Mathematik für Physiker II
SoSe 2007

nach Prof. Dr. Frank Loose
Katharina von Sturm, Vanessa Graber, Pascal Uter,
Konstantin Sering, Christoph Zimmermann, Matthias Körber

email: koerber.matthias@web.de

Version: 0.63

Inhaltsverzeichnis

1 Vektorräume

1.1 Definition: Gruppe

Ein Paar $(G, *)$ bestehend aus einer Menge G und einer Verknüpfung $* : G \times G \mapsto G,\ (a, b) \to a * b$, heißt eine **Gruppe**, wenn folgendes gilt:

a) Für alle $a, b, c \in G$ gilt: $(a * b) * c = a * (b * c)$ (**Assoziativgesetz**).

b) Es existiert ein Element $e \in G$, so dass gilt:

- **i)** $\forall\, a \in G$ ist $a * e = a = e * a$ (**Existenz des neutralen Elementes**)
- **ii)** Zu jedem $a \in G$ existiert ein $b \in G$, so dass gilt: $a * b = e = b * a$ (**Existenz des inversen Elements**).

1.2 Kommentar

a) Ähnlich wie bei der Körperdefinition (siehe Teil I) sieht man, dass das neutrale Element $e \in G$ und auch das inverse Element $a \in G$ eindeutig bestimmt ist. (Übung)

b) Oft wird die Verknüpfung einer Gruppe G multiplikativ geschrieben (oder auch der Punkt ganz weggelassen),

$$a * b = a \cdot b = ab.$$

In diesem Fall wird das neutrale Element mit $e = 1$ und das zu $a \in G$ inverse Element mit a^{-1} bezeichnet,

$$a \cdot 1 = a = 1 \cdot a,$$

$$a \cdot a^{-1} = 1 = a^{-1} \cdot a$$

(vgl. allerdings 1.4).

1.3 Definition: abelsch

Eine Gruppe $(G, *)$ heißt **abelsch** (oder auch kommutativ), wenn für alle $a, b \in G$ gilt:

$$a * b = b * a \quad \text{(Kommutativgesetzt)}$$

1.4 Kommentar: additive Verknüpfung

a) Ist eine Gruppe $(G, *)$ abelsch, so notiert man ihre Verknüpfung meist additiv:

$$a * b = a + b.$$

Das neutrale Element wird dann üblicher Weise mit $e = 0$ und das inverse Element zu $a \in G$ wird mit $-a$ bezeichnet,

$$a + 0 = a = 0 + a,$$

$$a + (-a) = (-a) + a = 0.$$

b) Ist $(K, +, \cdot)$ ein Körper, so ist $(K, +)$ eine abelsche Gruppe, auch $(K^*, \cdot)$ mit $K^* = K \setminus \{0\}$ ist eine abelsche Gruppe und das Distributivgesetz gilt:

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

für alle $a, b, c \in K$.

1.5 Beispiele

a) Eine Gruppe G besitzt nach (1.1 b) mindestens ein Element. Die einfachste aller Gruppen (die **triviale Gruppe**) ist daher sicher:

$$G = \{0\}$$

(mit der offensichtlichen Verknüpfung "$0 + 0 = 0$").

b) Jeder Körper K zusammen mit seiner Addition ist offenbar eine abelsche Gruppe, z.B. $K = \mathbb{R}, \mathbb{C}, \mathbb{Q}, F_2$.

c) Ist K ein Körper, so ist auch K^* zusammmen mit der Multiplikation des Körpers eine abelsche Gruppe, z.B.

$$\mathbb{R}^*, \mathbb{C}^*, \mathbb{Q}^*, F_2^*.$$

d) $G = \mathbb{Z}$ zusammen mit der Addition ist eine (abelsche) Gruppe. (Übung)

e) Sei X eine beliebige Menge. Auf der folgenden Menge

$$G := \mathrm{Bij}(X) := \{f : X \to X |\ f \text{ ist bijektiv}\}$$

betrachten wir die Verkettung

$$\circ : G \times G \to G,\ (f, g) \mapsto f \circ g$$

als Verknüpfung. Es ist dann G eine (wenn X mindestens drei Elemente hat nicht abelsche) Gruppe (Übung). Ist speziell X endlich mit n Elementen, sagen wir $X = \{1, ..., n\}$, so wird G mit:

$$\mathcal{S}_n := \mathrm{Bij}(\{1, ..., n\})$$

bezeichnet. Sie heißt die symmetrische Gruppe in n Einträgen und hat $n!$ Elemente.

1.6 Definition: Untergruppe

Sei G eine Gruppe. Eine nicht-leere Teilmenge $H \subseteq G$ heißt eine **Untergruppe** von G, wenn gilt:

a) Für alle $a, b \in H$ ist auch $ab \in H$;

b) für alle $a \in H$ ist auch $a^{-1} \in H$.

1.7 Kommentar: jede Untergruppe ist Gruppe

Ist $H \subseteq G$ eine Untergruppe, so enthält H nach Definition mindestens ein Element $a \in H$. Nach a) und b) ist dann aber auch $1 = a \cdot a^{-1} \in H$. Die Einschränkungen der Verknüpfung mit $* : H \times H \to H$ macht dann $(H, *)$ selbst zu einer Grupppe.

