $(n-1)$-quadratische Matrix, die aus A durch Weglassen der i-ten Zeile und j-ten Spalte entsteht.

Induktionsschritt: Sei $A = (x_{ij}) \in \mathbb{K}^{n,n}$. Wir fixieren ein j mit $1 \le j \le n$ und definieren

11.19

$$\boxed{D_n(A) = \sum_{i=1}^{n} (-1)^{i+j} x_{ij} \cdot D_{n-1}(A_{ij})}$$

Behauptung: Ist D_{n-1} eine Determinante, dann ist auch D_n eine Determinante.

Nachweis von (D1): $D_n(e^1, \dots, e^n) = D_n(E_n)$. Es gilt

$$\begin{aligned} D_n(E_n) &= \sum_{i=1}^{n} (-1)^{i+j} \delta_{ij} \cdot D_{n-1}((E_n)_{ij}) = (-1)^{j+j} D_{n-1}((E_n)_{jj}) \\ &= (-1)^{2j} D_{n-1}(E_{n-1}) = 1 \end{aligned}$$

Nachweis von (D2): Wir müssen zeigen, dass D_n in den Zeilenvektoren von $A = (x_{ij})$ multilinear ist. Sei $r \in \mathbb{K}$, sei $B \in \mathbb{K}^{n,n}$ die Matrix mit den Zeilenvektoren $x^1, \dots, y^k, \dots, x^n$ und C die Matrix mit den Zeilenvektoren $x^1, \dots, r \cdot x^k + y^k, \dots, x^n$. Wir müssen zeigen:

$$D_n(C) = r \cdot D_n(A) + D_n(B)$$

(Warum folgt daraus die Linearität in der k-ten Zeile?)

$$A = \begin{pmatrix} x_{11} \cdots x_{1n} \\ \vdots \quad \vdots \\ x_{k1} \cdots x_{kn} \\ \vdots \quad \vdots \\ x_{n1} \cdots x_{nn} \end{pmatrix} B = \begin{pmatrix} x_{11} \cdots x_{1n} \\ \vdots \quad \vdots \\ y_{k1} \cdots y_{kn} \\ \vdots \quad \vdots \\ x_{n1} \cdots x_{nn} \end{pmatrix} C = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ rx_{k1} + y_{k1} & \cdots & rx_{kn} + y_{kn} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}$$

Zur Unterscheidung setzen wir $B = (b_{ij})$, $C = (c_{ij})$. Wir untersuchen

$$(-1)^{i+j} c_{ij} \cdot D_{n-1}(C_{ij})$$

Ist $\mathbf{i} \neq \mathbf{k}$, so hat C_{ij} noch die alte k-te Zeile (aber ohne das Element der j-ten Spalte). Dies Zeile ist das r-fache der entsprechenden Zeile von A_{ij} **plus** der entsprechenden Zeile von B_{ij}. Da D_{n-1} Determinante ist, folgt

$$D_{n-1}(C_{ij}) = r \cdot D_{n-1}(A_{ij}) + D_{n-1}(B_{ij})$$

Anderseits gilt $c_{ij} = x_{ij} = b_{ij}$. Also

$$(*) \quad (-1)^{i+j} c_{ij} D_{n-1}(C_{ij}) = (-1)^{i+j} \cdot r \cdot c_{ij} \cdot D_{n-1}(A_{ij}) + (-1)^{i+j} c_{ij} D_{n-1}(B_{ij})$$
$$= (-1)^{i+j} r \cdot x_{ij} D_{n-1}(A_{ij}) + (-1)^{i+j} b_{ij} \cdot D_{n-1}(B_{ij})$$

Ist $\mathbf{i} = \mathbf{k}$, dann ist $A_{kj} = B_{kj} = C_{kj}$, aber $c_{kj} = rx_{kj} + y_{kj}$. Es folgt

$$(-1)^{i+j} \cdot c_{kj} \cdot D_{n-1}(C_{kj}) = (-1)^{i+j} \cdot r \cdot x_{kj} \cdot D_{n-1}(A_{kj}) + (-1)^{i+j} \cdot b_{kj} \cdot D_{n-1}(B_{kj})$$

D.h. die Gleichung $(*)$ gilt für alle $i = 1, \ldots, n$. Es folgt

$$\begin{aligned} D_n(C) = \sum_{i=1}^{n} (-1)^{i+j} c_{ij} D_{n-1}(C_{ij}) &= r \cdot \sum_{i=1}^{n} (-1)^{i+j} x_{ij} \cdot D_{n-1}(A_{ij}) \\ &\qquad + \sum_{i=1}^{n} (-1)^{i+j} b_{ij} D_{n-1}(B_{ij}) \\ &= r \cdot D_n(A) + D_n(B) \end{aligned}$$

Nachweis von (D3): Sei etwa $x^k = x^{k+1}$. Dann gilt $A_{kj} = A_{k+1,j}$ und $x_{kj} = x_{k+1,j}$. Für $i \neq k$, $k+1$ hat A_{ij} zwei gleiche Zeilen, so dass $D_{n-1}(A_{ij}) = 0$. Es folgt

$$\begin{aligned} D_n(A) &= \sum_{i=1}^{n} (-1)^{i+j} x_{ij} D_{n-1}(A_{ij}) \\ &= (-1)^{k+j} x_{kj} D_{n-1}(A_{kj}) + (-1)^{k+1+j} x_{k+1,j} D_{n-1}(A_{k+1,j}) \\ &= 0 \end{aligned}$$

11.20 Bezeichnung: Die induktive Berechnung der Determinante mit Hilfe der Formel (11.19) nennt man **Entwicklung nach der j-ten Spalte**.

11.21 Bemerkung: Sei $(j_1,\dots,j_n)$ eine Permutation. $\mathrm{sign}(j_1,\dots,j_n)$ haben wir über eine **spezielle** Folge von Vertauschungen definiert, die $(j_1,\dots,j_n)$ in $(1,\dots,$ $n)$ überführen. Wählen wir nun eine **beliebige** Folge von l Vertauschungen, die $(j_1,\dots,j_n)$ in $(1,\dots,n)$ überführen, so gilt nach (11.10) und (11.15), da D existiert:

$$\mathrm{sign}(j_1,\dots,j_n) = D(e^{j_1},\dots,e^{j_n}) = (-1)^l D(e^1,\dots,e^n) = (-1)^l.$$

Haben wir also zwei Folgen von k bzw. l Vertauschungen, die $(j_1,\dots,j_n)$ in $(1,\dots,n)$ überführen, so gilt

$$(-1)^k = \mathrm{sign}(j_1,\dots,j_n) = (-1)^l$$

d.h. k und l sind beide entweder gerade oder ungerade.

Beispiel:

$(3,2,1,) \longmapsto (1,2,3,)$ eine Vertauschung

$(3,2,1,) \longmapsto (3,1,2,) \longmapsto (1,3,2,) \longmapsto (1,2,3,)$ drei Vertauschungen

Da wir wissen, dass die Determinante existiert, können wir sie in den unteren Dimensionen leicht mit Hilfe der Formel (11.15) berechnen.

11.22 Beispiel: $D_1 : \mathbb{K} \to \mathbb{K}$, $D_1 = id$

$D_2 : \mathbb{K}^2 \times \mathbb{K}^2 \to \mathbb{K}$: Wir haben die Permutationen $(1,2)$ und $(2,1)$

$$D_2 : \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = x_{11} \cdot x_{22} - x_{12}x_{21}$$

$D_3 : \mathbb{K}^3 \times \mathbb{K}^3 \times \mathbb{K}^3 \to \mathbb{K}$: Wir haben 6 Permutationen:

$(1,2,3),(2,3,1),(3,1,2),(3,2,1),(1,3,2),(2,1,3)$

Die ersten drei haben positives, die letzten drei negatives Vorzeichen. Zur Berechnung von

$$D_3 \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} = D_3(A)$$

ergänzt man das Matrixschema zu

$$\begin{array}{ccccc} x_{11} & x_{12} & x_{13} & x_{11} & x_{12} \\ x_{21} & x_{22} & x_{23} & x_{21} & x_{22} \\ x_{31} & x_{32} & x_{33} & x_{31} & x_{32} \end{array}$$

und multipliziert entlang der Linien. Linien von links oben nach rechts unten geben positives, Linien von links unten nach rechts oben negatives Vorzeichen:

$$D_3 = x_{11}x_{22}x_{33} + x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} - x_{13}x_{22}x_{31} - x_{11}x_{23}x_{32} - x_{12}x_{21}x_{33}$$

11.23 Bezeichnung: Σ_n bezeichne die Menge aller Permutationen von $\{1,\ldots,n\}$.

11.24 Aufgabe: Beweisen Sie durch vollständige Induktion, dass Σ_n genau $n!$ Elemente besitzt.

Nach 11.24 gibt es 24 Permutationen $(j_1,\ldots,j_4)$. Damit wird die explizite Berechnung von $D_n(A)$ für $n \geq 4$ schon ziemlich aufwendig. Zum Glück können wir mit Hilfe der Linearität und des Lemmas (11.10) Zeilenoperationen und damit eine Art Gauß-Algorithmus zur Vereinfachung anwenden.

11.25 Lemma: Ist $f: V \times \ldots \times V \to W$ alternierende multilineare Abbildung und $r \in \mathbb{K}$, so gilt für $i \neq j$

$$f(x^1,\ldots,x^i\ldots,x^j,\ldots,x^n) = f(x',\ldots,x^i + r\cdot x^j,\ldots,x_j,\ldots,x^n)$$

Beweis:

$$\begin{aligned} & f(x^1,\ldots,x^i + r\cdot x^j,\ldots,x_j,\ldots,x^n) \\ & = f(x^1,\ldots,x^i\ldots,x^j,\ldots,x^n) + r\cdot \underbrace{f(x^1,\ldots,x^j\ldots,x^j,\ldots,x^n)}_{=0 \text{ nach } (11.10)} \end{aligned}$$

□

Haben wir eine Matrix so in obere Dreiecksform gebracht, sind wir fertig.

(**Vorsicht:** Multipliziert man eine Zeile mit $r \neq 0$, verändert sich die Determinante ebenfalls um den Faktor r. Vertauscht man zwei Zeilen, ändert die Determinante nach (11.10) ihr Vorzeichen!)

11.26

$$D\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & 0 & \ddots & \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} = a_{11}\cdot a_{22}\cdot\ldots\cdot a_{nn}$$

Beweis: Entwicklung nach der ersten Spalte:

$$D(A) = (-1)^{1+1} a_{11} \cdot D \begin{pmatrix} a_{22} & - & - & a_{2n} \\ & \diagdown & & | \\ & & \diagdown & | \\ 0 & & & a_{nn} \end{pmatrix} + 0 = a_{11} \cdot D \begin{pmatrix} a_{22} & - & - & a_{2n} \\ & \diagdown & & | \\ & & \diagdown & | \\ 0 & & & a_{nn} \end{pmatrix}$$

induktiv folgt das Ergebnis. □

Im Gegensatz zu den Gleichungssystemen hat die Determinante den Vorteil, dass man statt Zeilen- auch dieselben Spaltenmanipulationen durchführen darf, denn es gilt

11.27 Satz: Für eine n-quadratische Matrix gilt

$$D(A) = D(A^t).$$

Eine Permutation $(j_1, \dots, j_n)$ können wir auch als bijektive Abbildung

$$\sigma : \{1, \dots, n\} \to \{1, \dots, n\},\ k \mapsto j_k$$

auffassen. Da $j_k \neq j_l$ für $k \neq l$, ist σ injektiv und damit auch bijektiv. Da jede bijektive Abbildung genau eine Umkehrabbildung besitzt, ist die Zuordnung

$$\Sigma_n \to \Sigma_n,\ \sigma \mapsto \sigma^{-1}$$

bijektiv. Benutzen wir diese Interpretation einer Permutation, so gilt

$$x_{1 j_1} \cdot \ldots \cdot x_{n j_n} = x_{1\sigma(1)} \cdot \ldots \cdot x_{n\sigma(n)}.$$

Wir vertauschen nun die Faktoren in diesem Produkt so, dass sie nach dem 2. Index statt nach dem 1. geordnet sind. Man beachte dabei, dass jede Zahl $k \in \{1, \dots, n\}$ genau einmal unter $\sigma(1), \dots, \sigma(n)$ auftritt. Man beachte weiter, dass der Faktor mit dem 2. Index k gerade $x_{\sigma^{-1}(k),k}$ ist. Da die Zuordnung $\sigma \to \sigma^{-1}$ bijektiv ist, folgt aus (11.15)

$$\begin{aligned} D(A) &= \sum_{\sigma \in \Sigma_n} x_{1,\sigma(1)} \cdot \ldots \cdot x_{n,\sigma(n)} \\ &= \sum_{\sigma \in \Sigma_n} x_{\sigma^{-1}(1),1} \cdot \ldots \cdot x_{\sigma^{-1}(n),n} \cdot \operatorname{sign}(\sigma) \\ &\overset{(*)}{=} \sum_{\sigma^{-1} \in \Sigma_n} x_{\sigma(1),1} \cdot \ldots \cdot x_{\sigma(n),n} \cdot \operatorname{sign}(\sigma^{-1}) \\ &\overset{(**)}{=} \sum_{\sigma \in \Sigma} x_{\sigma(1),1} \cdot \ldots \cdot x_{n,\sigma(n)} \cdot \operatorname{sign}(\sigma) \\ &= D(A^t) \end{aligned}$$

Gleichung $(*)$ ist Umbenennung von σ nach σ^{-1}. Gleichung $(**)$ benutzt, dass $\sigma^{-1} \mapsto \sigma$ bijektiv ist und $\text{sign}(\sigma) = \text{sign}(\sigma^{-1})$. Letzteres ist klar, denn die Vertauschungen, die man für den Übergang von σ nach $id = (1, \ldots, n)$ braucht, muss man bei σ^{-1} gerade rückgängig machen. $\square$

11.28 Folgerung: Entwicklung nach der i-ten Zeile. Sei i fest gewählt.

$$D_n(A) = \sum_{j=1}^{n} (-1)^{i+j} x_{ij} \cdot D_{n-1}(A_{ij}) \quad \text{für } A = (x_{ij}) \in \mathbb{K}^{n,n}$$

Als weiteres wichtiges Ergebnis zeigen wir:

11.29 Satz: $D : \mathbb{K}^{n,n} \to \mathbb{K}$ ist multiplikativ, d.h. $D(A \cdot B) = D(A) \cdot D(B)$.

Beweis: Die Abbildung

$$f : \mathbb{K}^n \times \ldots \times \mathbb{K}^n \to \mathbb{K}, \; f(x^1, \ldots, x^n) = D(x^1 \cdot B, \ldots, x^n \cdot B)$$

ist eine alternierende n-Form, denn für $r \in \mathbb{K}$ und $y^i \in \mathbb{K}^n$ gilt

$$\begin{aligned}
f(x^1, \ldots, rx^i + y^i, \ldots, x^n) &= D(x^1 \cdot B, \ldots, (rx^i + y^i) \cdot B, x^n \cdot B) \\
&= D(x^1 \cdot B, \ldots, r \cdot x^i \cdot B + y^i \cdot B, x^n \cdot B) \\
&= r \cdot D(x^1 \cdot B, \ldots, x^i \cdot B, \ldots, x^n \cdot B) + D(x^1 \cdot B, \ldots, y^i \cdot B, \ldots, x^n \cdot B) \\
&= r \cdot f(x^1, \ldots, x^i, \ldots, x^n) + f(x^1, \ldots, y^i, \ldots, x^n)
\end{aligned}$$

Gilt $x^i = x^{i+1}$, so folgt $x^i \cdot B = x^{i+1} \cdot B$, so dass $f(x^1, \ldots, x^n) = 0$.
Ist A die Matrix mit den Zeilenvektoren $x^1, \ldots, x^n$, dann sind $x^1 \cdot B, \ldots, x^n \cdot B$ die Zeilenvektoren von $A \cdot B$, so dass $f(A) = D(A \cdot B)$. Da

$$f(A) = \sum_{(j_1, \ldots, j_n) \in \Sigma_n} \text{sign}(j_1, \ldots, j_n) \cdot x_{1j_1} \cdot \ldots \cdot x_{nj_n} \cdot f(E_n) = D(A) \cdot D(B)$$

nach (11.14), erhalten wir die Behauptung. $\square$
Kommen wir abschließend auf unser ursprüngliches Problem zurück:

11.30 Satz: $A \in \mathbb{K}^{n,n}$ ist invertierbar $\iff D(A) \neq 0$.

Beweis: Ist A invertierbar, dann folgt aus (11.29)

$$1 = D(E_n) = D(A \cdot A^{-1}) = D(A) \cdot D(A^{-1})$$

Also $D(A) \neq 0$.

Ist A nicht invertierbar, so ist $rang(A) = r < n$. Die aus A mit dem Gauß-Algorithmus gewonnene obere Dreiecksmatrix hat eine Nullzeile. Aus (11.26) folgt somit $D(A) = 0$. □

Wir können aber noch erheblich mehr sagen. Dafür beweisen wir zunächst

11.31 Lemma: Sei $A \in \mathbb{K}^{n,n}$ eine Matrix mit dem Zeilenvektoren $z^1, \ldots, z^n$ und den Spaltenvektoren $s^1, \ldots, s^n$. Dann gilt

$$\begin{aligned} D(z^1, \ldots, z^{i-1}, e^j, z^{i+1}, \ldots, z^n) &= (-1)^{i+j} D(A_{ij}) \\ &= D(s^1, \ldots, s^{j-1}, e^i, s^{j+1}, \ldots, s^n) \end{aligned}$$

Beweis: Entwickeln wir nach der i-ten Zeile, erhalten wir die linke Gleichung, entwickeln wir nach der j-ten Spalte, erhalten wir die rechte. □

11.32 Definition: Sei $A \in \mathbb{K}^{n,n}$. Dann heißt die Matrix $A^* = (a^*_{ij}) \in \mathbb{K}^{n,n}$ mit

$$a^*_{ij} = (-1)^{i+j} D(A_{ji})$$

(beachte die Reihenfolge der Indices) die zu A **ADJUNGIERTE** Matrix.

11.33 Satz: Für $A = (a_{ij}) \in \mathbb{K}^{n,n}$ gilt

$$A \cdot A^* = A^* \cdot A = D(A) \cdot E_n$$

Insbesondere gilt: Ist $D(A) \neq 0$, so ist A invertierbar und

$$A^{-1} = \frac{1}{D(A)} \cdot A^*$$

Sei z^i der i-te Zeilenvektor von A und $A \cdot A^* = C = (c_{ik})$. Es gilt mit (11.31)

$$\begin{aligned} c_{ik} &= \sum_{j=1}^{n} a_{ij} \cdot a^*_{jk} = \sum_{j=1}^{n} a_{ij} \cdot (-1^{j+k} D(A_{kj}) \\ &= \sum_{j=1}^{n} a_{ij} \cdot D(z^1, \ldots, z^{k-1}, e^j, z^{k+1}, \ldots, z^n) \\ &= D(z^1, \ldots, z^{k-1}, \sum_{j=1}^{n} a_{ij} e^j, z^{k+1}, \ldots, z^n) \\ &= D(z^1, \ldots, z^{k-1}, z^i, z^{k+1}, \ldots, z^n) = \begin{cases} 0 & i \neq k \quad \text{nach (11.10.2)} \\ D(A) & i = k \end{cases} \end{aligned}$$

Da Matrizenmultiplikation nicht kommutativ ist, muss man auch $A^* \cdot A = D(A) \cdot E_n$ nachweisen. Das macht man genauso mit Hilfe der Spaltenvektoren s^j von A und der 2. Formel aus (11.31).

Alternativ kann man auch folgendes Resultat benutzen.

11.34 Aufgabe: $(A^t)^* = (A^*)^t$

Als weitere wichtige Anwendung zeigen wir

11.35 Cramer'sche Regel: Sei A eine n-quadratische invertierbare Matrix mit den **Spalten**vektoren $s^1, \dots, s^n$. Dann hat das lineare Gleichungssystem

$$A \cdot x = b$$

für jedes $b \in \mathbb{K}^n$ genau eine Lösung $x = A^{-1} \cdot b$. Für die i-te Komponente x_i von x gilt

$$x_i = \frac{D(s^1, \dots, s^{i-1}, b, s^{i+1}, \dots, s^n)}{D(A)}$$

Beweis: $x = A^{-1} \cdot b = \frac{1}{D(A)} \cdot (A^* \cdot b)$. Ist c_i die i-te Komponente von A^*b, gilt

$$\begin{aligned} c_i &= \sum_{j=1}^{n} a^*_{ij} \cdot b_j = \sum_{j=1}^{n} b_j \cdot D(s^1, \dots, s^{i-1}, e^j, s^{i+1}, \dots, s^n) \\ &= D(s^1, \dots, s^{i-1}, \sum_{j=1}^{n} b_j e^j, s^{i+1}, \dots, s^n) = D(s^1, \dots, s^{i-1}, b, s^{i+1}, \dots, s^n). \end{aligned}$$

□

Wir wollen auch noch zwei Beispiele exemplarisch behandeln. Zur Berechnung von Determinanten stehen uns die Entwicklungssätze (11.18) und (11.28), die Zeilenmanipulationen (11.5 (D3)) und (11.25) zur Verfügung, und wegen (11.27) auch die entsprechenden **Spaltenmanipulationen**.

11.36 Beispiel:

$$\begin{aligned} D\begin{pmatrix} 1 & -4 & 1 & 6 \\ 2 & 0 & 3 & 3 \\ 0 & 2 & 1 & 5 \\ 1 & 3 & 2 & 2 \end{pmatrix} &\overset{(1)}{=} D\begin{pmatrix} 1 & -4 & 1 & 6 \\ 0 & 8 & 1 & -9 \\ 0 & 2 & 1 & 5 \\ 0 & 7 & 1 & -4 \end{pmatrix} \overset{(2)}{=} D\begin{pmatrix} 8 & 1 & -9 \\ 2 & 1 & 5 \\ 7 & 1 & -4 \end{pmatrix} \\ &\overset{(3)}{=} D\begin{pmatrix} 8 & 1 & -9 \\ -6 & 0 & 14 \\ -1 & 0 & 5 \end{pmatrix} \overset{(4)}{=} -D\begin{pmatrix} -6 & 14 \\ -1 & 5 \end{pmatrix} \overset{(4)}{=} -(-6 \cdot 5 + 14) = 16 \end{aligned}$$

In (1) subtrahieren wir Vielfache der ersten Zeile von den übrigen, in (2) entwickeln wir nach der erste Spalte, in (3) subtrahieren wir die erste Zeile von den übrigen, in (4) entwickeln wir nach der 2. Spalte, in (5) benutzen wir (11.22).