1.8 Beispiel

a) Ist G eine Gruppe so sind $H := \{1\}$ und $H := G$ offenbar Untergruppen (die **trivialen Untergruppen**).

b) Sei $G = \mathbb{Z}$ und $n \in \mathbb{N}$. Dann ist offenbar

$$n\mathbb{Z} = \{nk\ \in\ \mathbb{Z} : k \in \mathbb{Z}\}$$

für jedes $n \in \mathbb{N}$ eine Untergruppe von $\mathbb{Z}$.

1.9 Definition: K-Vektorraum

Sei K ein Körper. Ein Tripel $(V, +, \cdot)$ heißt ein **K-Vektorraum** (Vektorraum über K), wenn V eine Menge, $+ : V \times V \to V$ eine (innere) Verknüpfung und $\cdot : K \times V \to V$ eine (äußere) Verknüpfung, so dass gilt:

a) $(V, +)$ ist eine abelsche Gruppe;

b)

- **i)** Für alle $v \in V$ gilt: $1 \cdot v = v$;
- **ii)** für alle $v, w \in V$, und $\lambda \in K$ gilt:
$$\lambda \cdot (v + w) = (\lambda \cdot v) + (\lambda \cdot w);$$
- **iii)** für alle $v \in V$ und $\lambda, \mu \in K$ gilt:
$$(\lambda + \mu) \cdot v = (\lambda \cdot v) + (\mu \cdot v);$$
- **iv)** für alle $\lambda, \mu \in K$ und $v \in V$ gilt:
$$(\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v).$$

1.10 Kommentar

a) Ein K-Vektorraum V hat also stets mindestens ein Element: dass Nullement 0_V, welches man nicht mit dem Nullelement des Körpers 0_K verwechseln sollte. Der einfachste K-Vektorraum ist deshalb durch:
$$V = \{0\}$$
(und $+$ und $\cdot$ in offensichtlicher Weise definiert) gegeben. Er heißt der **Null-Vektorraum**.

b) Die innere Verknüpfung $+$ eines K-Vektorraums $(V, +, \cdot)$ wird überlicher Weise als die Addition in V und die äußere Verknüpfung $\cdot$ als skalare Multiplikation in V bezeichnet. Die Elemente von V heißen **Vektoren** und die Elemente aus dem Körper K nennt man üblicher Weise **Skalare**.

c) Oft spricht man (etwas ungenau) nur von einem Vektorraum V (unterdrückt also die Verknüpfungen $+$ und $\cdot$ und auch den Grundkörper K), wenn klar ist, über welchem Körper der Vektorraum liegt. Der Punkt für die skalare Multiplikation wird häufig weggelassen und Punktrechnung geht wie üblich vor Strichrechnung, z.B. bedeutet:
$$\lambda v + \mu w = (\lambda \cdot v) + (\mu \cdot w)$$
für $\lambda,\ \mu \in K,\ v, w \in V$.

1.11 Beispiele

Sei K ein Körper.

a) Für $n \in \mathbb{N}$ definieren wir auf dem cartesichen Produkt:

$$V = K^n = \{(x_1, ..., x_n) : x_i \in K, i = 1, ..., n\}$$

eine innere Verknüpfung $+ : V \times V \to V$ durch

$$(x_1, ..., x_n) + (y_1, ..., y_n) := (x_1 + y_1, ..., x_n + y_n)$$

und eine äußere Verknüpfung $\cdot : K \times V \mapsto V$ durch

$$\lambda(x_1, ..., x_n) := (\lambda x_1, ..., \lambda x_n).$$

Dann wird $(V, +, \cdot)$ zu einem K-Vektorraum, den wir mit K^n bezeichnen (Übung).

b) Speziell im Fall $n = 1$ sehen wir, dass jeder Körper ein Vektorraum über sich selbst ist. (Die innere Multiplikation des Körpers $* : K \times K \mapsto K$ von K nun eben als skalare Multiplikation von K auf $K = V$ aufgefasst.)

c) Sei X (irgend) eine Menge. Auf

$$V := \text{Abb}(X, K) := \{f : X \mapsto K \text{ Abbildung}\}$$

definieren wir innere und äußere Verknüpfung so:

$$(f + g)(x) := f(x) + g(x)$$

$$(\lambda f)(x) := \lambda f(x)$$

für $\lambda \in K$, $f, g \in V$, $x \in X$. Es wird dann V damit zu einem K-Vektorraum (Übung).

1.12 Kommentar: geometrische Interpretation

a) Für $K = \mathbb{R}$ und $n = 2$ bzw. $n = 3$ bietet Beispiel (1.11.a) ein Modell für die Anschauungsebene und den Anschauungsram. In ihm kann man die Addition und skalare Multiplikation durch die folgende **"Parallelogrammregel"** und **Streckung bzw. Stauchung** (geometrisch) interpretieren.

Abbildung fehlt

Die Theorie der Vektorräume (und ihrer linearen Abbildungen) wird deshalb auch geeignet sein, Probleme der **analytischen Geometrie** zu behandeln.

b) Die Axiome (a) und (b) aus (1.9) implizieren nun eine Reihe von Rechenregeln (ähnlich wie dies die Köperaxiome für das Rechnen in K tun (vgl. Teil 1)), z.B. gilt:

$$\forall\ v \in V : 0_K \cdot v = 0_V,$$

denn aus $0_K v = (0_K + 0_K)v = 0_K v + 0_K v$ folgt durch Addition des Negativen $-0_K v$ auf beiden Seiten, dass $0_V = 0_K v$. Im Folgenden lassen wir den Index 'K' oder 'V' weg, weil in der Regel klar sein wird, welche Null gemeint sein wird.

1.13 Definition: Untervektorraum

Sei K ein Körper und V ein K-Vektorraum. Eine nicht-leere Teilmenge $U \subseteq V$ heißt ein **Untervektorraum** von V (kurz: Unterraum), wenn folgendes gilt:

a) $\forall\ v, w \in U$ gilt: $v + w \in U$;

b) $\forall\ \lambda \in K, v \in U$ gilt: $\lambda \cdot v \in U$.

1.14 Kommentar: Unterräume

a) Jeder Unterraum $U \subseteq V$ enthält das Nullelement von V, denn es gibt ein $v \in U$. Wegen (b) ist dann auch $-v = (-1)v \in U$ und wegen (a) dann auch $0 = v + (-v) \in U$.

b) Die Einschränkungen von $+$ und $\cdot$ auf $U \times U$ und $K \times U$ machen dann $(U, +, \cdot)$ selbst zu einem K-Vektorraum.

c) $U = \{0\}$ und $U = V$ sind offenbar Unterräume. Sie heißen die **trivialen Unterräume**.

1.15 Beispiele: Unterräume

Sei K ein Körper, $n \in \mathbb{N}$.

a) Sei $v \in K^n$ beliebig. Dann ist

$$U := \{\lambda \cdot v \in K^n : \lambda \in K\} \quad (=: Kv)$$

ein Unterraum von V (denn $\lambda v + \mu v = (\lambda + \mu)v$ und $\lambda(\mu v)) = (\lambda\mu)v$. Für $v \neq 0$ nennen wir U **eine Gerade in** K^n.

b) Seien $v, w \in K^n$ beliebig. Dann ist

$$U = \{\lambda v + \mu w \in K^n : \lambda, \mu \in K\}$$

Unterraum. (Übungsaufgabe) Für $v, w \neq 0$ und $w \notin Kv$ heißt **U eine Ebene** in K^n.

c) Sei V ein K-Vektorraum und $v_1, ..., v_r \in V$ $(r \in \mathbb{N})$. Dann ist

$$U = \{\lambda_1 v_1 + ... + \lambda_r v_r \in V : \lambda_i \in K, i = 1, ..., r\}$$

ein Unterraum von V (Übung).

d) Sei $X = [0, 1] \subseteq \mathbb{R}$ und $k \in \mathbb{N}$ (oder $k = \infty$ oder $k = \omega$). Dann ist

$U = \mathfrak{C}^k(X) = \{f : [0, 1] \mapsto \mathbb{R} : f \text{ ist k-mal stetig diff'bar}\} \subseteq \mathrm{Abb}(X, \mathbb{R})$

(vgl 1.11 c) (nach Teil I) ein Unterraum.

e) Sei

$$K[X] := \{a_0 + a_1 X + ... + a_r X^r : r \in \mathbb{N}_0,\ a_i \in K,\ i = 0, ..., r\}$$

die Menge der (**formalen) Polynome** mit Koeffizienten in K. Wir definieren

$$\begin{aligned}&(a_0 + a_1 X + ... + a_r X^r) + (b_0 + b_1 X + ... + b_s X^s)\\ := &(a_0 + b_0) + (a_1 + b_1)X + ... + (a_{\max(r,s)} + b_{\max(r,s)})X^{\max(r,s)}\end{aligned}$$

(wobei $a_i = 0$ für $i > r$ bzw. $b_j = 0$ für $j > s$ sei). Außerdem setzen wir:

$$\lambda(a_0 + a_1 X + ... + a_r X^r) := \lambda a_0 + (\lambda a_1)X + ... + (\lambda a_r)X^r \ .$$

Dann wird K[X] zusammen mit + und $\cdot$ zu einem K-Vektorraum (Übung).

f) Sei weiter für jedes $n \in \mathbb{N}$

$$K[X]^{(n)} := \{P \in K[X] : \deg(P) \leq n\}$$

(mit $P = \sum_{i=0}^r a_i X^i$ und $\deg(P) = r :\Leftrightarrow r = \max\{k \in \mathbb{N}_0 : a_k \neq 0\}$ sowie $\deg(0) := -\infty$). Dann ist $K[X]^{(n)}$ ein Unterraum von $K[X]$.

1.16 Definition: Familie

Seien I und X Mengen. Eine **(I-indizierte) Familie** (von Mitgliedern) in X ist eine Abbildung $f : I \mapsto X$. Ist $a_i = f(i)$, für $i \in I$, so wird sie üblicher Weise mit

$$\mathfrak{a} = (a_i)_{i \in I}$$

bezeichnet. I heißt **Indexmenge** der Familie.