11.37 Die Vandermondeschen Determinante: Für $x_1,\dots,x_n$ mit $n\geq 2$ sei

$$V(x_1,\dots,x_n)=D\begin{pmatrix}1 & x_1 & x_1^2 & \dots & x_1^{n-1}\\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1}\\ \vdots & \vdots & & \vdots & \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1}\end{pmatrix}$$

Behauptung: $V(x_1,\dots,x_n)=\prod\limits_{1\leq i<j\leq n}(x_j-x_i)$:= Produkt aller Faktoren (x_j-x_i) mit $1\leq i<j\leq n$

Beweis: durch Induktion nach n: Für $n=2$ ist $D\begin{pmatrix}1 & x_1\\ 1 & x_2\end{pmatrix}=x_2-x_1$ nach (11.22).

Induktionsschritt: Beginnend mit der letzten Spalte subtrahieren wir nacheinander das x_1-fache der $(j-1)$-ten Spalte von der j-ten Spalte:

$$V(x_1,\dots,x_n)=D\begin{pmatrix}1 & 0 & 0 & \dots & 0\\ 1 & x_2-x_1 & x_2\cdot(x_2-x_1) & \dots & x_2^{n-2}(x_2-x_1)\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_n-x_1 & x_n\cdot(x_n-x_1) & \dots & x_n^{n-2}(x_n-x_1)\end{pmatrix}$$

$$\overset{(1)}{=}D\begin{pmatrix}x_2-x_1 & x_2\cdot(x_2-x_1) & \dots & x_2^{n-2}(x_2-x_1)\\ \vdots & \vdots & & \vdots\\ x_n-x_1 & x_n\cdot(x_n-x_1) & \dots & x_n^{n-2}(x_n-x_1)\end{pmatrix}$$

$$\overset{(2)}{=}(x_2-x_1)\cdot\ldots\cdot(x_n-x_1)\cdot V(x_2,x_3,\dots,x_n)$$

(1) Entwicklung nach de ersten Zeile

(2) Herausziehen der Faktoren $(x_2-x_1),\dots,(x_n-x_1)$ aus den Zeilen.

Nach Induktionsvoraussetzung folgt die Behauptung. $\square$

Teil III

Lineare Endomorphismen

In diesem Kapitel untersuchen wir Endomorphismen $f: V \to V$ endlich-dimensionaler Vektorräume V. Wir wollen dabei f in möglichst einfacher Form beschreiben, indem wir eine geeignete Basis auswählen. D.h. wir suchen eine Basis $\mathscr{B}$ von V, so dass die Matrix $M^{\mathscr{B}}_{\mathscr{B}}(f)$ eine übersichtliche Form erhält. Aus (7.11) wissen wir, dass es solche Formen gibt, wenn man die Basen im Definitions- und im Wertebereich **unabhängig** voneinander wählen darf. Jetzt soll die Basis im Definitionsbereich **dieselbe** wie im Wertebereich sein.

12 Eigenwerte und Eigenvektoren

Sei $\mathbb{K}$ ein Körper, V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und

$$f: V \to V$$

eine $\mathbb{K}$-lineare Abbildung.

12.1 Definition: Ein Untervektorraum $U \subset V$ heißt f-**INVARIANT**, wenn

$$f(U) \subset U$$

12.2 Beispiel: (1) Da $f(0) = 0$, sind die Untervektorräume $\{0\}$ und V stets f-invariant.

(2) Sei $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Dann besitzt $f_A: \mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto A \cdot x$ außer $\{0\}$ und $\mathbb{R}^2$ keine f_A-invarianten Untervektorräume. Sei etwa $U \neq \{0\}$ ein f_A-invarianter Untervektorraum und $x = \binom{x_1}{x_2} \in U$, so gilt $f_A(x) = \binom{x_2}{-x_1}$. Ist $x \neq 0$, so sind x und $f_A(x)$ linear unabhängig, denn

$$\det \begin{pmatrix} x_1 & x_2 \\ x_2 & -x_1 \end{pmatrix} = -x_1^2 - x_2^2 \neq 0.$$

Da $f_A(x) \in U$ ist, ist $\dim U = 2$, also $U = \mathbb{R}^2$.

(3) Sei A wie in (2), wir betrachten aber f_A als Abbildung

$$f_A: \mathbb{C}^2 \to \mathbb{C}^2$$

Hier ist die Situation wesentlich anders:

$$U_1 = \mathbb{C} \cdot \begin{pmatrix} 1 \\ i \end{pmatrix} \text{ und } U_2 = \mathbb{C} \cdot \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

sind f_A-invariante, komplexe 1-dimensionale Untervektorräume von $\mathbb{C}^2$:

$$\begin{aligned} A \cdot \binom{1}{i} &= \binom{i}{-1} = i \cdot \binom{1}{i} = f_A\binom{1}{i} \in U_1 \\ A \cdot \binom{1}{-i} &= \binom{-i}{-1} = -i \cdot \binom{1}{-i} = f_A\binom{1}{-i} \in U_2 \end{aligned}$$

Da $\binom{1}{i}$ und $\binom{1}{-i}$ linear unabhängig sind, folgt

$$\mathbb{C}^2 = U_1 \oplus U_2.$$

Nehmen wir $\mathscr{B} = \{\binom{1}{i}, \binom{1}{-i}\}$ als Basis von $\mathbb{C}^2$, erhalten wir aus unseren Berechnungen

$$M^{\mathscr{B}}_{\mathscr{B}}(f_A) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Bzgl. dieses “Koordinatensystems” ist f_A besonders einfach zu beschreiben: Ist $\binom{x_1}{x_2}$ der Koordinatenvektor von $v \in \mathbb{C}^2$ bzgl. $\mathscr{B}$, so erhält man den Koordinatenvektor $\binom{y_1}{y_2}$ von $f_A(v)$ bzgl. $\mathscr{B}$ durch einfache Multiplikation jeder Komponenten:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = M^{\mathscr{B}}_{\mathscr{B}}(f_A) \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} i \cdot x_1 \\ -i \cdot x_2 \end{pmatrix}.$$

Beispiel (3) zeigt, dass f-invariante Unterräume entscheidend zum globalen Verständnis der Abbildung f beitragen können. Das folgende Ergebnis zeigt, dass das auch im allgemeinen Fall so ist:

Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum $f : V \to V$ ein Endomorphismus und $U \subset V$ ein f-invarianter Unterraum. Dann können wir f zu einem Endomorphismus

$$\hat{f} : U \to U, \quad u \mapsto f(u)$$

einschränken. Wir wählen eine Basis $\hat{\mathscr{B}} = \{v^1, \ldots, v^k\}$ von U und ergänzen diese zu einer Basis $\mathscr{B} = \{v^1, \ldots, v^k, v^{k+1}, \ldots, v^n\}$ von V.

12.3 Satz: Mit diesen Beziehungen gilt: $M^{\mathscr{B}}_{\mathscr{B}}(f_A)$ ist eine **BLOCKMATRIX** der Form

$$M^{\mathscr{B}}_{\mathscr{B}}(f) = \begin{pmatrix} \hat{A} & B \\ 0 & \overline{A} \end{pmatrix} \text{ mit } \hat{A} = M^{\hat{\mathscr{B}}}_{\hat{\mathscr{B}}}(\hat{f}),$$

$\overline{A} \in \mathbb{K}^{n-k,n-k}$, $B \in \mathbb{K}^{k,n-k}$.

Beweis: Der j-te Spaltenvektor von $M_{\mathscr{B}}^{\mathscr{B}'}(f)$ ist der Koordinatenvektor von $f(v^j)$ bzgl. $\mathscr{B}$. Nun gilt $1 \leq j \leq k$, dass $f(v^j) = \hat{f}(v^j) \in U$, also

$$f(v^j) = a_{1j}v^1 + \ldots + a_{kj}v^k + 0 \cdot v^{k+1} + \ldots + 0 \cdot v^n$$

und $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{kj} \end{pmatrix}$ ist der Koordinatenvektor von $\hat{f}(v^j)$ bzgl. $\hat{\mathscr{B}}$. $\square$

In der betrachteten Situation gilt

$$V = U \oplus Span\{v^{k+1}, \ldots, v^n\} = U \oplus \overline{U}$$

mit $\overline{U} = Span\{v^{k+1}, \ldots, v^n\}$.

12.4 Definition: Ist U Untervektorraum eines $\mathbb{K}$-Vektorraumes V, so nennt man einen Untervektorraum U' von V **KOMPLEMENTÄRRAUM** von U, falls $V = U \oplus U'$.

12.5 Warnung: Zu einem Untervektorraum $U \subset V$ gibt es i.a. viele Komplementärräume. Betrachte z.B. $U = \mathbb{R} \cdot e^1 \subset \mathbb{R}^2$, dann ist jeder Raum $U' = \mathbb{R} \cdot \binom{x}{y}$ mit $y \neq 0$ ein Komplementärraum von U.

Wir könnten nun das Glück haben, das unser f-invariantes U einen f-invarianten Komplementärraum $\overline{U}$ besitzt. Wir wählen eine Basis $\overline{\mathscr{B}} = \{v^{k+1}, \ldots, v^n\}$ von $\overline{U}$. Dann ist nach Def. 12.4 $\mathscr{B} = \{v^1, \ldots, v^n\}$ eine Basis von $\mathscr{B}$. Der Endomorphismus f schränkt zu einem Endomorphismus

$$\overline{f} : \overline{U} \to \overline{U}, \quad \overline{u} \mapsto f(\overline{u})$$

ein. Für $j > k$ gilt dann $f(v^j) = \overline{f}(v^j) \in \overline{U}$, also

$$f(v^j) = 0 \cdot v^1 + \ldots + 0 \cdot v^k + a_{k+1,j}v^{k+1} + \ldots + a_{n,j}v^n.$$

Wir erhalten zusammen mit (12.3):

12.6 Satz: Ist U ein f-invarianter Teilraum von V und $\overline{U}$ ein f-invarianter Komplementärraum von U, dann gilt bzgl. der getroffenen Basiswahlen

$$M_{\mathscr{B}}^{\mathscr{B}}(f) = \begin{pmatrix} \hat{A} & 0 \\ 0 & \overline{A} \end{pmatrix} \text{ mit } \hat{A} = M_{\hat{\mathscr{B}}}^{\hat{\mathscr{B}}}(\hat{f}) \text{ und } \overline{A} = M_{\overline{\mathscr{B}}}^{\overline{\mathscr{B}}}(\overline{f})$$

$\square$

Wie wir in (12.2.3) gesehen haben, können wir einen Endomorphismus f dann am einfachsten beschreiben, wenn wir eine Basis von V finden, bzgl. der f durch eine **DIAGONALMATRIX** $A = (a_{ij})$ beschrieben wird, d.h. $a_{ij} = 0$ für $i \neq j$.

12.7 Definition: Ein Endomorphismus $f : V \to V$ heißt **DIAGONALISIERBAR**, wenn V eine Basis $\mathscr{B}$ besitzt, so dass $M^{\mathscr{B}}_{\mathscr{B}}(f)$ eine Diagonalmatrix ist. Eine Matrix $A \in \mathbb{K}^{n,n}$ heißt **DIAGONALISIERBAR**, wenn $f_A : \mathbb{K}^n \to \mathbb{K}^n$ diagonalisierbar ist.

12.8 Satz: (1) Für einen Endomorphismus $f : V \to V$ sind äquivalent:

(i) f ist diagonalisierbar

(ii) $V = U_1 \oplus \ldots \oplus U_n$, wobei $\dim U_i = 1$ und U_i f-invariant ist für $i = 1, \ldots, n$.

(2) $A \in \mathbb{K}^{n,n}$ ist diagonalisierbar $\iff$ es gibt eine invertierbare Matrix $T \in \mathbb{K}^{n,n}$ und eine Diagonalmatrix D, so dass $T \cdot A \cdot T^{-1} = D$

Beweis 1: (i) $\Rightarrow$ (ii): Sei $\mathscr{B} = \{v^1, \ldots, v^n\}$ Basis von V, so dass

$$M^{\mathscr{B}}_{\mathscr{B}}(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Der j-te Spaltenvektor $\begin{pmatrix} 0 \\ \vdots \\ \lambda_j \\ 0 \\ 0 \end{pmatrix}$ von $M^{\mathscr{B}}_{\mathscr{B}}(f)$ ist der Koordinatenvektor von $f(v^j)$ bzgl. $\mathscr{B}$. Es folgt

$$f(v^j) = \lambda_j v^j \in \mathbb{K} \cdot v^j.$$

Also ist $\mathbb{K} \cdot v^j$ ein 1-eindimensionaler f-invarianter Teilraum und

$$V = \mathbb{K} \cdot v^1 \oplus \ldots \oplus \mathbb{K} \cdot v^n$$

(ii) $\Rightarrow$ (i): Sei $v^j \neq 0$ ein beliebiger Vektor aus U_j. Dann ist $\mathscr{B} = \{v^1, \ldots, v^n\}$ eine Basis von V, da

$$V = U_1 \oplus \ldots \oplus U_n = \mathbb{K} \cdot v^1 \oplus \ldots \oplus \mathbb{K} \cdot v^n$$

Da $\{v^j\}$ Basis von U_j ist und U_j f-invariant ist, gibt es ein $\lambda_j \in \mathbb{K}$ mit $f(v^j) = \lambda_j \cdot v^j$. Damit folgt

$$M^{\mathscr{B}}_{\mathscr{B}}(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

$\square$

Beweis 2: Sei $\mathscr{S}$ die Standardbasis von $\mathbb{K}^n$. Wir erinnern: $M^{\mathscr{S}}_{\mathscr{S}}(f_A) = A$. Es gilt: A diagonalisierbar $\iff$ f_A diagonalisierbar $\iff$ $\exists$ Basis $\mathscr{B}$ von $\mathbb{K}^n$, so dass $M^{\mathscr{B}}_{\mathscr{B}}(f_A) = D$ Diagonalmatrix ist.
Wir betrachten nun $\mathbb{K}^n$ mit folgenden Basen

$$f_A : \begin{array}{ccccccc} \mathbb{K}^n & \xrightarrow{id} & \mathbb{K}^n & \xrightarrow{f_A} & \mathbb{K}^n & \xrightarrow{id} & \mathbb{K}^n \\ \mathscr{B} & & \mathscr{S} & & \mathscr{S} & & \mathscr{B} \end{array}$$

Aus (7.8) folgt: $D = M^{\mathscr{B}}_{\mathscr{B}}(f_A) = M^{\mathscr{B}}_{\mathscr{S}}(id) \cdot M^{\mathscr{S}}_{\mathscr{S}}(f_A) \cdot M^{\mathscr{S}}_{\mathscr{B}}$.
Wir setzen $T := M^{\mathscr{B}}_{\mathscr{S}}(id)$. Da id ein linearer Isomorphismus ist und

$$M^{\mathscr{S}}_{\mathscr{B}}(id) \cdot M^{\mathscr{B}}_{\mathscr{S}}(id) = M^{\mathscr{S}}_{\mathscr{S}}(id) = E_n,$$

ist $M^{\mathscr{S}}_{\mathscr{B}}(id) = T^{-1}$. Also $D = T \cdot A \cdot T^{-1}$. $\square$

12.9 Beispiel: Aus den Beispielen (12.2.2) und (12.2.3) folgt mit ((12.8.1):

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \text{ ist über } \mathbb{C} \text{ diagonalisierbar, aber nicht über } \mathbb{R}.$$

Wir wollen die Matrix $T \in \mathbb{C}^{2,2}$ bestimmen, für die $T \cdot A \cdot T^{-1}$ von Diagonalgestalt ist. Wie wir gezeigt haben, hat $M^{\mathscr{B}}_{\mathscr{B}}(f_A)$ für $\mathscr{B} = \{\binom{1}{i}, \binom{1}{-i}\}$ Diagonalgestalt $D = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. Aus den Gleichungen

$$D = M^{\mathscr{B}}_{\mathscr{B}}(f_A) = M^{\mathscr{B}}_{\mathscr{S}}(id) \circ M^{\mathscr{S}}_{\mathscr{S}}(f_A) \circ M^{\mathscr{S}}_{\mathscr{B}}(id) \quad \text{und} \quad M^{\mathscr{B}}_{\mathscr{S}}(id) = (M^{\mathscr{S}}_{\mathscr{B}}(id))^{-1}$$

eralten wir $T = (M^{\mathscr{S}}_{\mathscr{B}}(id))^{-1}$. Die gesuchte Matrix ist also invers zu

$$M^{\mathscr{S}}_{\mathscr{B}}(id) = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}.$$

Man sieht aber auch sofort, dass

$$e^1 = \frac{1}{2}\begin{pmatrix} 1 \\ i \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ -i \end{pmatrix}, \quad e^2 = -\frac{i}{2}\begin{pmatrix} 1 \\ i \end{pmatrix} + \frac{i}{2}\begin{pmatrix} 1 \\ -i \end{pmatrix},$$

so dass

$$T = \begin{pmatrix} \frac{1}{2} & -\frac{i}{2} \\ \frac{1}{2} & \frac{i}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

$\square$

Um einen Endomorphismus zu diagonalisieren, müssen wir also f-invariante 1-dimensionale Unterräume finden. Dies ist unser nächstes Ziel:

12.10 Definition: Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$

(1) $\lambda \in \mathbb{K}$ heißt **EIGENWERT** von f, wenn es einen Vektor $v \neq 0$ in V gibt, so dass $f(v) = \lambda \cdot v$,

(2) ein solches v heißt **EIGENVEKTOR** von f zum Eigenwert λ.

(3) Ist $\lambda \in \mathbb{K}$ ein Eigenwert, so heißt

$$E(\lambda) = \mathrm{Kern}(f - \lambda \cdot id) = \mathrm{Kern}(\lambda \cdot id - f)$$

EIGENRAUM von f zum Eigenwert λ.

12.11 Erste Beobachtungen:

(1) Ist v eine Eigenvektor von f (zum Eigenwert λ), dann ist $\mathbb{K} \cdot v$ ein 1-dimensonaler f-invarianter Teilraum von V.

Denn ist $a \cdot v \in \mathbb{K} \cdot v$, gilt $f(a \cdot v) = a \cdot f(v) = (a \cdot \lambda) \cdot v \in \mathbb{K} \cdot v$.

(2) $E(\lambda)$ enthält außer dem Nullvektor genau die Eigenvektoren zum Eigenwert λ.

Beweis: $v \in E(\lambda) \iff (f - \lambda \cdot id)(v) = 0 \iff f(v) = \lambda \cdot v$.

(3) $E(\lambda)$ ist f-invarianter Teilraum von V (folgt aus (1) und (2)). Da es ein $v \neq 0$ in $E(\lambda)$ gibt, ist $\dim E(\lambda) \geq 1$

(4) Aus (2) folgt: Die Einschränkung von f auf $E(\lambda)$ ist gegeben durch

$$\lambda \cdot id : E(\lambda) \to E(\lambda), \quad v \mapsto f(v) = \lambda \cdot v$$

(5) Aus (12.8.1) folgt: f ist diagonalisierbar $\iff$ V besitzt eine Basis aus Eigenvektoren von f.

12.12 Beispiel: $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ aus (12.2.2) besitzt keinen Eigenwert und keinen Eigenvektor, da es keinen 1-dimensionalen invarianten Teilraum hat.

Im Beispiel (12.2.3) besitzt $f_A : \mathbb{C}^2 \to \mathbb{C}^2$ die Eigenwerte i und $-i$. Die Eigenräume $E(i)$ und $E(-i)$ sind beide 1-dimensional. Der Vektor $\binom{1}{i}$ ist Eigenvektor zum Eigenwert i und $\binom{1}{-i}$ ist Eigenvektor zum Eigenwert $-i$.

12.13 Satz: Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$. Seien $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$ paarweise verschiedene Eigenwerte von f. Dann gilt

(1) Ist v^i Eigenvektor zu λ_i, $i = 1, \ldots, m$, dann sind $v^1, \ldots, v^m$ linear unabhängig.

(2) Die Summe der Eigenräume $E(\lambda_1), \ldots, E(\lambda_m)$ ist direkt, d.h.

$$E(\lambda_1) \oplus E(\lambda_2) \oplus \ldots \oplus E(\lambda_m) \subset V$$

Es folgt: $\Sigma_{i=1}^{n} \dim E(\lambda_i) \leq \dim V = n$.

Beweis 1: Induktion nach m. Da $v^1 \neq 0$, ist v^1 linear unabhängig.
Induktionsschritt: Vor: $v^1, \ldots, v^{m-1}$, $m \geq 2$, sind linear unabhängig.
Beh.: $v^1, \ldots, v^{m-1}, v^m$ sind linear unabhängig.
Sei $a_1 \cdot v^1 + a_2 \cdot v^2 + \ldots + a_m \cdot v^m = 0$. Es folgt

$$\begin{aligned} \lambda_m \cdot a_1 \cdot v^1 + \ldots + \lambda_m \cdot a_m \cdot v^m &= \lambda_m \cdot 0 = 0 \\ = f(a_1 \cdot v^1 + \ldots + a_m \cdot v^m) &= \lambda_1 \cdot a_1 \cdot v^1 + \ldots + \lambda_m \cdot a_m \cdot v^m \end{aligned}$$

Also: $a_1 \cdot (\lambda_m - \lambda_1) v^1 + \ldots + a_{m-1}(\lambda_m - \lambda_{m-a}) \cdot v^{m-1} = 0$.
Da $v^1, \ldots v^{m-1}$ linear unabhängig sind, folgt

$$a_i \cdot (\lambda_m - \lambda_i) = 0 \quad \text{für } i = 1, \ldots, m-1.$$

Da $\lambda_m - \lambda_i \neq 0$ für $i = 1, \ldots, m-1$ nach Voraussetzung, folgt $a_i = 0$ für $i = 1, \ldots, m-1$. Also gilt $a_m \cdot v^m = 0$. Da $v^m \neq 0$, folgt auch $a_m = 0$.
Beweis 2: Indem wir entsprechend umnummerieren, genügt es zu zeigen, dass

$$(E(\lambda_1) + \ldots + E(\lambda_{m-1})) \cap E(\lambda_m) = \{0\}$$

Sei v aus diesem Schnitt, dann hat v eine Darstellung

$$v = a_1 \cdot w^1 + \ldots + a_{m-i} \cdot w^{m-1} = a_m \cdot w^m$$

mit $w^i \in E(\lambda_i)$, $w^i \neq 0$ und $a_i \in \mathbb{K}$. Es folgt

$$0 = a_1 w^1 + \ldots + a_{m-1} w^{m-1} - a_m w^m.$$

Nach Teil (1) sind $w^1, \ldots, w^m$ linear unabhängig. Also ist $a_m = 0$ und damit $v = 0$.
$\square$

Unser nächstes Problem ist es, Eigenwerte und ihre zugehörigen Eigenräume zu finden. $\lambda \in \mathbb{K}$ ist ein Eigenwert, wenn die Abbildung

$$\lambda \cdot id - f : V \to V$$

einen nicht-trivialen Kern hat. Da V endlich-dimensional ist, ist das gleichbedeutend damit, dass $\lambda \cdot id - f$ nicht invertierbar ist. Dafür genügt es nachzuweisen, dass

$$M_{\mathscr{B}}^{\mathscr{B}}(\lambda \cdot id - f) = \lambda \cdot M_{\mathscr{B}}^{\mathscr{B}}(id) - M_{\mathscr{B}}^{\mathscr{B}}(f) = \lambda \cdot E_n - M_{\mathscr{B}}^{\mathscr{B}}(f)$$

nicht-invertierbar ist, wobei $\mathscr{B}$ eine beliebige Basis von V ist (vergl. §7). Die Nicht-Invertierbarkeit von Matrizen kann man aber mit Hilfe ihrer Determinante testen (11.30).

Sei also $A = (a_{ij}) \in \mathbb{K}^{n,n}$. Wir betrachten

$$\operatorname{Det}(\lambda \cdot E_n - A).$$

12.14 Lemma: $P_A := \operatorname{Det}(X \cdot E_n - A)$ ist ein Polynom von Grad n in der Unbestimmten X mit Koeffizienten in $\mathbb{K}$ von der Form

$$P_A = X^n + c_{n-1}X^{n-1} + \ldots + c_1 \cdot X + c_0 \quad c_i \in \mathbb{K}$$

mit $c_{n-1} = -a_{11} - a_{22} - \ldots - a_{nn}$ und $c_0 = (-1)^n \cdot \operatorname{Det} A$.

Beweis: Nach (11.15) gilt für $X \cdot E_n - A = B = (b_{ij})$

$$\operatorname{Det} B = \sum_{\sigma \in \Sigma_n} sign(\sigma) \cdot b_{1\sigma(1)} \cdot \ldots \cdot b_{n\sigma(n)} = b_{11} \cdot b_{22} \cdot \ldots \cdot b_{nn} + C$$

mit $C = \sum\limits_{\substack{\sigma \in \Sigma_n \\ \sigma \neq id}} b_{1\sigma(1)} \cdot \ldots \cdot b_{n\sigma(n)}$. Für $\sigma \neq id$ gilt $\sigma(i) \neq i$ für mindestens zwei $i \in \{1, \ldots, n\}$. Damit enthält jeder Summand von C mindestens zwei Faktoren b_{ij} mit $i \neq j$. Für diese gilt $b_{ij} = -a_{ij}$. Damit können höchstens $n-2$ Faktoren von der Form $b_{ii} = X - a_{ii}$ sein, d.h. C ist ein Polynom in X in Grad $\leq n-2$.

$$b_{11} \ldots b_{nn} = (X - a_{11}) \cdot (X - a_{22}) \cdot \ldots \cdot (X - a_{nn})$$

Multiplizieren wir aus erhalten wir

$$\operatorname{Det} B = X^n - (a_{11} + a_{22} + \ldots + a_{nn}) \cdot X^{n-1} + d_{n-2}X^{n-2} + \ldots + d_1 \cdot X + d_0 + C.$$

Es folgt: P_A ist eine Polynom n-ten Grades in X. Der Koeffizient von X^{n-1} ist $-(a_{11} + a_{22} + \ldots + a_{nn})$.

Setzen wir $X = 0$, erhalten wir

$$P_A(0) = c_0 = \operatorname{Det}(-A) = (-1)^n \cdot \operatorname{Det} A$$

$\square$

12.15 Definition: (1) Für $A \in \mathbb{K}^{n,n}$ heißt $P_A = \mathrm{Det}(X \cdot E_n - A)$ das **CHARAKTERISTISCHE POLYNOM** von A in der Unbestimmtem X

(2) Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ und $\mathscr{B}$ eine Basis von V, dann heißt

$$P_f := P_A \quad \text{mit} \quad A = M^{\mathscr{B}}_{\mathscr{B}}(f) = (a_{ij})$$

das **CHARAKTERISTISCHE POLYNOM** von f,

$$\mathrm{Det}(f) := \mathrm{Det}\, M^{\mathscr{B}}_{\mathscr{B}}(f)$$

die **DETERMINANTE** von f und

$$Tr(f) := a_{11} + a_{22} + \ldots + a_{nn} = Tr(M^{\mathscr{B}}_{\mathscr{B}}(f))$$

die **SPUR** von f bzw. von A.

Diese Definition ist sinnvoll, denn

12.16 Lemma: $P_f, \mathrm{Det}(f)$ und $Tr(f)$ hängen nicht von der Wahl von $\mathscr{B}$ ab.

Beweis: Seien $\mathscr{B}$ und $\overline{\mathscr{B}}$ zwei Basen von V. Dann gilt $M^{\overline{\mathscr{B}}}_{\overline{\mathscr{B}}}(f) = M^{\overline{\mathscr{B}}}_{\mathscr{B}}(id) \cdot M^{\mathscr{B}}_{\mathscr{B}}(f) \cdot M^{\mathscr{B}}_{\overline{\mathscr{B}}}(id)$. Wir setzen $A = M^{\mathscr{B}}_{\mathscr{B}}(f)$ und $T = M^{\overline{\mathscr{B}}}_{\mathscr{B}}(id)$. Dann ist $M^{\mathscr{B}}_{\overline{\mathscr{B}}}(id) = T^{-1}$ (siehe Beweis von 12.8). Für $B = M^{\overline{\mathscr{B}}}_{\overline{\mathscr{B}}}(f)$ gilt also $B = T \cdot A \cdot T^{-1}$. Es folgt

$$\begin{aligned}
\mathrm{Det}(X \cdot E_n - B) &= \mathrm{Det}(T \cdot (X \cdot E_n) \cdot T^{-1} - T \cdot A \cdot T^{-1}) \\
&= \mathrm{Det}(T \cdot (X \cdot E_n - A) \cdot T^{-1}) = \mathrm{Det}(T) \cdot \mathrm{Det}(X \cdot E_n - A) \cdot (\mathrm{Det}\, T)^{-1} \\
&= \mathrm{Det}(X \cdot E_n - A)
\end{aligned}$$

Das zeigt, dass P_f unabhängig von der Basiswahl ist. Da $-Tr(A)$ der Koeffizient von X^{n-1} in P_f und $(-1)^n \cdot \mathrm{Det}(A)$ der Koeffizient von X^0 in P_f ist, folgen die anderen Aussagen. $\square$

Nach unseren Überlegungen ist $\lambda \in \mathbb{K}$ genau dann Eigenwert von f, wenn $\lambda \cdot id - f$ nicht bijektiv ist, d.h. wenn $\mathrm{Det}(\lambda \cdot id - f) = 0$ ist. Also gilt

12.17 Satz: $\lambda \in \mathbb{K}$ ist genau dann Eigenwert von $f \in \mathrm{End}_{\mathbb{K}}(V)$, wenn λ Nullstelle des Polynoms P_f ist.

12.18 Beispiel: Sei $f : V \to V$ eine $\mathbb{R}$-lineare Abbildung von $\mathbb{R}$-Vektorräumen, beschrieben bzgl. einer Basis $\mathscr{B}$ von V durch die Matrix

$$M^{\mathscr{B}}_{\mathscr{B}}(f) = A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}$$

$$
\begin{aligned}
P_f &= \operatorname{Det}\begin{pmatrix} X & 1 & -1 \\ 3 & X+2 & -3 \\ 2 & 2 & X-3 \end{pmatrix} \\
&= X\cdot(X+2)(X-3)-6-6+2(X+2)-3(X-3)+6X \\
&= X^3-X^2-X+1=(X-1)^2\cdot(X+1)
\end{aligned}
$$

Also hat f die Eigenwerte $\lambda = \pm 1$.
Wir bestimmen die Eigenräume:
$E(\lambda)$ ist der Lösungsraum von $(\lambda \cdot E_3 - A)\cdot x = 0$.
$E(1)$: zu lösen ist das Gleichungssystem

$$
(E_3 - A)\cdot x = \begin{pmatrix} 1 & 1 & -1 \\ 3 & 3 & -3 \\ 2 & 2 & -2 \end{pmatrix}\cdot x = 0
$$

x_1	x_2	x_3	$b=0$	
1	1	-1		
3	3	-3		
2	2	-2		
1	1	-1		I
0	0	0		$\mathrm{II}-3\cdot\mathrm{I}$
0	0	0		$\mathrm{III}-2\cdot\mathrm{I}$

Damit hat $E(1)$ die Basis

$$
v^1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \qquad v^2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}
$$

$E(-1)$: zu lösen ist das Gleichungssystem

$$
(-E_3 - A)\cdot X = \begin{pmatrix} -1 & 1 & -1 \\ 3 & 1 & -3 \\ 2 & 2 & -4 \end{pmatrix}\cdot x = 0
$$

$$
\begin{array}{rrr|cl}
x_1 & x_2 & x_3 & b=0 & \\
\hline
-1 & 1 & -1 & & \\
3 & 1 & -3 & & \\
2 & 2 & -4 & & \\
\hline
1 & -1 & 1 & & -\mathrm{I} \\
0 & 4 & -6 & & \mathrm{II}+3\cdot\mathrm{I} \\
0 & 4 & -6 & & \mathrm{III}+2\cdot\mathrm{I} \\
\hline
1 & -1 & 1 & & \\
0 & 1 & -\frac{3}{2} & & \frac{1}{4}\cdot\mathrm{II} \\
0 & 0 & 0 & & \mathrm{III}-\mathrm{II} \\
\hline
1 & 0 & -\frac{1}{2} & & \mathrm{I}+\mathrm{II} \\
0 & 1 & -\frac{3}{2} & &
\end{array}
$$

Damit ist $w' = \begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \\ 1 \end{pmatrix}$ und damit auch $w = 2 \cdot w' = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$ Basis von $E(-1)$.

Also besitzt V eine Basis aus Eigenvektoren von f, nämlich $\mathscr{B}' = \{v^1, v^2, w\}$, und ist damit diagonalisierbar. Genauer gilt

$$M^{\mathscr{B}'}_{\mathscr{B}'}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$\square$

12.19 Beispiel: Wir betrachten dieselbe Matrix über den Körper $\mathbb{F}_2$. Dann ist sie von der Form

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

$$P_f = \mathrm{Det} \begin{pmatrix} X & 1 & 1 \\ 1 & X & 1 \\ 0 & 0 & X+1 \end{pmatrix} = X^2 \cdot (X+1) + (X+1) = (X+1)^3$$

(Das hätten wir auch direkt aus P_f von Beispiel 12.18 herleiten können).

Wir haben genau einen Eigenwert, nämlich $\lambda = 1$

$$E_3 - A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

hat Rang 1. Damit ist $\dim E(1) = 2 \neq \dim V = 3$. Also ist f über $\mathbb{F}_2$ nicht diagonalisierbar.

In beiden Beispielen ist das charakteristische Polynom ein Produkt linearer Polynome. Es stellt sich gleich heraus, dass das ein notwendiges Kriterium für die Diagonalisierbarkeit ist.

12.20 Definition: Sei $P = X^n + a_{n-1}X^{n-1} + \ldots + a_0$ ein Polynom über dem Körper $\mathbb{K}$, d.h. die Koeffizienten a_i liegen in $\mathbb{K}$. Wir sagen, P zerfällt in **LINEARFAKTOREN**, wenn

$$P = (X - \lambda_1) \cdot (X - \lambda_2) \cdot \ldots \cdot (X - \lambda_n)$$

mit $\lambda_1 \cdot \ldots \cdot \lambda_n \in \mathbb{K}$. Gilt

$$P = (X - \lambda_1)^{r_1} \cdot \ldots \cdot (X - \lambda_k)^{r_k},$$

wobei $\lambda_i \neq \lambda_j$ für $i \neq j$ und $r_i > 0$, nennt man r_i die **ORDNUNG** oder **VIELFACHHEIT** der Nullstelle λ_i von P und schreibt auch $r_i = \nu_{\lambda_i}(P)$.

12.21 Lemma: Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ und λ Eigenwert von f. Dann gilt

$$\dim E(\lambda) \leq \nu_\lambda(P_f)$$

Beweis: $E(\lambda)$ ist f-invariant. Sei $\{v^1, \ldots, v^k\}$ Basis von $E(\lambda)$ und $\{v^1, \ldots, v^n\}$ die Ergänzung dieser Basis zu einer Basis $\mathscr{B}$ von V. Da $f(v^j) = \lambda \cdot v^j$ für $1 \leq j \leq k$, hat $A = M^{\mathscr{B}}_{\mathscr{B}}(f)$ die Form

$$A = \begin{pmatrix} \lambda \cdot E_k & B \\ 0 & A' \end{pmatrix}$$

Nach dem Entwicklungssatz (11.19) gilt

$$P_f(x) = \mathrm{Det}(x \cdot E_n - A) = \mathrm{Det}\begin{pmatrix} (x-\lambda)E_k & B \\ 0 & x \cdot E_{n-k} - A' \end{pmatrix} = (x-\lambda)^k \cdot P_{A'}(x)$$

Damit ist $\nu_\lambda(P_f) \geq k = \dim E(\lambda)$. □

Wir fassen zusammen

12.22 Satz: $f \in \mathrm{End}_{\mathbb{K}}(V)$ ist genau dann diagonalisierbar, wenn folgende beiden Bedingungen erfüllt sind:

(1) P_f zerfällt in Linearfaktoren

(2) Für jeden Eigenwert λ von f gilt $\nu_\lambda(P_f) = \dim E(\lambda)$.

Beweis: Nach (12.11.5) ist f genau dann diagonalisierbar, wenn V eine Basis $\mathscr{B} = \{v^1, \ldots, v^n\}$ von Eigenvektoren besitzt. Sind $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von f und ordnen wir die v^i nach diesen Eigenwerten, d.h. zuerst die Basisvektoren zum Eigenwert λ_1, dann die zum Eigenwert λ_2 usw., so gilt

$$M^{\mathscr{B}}_{\mathscr{B}}(f) = \begin{array}{c} \left(\begin{array}{cccc} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_1 & & 0 \\ & & & \ddots & \\ & 0 & & & \lambda_k \\ & & & & & \ddots \\ & & & & & & \lambda_k \end{array} \right) \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\1\\1\end{array}}\right\} r_1 \\ \\ \left.\vphantom{\begin{array}{c}1\\1\\1\end{array}}\right\} r_k \end{array}$$

wobei r_i Basisvektoren Eigenvektoren zum Eigenwert λ_i sind. Insbesondere gilt $\dim E(\lambda_i) \geq r_i$ und

$$P_f = \mathrm{Det}(X \cdot E_n - M^{\mathscr{B}}_{\mathscr{B}}(f)) = (X - \lambda_1)^{r_1} \cdot \ldots \cdot (x - \lambda_k)^{r_k}.$$

Da $\dim E(\lambda_i) \leq \nu_{\lambda_i}(P_f) = r_i$, folgt $\dim E(\lambda_i) = \nu_{\lambda_i}(P_f)$.
Gelten umgekehrt (1) und (2), hat P_f die Form

$$P_f(x) = (x - \lambda_1)^{r_1} \cdot \ldots \cdot (x - \lambda_k)^{r_k}$$

mit verschiedenen $\lambda_1, \ldots, \lambda_k$, und es gilt $r_1 + r_2 + \ldots + r_k = n$. Wegen (2) ist $\dim E(\lambda_i) = r_i$. Nach (12.13.2) ist

$$E(\lambda_i) \oplus \cdots \oplus E(\lambda_k) \subset V$$

Da aber $\dim(E(\lambda_1) \oplus \cdots \oplus E(\lambda_k)) = r_1 + \ldots r_k = n$, ist

$$E(\lambda_i) \oplus \cdots \oplus E(\lambda_k) = V$$

Wählen wir nun für jedes $i = 1, \ldots, k$ eine Basis $\mathscr{B}_i$ von $E(\lambda_i)$, so ist $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2 \cup \ldots \cup \mathscr{B}_k$ eine Basis von Eigenvektoren von V und damit f diagonalisierbar.
□

12.23 Korollar: Besitzt $f \in \mathrm{End}_{\mathbb{K}}(V)$ n verschiedene Eigenwerte $(n = \dim V)$, so ist f diagonalisierbar. □

13 Polynomringe

Die Ergebnisse des Abschnitts 12 sind für uns Veranlassung, uns näher mit Polynomen zu befassen. In den Übungen haben wir Polynome als Funktionen $\mathbb{K} \to \mathbb{K}$ behandelt. In der Algebra sind aber Polynome formale Ausdrücke, die u.a. “Polynomfunktionen” definieren.

13.1 Beispiel: Die als formale Ausdrücke **verschiedenen** Polynome über $\mathbb{F}_3$

$$X^3 + X + [2] \qquad \text{und} \qquad [2] \cdot X + [2]$$

sind als “Funktionen” $\mathbb{F}_3 \to \mathbb{F}_3$ gleich. Als algebraische Objekte werden sie aber als verschieden angesehen.

13.2 Definition: Sei $\mathbb{K}$ ein Körper (die Definition gilt wörtlich auch für Ringe). Ein **POLYNOM** in der **UNBESTIMMTEN** X mit **KOEFFIZIENTEN** in $\mathbb{K}$ ist ein formaler Ausdruck

$$P = a_0 + a_1X + a_2X^2 + \ldots + a_nX^n = \sum_{i=0}^{n} a_iX^i$$

mit $a_0, \ldots, a_n \in \mathbb{K}$, $n \in \mathbb{N}$, $X^0 = 1 \in \mathbb{K}$, $X^1 := X$. Wir nennen $(a_0, a_1, \ldots, a_n, 0, 0, \ldots)$ die **KOEFFIZIENTENFOLGE** (sie besteht ab einer gewissen Stelle nur aus Nullen) $\operatorname{grad} P = \max\{k \in \mathbb{N}; a_k \neq 0\}$ heißt der **GRAD** von P, d.h. die Koeffizienten a_i von P sind für $i > \operatorname{grad} P$ stets 0.

Zwei Polynome P und Q heißen gleich, wenn ihre Koeffizientenfolgen gleich sind. Mit $\mathbb{K}[X]$ bezeichnen wir die Menge aller Polynome “über $\mathbb{K}$”.

13.3 Addition und Multiplikation: Seien $P, Q \in \mathbb{K}[X]$ und seien $(a_0, a_1, \ldots)$ und $(b_0, b_1, \ldots)$ die Koeffizientenfolgen von P bzw. Q, d.h.

$$P = \sum_{i=0}^{n} a_i \cdot X^i, \quad n = \operatorname{Grad} P, \quad Q = \sum_{i=0}^{n} b_i \cdot X^i, \quad m = \operatorname{Grad} Q.$$

Wir definieren:

$P + Q$ ist das Polynom mit Koeffizientenfolge

$$(a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots)$$

$P \cdot Q$ ist das Polynom mit der Koeffizientenfolge $(c_0, c_1, c_2, \ldots)$, wobei

$$c_0 = a_0 \cdot b_0, \quad c_1 = a_1 \cdot b_0 + a_0 \cdot b_1, \; \ldots, \; c_k = \sum_{i+j=k}^{n} a_i \cdot b_j$$

13.4 Satz: Sei $\mathbb{K}$ ein Körper

(1) $\mathbb{K}[X]$ ist ein kommutativer Ring und $\mathbb{K} \subset \mathbb{K}[X]$, die Teilmenge der Polynome mit Koeffizientenfolge $(a_0, 0, 0, \ldots)$, ein Unterkörper.

(2) $\mathbb{K}[X]$ ist bzgl. der Addition und der Multiplikation mit dem Unterkörper $\mathbb{K} \subset \mathbb{K}[X]$ ein $\mathbb{K}$-Vektorraum mit Basis $\{1, X, X^2, \ldots)$.

Diesen Satz haben wir im Rahmen der Übungen bereits bewiesen. Die Elemente aus $\mathbb{K} \subset \mathbb{K}[X]$ nennt man die **KONSTANTEN** Polynome.

13.5 Bezeichnung: Sei P ein Polynom von Grad n und $(a_0, a_1, \ldots, a_n, 0, 0, \ldots)$ seine Koeffizientenfolge, wobei $a_n \neq 0$ und $a_i = 0$ für $i > n$. Dann heißt a_n der **LEITKOEFFIZIENT** und a_0 der **KONSTANTE TERM** von P.

13.6 Sei $\mathbb{K}$ ein Körper. Für $P, Q \in \mathbb{K}[X] \backslash 0$ gilt (das Nullpolynom hat keinen Grad)

(1) $\operatorname{grad}(P+Q) \leq \max(\operatorname{grad} P, \operatorname{grad} Q)$

(2) $\operatorname{grad}(P \cdot Q) = \operatorname{grad} P + \operatorname{grad} Q$. Genauer gilt: Ist a_n Leitkoeffizient von P und b_m Leitkoeffizient von Q, so ist $a_n \cdot b_m$ Leitkoeffizient von $P \cdot Q$.

Beweis: trivial $\square$

13.7 Sei $P \in \mathbb{K}[X]$. Dann definiert P durch **Einsetzen** eine Abbildung

$$E_P : \mathbb{K} \to \mathbb{K}, \qquad E_P(r) = \sum_{i=0} a_i \cdot r^i,$$

falls $P = \sum\limits_{i=0}^{n} a_i X^i$. Wir schreiben kürzer $P(r)$ für $E_P(r)$.