1.17 Kommentar: Familien und Folgen

a) Eine Familie in X ist etwas anderes als eine Teilmenge von X. Ist zum Beispiel I endlich mit $I = \{1, ..., n\}$, so ist eine Familie in X indiziert über I ein **n-Tupel**

$$\mathfrak{a} = (a_1, a_2, ..., a_n) \ .$$

Dabei kommt es durchaus auf die Reihenfolge der Mitglieder an und es können auch Mitglieder mehrfach auftreten. Zu unterscheiden ist deshalb $\mathfrak{a}$ von der Teilmenge

$$A = \{a_1, a_2, ..., a_n\} \subseteq X \ .$$

b) Ist $I = \mathbb{N}$, so heißt eine I-indizierte Familie in X eine **Folge in** X.

Im Folgenden sei K stets ein Körper.

1.18 Bemerkung: Unterräume

Sei V ein K-Vektorraum und sei $(U_i)_{i \in I}$ eine Familie von Untervektorräumen von V. Dann ist auch die Schnittmenge der Familie

$$U := \bigcap_{i \in I} U_i$$

ein Unterraum von V.

Beweis:
Da $0 \in U_i, \ \forall \ i \in I$, ist $0 \in U$ und damit $U \neq \varnothing$. Seien $v, w \in U \Rightarrow v, w \in U_i, \ \forall \ i \in I \Rightarrow v + w \in U_i, \ \forall \ i \in I \Rightarrow v + w \in U$. Ähnlich: $v \in U, \ \lambda \in K \Rightarrow \lambda \cdot v \in U$.

$\square$

1.19 Definition: Erzeugendensystem

Sei V ein K-Vektorraum und $\mathfrak{a} = (v_i)_{i\in I}$ eine Familie in V. Der kleinste Untervektorraum, der alle Mitgleider von $\mathfrak{a}$ enthält, heißt der von $\mathfrak{a}$ erzeugte Unterraum und wird mit $\langle\mathfrak{a}\rangle$ bezeichnet, also:

$$\langle\mathfrak{a}\rangle := \bigcap_{U\subseteq V \text{ Unterraum },\quad U\ni v_i\ \forall\ i\in I} U\ .$$

Ist $W = \langle\mathfrak{a}\rangle \subseteq V$, so heißt $\mathfrak{a}$ ein **Erzeugendensystem von** W.

1.20 Bemerkung

Sei V ein K-Vektorraum und $\mathfrak{a} = (v_i)_{i\in I}$ eine Familie in V. Dann gilt:

$$\langle\mathfrak{a}\rangle = \left\{\sum_{i\in I}\lambda_i v_i \in V :\ \lambda_i \in K \text{ für } i\in I \text{ , und } \lambda_i = 0 \text{ für fast alle } i \in I\right\}\ .$$

Beweis:
Jeder Unterraum $U \subseteq V$, der alle $v_i \in V$ enthält, $i \in I$, muss auch jede Linearkombination $\lambda_{i_1}v_{i_1} + ... + \lambda_{i_r}v_{i_r}$ enthalten. Setzen wir $U := \left\{\sum_{i\in I}\lambda_i v_i\right\}$ so ist also: $U \subseteq \langle\mathfrak{a}\rangle$ Es ist aber U offenbar auch eine Unterraum mit $v_i \in I$, $\forall\ i \in I$. Daher gilt auch $U \supseteq \langle\mathfrak{a}\rangle$, also:

$$U = \langle\mathfrak{a}\rangle\ .$$

□

1.21 Kommentar: Linearkombination

a) Ist I eine unendliche Menge, so bedeutet (wie in Teil 1) "für fast alle $i \in I$", dass für höchstens endlich viele $i \in I$ die Bedingung nicht gilt. In unserem Fall ist also :

$$\sum_{i\in I}\lambda_i v_i = \lambda_{i_1}v_{i_1} + ... + \lambda_{i_r}v_{i_r}$$

nur eine endliche Summe (wo $i_1, ..., i_r \in I$ gerade die Indizies i seien, wo $\lambda_i \neq 0$ ist). Damit ist auch erklärt, was die Summe "$\sum_{i\in I}$" bedeutet.

b) Wie in Teil 1 vereinbaren wir, dass für $I = \varnothing$:

$$\sum_{i\in\varnothing} := 0 \text{ (in } V)$$

sei.

c) Sind $v_1, ..., v_r \in V$ und $\lambda_1, .., \lambda_r \in K$ so nennt man:

$$\lambda_1 v_1 + ... + \lambda_r v_r = \sum_{i=1}^{r} \lambda_i v_i$$

eine **Linearkombination** von $(v_1, ..., v_r)$ in V.

1.22 Kommentar

Sei V eine K-Vektorraum.

a) Setzt man $\mathfrak{a} = (v)_{v \in V}$ mit $I = V$, so ist natürlich $\langle \mathfrak{a} \rangle = V$, also $\mathfrak{a}$ Erzeugendensytem für V. Oft ist man allerdings interessiert, möglichst kleine Erzeugendensysteme für V zu finden.

b) Für $V = \{0\}$ ist offenbar die leere Familie (), das heißt $(I = \varnothing)$ ein Erzeugendensystem.