Wichtig für unsere Anwendungen ist das folgende Ergebnis mit seinen Folgerungen.

13.8 Division mit Rest: Sei $\mathbb{K}$ ein Körper und seien $F = a_0 + a_n X + \ldots + a_n X^n$ und $G = b_0 + b_1 X + \ldots + b_m X^m$ mit $b_m \neq 0$ Polynome in $\mathbb{K}[X]$. Dann gibt es Polynome $Q, R \in \mathbb{K}[X]$, so dass

$$F = Q \cdot G + R \qquad \text{mit } R = 0 \text{ oder } \operatorname{grad} R < \operatorname{grad} G$$

Beweis: Ist $n < m$, setzen wir $Q = 0$ und $R = F$. Sei also $n \geq m$. Wir beweisen diesen Teil durch Induktion nach n.

$\mathbf{n=0}$: $F = a_0$, $G = b_0$ (wegen $n \geq m$). Da $b_0 \neq 0$ nach Voraussetzung, gilt mit $Q = \frac{a_0}{b_0}$: $F = a_0 = Q \cdot G + 0$.

Induktionsschritt: Sei $Q_1 = \frac{a_n}{b_m} \cdot X^{n-m}$

$$\begin{array}{rcl} F & = & a_n X^n + a_{n-1} X^{n-1} + \ldots + a_0 \\ Q_1 \cdot G & = & a_n X^n + \frac{a_n}{b_m} b_{m-1} X^{n-1} + \ldots + \frac{a_n}{b_m} b_0 \cdot X^{n-m} \end{array}$$

$F - Q_1 \cdot G = F_1$ ist ein Polynom von Grad $\operatorname{grad} F_1 < n$. Nach Induktionsvoraussetzung gilt es Polynome $Q_2, R \in \mathbb{K}[X]$, so dass

$$F_1 = Q_2 \cdot G + R \qquad \text{mit } \operatorname{grad} R < \operatorname{grad} G \text{ oder } R = 0$$

Es folgt:

$$F = Q_1 \cdot G + F_1 = (Q_1 + Q_2) \cdot G + R$$

Setze $Q = Q_1 + Q_2$. □

13.9 Definition: Sei $P \in \mathbb{K}[X]$. Ein Element $r \in \mathbb{K}$ heißt **NULLSTELLE** von P, wenn $P(r) = 0$.

13.10 Satz: Ist $\mathbb{K}$ ein Körper und $r \in \mathbb{K}$ eine Nullstelle von $P \in \mathbb{K}[X]$, dann ist das lineare Polynom $X - r$ ein **TEILER** von P. D.h. es gibt ein Polynom $Q \in \mathbb{K}[X]$, so dass $P = Q \cdot (X - r)$.

Beweis: Wir teilen P durch $X - r$ mit Rest:

$$P = Q(X - r) + R \qquad \text{mit } R = 0 \text{ oder } \operatorname{grad} R = 0$$

Im zweiten Fall ist R ein konstantes Polynom $R = c$ mit $c \neq 0$. Nach Voraussetzung gilt $P(r) = 0$. Also

$$0 = P(r) = Q(r) \cdot (r - r) + c = 0 + c$$

Also ist $c = 0$, d.h. $R = 0$. □

13.11 Folgerung: Ist $\mathbb{K}$ ein Körper und $P \in \mathbb{K}[X]$ ein Polynom von Grad n, so hat P höchstens n Nullstellen.

Beweis: Induktion nach n. Ist $n = 0$, so ist $P = a_0$ mit $a_0 \neq 0$. Also hat P keine Nullstelle.

Induktionsschritt: Sei r Nullstelle von P. Dann gibt es $Q \in \mathbb{K}[X]$ mit $P = Q \cdot (X - r)$. Ist $s \neq r$ eine weitere Nullstelle, gilt

$$0 = P(s) = Q(s) \cdot (s - r).$$

Da $s - r \neq 0$, ist s Nullstelle von Q. Da $\operatorname{grad} Q = \operatorname{grad} P - 1 = n - 1$, hat Q nach Induktion höchstens $n - 1$ verschiedene Nullstellen. $\square$

Besonders schön ist die Theorie der komplexen Polynome, denn es gilt

13.12 Fundamentalsatz der Algebra: Jedes komplexe Polynom $P = a_n X^n + \ldots + a_1 X + a_0$ vom Grad $n \geq 1$ zerfällt in Linearfaktoren, d.h.

$$P = a_n \cdot (X - \lambda_n) \cdot (X - \lambda_{n-1}) \cdot \ldots \cdot (X - \lambda_1) \qquad \text{mit } \lambda_i \in \mathbb{C} \text{ für } i = 1, \ldots, n.$$

Für dieses Resultat gibt es keinen rein algebraischen Beweis. Deshalb müssen wir im Rahmen dieser Vorlesung auf einen Beweis verzichten.

Der Fundamentalsatz der Algebra hat Konsequenzen für die Behandlung sowohl $\mathbb{C}$- als auch $\mathbb{R}$-linearer Abbildungen, wie wir in den kommenden Abschnitten sehen werden.

14 Trigonalisierbare Endomorphismen

Selbst wenn das charakteristische Polynom eines Endomorphismus' $f : V \to V$ in Linearfaktoren zerfällt, braucht F nicht diagonalisierbar zu sein. D.h. die Bedingung $\dim E(\lambda) = \nu_\lambda(P_f)$ ist nicht immer erfüllt.

14.1 Beispiel: $f : \mathbb{R}^2 \to \mathbb{R}^2$, $\binom{x_1}{x_2} \mapsto \binom{x_1+x_2}{x_2}$
Die Matrix von f bzgl. der Standardbasis ist

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

$P_f = (X-1)^2$
Damit zerfällt das charakteristische Polynom in Linearfaktoren, 1 ist der einzige Eigenwert. $\binom{x_1}{x_2}$ ist genau dann Eigenwert zum einzigen Eigenwert 1, falls

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_2 \end{pmatrix}$$

d.h. wenn $x_2 = 0$. Damit ist $E(1) = \mathbb{R} \times 0 \subset \mathbb{R}^2$, also 1-dimensional.

Wir fragen uns, was wir in einem solchen Fall erreichen können.

14.2 Definition: $f \in \mathrm{End}_{\mathbb{K}}(V)$ heißt **TRIGONALISIERBAR**, wenn V eine Basis $\mathscr{B}$ besitzt, so dass $M^{\mathscr{B}}_{\mathscr{B}}(f)$ eine obere Dreiecksmatrix ist.

14.3 Satz: $f \in \mathrm{End}_{\mathbb{K}}(V)$ ist genau dann trigonalisierbar, wenn das charakteristischen Polynom P_f von f in Linearfaktoren zerfällt.

Beweis: Sei f trigonalisierbar. Dann hat V eine Basis $\mathscr{B}$, so dass

$$A = M^{\mathscr{B}}_{\mathscr{B}}(f) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & & \vdots \\ & & \ddots & \\ & 0 & & a_{nn} \end{pmatrix}$$

Aus (11.26) folgt

$$P_f(x) = \mathrm{Det} \begin{pmatrix} x - a_{11} & -a_{12} & \cdots & -a_{1n} \\ 0 & x - a_{22} & & \vdots \\ & & \ddots & \\ & 0 & & x - a_{nn} \end{pmatrix} = (x - a_{11}) \cdot (x - a_{22}) \cdot \ldots \cdot (x - a_{nn}).$$

Also zerfällt P_f in Linearfaktoren.
Die Umkehrung beweisen wir durch Induktion nach $n = \dim V = \operatorname{grad} P_f$.
$n = 1$: $M^{\mathscr{B}}_{\mathscr{B}}(f) = (a_{11})$ ist obere Dreiecksmatrix. Also ist nichts zu zeigen.
<u>Induktionsschritt:</u> $P_f(x) = \prod_{i=1}^{n}(x - \lambda_i)$
Dann sind $\lambda_i, \ldots, \lambda_n$ die nicht notwendig verschiedenen Eigenwerte von f. Sei v^1 ein Eigenvektor zu Eigenwert λ_1 und

$$\mathscr{B} = \{v^1, w^2, \ldots, w^n\}$$

die Ergänzung von v^1 zu einer Basis von V. Dann gilt

$$A = M^{\mathscr{B}}_{\mathscr{B}}(f) = \begin{pmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & b_{32} & \cdots & b_{3n} \\ \vdots & \vdots & & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

Sei $W = \operatorname{Span}\{w^2, \ldots, w^n\}$ und $p : V \to W$ die lineare Abbildung, die durch $p(v^1) = 0$ und $p(w^i) = w^i$, $i = 2, \ldots, n$, gegeben ist. Insbesondere gilt $p(w) = w$ für alle w aus W. Sei

$$g : W \to W, \qquad g(w) = p \circ f(w).$$

Die Matrix von g bzgl. der Basis $\mathscr{B}' = \{w^2, \ldots, w^n\}$ von W ist

$$B = M^{\mathscr{B}'}_{\mathscr{B}'}(g) = \begin{pmatrix} b_{22} & \cdots & b_{2n} \\ \vdots & & \vdots \\ b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

Entwicklung nach der 1. Spalte ergibt

$$P_f(x) = \operatorname{Det}(x \cdot E_n - A) = (x - \lambda_1) \cdot \operatorname{Det}(x \cdot E_{n-1} - B) = (x - \lambda_1) \cdot P_g(x).$$

Also zerfällt $P_g(x) = \prod_{i=2}^{n}(x - \lambda_i)$ in Linearfaktoren. Nach Induktionsannahme hat W eine Basis $\overline{\mathscr{B}} = \{v^2, \ldots, v^n\}$, so dass $C = M^{\overline{\mathscr{B}}}_{\overline{\mathscr{B}}}(g)$ eine obere Dreiecksmatrix ist, d.h.

$$g(v^j) = \sum_{i=2}^{n} c_{ij} v^i \quad \text{mit } c_{ij} = 0 \text{ für } i > j,\ j = 2, \ldots, n.$$

$\hat{\mathscr{B}} = \{v^1, v^2, \ldots, v^n\}$ ist Basis von V. Ist $R = (r_{ij}) = M^{\hat{\mathscr{B}}}_{\hat{\mathscr{B}}}(f)$, dann gilt

$$f(v^1) = \lambda_1 \cdot v^1 = r_{11} \cdot v^1.$$

Also $r_{11} = \lambda_1$ und $r_{i1} = 0$ für $i > 1$. Für $j \geq 2$ erhalten wir

$$p \circ f(v^j) = p\left(\sum_{i=1}^{n} r_{ij} v^i\right) \overset{(1)}{=} p\left(\sum_{i=2}^{n} r_{ij} v^i\right) \overset{(2)}{=} \sum_{i=2}^{n} r_{ij} v^i = g(v^j) = \sum_{i=2}^{n} c_{ij} v^i.$$

Es gelten (1) und (2), weil $p(v^1) = 0$ und $p(w) = w$ für $w \in W$. Es folgt $r_{ij} = c_{ij}$ für $r \geq 2$ und damit $r_{ij} = 0$ für $i > j$. Also hat R obere Dreiecksform. $\square$

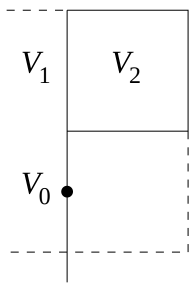
Mit Hilfe des Fundamentalsatzes der Algebra (13.12) erhalten wir als Folgerung

14.4 Satz: Ist V ein $\mathbb{C}$-Vektorraum, dann ist jeder $\mathbb{C}$-Endomorphismus $f : V \to V$ trigonalisierbar.

Wir schließen mit einer "geometrischen" Interpretation der Trigonalisierbarkeit.

14.5 Definition: Eine Sequenz $V_0 \subset V_1 \subset V_2 \subset \ldots \subset V_n$ von Untervektorräumen eines n-dimensionalen $\mathbb{K}$-Vektorraumes V heißt **FAHNE** in V, wenn $\dim V_i = i$. Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$, heißt eine Fahne f-**INVARIANT**, falls jedes V_i f-invariant ist.

Der Name "Fahne" ist suggestiv: $V_0 = \{0\}$ wird als Fußpunkt, $V_1 \cong \mathbb{K}$ wird als "Fahnenstange", $V_2 \cong \mathbb{K} \times \mathbb{K}$ als Fahnentuch interpretiert usw.



14.6 Fahnensatz: $f \in \mathrm{End}_{\mathbb{K}}(V)$ mit endlich-dimensionalem V ist genau dann trigonalisierbar, wenn V eine f-invariante Fahne besitzt.

Beweis: Ist f trigonalisierbar, dann hat V eine Basis $\mathscr{B} = \{v^1, \ldots, v^n\}$, so dass

$$A = M^{\mathscr{B}}_{\mathscr{B}}(f) \begin{pmatrix} a_{11} & a_{12} \cdots a_{1n} \\ 0 & a_{22} \cdots a_{2n} \\ & 0 \quad \ddots \quad \vdots \\ 0 & a_{nn} \end{pmatrix}$$

Sei $V_0 = \{0\}$ und $V_j = \{v^1, \dots, v^j\}$ für $j = 1, \dots, n$. Da

$$f(v^j) = \sum_{i=1}^{n} a_{ij} v^i = a_{1j} v^1 + \dots + a_{ij} v^j \in V_j \subset V_k \text{ für } j \leq k,$$

ist jedes V_k f-invariant. Da $\dim V_j = j$, erhalten wir eine Fahne.
Sei umgekehrt $V_0 \subset V_1 \subset \dots \subset V_n$ eine f-invariante Fahne. Beginnend mit 0 konstruieren wir durch Basisergänzung eine Basis $\{v^1, \dots, v^n\}$ von $V_n = V$, so dass $\{v^1, \dots, v^j\}$ Basis $\mathscr{B}$ von V_j ist. Da $f(v^j) \in V_j$, folgt

$$f(v^j) = \sum_{i=1}^{j} a_{ij} v^i.$$

Damit ist $M_{\mathscr{B}}^{\mathscr{B}}(f) = (a_{ij})$ eine obere Dreiecksmatrix. $\square$

Teil IV

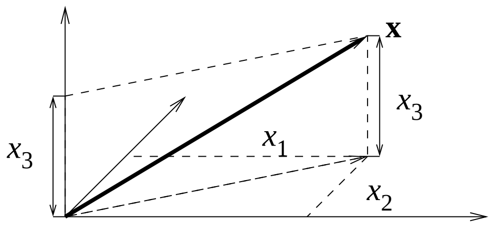
Vektorräume mit Skalarprodukt

Für geometrische Betrachtungen im $\mathbb{R}^n$ hätte man gerne so etwas wie Längen- und Winkelbegriffe. Die Volumenform in §11 ist eine Betrachtung in diese Richtung. Wir greifen die dort aufgelisteten Axiome auf, um den bekannten Längen- und Winkelbegriff des $\mathbb{R}^2$ und $\mathbb{R}^3$ auf $\mathbb{R}^n$ zu erweitern.

15 Skalarprodukte

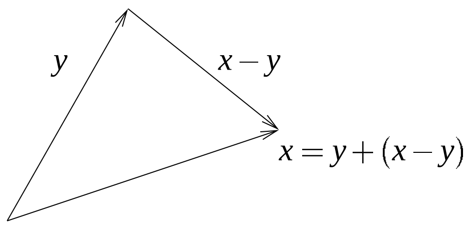
Sei $x = (x_1, x_2, x_3)$. Die “**Länge**” von x ist der “Abstand” von x vom Nullpunkt. Nach dem Satz von Pythagoras gilt

15.1

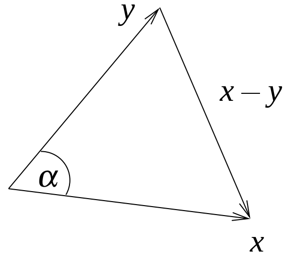


$$\begin{aligned} \| x \|: &= \text{Länge von } x \\ &= {}_+\sqrt{x_1^2 + x_2^2 + x_3^2} \end{aligned}$$

Betrachtet man zwei Punkte $x, y \in \mathbb{R}^3$, so ist ihr **Abstand** die Länge des Vektors $x - y$:



Kommen wir nun zur **Winkelmessung**: Sei $a = \| x - y \|$, $b = \| y \|$, $c = \| x \|$. Im $\mathbb{R}^2$



Der Cosinussatz besagt

$$a^2 = b^2 + c^2 - 2bc\cos\alpha$$

Also

$$\cos\alpha = \frac{b^2 + c^2 - a^2}{2bc}$$

$$\begin{aligned} b^2+c^2-a^2 &= \| y \|^2 + \| x \|^2 - \| x-y \|^2 \\ &= y_1^2+y_2^2+x_1^2+x_2^2-(x_1-y_1)^2-(x_2-y_2)^2 \\ &= 2x_1y_1+2x_2y_2 \end{aligned}$$

Es folgt

15.2 $\cos\alpha = \dfrac{x_1y_1+x_2y_2}{\| x \| \cdot \| y \|}$

Die Formeln 15.1 und 15.2 lassen sich nun in offensichtlicher Weise zu Längen- und Winkelmaßen im $\mathbb{R}^n$ erweitern. Dazu setzen wir

15.3 Definition: Das **EUKLIDISCHE** oder **STANDARDSKALARPRODUKT** von x und y aus $\mathbb{R}^n$ ist die Zahl

$$< x,y >= \sum_{i=1}^{n} x_iy_i = x_1y_1+x_2y_2+\ldots+x_ny_n$$

15.4 Rechenregeln: Für $x,y,z \in \mathbb{R}^n$ und $r \in \mathbb{R}$ gilt

(1) $< -,- >: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ ist eine **BILINEARFORM** (vgl. 11.6), d.h. für $x,y,z \in V$ und $r \in \mathbb{R}$ gilt:

$< x+y,z >=< x,z > + < y,z >, \quad < r\cdot x,z >= r\cdot < x,z >$

$< x,y+z >=< x,y > + < x,z >, \quad < x,r\cdot z >= r\cdot < x,z >$

(2) $< x,y >=< y,x >$, d.h $< -,- >$ ist **SYMMETRISCH**

(3) Für $x \neq 0$ gilt $\| x \|^2 =< x,x >> 0$, d.h. $< -,- >$ ist **POSITIV DEFINIT**.

Diese Regeln rechnet man leicht nach. □

Schon zur bequemeren Behandlung von Untervektorräumen, aber auch für die Untersuchung allgemeiner Vektorräume, etwa Räumen von Funktionen, empfiehlt sich, den Begriff des Skalarproduktes von Standardbeispielen $\mathbb{R}^n$ auf beliebige Vektorräume auszudehnen. Dazu erhebt man (wie schon früher praktiziert) die grundlegenden Rechenregeln zu Definitionen. Damit wir (15.4.3) mit in eine solche Definition aufnehmen können, beschränken wir uns auf folgenden Fall:

15.5 Konvention: Falls nicht ausdrücklich anders vermerkt, ist $\mathbb{K}$ ab jetzt $\mathbb{R}$ oder $\mathbb{C}$.

15.6 Definition: Sei V ein $\mathbb{K}$-Vektorraum.

Eine **BILINEARFORM** auf V ist eine Multilinearform $f : V \times V \to \mathbb{K}$. Eine Bilinearform f heißt **SYMMETRISCH**, falls $f(v,w) = f(w,v)$.

Für $\mathbb{C}$-Vektorräume ist man meist an anderen Formen interessiert.

15.7 Definition: Sei V ein $\mathbb{C}$-Vektorraum. Eine Funktion

$$f : V \times V \to \mathbb{C}$$

heißt **SESQUILINEARFORM** (="anderthalbfach linear"), wenn gilt

(1) f ist $\mathbb{C}$-linear im ersten Argument, d.h. $f(u+v,w) = f(u,w) + f(v,w)$ und $f(r \cdot u, w) = r \cdot f(u,w), \quad \forall u,v,w, \in V, \forall r \in \mathbb{C}.$

(2) f ist $\mathbb{C}$-**SEMILINEAR** im zweiten Argument, d.h. $\forall u,v,w \in V, \forall r \in \mathbb{C}$ gilt $f(u,v+w) = f(u,v) + f(u,w) \quad f(u,r\cdot v) = \overline{r} \cdot f(u,v)$

Eine Sesquilinearform f heißt **HERMITESCH**, wenn

$$f(u,v) = \overline{f(v,u)} \qquad \forall u,v \in V$$

15.8 Ist f eine hermitesche Sesquilinearform auf einem $\mathbb{C}$-Vektorraum V, so gilt: $f(v,v) \in \mathbb{R}\ \forall v \in V$

Beweis: $f(v,v) = \overline{f(v,v)}$, also ist $f(v,v) \in \mathbb{R}$. □

15.9 Definition: Sei V ein $\mathbb{K}$-Vektorraum. Eine symmetrische Bilinearform ($\mathbb{K} = \mathbb{R}$) bzw. eine hermitesche Sesquilinearform ($\mathbb{K} = \mathbb{C}$) heißt **POSITIV DEFINIT**, falls $f(v,v) > 0 \quad \forall v \neq 0$ aus V.

Ein **SKALARPRODUKT** $< -,- >: V \times V \to \mathbb{K}$ ist eine positiv definite symmetrische Bilinearform ($\mathbb{K} = \mathbb{R}$) bzw. hermitesche Sequilinearform ($\mathbb{K} = \mathbb{C}$).

Das Paar $(V, < -,- >)$ heißt dann **EUKLIDISCHER** ($\mathbb{K} = \mathbb{R}$) bzw. **UNITÄRER** ($\mathbb{K} = \mathbb{C}$) **RAUM**.

15.10 Beispiel: (1) $\mathbb{R}^n$ mit dem Standardskalarprodukt $< -,- >_{\mathbb{R}}$ ist ein euklidischer Raum.

(2) $\mathbb{C}^n$ mit dem Skalarprodukt

$$< x,y >_{\mathbb{C}} := \sum_{i=1}^{n} x_i \cdot \overline{y_i}$$

ist ein unitärer Vektorraum.

(3) Sei $C([0,1],\mathbb{K})$ der Vektorraum der stetigen Abbildungen

$$f : [0,1] \to \mathbb{K}$$

Dann ist

$$< f, g >:= \int_0^1 f \cdot \overline{g}\, dx$$

ein Skalarprodukt auf $C([0,1], \mathbb{K})$. Hierbei ist

$$\overline{g} : [0,1] \to \mathbb{K} \qquad \overline{g}(t) = \begin{cases} \overline{g(t)} & \mathbb{K} = \mathbb{C} \\ g(t) & \mathbb{K} = \mathbb{R} \end{cases}$$

Man prüft leicht nach, dass in allen Fällen die Axiome erfüllt sind.