1.23 Beispiel: Erzeugendensysteme

Sei K (wie immer) ein Körper.

a) Für $n \in \mathbb{N}$ und jedem $i \in \{1, ..., n\}$ setzen wir:

$$e_i = (0, .., 0, \underbrace{1}_{\text{i-te stelle}}, 0, ..., 0) \in K^n \ .$$

Setzen wir für $j, k \in \mathbb{N}$:

$$\delta_{jk} := \begin{cases} 1 & \text{wenn } j = k \\ 0 & \text{wenn } j \neq k \ , \end{cases}$$

das so genannte **Kronecker-Symbol**, so ist also:

$$e_i = (\delta_{1i}, \ldots, \delta_{ni}) \ \ (i = 1, \ldots, n) \ .$$

Es ist dann $\mathfrak{a} = (e_1, \ldots, e_n)$ ein endliches Erzeugendensystem für $V = K^n$, denn jeder Vektor $v = (x_1, \ldots, x_n) \in K^n$ ist offenbar Linearkombination von $e_1, \ldots, e_n$:

$$v = (x_1, \ldots, x_n) = x_1 e_1 + \ldots . + x_n e_n \ .$$

b) Die Familie der **Monome** $(1, X, X^2, \ldots)$ ist ein Erzeugendensystem von $K[X]$.

1.24 Definition: endlich erzeugt

Ein K-Vektorraum heißt **endlich erzeugt**, wenn es ein endliches Erzeugendensystem $\mathfrak{a} = (v_1, \ldots, v_n)$ von V gibt,

$$V = \langle (v_1, \ldots, v_n) \rangle =: \langle v_1, \ldots, v_n \rangle \ .$$

1.25 Beispiel: endlich erzeugt

Für jedes $n \in \mathbb{N}$ ist nach (1.23) $V = K^n$ endlich erzeugt. $K[X]$ oder $\mathfrak{C}^k(I)$ $(I = [0,1])$ (bei $K = \mathbb{R}$, $k \in \mathbb{N} \cup \{\infty, \omega\}$) sind es nicht (vgl. (1.42)).

1.26 Motivation: sparsame Erzeugung

Um einen Vektorraum „so sparsam wie möglich zu erzeugen“ führt man folgenden wichtigen Begriff ein:

1.27 Definition: linear unabhängig

Sei V ein K-Vektorraum. Eine endliche Familie $\mathfrak{a} = (v_1, \ldots v_n)$ bestehend aus n Vektoren heißt **linear unabhängig**, wenn folgendes gilt: Sind $\lambda_1, \ldots, \lambda_n \in K$ derart, dass

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$$

ist, so muss bereits $\lambda_1 = \ldots = \lambda_n = 0$ sein.

1.28 Kommentar: linear abhängig

a) $(v_1, \ldots, v_n)$ ist linear unabhängig, genau wenn keines der Mitglieder Linearkombination der anderen ist,

$$v_i \notin \langle v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \rangle \ ,$$

$\forall\ i = 1, \ldots, n$. Ist nämlich $(v_1, \ldots, v_n)$ **linear abhängig**, d.h. nicht linear unabhängig, so gibt es also $\lambda_1, \ldots, \lambda_n \in K$ mit:

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0 \ ,$$

und nicht alle $\lambda_i = 0$. Ist etwa $\lambda_1 \neq 0$ so ist dann:

$$v_1 = -\frac{\lambda_2}{\lambda_1} v_2 - \ldots - \frac{\lambda_n}{\lambda_1} v_n \ ,$$

also Linearkombination von $v_2, \ldots, v_n$ Ist umgekehrt etwa v_n eine Linearkombination von $v_1, \ldots, v_{n-1}$,

$$v_n = \mu_1 v_1 + \ldots + \mu_{n-1} v_{n-1}$$

(mit $\mu_1, \ldots, \mu_{n-1} \in K$) so ist

$$\mu_1 v_1 + \ldots + \mu_{n-1} v_{n-1} + (-1) v_n = 0$$

also mit

$$\lambda_1 := \mu_1, \ldots, \lambda_{n-1} := \mu_{n-1}, \lambda_n := -1$$

auch

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0 \ ,$$

und dabei sind nicht alle λ_i's gleich Null.

b) Ist $\mathfrak{a} = (v_1, \ldots, v_n)$ linear unabhängig, so ist insbesondere $v_i \neq 0$ für alle $i = 1, \ldots, n$, Weiter ist $v_2 \notin \langle v_1 \rangle, v_3 \notin \langle v_1, v_2 \rangle$ usw., also $0 \subsetneq \langle v_1 \rangle \subsetneq \langle v_1, v_2 \rangle \subsetneq \langle v_1, v_2, v_3 \rangle \ldots \subsetneq \langle v_1, \ldots, v_n \rangle$.