Wegen Beispiel (1) und (2) haben wir ein Standardskalarprodukt auf jedem $\mathbb{K}^n$. Da jeder n-dimensionale $\mathbb{K}$-Vektorraum V isomorph zu $\mathbb{K}^n$ ist, kann man das Standardskalarprodukt $< -, - >_{\mathbb{K}}$ mit Hilfe des Isomorphismus' $\varphi : V \to \mathbb{K}^n$ auf V übertragen: Wir definieren

$$< -, - >_{\varphi}: V \times V \to \mathbb{K}, \quad < v, w >_{\varphi} = < \varphi(v), \varphi(w) >_{\mathbb{K}}$$

Die Axiome kann man "im Kopf" überprüfen. Wir erhalten

15.11 Satz: Jeder endlich-dimensionale $\mathbb{K}$-Vektorraum besitzt ein Skalarprodukt.

Bevor wir weitere Eigenschaften des Skalarproduktes untersuchen, wenden wir uns der Definition des Längenbegriffs zu.

15.12 Definition: Eine **Norm** auf einem $\mathbb{K}$-Vektorraum V ist eine Abbildung

$$V \to \mathbb{R}, \qquad v \mapsto \| v \|$$

so dass gilt:

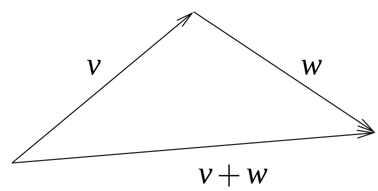
(1) $\| v \| = 0 \iff v = 0$

(2) $\| r \cdot v \| = |r| \cdot \| v \| \quad \forall v \in V, \forall r \in \mathbb{K}$

(3) $\| v + w \| \leq \| v \| + \| w \| \quad \forall v, w \in V$ (DREIECKSUNGLEICHUNG)

Das Paar $(V, \| - \|)$ heißt **NORMIERTER VEKTORRAUM**.

Alle drei Axiome sind Bedingungen, die wir von einem vernünftigen Längenbegriff erwarten. Die Dreiecksungleichung ist für den üblichen euklidischen Längenbegriff in der Ebene sofort einsichtig.:



Die folgende einfache Überlegung zeigt, dass nur positive Längen möglich sind:

15.13 Ist $\| - \|$ eine Norm auf V, dann gilt $\| v \| \geq 0 \; \forall v \in V$,

Beweis: $0 \overset{(1)}{=} \| v - v \| \overset{(3)}{\leq} \| v \| + \| -v \| \overset{(2)}{=} \| v \| + |(-1)| \cdot \| v \| = 2 \cdot \| v \|$
Also $\| v \| \geq 0$. □

15.14 Satz: Ist $< -, - >$ ein Skalarprodukt auf V, dann wird durch

$$\| v \| := {}_{+}\sqrt{< v, v >}$$

eine Norm auf V definiert, und es gilt die Cauchy-Schwartz'sche Ungleichung für alle $u, v \in V$

$$| < u, v > | \leq \| u \| \cdot \| v \|$$

Das Gleichheitszeichen gilt genau dann, wenn u und v linear abhängig sind.

Beweis: $v = 0 \Rightarrow v = 0 \cdot v \Rightarrow < v, v > = < 0 \cdot v, v > = 0 \cdot < v, v > = 0$
Weiter gilt: $\| v \| = 0 \iff < v, v > = 0 \Rightarrow v = 0$, da $< -, - >$ positiv definiert ist.
$\| r \cdot v \| = \sqrt{< r \cdot v, r \cdot v >} = \sqrt{r \cdot \bar{r} \cdot < v, v >} \overset{(2.13)}{=} |r| \cdot \| v \|$
Wir beweisen weiter die Cauchy-Schwarz'sche Ungleichung:
Für $v = 0$ erhalten wir auf beiden Seiten 0. Sei also $v \neq 0$. Es gilt für $s = \| v \|^2 \in \mathbb{R}$ und $t = < u, v > \in \mathbb{C}$

$$\begin{aligned}
0 &\leq < s \cdot u - t \cdot v, s \cdot u - t \cdot v > \\
&= s \cdot \bar{s} \cdot < u, u > - s \cdot \bar{t} \cdot < u, v > - t \cdot \bar{s} < v, u > + t \cdot \bar{t} < v, v > \\
&= s^2 \cdot \| u \|^2 - 2s \cdot t \cdot \bar{t} + t \cdot \bar{t} \cdot \| v \|^2 = s(\| v \|^2 \cdot \| u \|^2 - 2|t|^2 + |t|^2) \\
&= s \cdot (\| v \|^2 \cdot \| u \|^2 - | < u, v > |^2)
\end{aligned}$$

Da $s = \| v \|^2 \neq 0$, folgt die Ungleichung.
Das Gleichheitszeichen gilt genau dann, wenn $s \cdot u - t \cdot v = 0$. Da $s \neq 0$, sind u und v linear abhängig.
Sind umgekehrt u und v linear abhängig, gilt $u = r \cdot v$, da $v \neq 0$. Also

$$| < u, v > | = |r \cdot < v, v > | = |r| \cdot \| v \|^2 = \| r \cdot v \| \cdot \| v \| = \| u \| \cdot \| v \|$$

Es bleibt, die Dreiecksungleichung zu zeigen:

$$\begin{aligned}
\| u + v \|^2 &= < u + v, u + v > = \| u \|^2 + < u, v > + < v, u > + \| v \|^2 \\
&\leq \| u \|^2 + 2 \cdot | < u, v > | + \| v \|^2 \leq \| u \|^2 + 2 \| u \| \cdot \| v \| + \| v \|^2 \\
&= (\| u \| \cdot \| v \|)^2
\end{aligned}$$

Denn ist $< u, v > = a + bi$, so ist $< v, u > = \overline{< u, v >} = a - bi$ und $| < u, v > | = \sqrt{a^2 + b^2}$. Also $< u, v > + < v, u > = 2a \leq 2\sqrt{a^2 + b^2}$. □

Eine Norm erlaubt die Definition eines Abstandsbegriffes:

15.15 Definition: Ist $(V, \| - \|)$ ein normierter Vektorraum, dann nennen wir

$$\| v - w \|$$

den **ABSTAND** von v und w aus V (bzgl. $\| - \|$)

15.16 Rechenregeln für den Abstand: Für $u, v, w \in V$ gilt

(1) $\| u - v \| = 0 \iff u = v$

(2) $\| u - v \| = \| v - u \|$

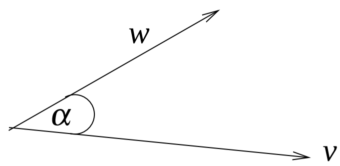
(3) $\| u - w \| \leq \| u - v \| + \| v - w \|$ (Dreiecksgleichung)

Diese Regeln folgen sofort aus (15.12). $\square$

In einem euklidischen Raum können wir ein vernünftiges Winkelmaß einführen, indem wir (15.2) erweitern

15.17 Definition: Ist $(V, < -, - >)$ ein euklidischer Raum, definieren wir als Maß des Winkels zwischen zwei Vektoren v, w

$$\cos \alpha = \frac{< v, w >}{\| v \| \cdot \| w \|}$$



Im unitären Fall ist $< v, w > \in \mathbb{C}$ und außerdem $< v, w > = \overline{< w, v >}$, so dass wir mit der Definition (15.17) wenig anfangen können. Im Spezialfall $< v, w > = 0$ ist das natürlich anders. Im diesem Fall ist $\alpha = \frac{\pi}{2}$, d.h. v und w stehen senkrecht aufeinander. Damit ist folgende Definition in jedem Vektorraum mit Skalarprodukt sinnvoll:

15.18 Definition: Sei $(V, < -, - >)$ ein Vektorraum mit Skalarprodukt

(1) $v, w \in V$ **STEHEN SENKRECHT AUFEINANDER**, wenn $< v, w > = 0$. Wir schreiben $u \perp w$.

(2) $v \in V$ **STEHT SENKRECHT AUF EINEM UNTERVEKTORRAUM** U, wenn $< v, u > = 0 \quad \forall u \in U$. Wir schreiben $u \perp U$.

(3) Zwei Untervektorräume $U_1, U_2 \subset V$ sind **ORTHOGONAL ZUEINANDER**, wenn $< u_1, u_2 > = 0 \quad \forall u_1 \in U_1,\ \forall u_2 \in U_2$. Wir schreiben $U_1 \perp U_2$.

(4) Für einen Untervektorraum $U \subset V$ heißt

$$U^{\perp} := \{v \in V; v \perp U\}$$

das **ORTHOGONALE KOMPLEMENT** von U.

15.19 Satz: Sei $(V, < -, - >)$ ein Vektorraum mit Skalarprodukt. Dann gilt

(1) Ist $M \subset V$ eine nicht-leere Teil**menge** und

$$M^{\perp} = \{v \in V; < v, w >= 0 \; \forall w \in M\},$$

dann ist $M^{\perp}$ ein Untervektorraum von V und $M^{\perp} = (\operatorname{Span} M)^{\perp}$

(2) Ist $U \subset V$ Untervektorraum, dann ist die Summe $U + U^{\perp}$ direkt.

(3) Ist $U \subset V$ Untervektorraum und $\dim U < \infty$, so ist $V = U \oplus U^{\perp}$.

Beweis 1:

(i) $0 \in M^{\perp}$, so dass $M^{\perp} \neq \emptyset$

(ii) $v^1, v^2 \in M^{\perp} \Rightarrow < v^1 + v^2, w >=< v^1, w > + < v^2, w >= 0 + 0 = 0 \; \forall w \in M$. Also ist $v^1 + v^2 \in M^{\perp}$.

(iii) $v \in M^{\perp}$, $r \in \mathbb{K} \Rightarrow < r \cdot v, w >= r \cdot < v, w >= r \cdot 0 = 0 \; \forall w \in M$. Also ist $r \cdot v \in M^{\perp}$.

Damit ist $M^{\perp}$ ein Untervektorraum. Da $M \subset \operatorname{Span} M$, folgt $(\operatorname{Span} M)^{\perp} \subset M^{\perp}$. Sei nun $v \in M^{\perp}$ und $w = r_1 \cdot w^1 + \ldots + r_n \cdot w^n \in \operatorname{Span} M$, wobei $r_i \in \mathbb{K}$ und $w^i \in M$ für $i = 1, \ldots, n$. Dann gilt

$$< v, w >= \sum_{i=1}^{n} \overline{r_i} \cdot < v, w^i >= 0$$

da $< v, w^i >= 0 \quad \forall i = 1, \ldots, n$. Also ist $v \in (\operatorname{Span} M)^{\perp}$.

Beweis 2: Sei $v \in U \cap U^{\perp}$. Dann gilt $< v, v >= 0$. Es folgt $v = 0$.

Beweis 3 stellen wir noch etwas zurück. $\square$

Als nächstes wollen wir zeigen, dass V ein "rechtwinkliges Koordinatensystem" besitzt.

15.20 Definition: Sei $(V, < -, - >)$ ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. $\mathscr{B} = \{v^1, \dots, v^k)$ heißt **ORTHOGONALSYSTEM** in V, wenn

$$v^i \perp v^j \quad \forall i \neq j, \quad i, j = i \dots, n.$$

Gilt außerdem $\| v^i \| = \sqrt{< v^i, v^i >} = 1$ für $i = 1, \dots, n$ heißt $\mathscr{B}$ **ORTHONORMALSYSTEM** in V.

Ist $\mathscr{B}$ außerdem noch Basis, spricht man von einer **ORTHOGONAL-** bzw. **ORTHONORMALBASIS**.

15.21 Beispiel: Die Standardbasis ist eine Orthonormalbasis in $(\mathbb{K}, < -, - >_{\mathbb{K}})$. Denn $< e^i, e^j > = \delta_{ij}$.

Wir wollen jetzt ein Verfahren vorstellen, das aus einer gegebenen Basis eine Orthonormalbasis macht.

15.22 Der GRAM-SCHMIDT-Orthogonalisierungsprozess

Sei $\{v^1, \dots, v^n\}$ ein System linear unabhängiger Vektoren in V. Dann konstruieren wir **induktiv** ein Orthogonalsytem $\{u^1, \dots, u^n\}$, so dass

$$\mathrm{Span}\{v^1, \dots, v^i\} = \mathrm{Span}\{u^1, \dots, u^i\} \qquad i = 1, \dots, n$$

$\mathbf{n = 1}$: $v^1 \neq 0$. Setze $u^1 = \frac{v^1}{\|v^1\|}$. Dann gilt $< u^1, u^1 > = \frac{v}{\|v^1\|^2} \cdot \| v \|^2 = 1$.

Induktionsschritt: Sei $\{u^1, \dots, u^{n-1}\}$ ein Orthonormalsystem, so dass

$$\mathrm{Span}\{v^1, \dots, v^i\} = \mathrm{Span}\{u^1, \dots, u^i\} \qquad \text{für } 1 \leq i \leq n-1$$

Da $v^n \notin \mathrm{Span}\{v^1, \dots, v^{n-1}\}$, ist

$$w = v^n - \sum_{i=1}^{n-1} < v^n, u^i > \cdot u^i \neq 0.$$

Also ist $\| w \| \neq 0$, und wir setzen

$$u^n = \frac{w}{\| w \|}$$

Da $v^n \in \mathrm{Span}\{u^1, \dots, u^{n-1}, w\}$, folgt $\mathrm{Span}\{u^1, \dots, u^n\} = \mathrm{Span}\{v^1, \dots, v^n\}$. Weiter gilt $\| u^n \| = 1$, und für $i \leq n-1$

$$\begin{aligned} < u^n, u^i > &= \tfrac{1}{\|w\|} < v^n, u^i > - \tfrac{1}{\|w\|} \sum_{j=1}^{n-1} < v^n, u^j > \cdot \underbrace{< u^j, u^i >}_{=\delta_{ij}} \\ &= \tfrac{1}{\|w\|} < v^n, u^i > - \tfrac{1}{\|w\|} < v^n, u^i > = 0 \end{aligned}$$

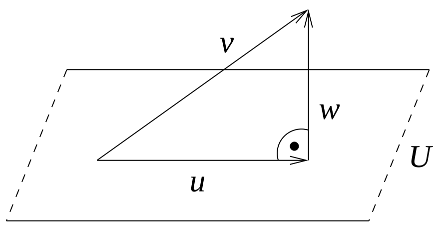
Also ist $\{u^1, \dots, u^n\}$ ein Orthonormalsystem, wie gewünscht. $\square$

Als Folgerung erhalten wir

15.23 Satz: Sei $(V, < -, - >)$ ein $\mathbb{K}$-Vektorraum endlicher Dimension mit Skalarprodukt. Dann lässt sich jedes Orthonormalsystem $\{u^1, \ldots, u^k\}$ zu einer Orthonormalbasis ergänzen. Insbesondere besitzt V eine Orthonormalbasis.

Wir benutzen diesen Satz, um einen einfachen Beweis von (15.19.3) nachzutragen.

15.24 Definition: Sei $U \subset V$ ein Untervektorraum und $v \in V$. Ein Vektor $u \in U$ heißt **ORTHOGONALE PROJEKTION** von $v \in V$ auf U, wenn es ein $w \in U^\perp$ gibt, so dass $v = u + w$, d.h. wenn $v - u \in U^\perp$



15.25 Lemma: (1) Sei $U \subset V$ Untervektorraum und $v \in V$. Dann gibt es höchstens eine orthogonale Projektion u von v auf U.

(2) Ist U endlich-dimensional und $\{u^1, \ldots, u^k\}$ eine Orthonormalbasis von U, dann ist
$$u = \sum_{j=1}^{k} < v, u^j > u^j \in U$$
orthogonale Projektion von v auf U.

Beweis 1: Sei $v = u^1 + w^1 = u^2 + w^2$ mit $u^1, u^2 \in U$, $w^1, w^2 \in U^\perp$. Wir müssen zeigen, dass $u^1 = u^2$. Es gilt, da $u^1 - u^2 \in U$ und $w^2 - w^1 \in U^\perp$,
$$u^1 - u^2 = w^2 - w^1 \in U \cap U^\perp = \{0\}.$$
Also $u^1 - u^2 = 0$.

Beweis 2: Sei $w = v - u$. Dann gilt
$$< w, u^j > = < v, u^j > - \sum_{i=1}^{n} < v, u^i > \underbrace{< u^i, u^j >}_{\delta_{ij}} = < v, u^j > - < v, u^j > = 0$$

Also folgt: $w \in \{u^1 \ldots, u^k\}^\perp = (\mathrm{Span}\{u^1, \ldots, u^k\})^\perp = U^\perp$. $\square$

Beweis 15.9.3: Wer müssen zeigen $V = U + U^{\perp}$. Das folgt aber aus (15.25.2). $\square$

Zum Abschluss beweisen wir noch zwei Ergebnisse von eigenständigem Interesse:

15.26 Satz des Pythagoras: Sei $\{v^1 \dots, v^n\}$ ein Ortho**gonal**system in V. Dann gilt:

$$\| v^1 + v^2 + \dots + v^n \|^2 = \| v^1 \|^2 + \| v^2 \|^2 + \dots + \| v^n \|^2$$

Beweis: $\| v^1 + \dots + v^n \|^2 = < v^1 + \dots + v^n, v^1 + \dots + v^n > = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} < v^i, v^j >$

$$= \sum_{i=1}^{n} < v^i, v^i > = \sum_{i=1}^{n} \| v^i \|^2, \text{ weil } < v^i, v^j > = 0 \text{ für } i \neq j.$$

$\square$

15.27 Bessel'sche Ungleichung: Sei $\{v^1, \dots, v^n\}$ ein Ortho**normal**system in V. Dann gilt für alle $w \in V$

$$\sum_{i=1}^{n} | < w, v^i > |^2 \leq \| w \|^2 .$$

Gleichheit gilt genau dann, wenn $w \in \mathrm{Span}\{v^1, \dots, v^n\}$.

Beweis: Sei $u = \sum\limits_{i=1}^{n} < w, v^i > \cdot v^i$ die orthogonale Projektion von w auf $U = \mathrm{Span}\{v^1, \dots v^n\}$. Dann ist $z = w - u \in U^{\perp}$, d.h. $z \perp u$. Aus dem Satz des Pythagoras folgt

$$\begin{aligned}
\| w \|^2 = \| z + u \|^2 &= \| z \|^2 + \| u \|^2 \\
&= \| z \|^2 + < \sum_{i=1}^{n} < w, v^i > \cdot v^i, \sum_{j=1}^{n} < v^i, v^j > \cdot v^j > \\
&= \| z \|^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} < w, v^i > \cdot \overline{< w, v >} \cdot \underbrace{< v^i, v^j >}_{\delta_{ij}} \\
&= \| z \|^2 + \sum_{i=1}^{n} | < w, v^i > |^2
\end{aligned}$$

Gleichheit in der Besselschen Ungleichung gilt genau dann, wenn $z = 0$, d.h. wenn $w \in U$. $\square$

Anwendung:

In (15.15) haben wir den Abstand zweier Elemente $v, w \in V$ durch $\| v - w \|$ definiert. Wir wollen nun den Abstand eines Elements v von einem Untervektorraum U definieren.

15.28 Definition: Sei $v \in V$ und U Untervektorraum von V. Dann heißt

$$d(v,U) = \mathrm{Min}\{\| v - u \|; u \in U\}$$

der **ABSTAND** von v und U.

15.29 Satz: Ist $\dim U < \infty$, dann existiert $d(v,U)$. Es gibt genau ein $u \in U$ mit

$$d(v,U) = \| v - u \|,$$

nämlich die orthogonale Projektion von v auf U.

Beweis: Nach (15.24) ist $v - u \in U^{\perp}$. Sei $w \in U$ beliebig und $z = w - u$. Dann ist $z \in U$, und es gilt

$$\begin{aligned} \| v - w \|^2 &= < v - u - z, v - u - z > \\ &= < v - u, v - u > - < v - u, z > - < z, v - u > + < z, z > \\ &= \| v - u \|^2 + \| z \|^2, \quad \text{da } v - u \in U^{\perp} \text{ und } z \in U. \end{aligned}$$

Für $z = 0$ erhalten wir das Minimum. □

15.30 Methode der kleinsten Quadrate

Sei $A \in \mathbb{K}^{m,n}$ eine Matrix von Rang n und sei $m > n$. Dann ist das lineare Gleichungssystem

$$A \cdot x = b \qquad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

nicht für jedes $b \in \mathbb{K}^m$ lösbar. Wir suchen in diesem Fall die **beste Näherungslösung**, d.h. $z \in \mathbb{K}^n$, für das (wir benutzen das Standardskalarprodukt)

$$\| A \cdot z - b \| \qquad \text{minimal ist.}$$

Sei $U = \text{Bild } f_A = \{A \cdot x;\ x \in \mathbb{K}^n\} = \mathrm{Span}\{s^1, \ldots, s^n\}$, wobei s^j der j-te Spaltenvektor von A ist. Sei u die orthogonale Projektion von b auf U nach (15.29) gilt

$$\| u - b \| = \mathrm{Min}\{\| u_1 - b \|, u_1 \in U\}.$$

Nach Voraussetzung ist f_A injektiv (denn $rgA = n < m$). Also gilt es genau ein $z \in \mathbb{R}^n$ mit $f_A(z) = A \cdot z = u$. Dieses z ist die gesuchte Näherungslösung.
Wir wollen z berechnen. Da $b - u \in U^{\perp}$ und $U = \mathrm{Span}\{s^1, \ldots, s^n\}$ ist, haben wir ein System von Gleichungen

$$< b - u, s^j >= 0 \quad \text{also } < A \cdot z, s^j >=< u, s^j >=< b, s^j > \forall j = 1, \ldots, n$$

Da nun

$$< A\cdot z, s^j >= (A\cdot z)^t \cdot \overline{s^j} = z^t \cdot A^t \cdot \overline{s^j} \qquad \text{und}$$

$$A^t \cdot \overline{s^j} = \begin{pmatrix} < s^1, s^j > \\ \vdots \\ < s^n, s^j > \end{pmatrix},$$

folgt $< A\cdot z, s^j >= \sum_{k=1}^n z_k \cdot < s^k, s^j >$. Also ist das Gleichungssystem durch

$$\sum_{k=1}^n c_{jk} z_k = \sum_{k=1}^n < s^k, s^j > z_k =< b, s^j > \qquad j = 1, \ldots, n$$

gegeben. Die Koeffizientenmatrix des Systems ist $C = \overline{A}^t \cdot A$, denn der (j,k)-te Eintrag von $\overline{A}^t \cdot A$ ist

$$(\overline{s^j})^t \cdot s^k = (s^k)^t \cdot \overline{s^j} =< s^k, s^i >= c_{jk}$$

In Matrixschreibweise ist z die eindeutige Lösung des Gleichungssystems

$$\overline{A}^t \cdot A \cdot z = \begin{pmatrix} < b, s^1 > \\ \vdots \\ < b, s^n > \end{pmatrix} = \overline{A}^t \cdot b$$

(die zweite Gleichung rechnet man sofort nach). Da z eindeutige Lösung ist, muss $\overline{A}^t \cdot A$ invertierbar sein. Wir erhalten somit

$$\boxed{z = (\overline{A}^t \cdot A)^{-1} \cdot \overline{A}^t \cdot b}$$

Im letzten Schritt der Herleitung haben wir außerdem gezeigt:

15.31 Satz: Ist $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$, $A \in \mathbb{K}^{m,n}$ eine Matrix vom Rang $n < m$, dann ist $\overline{A}^t \cdot A$ invertierbar. $\square$

16 Bilinearformen und Matrizen

Sei V ein $\mathbb{K}$-Vektorraum mit Basis $\mathscr{B} = \{v^1, \ldots, v^n\}$ und $f : V \times V \longrightarrow \mathbb{K}$ eine Bilinearform bzw. Sesquilinearform. Da für $x \in \mathbb{R}$ stets $\overline{x} = x$ gilt, behandeln wir die Fälle parallel, indem wir die Beweise für die Sesquilinearform (Fall $\mathbb{K} = \mathbb{C}$) führen.