1.29 Beispiel: linear unabhängig

a) Die leere Familie () gelte als linear unabhängig.

b) Sei $V = K^n$ $(n \in \mathbb{N})$. Dann ist $\mathfrak{a} = (e_1, \ldots, e_n)$ linear unabhängig, denn ist:

$$\lambda_1 e_1 + \ldots + \lambda_n e_n = 0$$

so ist $(\lambda_1, \ldots, \lambda_n) = 0$, also $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$.

c) Die Familie $(1, X, X^2, \ldots)$ in $V = K[X]$ ist linear unabhängig (siehe (1.30), Übung).

d) Die Familie $\mathfrak{a} = ((1,0), (0,1), (1,1))$ in $V = K^2$ ist linear abhängig, denn

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ ,$$

also

$$v_3 = v_1 + v_2$$

mit $v_1 = (1,0), v_2 = (0,1)$ und $v_3 = (1,1)$. Beachte aber: Jede der Teilfamilien $(v_1, v_2), (v_1, v_3), (v_2, v_3)$ ist linear unabhängig.

1.30 Kommentar: linear unabhängig

a) Für eine unendliche Familie $(v_i)_{i\in I}$ in V sagt man, dass sie linear unabhängig ist, wenn jede endliche Teilfamilie linear unabhängig ist.

b) Wenn Erzeugendensysteme eines K-Vektrorraums V nicht zu klein sein können, können linear unabhängige Familien nicht zu groß sein. Gewissermaßen optional ist die Situation, wenn Folgendes vorliegt:

1.31 Definition: Basis

Sei V ein K-Vektorraum. Eine Familie $\mathfrak{a} = (v_1, \dots, v_n)$ heißt eine **Basis** von V, wenn folgendes gilt:

a) $\mathfrak{a}$ ist Erzeugendensystem,

b) $\mathfrak{a}$ ist linear unabhängig.

1.32 Beispiel: (kanonische) Basis

a) Die leere Familie () ist also eine Basis des Nullraums.

b) Sei $n \in \mathbb{N}$. Dann ist offenbar $\mathfrak{a} = (e_1, .., e_n)$ eine Basis von $V = K^n$. Wir nennen sie die **kanonische Basis** von K^n.

c) Auch $((1,1),(1,-1))$ ist z.B. eine Basis von $\mathbb{R}^2$ (oder allgemeiner K^2, wenn $1+1 \neq 0$ in K ist, Übung).

d) $(1, X, \dots, X^n)$ ist eine Basis von $K[X]^{(n)}$ (Übung).

1.33 Kommentar: Basis

a) Allgemein sagt man, dass eine Familie $\mathfrak{a} = (v_i)_{i\in I}$ von Vektoren in V eine Basis heißt, wenn sie Erzeugendensystem und linear unabhängig ist.

b) Zum Beispiel ist $(1, X, X^2, X^3, \dots)$ in $K[X]$ eine Basis.

c) Es ist überhaupt nicht einfach zu sehen, dass jeder (endlich erzeugte) Vektorraum eine Basis hat, und dass je zwei Basen jeweils die gleiche Länge haben, d.h. die gleiche Anzahl von Mitgliedern. Dazu:

1.34 Lemma

Sei V ein K-Vektorraum, $n \in \mathbb{N}$ und $0 \leq r \leq n-1$. Seien $v_1, \ldots, v_n \in V$, derart, dass $(v_1, .., v_r)$ linear unabhängig, aber $(v_1, \ldots, v_n)$ linear abhängig ist. Dann existiert ein
$p \in \{r+1, \ldots, n\}$, so dass gilt:

$$\langle v_1, \ldots, v_n \rangle = \langle v_1, \ldots, v_{p-1}, v_{p+1}, \ldots, v_n \rangle \ .$$

Beweis:
Aus den Voraussetzungen folgt, dass es $\lambda_1, \ldots, \lambda_n \in K$ und ein $p \in \{r+1, \ldots, n\}$ mit $\lambda_p \neq 0$ gibt, so dass gilt:

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$$

$$\Rightarrow v_p = -\frac{\lambda_1}{\lambda_p} v_1 - \ldots - \frac{\lambda_{p-1}}{\lambda_p} v_{p-1} - \frac{\lambda_{p+1}}{\lambda_p} v_{p+1} - \ldots - \frac{\lambda_n}{\lambda_p} v_n \ \in \langle v_1, \ldots, v_{p-1}, v_{p+1}, \ldots, v_n \rangle$$

$$\Rightarrow \langle v_1, \ldots, v_n \rangle = \langle v_1, \ldots, v_{p-1}, v_{p+1}, \ldots, v_n \rangle \ .$$

$\square$

1.35 Satz: Basen endlicher Länge

Sei V ein endlich erzeugter K-Vektorraum. Dann besitzt V eine Basis. Jede Basis von V hat endliche Länge.

Beweis:
Sei $\mathfrak{a} = (v_1, \ldots, v_n)$ ein Erzeugendensystem von V. Durch Weglassen von Mitgliedern können wir $\mathfrak{a}$ gegebenenfalls so verkleinern, dass das (verkleinerte) $\mathfrak{a}$ **minimales Erzeugendensystem** ist (d.h.: jede Teilfamilie $\mathfrak{a}'$ von $\mathfrak{a}$ mit $\mathfrak{a}' \neq \mathfrak{a}$ ist kein Erz-System mehr). Sei also oBdA: $\mathfrak{a}$ ein min. Erzeugendensystem. Nach Lemma (1.34) muss $\mathfrak{a}$ dann schon linear unabhängig und damit eine Basis sein.
Ist $(v_1, \ldots, v_n)$ eine endliche Basis, $(w_i)_{i \in I}$ eine beliebige Basis, so existiert eine endliche Teilmenge $I' \subseteq I$, so dass:

$$V = \langle v_1, \ldots, v_n \rangle \subseteq \langle (w_i)_{i \in I'} \rangle$$

denn jedes $v_k \ \ (k = 1, \ldots, n)$ ist Linearkombination von nur endlich vielen w_i's. Dann ist bereits $(w_i)_{i \in I'}$ erzeugend und damit Basis, also $I' = I$.

$\square$

1.36 Kommentar: Basis

a) Hier erst sieht man z.B., dass $V = K[X]$ nicht endlich erzeugt ist, denn $(1, X, X^2, \ldots)$ ist eine unendliche Basis.

b) Auch ist hier erst klar, dass z.B. $V = K^n$ keine unendliche Basis haben kann.

c) Noch nicht klar ist, dass für einen endlich erzeugten Vektorraum V je zwei Basen $\mathfrak{a} = (v_1, \ldots, v_m)$ und $\mathfrak{b} = (w_1, \ldots, w_m)$ gleiche Länge haben, also $m = n$ ist.

1.37 Satz: Basisergänzungssatz

Sei V ein K-Vektorraum, $v_1, \ldots, v_r, w_1, \ldots, w_m \in V$ derart, dass $(v_1, \ldots, v_r)$ linear unabhängig und $(w_1, \ldots, w_m)$ ein Erzeugendensystem ist. Dann existiert eine Zahl $r \leq n \leq r + m$ und paarweise verschiedene $i_{r+1}, \ldots, i_n \in \{1, \ldots, m\}$, so dass:

$$(v_1, \ldots, v_r, w_{i_{r+1}}, \ldots, w_{i_n})$$

eine Basis von V ist.

Beweis:
Es gibt sicher paarweise verschiedene Indizes $i_{r+1}, \ldots, i_n$ (mit $r \leq n \leq r + m$), so dass $(v_1, \ldots, v_r, w_{i_{r+1}}, \ldots, w_{i_n})$ Erzeugendensystem wird, z.B. wenn man $n := r + m$ und $i_{r+1} := 1, \ldots, i_{r+m} = m$ setzt. Anwendung von (1.34) liefert wiederrum, dass man durch Weglassen einiger w_{i_k}'s erreichen kann, dass $(v_1, \ldots, v_r, w_{i_1}, \ldots, w_{i_n})$ nicht nur erzeugend bleibt, sondern auch linear unabhängig, also eine Basis wird.

$\square$

1.38 Satz: Länge von Basen

Sei V ein K-Vektorraum und seien $\mathfrak{a} = (v_1, \ldots, v_n)$ und $\mathfrak{b} = (w_1, \ldots, w_m)$ Basen von V. Dann gilt: $n = m$.

Beweis:
Die linear unabhänige Familie $(v_2, \ldots, v_n)$ kann man nach (1.37) durch Hinzunahme mindestens eines Mitgliedes aus $\mathfrak{b}$ zu einer Basis von V ergänzen: $\exists\, i_1, \ldots, i_{r_1} \in \{1, \ldots, m\}$ paarweise verschieden mit $(w_{i_1}, \ldots, w_{i_{r_1}}, v_2, \ldots, v_n)$ ist Basis, $r_1 \geq 1$. Das gleiche Argument liefert $i_{r_1+1}, \ldots, i_{r_1+r_2} \in \{1, \ldots, m\}$, so dass $i_1, \ldots, i_{r_1+r_2}$ paarweise verschieden sind mit: $(w_{i_1}, \ldots, w_{i_{r_1+r_2}}, v_3, \ldots, v_n)$ ist Basis, $r_2 \geq 1$. Nach n Schritten gibt es also $i_1, \ldots, i_{r_1+r_2+\ldots+r_n} \in \{1, \ldots, m\}$ paarweise verschieden, so dass $(w_{i_1}, \ldots, w_{i_{r_1+\ldots+r_n}})$ Basis von V ist. Also:

$$m \geq r_1 + \ldots + r_n \geq \underbrace{1 + \ldots + 1}_{n-mal} = n \ .$$

Aus Symmetriegründen folgt auch $n \geq m$, also $n = m$.

$\square$

1.39 Definition: Dimension von V

Sei V ein K-Vektorraum.

a) Ist V endlich erzeugt und $(v_1, \ldots, v_n)$ eine Basis und V, so heißt:

$$\dim_K V := n$$

die **Dimension von** V (über K).

b) Ist V nicht endlich erzeugt, so setzen wir:

$$\dim_K V := \infty \ .$$

1.40 Kommentar: wohldefinierte Dimensionen

a) Man beachte, dass erst Satz (1.38) (und (1.35)) sicher stellt, dass die Dimension eines Vektorraum **wohldefiniert** ist (d.h. unabhängig von der gewählten Basis).

b) Man kann zeigen, dass auch nicht endlich erzeugte Vektorräume stets eine Basis haben (siehe z.B. Bosch). Jeder Vektorraum hat also eine Basis!