16.1 Sind $v = \sum_{i=1}^n x_i \cdot v^i$ und $w = \sum_{j=1}^n y_j v^j$ zwei Vektoren in V. Dann gilt

$$f(v,w) = \sum_{i=1}^n x_i \cdot f(v^i, \sum_{j=1}^n y_j v^j) = \sum_{i=1}^n \sum_{j=1}^n x_i \cdot \overline{y_j} \cdot f(v^i, v^j)$$

Wir setzen

$$\boxed{A = (a_{ij}) \in \mathbb{K}^{n,n} \text{ mit } a_{ij} = f(v^i, v^j)}$$

Dann folgt (beachte: $\sum_{i=1}^n f(v^i, v^j) \cdot \overline{y_j}$ ist die i-te Komponente von $A \cdot \overline{y}$)

$$f(v,w) = (x_1, \ldots, x_n) \cdot A \cdot \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix} = x^t \cdot A \cdot \overline{y}.$$

Hierbei fassen wir die Koordinatenvektoren x von v und y von w als Spaltenvektoren auf und definieren

$$\overline{y} = \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix} \text{ für } y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

16.2 A heißt Matrix von f bzgl. der Basis $\mathscr{B}$. Wir schreiben $A = Q_{\mathscr{B}}(f)$.

An Hand der Matrix A kann man sofort feststellen, wann f symmetrisch bzw. hermitesch ist.

16.3 Satz: Sei V ein $\mathbb{K}$-Vektorraum mit Basis $\mathscr{B} = \{v^1, \ldots, v^n\}$, f eine Bi- bzw. Sesquilinearform auf V und $A = Q_{\mathscr{B}}(f)$. Dann gilt

(1) **Fall $\mathbb{K} = \mathbb{R}$:** f ist symmetrisch $\Longleftrightarrow$ A ist **SYMMETRISCH**, d.h.

$$A = A^t.$$

(2) **Fall $\mathbb{K} = \mathbb{C}$:** f ist hermitesch $\Longleftrightarrow$ A ist hermitesch, d.h.

$$\overline{A}^t = A, \quad \text{wobei } \overline{A} = (\overline{a_{ij}}), \text{ falls } A = (a_{ij})$$

Beweis: f hermitesch $\Longrightarrow a_{ij} = f(v^i, v^j) = \overline{f(v^j, v^i)} = \overline{a}_{ji}$.
A hermitesch: Dann gilt mit den Bezeichnungen aus (16.1)

$$\begin{aligned} f(v,w) &= x^t \cdot A \cdot \overline{y} \overset{(*)}{=} (A \cdot \overline{y})^t \cdot x = \overline{y}^t \cdot A^t \cdot x \\ \overline{f(w,v)} &= \overline{y^t \cdot A \cdot \overline{x}} = \overline{y}^t \overline{A} \cdot x = f(v,w,), \end{aligned}$$

da $\overline{A}^t = A$, also $\overline{A} = A^t$. □

16.4 Im Beweis haben wir in Gleichung (*) benutzt, dass für zwei Spaltenvektoren $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ stets gilt

$$x^t \cdot y = \sum_{i=1}^n x_i \cdot y_i = y^t \cdot x.$$

Außerdem haben wir die erste der folgenden Regeln verwandt:
Für Matrizen A, B gilt (falls die Seiten der Gleichungen definiert sind)

(1) $\overline{A \cdot B} = \overline{A} \cdot \overline{B}$

(2) $\overline{A + B} = \overline{A} + \overline{B}$

(3) $\overline{A^{-1}} = (\overline{A})^{-1}$.

Die ersten beiden Regeln gelten, weil $\overline{a \cdot b} = \overline{a}\,\overline{b}$ und $\overline{a+b} = \overline{a} + \overline{b}$ für $a, b \in \mathbb{C}$. Weiter gilt wegen (1)

$$\overline{A} \cdot \overline{A^{-1}} = \overline{A \cdot A^{-1}} = \overline{E_n} = E_n, \quad \text{also } \overline{A^{-1}} = (\overline{A})^{-1}.$$

Jetzt bleibt nur noch zu klären, wie man $Q_{\mathscr{B}}(f)$ ansieht, ob f positiv definit ist oder nicht. Unser erstes Ziel ist der Beweis des folgenden Satzes:

16.5 Satz: Sei f eine symmetrische Bilinearform ($\mathbb{K} = \mathbb{R}$) bzw. hermitesche Sesquilinearform ($\mathbb{K} = \mathbb{C}$) auf einem n-dimensionalen $\mathbb{K}$-Vektorraum V. Genau dann ist f positiv definit, wenn V eine Basis $\mathscr{B}$ besitzt, so dass

$$Q_{\mathscr{B}}(f) = E_n.$$

Beweis: Ist f positiv definit, so ist f ein Skalarprodukt: $f(v,w) = \langle v, w \rangle$. Nach (15.23) besitzt V dann eine Orthonormalbasis $\mathscr{B} = \{v^1, \ldots, v^n\}$. Für diese gilt:

$$f(v^i, v^j) = \langle v^i, v^j \rangle = \delta_{ij}.$$

Also $Q_{\mathscr{B}}(f) = E_n$.
Sei umgekehrt $\mathscr{B}$ eine Basis von V, so dass $Q_{\mathscr{B}}(f) = E_n$. Dann gilt nach (16.1) für $v = \sum_{i=1}^n x_i v^i$ ($\mathscr{B} = \{v^1, \ldots, v^n\}$)

$$f(v,v) = x^t \cdot E_n \cdot \overline{x} = \sum_{i=1}^n x_i \cdot \overline{x_i} = \sum_{i=1}^n |x_i|^2 > 0,$$

falls $v \neq 0$. □

Dieser Satz nützt recht wenig, weil wir nicht wissen, wie wir eine solche Basis finden können. Aber er hilft doch, Rückschlüsse auf die Form der Matrizen $Q_{\mathscr{B}}(f)$ zu ziehen: Sind $\mathscr{B}$ und $\mathscr{B}'$ verschiedene Basen von V und ist $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ der Koordinatenvektor von v bzgl. $\mathscr{B}$, dann ist nach § 7

$$T \cdot x \text{ mit } T = M_{\mathscr{B}}^{\mathscr{B}'}(id)$$

der Koordinatenvektor von x bzgl. $\mathscr{B}'$. Wir erhalten somit mit den Bezeichnungen aus (16.1)

$$\begin{aligned} f(v,w) &= x^t \cdot Q_{\mathscr{B}}(f) \cdot \overline{y} = (T \cdot x)^t \cdot Q_{\mathscr{B}'}(f) \cdot \overline{T \cdot y} \\ &= x^t \cdot T^t \cdot Q_{\mathscr{B}'}(f) \cdot \overline{T} \cdot \overline{y}. \end{aligned}$$

Es folgt, indem man die Basisvektoren von $\mathscr{B}$ einsetzt,

16.6 Transformationsformel: Ist V ein endlichdimensionaler $\mathbb{K}$-Vektorraum, f eine Bi- bzw. Sesquilinearform auf V, und sind $\mathscr{B}$ und $\mathscr{B}'$ zwei Basen von V, dann gilt

$$Q_{\mathscr{B}}(f) = T^t \cdot Q_{\mathscr{B}'}(f) \cdot \overline{T} \text{ mit } T = M_{\mathscr{B}}^{\mathscr{B}'}(id).$$

16.7 Folgerung: Ist $f : V \times V \longrightarrow \mathbb{K}$ ein Skalarprodukt auf einem n-dimensionalen Vektorraum V und ist $\mathscr{B}$ eine beliebige Basis, dann gilt für $A = Q_{\mathscr{B}}(f)$:

(1) A ist invertierbar

(2) $\operatorname{Det} A \in \mathbb{R}$ und $\operatorname{Det}(A) > 0$.

Beweis: Sei $\mathscr{B}'$ eine Orthonormalbasis von V und $T = M_{\mathscr{B}}^{\mathscr{B}'}(id)$. Wir haben in (16.5) gesehen, dass $Q_{\mathscr{B}'}(f) = E_n$, und aus (16.6) folgt $A = T^t \cdot \overline{T}$. Wir erhalten

$$\operatorname{Det} A = \operatorname{Det}(T^t \cdot \overline{T}) = \operatorname{Det}(T^t) \cdot \operatorname{Det}(\overline{T}) = \operatorname{Det}(T) \cdot \overline{\operatorname{Det}(T)} = |\operatorname{Det}(T)|^2 > 0.$$

Die Gleichung $\operatorname{Det}(\overline{T}) = \overline{\operatorname{Det} T}$ folgt aus der Formel (11.15) für die Determinante. Da T invertierbar ist, ist $\operatorname{Det}(T) \neq 0$. □

Mit Hilfe dieser Folgerung können wir nun ein brauchbares Kriterium beweisen.

16.8 Definition: Eine hermitesche (bzw. symmetrische) Matrix A aus $\mathbb{C}^{n,n}$ (bzw. $\mathbb{R}^{n,n}$) heißt **POSITIV DEFINIT**, wenn die Sesquilinearform (Bilinearform)

$$\begin{array}{ll} \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C} & \quad (x,y) \longmapsto x^t \cdot A \cdot \overline{y} \\ (\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}) & \end{array}$$

positiv definit ist, d.h. wenn für $x \neq 0$ gilt $x^t \cdot A \cdot \overline{x} > 0$.

16.9 Satz: (Hauptminorenkriterium): Sei $f : V \times V \longrightarrow \mathbb{K}$ eine symmetrische Bi- bzw. hermitesche Sesquilinearform auf dem n-dimensionalen $\mathbb{K}$-Vektorraum V und $\mathscr{B}$ eine Basis von V. Sei $A = Q_{\mathscr{B}}(f) = (a_{ij})$. Dann gilt

$$f \text{ positiv definit } \Longleftrightarrow A \text{ positiv definit } \Longleftrightarrow \operatorname{Det} A_k > 0 \text{ für } k = 1, \ldots, n,$$

wobei $A_k = (a_{ij})_{1 \leq i,j \leq k}$. Insbesondere ist $\operatorname{Det} A_k \in \mathbb{R}$.

Beweis: (f positiv definit $\Longleftrightarrow$ A positiv definit) folgt aus (16.1). Beachte

$$A_k = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ a_{21} & \cdots & a_{2k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} \qquad \text{ist symmetrisch.}$$

Sei A positiv definit und $x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} \in \mathbb{C}^k$. Wir ergänzen x durch Nullen zu einem Vektor $\binom{x}{0} \in \mathbb{C}^n$. Ist $x \neq 0$, so ist $\binom{x}{0} \neq 0$. Also

$$x^t \cdot A_k \cdot \overline{x} = \begin{pmatrix} x \\ 0 \end{pmatrix}^T \cdot A \cdot \begin{pmatrix} \overline{x} \\ 0 \end{pmatrix} > 0.$$

Also ist A_k positiv definit und damit $\operatorname{Det} A_k > 0$ nach (16.7). Sei umgekehrt $\operatorname{Det} A_k > 0$ für $k = 1, \ldots, n$. Wir zeigen induktiv, dass A_k positiv definit ist.

Anfang: $k = 1$, $A_1 = (a_{11})$. $\operatorname{Det} A_1 = a_{11} > 0$ und $x_1 \cdot a_{11} \cdot \overline{x}_1 = a_{11}|x_1|^2 > 0$, falls $x_1 \neq 0$.

Induktionsschritt: Sei $k > 1$ und A_{k-1} positiv definit. Wir zeigen, dass auch A_k positiv definit ist.

Da A_{k-1} ein Skalarprodukt auf $\mathbb{K}^{k-1}$ definiert, gibt es nach (15.23) eine Orthonormalbasis $\mathscr{B}' = \{w^1, \ldots, w^{k-1}\}$ von $\mathbb{K}^{k-1}$, d.h.

$$(w^i)^t \cdot A_{k-1} \overline{w^j} = \delta_{ij} \qquad 1 \leq i,j \leq k-1.$$

Wir betrachten die Form

$$f : \mathbb{K}^k \times \mathbb{K}^k \longrightarrow \mathbb{K}, \qquad (u,v) \longmapsto u^t \cdot A_k \cdot \overline{v}.$$

Sei $v^j = \binom{w^j}{0} \in \mathbb{K}^k$, d.h. wir machen aus w^j einen Vektor in $\mathbb{K}^k$, indem wir die k-te Komponente 0 hinzufügen. Sei $\langle -, - \rangle_{\mathbb{K}}$ das Standardskalarprodukt von $\mathbb{K}^k$. Dann ist

$$\mathscr{B} = \{v^1, \dots, v^{k-1}, w\}$$

mit $w = e^k - \sum_{j=1}^{k-1} \langle z^k, v^j \rangle_{\mathbb{K}} \cdot v^j$ eine Basis. Hier ist e^k der k-te Einheitsvektor und z^k der k-te Zeilenvektor von A_k. Wir berechnen $Q_{\mathscr{B}}(f)$. Da $v^j = \binom{w^j}{0}$, gilt offensichtlich

$$f(v^i, v^j) = (v^i)^t \cdot A_k \cdot \overline{v^j} = (w^i)^t \cdot A_{k-1} \cdot \overline{w^j} = \delta_{ij}.$$

Weiter gilt $f(w, v^j) =$

$$\begin{aligned} w^t \cdot A_k \cdot \overline{v^i} &= (z^k)^t \cdot \overline{v^i} - \sum_{j=1}^{k-1} \langle z^k, v^j \rangle_{\mathbb{K}} \cdot \underbrace{(v^j)^t \cdot A_k \cdot \overline{v^i}}_{\delta_{ij}} \\ &= \langle z^k, v^i \rangle_{\mathbb{K}} - \langle z^k, v^i \rangle_{\mathbb{K}} = 0. \end{aligned}$$

Da $Q_{\mathscr{B}}(f)$ hermitesch (bzw. symmetrisch) ist, gilt

$$Q_{\mathscr{B}}(f) = \begin{pmatrix} & \vdots & 0 \\ E_{k-1} & \vdots & \vdots \\ \dots\dots\dots & & 0 \\ 0 \quad \cdots \quad & 0 & c \end{pmatrix} \text{ mit } c = f(w,w).$$

Beachte nun, dass f bzgl. der Standardbasis $\mathscr{S}$ durch A_k beschrieben wird. Ist $T = M_{\mathscr{B}}^{\mathscr{S}}(id)$, so gilt $Q_{\mathscr{B}}(f) = T^t \cdot A_k \cdot \overline{T}$. Also (vergl. Beweis 16.7)

$$c = \operatorname{Det} Q_{\mathscr{B}}(f) = \operatorname{Det} A_k \cdot |\operatorname{Det} T|^2 > 0.$$

Setzen wir $v^k = \frac{1}{\sqrt{c}} w$, so ist f bzgl. der Basis $\{v^1, \dots, v^k\}$ durch E_k gegeben. Also ist f nach (16.5) positiv definit. $\square$

Um noch mehr über die Matrizen $Q_{\mathscr{B}}(f)$ einer symmetrischen Bilinearform oder hermiteschen Sesquilinearform zu erfahren, wenden wir die Eigenwertergebnisse auf den Endomorphismus

$$\mathbb{K}^n \longrightarrow \mathbb{K}^n, \quad x \longmapsto Q_{\mathscr{B}}(f) \cdot x$$

an.

16.10 Satz: Jede hermitesche Matrix $A \in \mathbb{C}^{n,n}$ und damit jede reelle symmetrische Matrix hat nur reelle Eigenwerte. (D.h. das charakteristische Polynom hat nur reelle Nullstellen.)

Beweis: Wir fassen A im reellen wie im komplexen Fall als komplexe Matrix auf. Sei also $\lambda \in \mathbb{C}$ ein komplexer Eigenwert und $v \neq 0$ aus $\mathbb{C}^n$ ein zugehöriger Eigenvektor. Sei $\langle -,- \rangle$ das Standardskalarprodukt auf $\mathbb{C}^n$. Dann gilt

$$\begin{aligned}
\overline{\lambda} \cdot \langle v,v \rangle &= \langle v, \lambda \cdot v \rangle = \langle v, A \cdot v \rangle = v^t \cdot \overline{A \cdot v} = v^t \cdot \overline{A} \cdot \overline{v} \\
\lambda \cdot \langle v,v \rangle &= \langle \lambda \cdot v, v \rangle = (A \cdot v)^t \cdot \overline{v} = v^t \cdot A^t \cdot \overline{v}.
\end{aligned}$$

Da A hermitesch ist, gilt $\overline{A} = A^t$, also $\lambda \cdot \langle v,v \rangle = \overline{\lambda} \cdot \langle v,v \rangle$. Da $\langle v,v \rangle \neq 0$, folgt $\lambda = \overline{\lambda}$. □

Wir kommen nun zu einem für zahlreiche Anwendungen wichtigen Ergebnis.

16.11 Satz: Sei $A \in \mathbb{K}^{n,n}$ eine symmetrische $(\mathbb{K} = \mathbb{R})$ bzw. hermitesche Matrix $(\mathbb{K} = \mathbb{C})$. Dann besitzt $\mathbb{K}^n$ eine bzgl. des Standardskalarproduktes $\langle -,- \rangle_{\mathbb{K}}$ orthonormale Basis aus Eigenvektoren von A. Insbesondere ist A diagonalisierbar.

Beweis: Sei $f = f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$. Wir zeigen induktiv:

Jeder k-dimensionale f-invariante Untervektorraum U von $\mathbb{K}^n$ besitzt eine Orthonormalbasis von Eigenvektoren von A. Mit $k = n$ ist der Satz bewiesen.

$\underline{k = 1}$: $U = \operatorname{Span}\{u\}$. Da U f-invariant ist, ist u Eigenvektor von f, und $\frac{u}{\|u\|}$ hat die Länge 1.

Induktionsschritt: Sei U k-dimensionaler f-invarianter Untervektorraum von $\mathbb{K}^n$. Dann können wir die Einschränkung g von f auf U als Endomorphismus von U auffassen:

$$g := f|U : U \longrightarrow U, \quad u \longmapsto f(u).$$

Jeder Eigenwert von g ist natürlich auch Eigenwert von f.

Nach dem Fundamentalsatz der Algebra (13.12) hat das charakteristische Polynom von g mindestens eine (komplexe) Nullstelle, also g einen komplexen Eigenwert λ. Nach (16.10) ist λ sogar reell. Sei $u \in U$ ein Eigenvektor zu λ. Wir wählen u so, dass $\|u\| = 1$. Sei V das orthogonale Komplement von u in U, d.h. $U = \mathbb{K} \cdot u \oplus V$ mit

$$V = \{v \in U;\ \langle v,u \rangle = 0\}.$$

Dann gilt $\dim V = k - 1$. Weiter gilt für jedes $v \in V$

$$\begin{aligned}
\langle f(v),u \rangle &= \langle A \cdot v,u \rangle = (A \cdot v)^t \cdot \overline{u} = v^t \cdot A^t \cdot \overline{u} = v^t \cdot \overline{A} \cdot \overline{u} \\
&= v^t \cdot \overline{A \cdot u} = v^t \cdot \overline{\lambda \cdot u} = \lambda \cdot \langle v,u \rangle = 0.
\end{aligned}$$

Also ist $f(v) \in V$ und damit V f-invariant.

Nach Induktionsannahme besitzt V eine Orthonormalbasis $\{v^1, \dots, v^{k-1}\}$ aus Eigenvektoren. Damit ist $\{v^1, \dots, v^{k-1}, u\}$ eine Orthonormalbasis von U aus Eigenvektoren. $\square$

Für beliebige endlichdimensionale Vektorräume V gilt ein ähnlicher Satz. Da wir auf V aber kein kanonisches Skalarprodukt haben, macht es wenig Sinn von Orthonormalbasen zu sprechen. Genauer: Zu jeder Basis $\mathscr{B}$ von V gibt es ein Skalarprodukt, bzgl. dem $\mathscr{B}$ Orthonormalbasis ist.

16.12 Die HAUPTACHSENTRANSFORMATION: Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum, $f : V \times V \longrightarrow \mathbb{K}$ eine symmetrische Bilinearform ($\mathbb{K} = \mathbb{R}$) bzw. hermitesche Sesquilinearform ($\mathbb{K} = \mathbb{C}$). Dann gilt:

(1) Für alle Basen $\mathscr{B}$ von V haben die Matrizen $Q_{\mathscr{B}}(f)$

   (a) denselben Rang

   (b) dieselbe Anzahl p von Eigenwerten $\lambda > 0$

   (c) dieselbe Anzahl q von Eigenwerten $\lambda < 0$

(2) V besitzt eine Basis $\mathscr{B}$, so dass $Q_{\mathscr{B}}(f)$ folgende Diagonalgestalt hat:

$$Q_{\mathscr{B}}(f) = \left(\begin{array}{ccccccc} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & 0 & \\ & & & -1 & & & \\ & & & & \ddots & & \\ & & & & & -1 & \\ & 0 & & & & 0 & \\ & & & & & & \ddots \\ & & & & & & & 0 \end{array}\right) \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\ \ddots \\ 1\end{array}}\right\} p \\ \left.\vphantom{\begin{array}{c}1\\ \ddots \\ 1\end{array}}\right\} q \\ \left.\vphantom{\begin{array}{c}1\\ \ddots \\ 1\end{array}}\right\} n-p-q \end{array} = \begin{pmatrix} E_p & 0 & 0 \\ 0 & -E_q & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Beweis: Sei $\mathscr{B}' = \{v^1, \dots, v^n\}$ Basis von V und $\hat{\mathscr{B}}$ eine weitere Basis. Mit $T = M^{\mathscr{B}'}_{\hat{\mathscr{B}}}(id)$ gilt $Q_{\hat{\mathscr{B}}}(f) = T^t \cdot Q_{\mathscr{B}'}(f) \cdot \overline{T}$ nach (16.6). Da T invertierbar ist, sind auch T^t und $\overline{T}$ invertierbar. Also haben $Q_{\hat{\mathscr{B}}}(f)$ und $Q_{\mathscr{B}'}(f)$ denselben Rang.

Sei nun $A = Q_{\mathscr{B}'}(f)$. Nach (16.11) gibt es eine Orthonormalbasis $\{z^1, \dots, z^n\} = \mathfrak{M}$ von $\mathbb{K}^n$ (bzgl. des Skalarprodukts $\langle -, - \rangle_{\mathbb{K}}$) aus Eigenvektoren von A. Nach dem Fundamentalsatz der Algebra (13.12) gilt für das charakteristische Polynom

$$P_A(x) = \mathrm{Det}(x \cdot E_n - A) = \prod_{i=1}^n (x - \lambda_i),$$

wobei $\lambda_1, \ldots, \lambda_n$ die (nicht notwendig verschiedenen, aber reellen) Eigenwerte von A sind. Wir ordnen die Eigenwerte so, dass $\lambda_1, \ldots, \lambda_p > 0$, $\lambda_{p+1}, \ldots, \lambda_{p+q} < 0$ und $\lambda_{p+q+1}, \ldots, \lambda_n = 0$. Weiter sei z^i Eigenvektor zu λ_i.

Die Matrix $Z = (z_{ij})$ mit den Spaltenvektoren $z^1, \ldots, z^n$ hat den Rang n, ist daher invertierbar. Damit ist

$$\mathscr{B} = \{w^1, \ldots, w^n\}$$

mit $w^j = \sum_{i=1}^n \overline{z_{ij}} \cdot v^i$ ebenfalls eine Basis und $\overline{Z} = M_{\mathscr{B}}^{\mathscr{B}'}(id_V)$. Also ist $\overline{z}$ der Koordinatenvektor von w^j bzgl. $\mathscr{B}'$, und wir erhalten aus (16.1)

$$\begin{aligned} f(w^i, w^j) &= \overline{z^i}^t \cdot A \cdot \overline{\overline{z^j}} = \overline{z^i}^t \cdot A \cdot z^j = \lambda_j \cdot \overline{z^i}^t \cdot z^j \\ &= \lambda_j \cdot \langle z^j, z^i \rangle_{\mathbb{K}} = \lambda_j \delta_{ij} \end{aligned} \qquad (*)$$

Damit hat $Q_{\mathscr{B}}(f)$ Diagonalgestalt. Ersetze nun w^i durch $u_i = \frac{1}{\sqrt{|\lambda_i|}} w^i$ für $i \leq p+q$. Dann ist $\{u^1, \ldots, u^{p+q}, w^{p+q+1}, \ldots, w^n\}$ die gesuchte Basis $\mathscr{B}$ bzgl. der $Q_{\mathscr{B}}(f)$ die in Teil (2) geforderte Gestalt hat.

Es bleibt der Nachweis von (1. (a), (b)): Wir setzen $V^+ = \text{Span}(w^1, \ldots, w^p)$, $V^- = \text{Span}(w^{p+1}, \ldots, w^{p+q})$ und $V^0 = \text{Span}(w^{p+q+1}, \ldots, w^n)$.

Sei nun $v = \sum_{i=1}^p x_i w^i \in V^+$, $v \neq 0$. Dann gilt wegen $(*)$

$$f(v,v) = \sum_{i=1}^{p} \sum_{j=1}^{p} x_i \cdot \overline{x_j} \underbrace{f(w^i, w^j)}_{\lambda_j \delta_{ij}} = \sum_{j=1}^p x_j \overline{x_j} \cdot \lambda_j = \sum_{j=1}^p \lambda_j |x_j|^2 > 0.$$

Analog gilt für $v \neq 0$ aus $V^- : f(v,v) < 0$.

Für V^0 können wir mehr zeigen:

Behauptung: $V^0 = \{v \in V;\ f(v,w) = 0 \quad \forall\, w \in V\}$

Beweis: Sei $v = \sum_{j=1}^n x_i \cdot w^i$ und $w = \sum_{j=1}^n y_j w^j$. Dann gilt nach $(*)$

$$f(v,w) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \overline{y_j} \cdot \underbrace{f(w^i, w^j)}_{\lambda_j \delta_{ij}} = \sum_{i=1}^{n} x_i \cdot \overline{y_i} \cdot \lambda_i = \sum_{i=1}^{p+q} x_i \cdot \overline{y_i} \lambda_i,$$

da $\lambda_i = 0$ für $i > p+q$. Es folgt:

$v \in V^0 \Longleftrightarrow x_i = 0$ für $1 \leq i \leq p+q \Longrightarrow f(v,w) = 0$.

Sei umgekehrt $f(v,w) = 0 \quad \forall\, w \in V$, dann gilt insbesondere $f(v, w^i) = 0$ für $i = 1, \ldots, n$. Aber

$$0 = f(v, w^i) = x_i \cdot \lambda_i.$$

Da $\lambda_i \neq 0$ für $i \leq p+q$, folgt $x_i = 0$ für $i \leq p+q$, also $v \in V^0$.

Aus (12.13.2) wissen wir

$$V = V^+ \oplus V^- \oplus V^0, \quad \dim V^+ = p, \quad \dim V^- = q, \quad \operatorname{rg} Q_{\mathscr{B}}(f) = p + q.$$

Für V^0 haben wir eine von der Basis $\mathscr{B}$ unabhängige Charakterisierung gefunden. Ist also $\hat{\mathscr{B}}$ eine andere Basis, erhalten wir eine entsprechende Zerlegung mit demselben V^0:

$$V = \hat{V}^+ \oplus \hat{V}^- \oplus V^0, \quad \dim \hat{V}^+ = \hat{p}, \ \dim \hat{V}^- = \hat{q}$$

mit $p + q = \hat{p} + \hat{q}$. Können wir zeigen, dass $p = \hat{p}$, ist der Beweis erbracht. Sei $\hat{v} \in \hat{V}^+ \cap (V^- \oplus V^0)$. Dann hat $\hat{v}$ eine Darstellung

$$\hat{v} = v_- + v_0 \ \text{ mit } v_- \in V^-, \ v_0 \in V^0.$$

Da $\hat{v} \in \hat{V}^+$, gilt $f(\hat{v},\hat{v}) \geq 0$. Aber aus der Charakterisierung von V^0 folgt

$$\begin{aligned} f(\hat{v},\hat{v}) &= f(v_- + v_0, v_- + v_0) = f(v_-,v_-) + f(v_0,v_-) + f(v_-,v_0) + f(v_0,v_0) \\ &= f(v_-,v_-) \leq 0. \end{aligned}$$

Es folgt $f(\hat{v},\hat{v}) = 0$. Da f auf $\hat{V}^+$ positiv definit ist, folgt $\hat{v} = 0$. Also ist die Summe $\hat{V}^+ + (V^- \oplus V^0)$ direkt, so dass $\hat{p} + q + \dim V^0 \leq n$. Da $p + q + \dim V^0 = n$, folgt $\hat{p} \leq p$. Vertauschen wir die Rollen von $\hat{V}^\pm$ und $V^\pm$, erhalten wir analog $p \leq \hat{p}$, so dass $p = \hat{p}$. □

Für $V = \mathbb{K}^n$ liefert der Beweis von (16.12)

16.13 Ergänzung: Ist $V = \mathbb{K}^n$, dann gibt es eine Ortho**normal**basis $\mathscr{B} = \{v^1, \dots, v^n\}$ von $\mathbb{K}^n$, so dass

$$Q_{\mathscr{B}}(f) = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & 0 \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \qquad \begin{array}{l} \lambda_1, \dots, \lambda_p > 0 \\ \lambda_{p+1}, \dots, \lambda_{p+q} < 0 \\ \lambda_{p+q+1} = \dots = \lambda_n = 0 \end{array}$$

Dabei ist jedes $\vec{v}$ Eigenvektor der Matrix $A = (a_{ij})$ mit $a_{ij} = f(e^i, e^j)$ und λ_i der zugehörige (reelle) Eigenwert. Ersetzen wir $\vec{v}$ durch $\frac{1}{\sqrt{|\lambda_i|}}\vec{v}$ für $1 \leq i \leq p+q$, erhalten wir eine Ortho**gonal**basis $\mathscr{B}'$ von $\mathbb{K}^n$, so dass

$$Q_{\mathscr{B}'}(f) = \left(\begin{array}{c:c:c} E_p & & 0 \\ \hdashline & -E_q & \\ \hdashline 0 & & 0 \end{array}\right)$$

Beweis: Wir nehmen im Beweis von (16.12) für $\mathscr{B}'$ die Standardbasis. Sei wieder $\{z^1, \dots, z^n\}$ eine Orthonormalbasis bzgl. des Standardskalarproduktes von Eigenvektoren von A. Wir ordnen sie wie im Beweis von (16.12). Dann ist $\mathscr{B} = \{\overline{z}^1, \dots, \overline{z}^n\}$ die gesuchte Orthonormalbasis. Da A die Matrix von f bzgl. der Standardbasis ist, gilt für die (i, j)-Komponente von $Q_{\mathscr{B}}(f)$ wie in der Gleichung $(*)$ des Beweises von (16.12).

$$f(\overline{z}^i, \overline{z}^j) = (\overline{z}^i)^t \cdot A \cdot \overline{\overline{z}}_j = \lambda_j \delta_{ij}$$

$\square$

16.14 Folgerung: Die Form f im Satz 16.12 ist genau dann positiv definit, wenn $Q_{\mathscr{B}}(f)$ nur Eigenwerte > 0 besitzt. Dabei ist $\mathscr{B}$ eine beliebige Basis von V.

Beweis: f positiv definit $\Longleftrightarrow V = V^+$ im Beweis (16.12) $\Longleftrightarrow p = n$. $\square$

Ergänzung: Negativ definite und indefinite Formen

In der Praxis sind neben positiv definiten Formen auch negativ definite und indefinite Formen von Bedeutung.

16.15 Definition: Eine hermitesche Sesquilinearform ($\mathbb{K} = \mathbb{C}$) bzw. symmetrische Bilinearform ($\mathbb{K} = \mathbb{R}$) $f : V \times V \to \mathbb{K}$ heißt **NEGATIV DEFINIT**, wenn $f(v, v) < 0$ für alle $v \neq 0$ aus V. Wir nennen f **INDEFINIT**, wenn es Vektoren $v,\ w \in V$ gibt, so dass $f(v, v) > 0$ und $f(w, w) < 0$.
Eine hermitesche ($\mathbb{K} = \mathbb{C}$) bzw. symmetrische ($\mathbb{K} = \mathbb{R}$) Matrix $A \in \mathbb{K}^{n,n}$ heißt **NEGATIV DEFINIT** bzw. **INDEFINIT**, wenn die Sesquilinearform (Bilinearform)

$$\begin{array}{ll} \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C} \qquad & (x, y) \longmapsto x^t \cdot A \cdot \overline{y} \\ (\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}) & \end{array}$$

negativ definit bzw. indefinit ist, d.h. wenn für $x \neq 0$ gilt $x^t \cdot A \cdot \overline{x} < 0$.

Negativ und indefinite Formen haben eine zu (16.14) analoge Charakterisierung (der Beweis verläuft analog):

16.16 Satz: Eine hermitesche Sesquilinearform bzw. symmetrische Bilinearform f ist genau dann

(1) negativ definit, wenn $Q_{\mathscr{B}}(f)$ nur Eigenwerte < 0 besitzt.

(2) indefinit, wenn $Q_{\mathscr{B}}(f)$ sowohl Eigenwerte > 0 als auch Eigenwerte < 0 besitzt.

16.17 Definition: Eine indefinite hermitesche Sesquilinearform oder symmetrische Bilinearform heißt **NICHT ENTARTET**, wenn $Q_{\mathscr{B}}(f)$ invertierbar ist.

Für negativ definite und nicht entartete indefinite Formen haben wir ein Hauptminorenkriterium analog zu (16.9).

16.18 Satz: Sei $f : V \times V \longrightarrow \mathbb{K}$ eine symmetrische Bi- bzw. hermitesche Sesquilinearform auf dem n-dimensionalen $\mathbb{K}$-Vektorraum V und $\mathscr{B}$ eine Basis von V. Sei $A = Q_{\mathscr{B}}(f)$. Dann gilt mit den Bezeichnungen aus (16.9):

(1) f negativ definit $\Longleftrightarrow A$ negativ definit $\Longleftrightarrow \mathrm{Det} A_k < 0$ für ungerade k und $\mathrm{Det} A_k > 0$ für gerade k, $k = 1, \dots, n$.

(2) f nicht entartet indefinit $\Longleftrightarrow A$ nicht entartet indefinit $\Longleftrightarrow \mathrm{Det} A_k \neq 0$ für $k = 1, \dots, n$, und weder der Fall (1) noch (16.9) liegt vor.

Beweis: Ist f ist genau dann negativ definit, wenn $g = -f$ positiv definit ist. Letzteres ist nach (16.9) genau dann der Fall, wenn $Q_{\mathscr{B}}(g) = Q_{\mathscr{B}}(-f) = -A$ positiv definit ist, wenn also $\mathrm{Det}(-A_k) = (-1)^k \mathrm{Det} A_k > 0$ für $k = 1, \dots, n$ gilt. Damit folgt die Behauptung.

Den zweiten Teil überlassen wir dem Leser als Übungsaufgabe. $\square$

17 Orthogonale und unitäre Abbildungen

$(V_1, \langle -,- \rangle_1)$ und $(V_2, \langle -,- \rangle_2)$ seien $\mathbb{K}$-Vektorräume mit Skalarprodukt, $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$. Es seien $\| - \|_1$ und $\| - \|_2$ die zugehörigen Normen. Wir wollen skalarprodukterhaltende Abbildungen studieren.

17.1 Definition: Eine Abbildung $f : V_1 \longrightarrow V_2$ heißt **ISOMETRISCH**, wenn

$$\langle f(v), f(w) \rangle_2 = \langle v, w \rangle_1 \quad \forall\, v, w \in V_1.$$

Ist f außerdem bijektiv, nennen wir f eine **ISOMETRIE**.

Wenn wir Abbildungen dieser Art studieren, sollten sie die gesamte Struktur erhalten, also auch die Vektorraumstrukturen. D.h. wir sollten genauer fordern, dass f $\mathbb{K}$-linear ist. Das folgt aber bereits aus der Bedingung, dass f Skalarprodukte erhält.

17.2 Satz: Jede isometrische Abbildung $f : V_1 \longrightarrow V_2$ ist $\mathbb{K}$-linear und injektiv.

Beweis: Für $v, w \in V_1$ und $r \in \mathbb{K}$ gilt, da f Skalarprodukte erhält

$$\begin{aligned}
0 &\leq \langle f(rv+w) - r \cdot f(v) - f(w), f(rv+w) - r \cdot f(v) - f(w) \rangle_2 \\
&= \langle f(rv+w), f(rv+w) \rangle_2 - \overline{r} \langle f(rv+w), f(v) \rangle_2 - \langle f(rv+w), f(w) \rangle_2 \\
&\quad - r \langle f(v), f(rv+w) \rangle_2 + r \cdot \overline{r} \langle f(v), f(v) \rangle_2 + r \langle f(v), f(w) \rangle_2 - \langle f(w), \\
&\quad f(rv+w) \rangle_2 + \overline{r} \langle f(w), f(v) \rangle_2 + \langle f(w), f(w) \rangle_2 \\
&= \langle rv+w, rv+w \rangle_1 - \overline{r} \langle rv+w, v \rangle_1 - \langle rv+w, w \rangle_1 - r \langle v, rv+w \rangle_1 + r\,\overline{r} \langle v, v \rangle_1 \\
&\quad + r \langle v, w \rangle_1 - \langle w, rv+w \rangle_1 + \overline{r} \langle w, v \rangle_1 + \langle w, w \rangle_1 \\
&= \langle rv+w-rv-w, rv+w \rangle_1 - \overline{r} \langle rv+w-rv-w, v \rangle_1 - \langle rv+w-rv-w, w \rangle_1 \\
&= 0
\end{aligned}$$

Da $\langle -,- \rangle$ positiv definit ist, folgt $f(rv+w) = r \cdot f(v) + f(w)$, d.h. f ist $\mathbb{K}$-linear. Ist $v \neq 0$, folgt $\langle f(v), f(v) \rangle_2 = \langle v, v \rangle_1 > 0$. Also ist $f(v) \neq 0$ und damit f injektiv.

$\square$

17.3 Folgerung: Eine isometrische Abbildung $f : V_1 \longrightarrow V_2$ mit $\dim V_1 = \dim V_2 < \infty$ ist stets bijektiv, also eine Isometrie. $\square$

Da eine isometrische Abbildung f Skalarprodukte erhält, erhält f Längen und Winkel. Weiß man bereits, dass f linear ist, folgt aus der Längentreue auch die Skalarprodukttreue.

17.4 Satz: Für eine $\mathbb{K}$-lineare Abbildung $f : V_1 \longrightarrow V_2$ sind äquivalent:

(1) f ist **SKALARPRODUKTTREU**, d.h. $\langle f(v), f(w)\rangle_2 = \langle v,w\rangle_1 \quad \forall\, v,w \in V_1$

(2) f ist **NORM-** oder **LÄNGENTREU**, d.h. $\|f(v)\|_2 = \|v\|_1 \quad \forall\, v \in V_1$

(3) f ist **UNITÄR**, d.h. $\|f(v)\|_2 = 1 \quad \forall\, v \in V_1$ mit $\|v\|_1 = 1$.

Beweis: (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) ist trivial.
(3) $\Longrightarrow$ (2): Ist $v = 0$, so ist wegen der Linearität auch $f(v) = 0$, also $\|f(v)\|_2 = 0 = \|v\|_1$.
Ist $v \neq 0$, so ist $w = \frac{v}{\|v\|_1}$ ein **Einheitsvektor**, d.h. $\|w\| = 1$. Es folgt

$$\|f(v)\|_2 = \|v\|_1 \cdot \|f(w)\|_2 = \|v\|_1 \text{ (wieder wurde die Linearität benutzt).}$$

(2) $\Longrightarrow$ (1) folgt aus der Tatsache, dass sich Skalarprodukte mit Hilfe der Norm berechnen lassen und dass f linear ist:

17.5 Lemma: (1) Für $\mathbb{R}$-Vektorräume V mit Skalarprodukt $\langle -,-\rangle$ gilt

$$\langle v,w\rangle = \frac{1}{2}(\|v+w\|^2 - (\|v\|^2 + \|w\|^2)) \quad \forall\, v,w \in V.$$

(2) Für $\mathbb{C}$-Vektorräume V mit Skalarprodukt $\langle -,-\rangle$ gilt

$$\langle v,w\rangle = \frac{1}{2}(\|v+w\|^2 + i\|v+iw\|^2 - (1+i)(\|v\|^2 + \|w\|^2)) \quad \forall v,w \in V.$$

Beweis: Über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ gilt für $v,w \in V$

$$\begin{aligned}\|v+w\|^2 &= \langle v+w, v+w\rangle = \langle v,v\rangle + \langle v,w\rangle + \overline{\langle v,w\rangle} + \langle w,w\rangle\\ &= \|v\|^2 + \|w\|^2 + 2\,\mathrm{Re}\langle v,w\rangle \qquad\qquad (*)\end{aligned}$$

Für $\mathbb{K} = \mathbb{R}$ ist $\mathrm{Re}\langle v,w\rangle = \langle v,w\rangle$, und wir erhalten Aussage (1).
Im Falle $\mathbb{K} = \mathbb{C}$ wenden wir die Formel $(*)$ auf $i \cdot w$ statt w an und erhalten

$$\|v+iw\|^2 = \|v\|^2 + |i| \cdot \|w\|^2 + 2\,\mathrm{Re}(\overline{i} \cdot \langle v,w\rangle) = \|v\|^2 + \|w\|^2 + 2\,\mathrm{Im}\langle v,w\rangle$$

da $\mathrm{Re}(\overline{i} \cdot z) = \mathrm{Re}(-i \cdot z) = \mathrm{Im}\, z \quad \forall\, z \in \mathbb{C}$. Es folgt

$$\langle v,w\rangle = \mathrm{Re}\langle v,w\rangle + i \cdot \mathrm{Im}\langle v,w\rangle = \frac{1}{2}(\|v+w\|^2 + i\|v+iw\|^2 - (1+i)(\|v\|^2 + \|w\|^2))$$

$\square$

Ein weiteres Kriterium für Skalarprodukttreue ist folgender einfacher Satz:

17.6 Satz: Sei $f : V_1 \longrightarrow V_2$ $\mathbb{K}$-lineare Abbildung.

(1) Ist f skalarprodukttreu und $\{v^1, \dots, v^n\}$ ein Orthonormalsystem in V_1, dann ist $\{f(v^1), \dots, f(v^n)\}$ ein Orthonormalsystem in V_2.

(2) Ist $\mathscr{B} = \{v^1, \dots, v^n\}$ eine Orthonormal**basis** in V_1 derart, dass $\{f(v^1), \dots, f(v^n)\}$ Orthonormal**system** in V_2 ist, so ist f skalarprodukttreu.

Beweis: (1) ist klar.
(2) Sei $v = \sum_{i=1}^n x_i \cdot v^i \in V_1$. Dann gilt

$$(\|f(v)\|_2)^2 = \left\langle f\left(\sum_{i=1}^n x_i v^i\right), f\left(\sum_{j=1}^n x_j v^j\right)\right\rangle = \left\langle \sum_{i=1}^n x_i f(v^i), \sum_{j=1}^n x_j f(v^i)\right\rangle$$
$$= \sum_{i=1}^n \sum_{j=1}^n x_i \cdot \overline{x_j} \underbrace{\langle f(v^i), f(v^j)\rangle}_{\delta_{ij}} = \sum_{i=1}^n x_i \overline{x_i}$$
$$(\|v\|_1)^2 = \left\langle \sum_{i=1}^n x_i v^i, \sum_{j=1}^n x_j v^j\right\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i \overline{x_j} \cdot \underbrace{\langle v^i, v^j\rangle}_{\delta_{ij}} = \sum_{i=1}^n x_i \overline{x_i}.$$

Also gilt: $\|f(v)\|_2 = \|v\|_1$. Nach (17.4 (2)) ist f skalarprodukttreu. □
Wegen (17.4 (3)) haben sich folgende Bezeichnungen eingebürgert:

17.7 Definition: Eine skalarprodukttreue Abbildung $f : V_1 \longrightarrow V_2$ heißt im Fall $\mathbb{K} = \mathbb{C}$ **UNITÄR** und im Fall $\mathbb{K} = \mathbb{R}$ **ORTHOGONAL**.

Wir wollen nun die Matrizen unitärer (bzw. orthogonaler) Abbildungen bzgl. gegebener **Orthonormalbasen** $\mathscr{B}_i$ von V_i untersuchen.

17.8 Definition: Eine Matrix $A \in \mathbb{K}^{m,n}$ heißt **UNITÄR** (Fall $\mathbb{K} = \mathbb{C}$) bzw. **ORTHOGONAL** (Fall $\mathbb{K} = \mathbb{R}$), falls $A^t \cdot \overline{A} = E_n$ (bzw. $A^t \cdot A = E_n$).
Im Fall $m = n$ ist dann A invertierbar und $A^{-1} = \overline{A}^t$ (bzw. $A^{-1} = A^t$).

17.9 Satz: Sei $f : V_1 \longrightarrow V_2$ eine $\mathbb{K}$-lineare Abbildung $\mathscr{B}_1 = \{v^1, \dots, v^n\}$ **Orthonormalbasis** von V_1 und $\mathscr{B}_2 = \{w^1, \dots, w^m\}$ **Orthonormalbasis** von V_2. Dann gilt:

$$f \text{ unitär bzw. orthogonal } \Longleftrightarrow M_{\mathscr{B}_1}^{\mathscr{B}_2}(f) \text{ ist unitär bzw. orthogonal.}$$

Beweis: Sei $M^{\mathscr{B}_2}_{\mathscr{B}_1}(f) = A = (a_{ij}) \in \mathbb{K}^{m,n}$. Dann gilt $f(v^j) = \sum_{i=1}^m a_{ij} w^i$. Also

$$\langle f(v^i), f(v^j)\rangle_2 = \left\langle \textstyle\sum_{k=1}^m a_{ki} w^k, \sum_{\ell=1}^m a_{\ell j} w^\ell \right\rangle_2 = \sum_{k=1}^m \sum_{\ell=1}^m a_{ki}\, \overline{a}_{\ell j} \cdot \underbrace{\langle w^k, w^\ell\rangle}_{=\delta_{k\ell}}$$

$$= \sum_{k=1}^m a^t_{ik}\, \overline{a}_{kj} = (i,j)\text{-ter Eintrag von } A^t \cdot \overline{A}.$$

Ist f unitär (orthogonal), gilt $\langle f(v^i), f(v^j)\rangle_2 = \langle v^i, v^j\rangle_1 = \delta_{ij}$, also $A^t \cdot \overline{A} = E_n$. Gilt umgekehrt $A^t \cdot \overline{A} = E_n$, folgt $\langle f(v^i), f(v^j)\rangle_2 = \delta_{ij}$. Also ist $\{f(v^1, \ldots, f(v^n)\}$ ein Orthonormalsystem. Nach (17.6 (2)) ist f skalarprodukttreu. □

Im weiteren betrachten wir unitäre (orthogonale) Endomorphismen

$$f : V \longrightarrow V$$

von **endlichdimensionalen** $\mathbb{K}$-Vektorräumen V mit Skalarprodukt $\langle -, -\rangle$. Aus (17.3) wissen wir, dass f ein Isomorphismus ist.

17.10 Lemma: Für jeden Eigenwert λ von f gilt $|\lambda| = 1$. Es folgt

(1) Fall $\mathbb{K} = \mathbb{R} : \lambda \in \{\pm 1\}$

(2) Fall $\mathbb{K} = \mathbb{C} : \lambda \in S^1 = \{z \in \mathbb{C}; |z| = 1\} = \{\cos\alpha + i \sin\alpha;\ 0 \leq \alpha < 2\pi\}$.

Beweis: Sei v Einheitseigenvektor zum Eigenwert λ. Dann gilt

$$1 = \|v\| = \|f(v)\| = \|\lambda \cdot v\| = |\lambda| \cdot \|v\| = |\lambda|.$$

□

17.11 Lemma: Ist $f : V \longrightarrow V$ unitäre (orthogonale) Abbildung und $U \subset V$ ein endlich dimensionaler f-invarianter Untervektorraum, dann ist das orthogonale Komplement $U^\perp$ ebenfalls f-invariant ($\dim V = \infty$ ist zugelassen).

Beweis: Sei $v \in U^\perp$, d.h. $\langle v, u\rangle = 0 \quad \forall\, u \in U$. Zu zeigen ist $\langle f(v), u\rangle = 0 \quad \forall\, u \in U$. Da U f-invariant ist, können wir f als Abbildung

$$f|U : U \longrightarrow U$$

auffassen. Da f injektiv und U endlichdimensional ist, ist $f|U$ ein Isomorphismus (6.29). Also gibt es zu $u \in U$ ein $u' \in U$ mit $f(u') = u$. Es folgt

$$\langle f(v), u\rangle = \langle f(v), f(u')\rangle = \langle v, u'\rangle = 0.$$

□

Wir benutzen diese Ergebnisse, um alle unitären (orthogonalen) Endomorphismen zu klassifizieren. Wir beginnen mit dem Fall $\mathbb{K} = \mathbb{C}$.

17.12 Satz: Ist $f : V \longrightarrow V$ unitärer Endomorphismus eines endlich dimensionalen $\mathbb{C}$-Vektorraumes V mit Skalarprodukt, dann besitzt V eine **Orthonormal**basis $\mathscr{B}$ aus Eigenvektoren von f. Damit hat $M_{\mathscr{B}}^{\mathscr{B}}(f)$ die Form

$$M_{\mathscr{B}}^{\mathscr{B}}(f) = \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & \lambda_n \end{pmatrix} \text{ mit } \lambda_i = \cos\alpha_i + i \sin\alpha_i$$

Beweis: Induktion nach n (vergl. Beweis 16.11).

$\boldsymbol{n = 1}$: Da V 1-dimensional ist, gilt $f(v) = \lambda \cdot v$ mit $|\lambda| = 1$.

Induktionsschritt: Sei λ ein Eigenwert von f und v ein Einheitseigenvektor zu λ. Nach dem Fundamentalsatz der Algebra existiert λ. Sei $U = (\mathbb{C} \cdot v)^{\perp}$. Nach (17.11) ist U f-invariant und besitzt nach Induktion eine Orthonormalbasis $\{v^2, \ldots, v^n\}$ aus Eigenvektoren von f. Damit ist $\{v, v^2, \ldots, v^n\}$ die gesuchte Orthonormalbasis. $\square$

Der reelle Fall ist komplizierter, denn der Fundamentalsatz der Algebra steht uns nicht zur Verfügung. Wir lösen das Problem, indem wir (17.12) benutzen. Damit wir rechnen können, übertragen wir das Problem auf einen orthogonalen Endomorphismus des $\mathbb{R}^n$ mit dem Standardskalarprodukt. Sei also

$$f : V \longrightarrow V$$

ein orthogonaler Endomorphismus des $\mathbb{R}$-Vektorraumes V und $\dim V = n$. Wir wählen eine Orthonormalbasis $\mathscr{B}' = \{v^1, \ldots, v^n\}$ von V und betrachten die orthogonale Abbildung

$$\varphi_{\mathscr{B}'} : V \longrightarrow \mathbb{R}^n, \qquad v^i \longrightarrow e^i.$$

Ist $A = M_{\mathscr{B}'}^{\mathscr{B}'}(f)$, erhalten wir ein kommutatives Quadrat

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\varphi_{\mathscr{B}'}"] & V \arrow[d, "\varphi_{\mathscr{B}'}"] \\ \mathbb{R}^n \arrow[r, "f_A"] & \mathbb{R}^n \end{tikzcd}$$

von orthogonalen Abbildungen. Wir können jetzt unsere Untersuchungen auf den Spezialfall $(\mathbb{R}^n, \langle -,- \rangle_{\mathbb{R}})$ beschränken: Ist $\mathscr{B} = \{\vec{w}^1, \ldots, \vec{w}^n\}$ eine Orthonormalbasis von $\mathbb{R}^n$, bzgl. der die Matrix von f_A eine spezielle Form hat, dann ist $\mathscr{B} = \{w^1, \ldots, w^n\}$ mit $w^i = \varphi_{\mathscr{B}'}^{-1}(\vec{w}^i)$ ein Orthonormalbasis von V, für die $M_{\mathscr{B}}^{\mathscr{B}}(f)$ dieselbe spezielle Matrix ist.

Das Transformationsverhalten wird in folgendem Diagramm verdeutlicht:

$$\begin{tikzcd}
V \arrow[r, "id"] \arrow[d, "\varphi_{\mathscr{B}}"] & V \arrow[r, "f"] \arrow[d, "\varphi_{\mathscr{B}'}"] & V \arrow[r, "id"] \arrow[d, "\varphi_{\mathscr{B}'}"] & V \arrow[d, "\varphi_{\mathscr{B}}"] \\
\mathbb{R}^n \arrow[r, "M^{\mathscr{L}}_{\bar{\mathscr{B}}}(id)\cdot -"] & \mathbb{R}^n \arrow[r, "A\cdot -"] & \mathbb{R}^n \arrow[r, "M^{\bar{\mathscr{B}}}_{\mathscr{S}}(id)"] & \mathbb{R}^n
\end{tikzcd}$$

$$\begin{tikzcd}
e^i \arrow[r] & \vec{w}^i \arrow[r] & f_A(\vec{w}^i)
\end{tikzcd}$$

$M^{\bar{\mathscr{B}}}_{\mathscr{S}}(id) \cdot A \cdot M^{\mathscr{L}}_{\bar{\mathscr{B}}}(i) = M^{\mathscr{B}}_{\mathscr{B}}(f)$ hat spezielle Form.
Hierbei haben wir folgende evidente Fakten benutzt:

17.13 (1) Die Komposition isometrischer Abbildungen ist isometrisch.
(2) Ist f ein isometrischer Isomorphismus, so ist f^{-1} isometrisch.

17.14 Satz: Ist $f : V \longrightarrow V$ orthogonale Abbildung eines n-dimensionalen $\mathbb{R}$-Vektorraumes V mit Skalarprodukt, dann besitzt V eine **Orthonormalbasis** $\mathscr{B}$, so dass $M^{\mathscr{B}}_{\mathscr{B}}(f)$ folgende Blockdiagonalgestalt hat:

$$M^{\mathscr{B}}_{\mathscr{B}}(f) = \begin{pmatrix} E_m & & & & \\ & -E_\ell & & 0 & \\ & & A_1 & & \\ 0 & & & \ddots & \\ & & & & A_k \end{pmatrix},$$

wobei A_i von der Form $A_i = \begin{pmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{pmatrix}$ ist. Also $m + \ell + 2k = n$.

Beweis durch Induktion nach n: Nach unseren Überlegungen brauchen wir nur den Fall $(\mathbb{R}^n, \langle -, - \rangle_{\mathbb{R}})$ behandeln.
$\boldsymbol{n = 1:}$ $f(e^1) = \lambda \cdot e^1$ mit $\lambda \in \{\pm 1\}$.
Induktionsschritt: Ist λ ein Eigenwert von f, können wir wie im Beweis von (17.12) vorgehen. Es bleibt nur der Fall zu betrachten, dass f **keinen** Eigenwert hat. Sei A die Matrix von $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ bzgl. der Standardbasis. Es gilt $A^t \cdot A = E_n$, da A orthogonal ist. Wir fassen A als **komplexe** Matrix auf. Da $A = \bar{A}$, ist A unitär und

$$f_A : \mathbb{C}^n \longrightarrow \mathbb{C}^n, \quad x \longmapsto A \cdot x$$

ein unitärer Endomorphismus bzgl. $\langle -,- \rangle_{\mathbb{C}}$. Sei $\lambda = \alpha + i\beta \in \mathbb{C}$ ein Eigenwert von A und

$$v = (z_1, \ldots, z_n) = (x_1 + iy_1, \ldots, x_n + iy_n) = x + iy$$

ein Eigenvektor zu λ. Wir wählen v so, dass $\|v\|_{\mathbb{C}}^2 = 2$. Da f_A nach Annahme keinen reellen Eigenwert hat und $|\lambda| = 1$ nach (17.10), gilt

$$\alpha^2 + \beta^2 = 1 \quad \text{und} \quad \beta \neq 0.$$

Behauptung: (1) $U := \operatorname{Span}\{x, y\} \subset \mathbb{R}^n$ ist f-invariant.

(2) $\{x, y\}$ ist eine Orthonormalbasis von U.

(3) $f|U : U \longrightarrow U$ hat bzgl. der Basis $\{x, y\}$ eine Matrix der Form

$$\begin{pmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{pmatrix}.$$

Beweis:

(1) Da $A \cdot v = \lambda \cdot v$, gilt

$$\begin{aligned} A \cdot (x + iy) &= A \cdot x + i \cdot A \cdot y = \lambda \cdot v = (\alpha + i\beta) \cdot (x + iy) \\ &= \alpha \cdot x - \beta \cdot y + i \cdot (\beta \cdot x + \alpha \cdot y) \end{aligned}$$

Also

$$A \cdot x = \alpha x - \beta \cdot y \in U \quad \text{und} \quad A \cdot y = \beta \cdot x + \alpha \cdot y \in U.$$

(2) Da A skalarprodukttreu ist, gilt (wir schreiben kürzer $\langle -,- \rangle$ für $\langle -,- \rangle_{\mathbb{R}}$).

(i) $(\alpha^2 + \beta^2) \cdot \|x\|^2 = \|x\|^2 = \|Ax\|^2 = \langle Ax, Ax \rangle = \alpha^2 \|x\|^2 - 2\alpha\beta \langle x, y \rangle + \beta^2 \|y\|^2$. Also

$$-2 \cdot \alpha \cdot \beta \langle x, y \rangle = \beta^2 \cdot (\|x\|^2 - \|y\|^2)$$

(ii) $(\alpha^2 + \beta^2) \cdot \langle x, y \rangle = \langle x, y \rangle = \langle Ax, Ay \rangle \overset{(1)}{=} \langle \alpha x - \beta y, \beta x + \alpha y \rangle$
$= \alpha\beta \|x\|^2 - \alpha\beta \|y\|^2 + (\alpha^2 - \beta^2) \langle x, y \rangle$.
Wir erhalten $2 \cdot \beta^2 \cdot \langle x, y \rangle = \alpha\beta \cdot (\|x\|^2 - \|y\|^2)$, und da $\beta \neq 0$,

$$2 \cdot \beta \cdot \langle x, y \rangle = \alpha \cdot (\|x\|^2 - \|y\|^2).$$

Addieren wir das α-fache des Ergebnisses (ii) zum Ergebnis (i), erhalten wir

$$0 = (\alpha^2 + \beta^2) \cdot (\|x\|^2 - \|y\|^2), \quad \text{so dass} \quad \|x\|^2 = \|y\|^2.$$

Da $\beta \neq 0$, folgt dann aus (ii), dass $\langle x,y\rangle = 0$.
Da $x,\ y \in \mathbb{R}^n$, haben wir weiter
$2 = \|v\|_{\mathbb{C}}^2 = \langle x+iy, x+iy\rangle_{\mathbb{C}} = \|x\|_{\mathbb{C}}^2 + i\cdot\overline{i}\|y\|_{\mathbb{C}}^2 + i\langle x,y\rangle_{\mathbb{C}} + \overline{i}\langle x,y\rangle_{\mathbb{C}}$
$= \|x\|^2 + \|y\|^2 + i\langle x,y\rangle + i\langle x,y\rangle = \|x\|^2 + \|y\|^2 = 2\cdot\|x\|^2$.
Es folgt $\|x\|^2 = \|y\|^2 = 1$. Also ist $\{x,y\}$ eine Orthonormalbasis von U.

(3) Nach (1) ist die Matrix der Einschränkung von f_A auf U bzgl. $\{x,y\}$:

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

Da $|\alpha| < 1$, gibt es ein $\varphi \in [0,2\pi[$ mit $\cos\varphi = \alpha$. Dann ist

$$\beta^2 = 1-\alpha^2 = 1-\cos^2\varphi = \sin^2\varphi.$$

Ist $\beta = \sin\varphi$, nehmen wir $\gamma = \varphi$, ist $\beta = -\sin\varphi$, nehmen wir $\gamma = -\varphi$ und erhalten eine (2×2)-Matrix, wie im Satz gefordert.

Nach (17.11) ist $U^\perp$ f-invariant. Da $\dim U^\perp = n-2$, hat $U^\perp$ nach Induktion eine Basis $\{v^1,\dots,v^{n-2}\}$ der angegebenen Form. Also ist $\mathscr{B} = \{v^1,\dots,v^{n-2},x,y\}$ die gesuchte Basis. $\square$

17.15 Ergänzung: (1) Sind $\mathscr{B}$ und $\mathscr{B}'$ Orthonormalbasen, dann ist die Transformationsmatrix $M_{\mathscr{B}}^{\mathscr{B}'}(id_V)$ unitär (bzw. orthogonal).

(2) Für jede orthogonale oder unitäre Matrix A gilt $|\operatorname{Det}(A)| = 1$.

Beweis: (1) $id: V \longrightarrow V$ ist unitär, also auch $M_{\mathscr{B}}^{\mathscr{B}'}(id)$ nach (17.9).

(2) Aus $A^t\cdot\overline{A} = E_n$ folgt: $1 = \operatorname{Det}(E_n) = \operatorname{Det}(A^t)\cdot\operatorname{Det}(\overline{A}) = \operatorname{Det}(A)\cdot\overline{\operatorname{Det}(A)} = |\operatorname{Det}(A)|^2$. $\square$

Sei $O(n)$ die Menge aller orthogonalen und $U(n)$ die aller unitären n-quadratischen Matrizen. Da die Komposition linearer Abbildungen der Multiplikation von Matrizen entspricht, ist das Produkt zweier orthogonaler bzw. zweier unitärer Matrizen wieder orthogonal. Da jede isometrische Abbildung $\mathbb{K}^n \longrightarrow \mathbb{K}^n$ nach (17.3) und (17.13) eine isometrische Umkehrabbildung besitzt, besitzt eine orthogonale (unitäre) Matrix eine orthogonale (unitäre) inverse Matrix. Damit definiert die Matrizenmultiplikation auf $O(n)$ und $U(n)$ eine sog. **GRUPPENSTRUKTUR**.

17.16 Definition: Eine **GRUPPE** G ist eine Menge mit einer „Verknüpfung", d.h. Abbildung

$$G\times G \longrightarrow G, \quad (a,b) \longmapsto a*b,$$

so dass folgende Axiome gelten:

(1) Assoziativität: $a*(b*c)=(a*b)*c \quad \forall\, a,b,c \in G$

(2) Neutrales Element: Es gibt ein Element $e \in G$, so dass $a*e=a \quad \forall\, a \in G$

(3) Inverse Elemente: Zu jedem $a \in G$ gibt es ein $\bar{a} \in G$, so dass $a*\bar{a}=e$.

Gilt außerdem

$$a*b=b*a \quad \forall\, a,b \in G$$

heißt G **ABELSCH** oder **KOMMUTATIV**.

17.17 Satz: $O(n)$ und $U(n)$ mit der Komposition als Verknüpfung sind Gruppen mit E_n als neutralem Element. □

17.18 Weitere Beispiele:

(1) V $\mathbb{K}$-Vektorraum $\Longrightarrow (V,+)$ ist abelsche Gruppe.

(2) $\mathbb{K}$ Körper $\Longrightarrow (\mathbb{K}\backslash\{0\},\cdot)$ ist Gruppe.

(3) V Vektorraum. Sei $Gl(V)$ die Menge der linearen **Isomorphismen** $f:V \longrightarrow V$. Dann ist $Gl(V)$ mit der Komposition als Verknüpfung eine Gruppe.

17.19 Definition: $O(n)$ und $U(n)$ nennt man **ORTHOGONALE** bzw. **UNITÄRE** Gruppe.

17.20 Satz und Definition: $SO(n)=\{A \in O(n); \mathrm{Det}A=1\}$ ist eine **UNTERGRUPPE** von $O(n)$, genannt **SPEZIELLE ORTHOGONALE GRUPPE**.
Wir haben zu zeigen, dass $E_n \in SO(n)$ ist, dass mit A,B auch $A \cdot B$ in $SO(n)$ ist und mit A auch A^{-1} in $SO(n)$ liegt, denn dann ist $SO(n)$ mit der Komposition als Verknüpfung eine Gruppe. Der Nachweis ist trivial. □

Da jede orthogonale Matrix als unitäre Matrix aufgefasst werden kann, erhalten wir folgende Sequenz von Untergruppen:

$$SO(n) \subset O(n) \subset U(n) \subset Gl(\mathbb{C}^n).$$