1.41 Korollar

a) Ist V endlich erzeugt und $U \subseteq V$ ein Unterraum, so ist U auch endlich erzeugt und es gilt:

$$\dim_K U \leq \dim_K V \ .$$

b) Ist V endlich erzeugt, $U \subseteq V$ ein Unterraum mit $\dim_K U = \dim_K V$, so ist schon $U = V$.

Beweis a) Sei $\mathfrak{a} = (v_1 \ldots, v_n)$ eine Basis von V. Sei weiter $\mathfrak{b} = (w_1, \ldots, w_k)$ linear unabhängig in U. Wegen (1.37) kann man $(w_1, \ldots, w_k)$ mit Mitglieder aus $\mathfrak{a}$ zu einer Basis von V ergänzen. Wegen (1.38) ist dann $k \leq n$. Insbesondere ist auch $U \subseteq V$ endlich erzeugt und $\dim_K U \leq \dim_K V$.

Beweis b) Ist nun $\dim_K U = \dim_K V$ und $\mathfrak{b} = (w_1, \ldots, w_n)$ Basis von U, so muss $\mathfrak{b}$ auch Basis von V sein, sonst könnte man es mit mindestens einem Mitglied aus $\mathfrak{a}$ zu einer Basis von V ergänzen, die aber mehr als n Mitglieder hätte.

□

1.42 Beispiel: Dimension

a) Sei $n \in \mathbb{N}_0$. Dann ist offenbar:

$$\dim_K K^n := n \ ,$$

weil $(e_1, \ldots, e_n)$ eine Basis ist. (Für $n = 0$ fassen wir K^0 als den Nullvektorraum auf.)

b) Die Polynome vom Grad kleiner gleich n erfüllen offensichtlich:

$$\dim_K K[X]^{(n)} = n + 1 \ ,$$

weil $(1, X, \ldots, X^n)$ eine Basis ist.

c) Fasst man $V = \mathbb{C}$ als $\mathbb{R}$-Vektorraum auf, so ist offenbar $(1, i)$ eine $\mathbb{R}$-Basis von $\mathbb{C}$. (Natürlich ist (1) eine $\mathbb{C}$-Basis von $\mathbb{C}$.) Es ist also:

$$\dim_{\mathbb{R}} \mathbb{C} = 2, \ \text{aber} \ \dim_{\mathbb{C}} \mathbb{C} = 1 \ .$$

d) Natürlich ist z.B.

$$\dim_{\mathbb{R}} \mathfrak{C}([a, b]) = \infty$$

(vgl. (1.25)), denn die Polynomfunktionen Pol $\subseteq \mathfrak{C}([a, b])$ sind bereits ein unendlich-dimensionaler Unterraum (Übung).

1.43 Definition: (direkte) Summe

Sei V eine K-Vektorraum und $U_1, U_2 \subseteq V$ Unterräume.

a) Es heißt dann:

$$U_1 + U_2 := \{v_1 + v_2 : v_1 \in U_1, v_2 \in U_2\}$$

die **Summe von U_1 und U_2**.

b) Ist zudem $U_1 \cap U_2 = \{0\}$, so nennen wir die Summe **direkt** und schreiben dann $U_1 \oplus U_2$.

1.44 Kommentar: Komplement

a) $U_1 + U_2$ ist offenbar wieder Unterraum von V. Er ist der kleinste Unterraum von V, der sowohl U_1 als auch U_2 enthält,

$$U_1 + U_2 = \langle U_1 \cup U_2 \rangle \ .$$

b) Jedes Element $v \in U_1 + U_2$ besitzt also eine Darstellung $v = v_1 + v_2$ mit $v_1 \in U_1$ und $v_2 \in U_2$. Ist die Summe direkt, so ist die Darstellung sogar eindeutig. Denn ist

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

mit $v_1, w_1 \in U_1$ und $v_2, w_2 \in U_2$, so ist

$$w_1 - v_1 = v_2 - w_2 \in U_1 \cap U_2 = \{0\} \ ,$$

also $v_1 = w_1$ und $v_2 = w_2$.

c) Ist $V = U_1 \oplus U_2$, so nennt man U_2 ein **Komplement** von U_1 in V.

1.45 Beispiel: direkte Summe

Sei $V = K^3$ und $U_1 = \langle e_1, e_2 \rangle$ und $U_2 = \langle e_1, e_3 \rangle$ (wo (e_1, e_2, e_3) die kanonische Basis von V sei). Dann ist:

$$U_1 + U_2 = \langle e_1, e_2, e_3 \rangle = V \ ,$$

aber die Summe ist nicht direkt, denn es ist $U_1 \cap U_2 = \langle e_1 \rangle \neq \{0\}$. Dagegen ist z.B. mit $U_2' := \langle (1, 1, 1) \rangle$:

$$V = U_1 \oplus U_2' \ .$$

1.46 Satz

Sei V ein endlich erzeugter Vektorraum und $U \subseteq V$ ein Unterraum. Dann gibt es ein Komplement $U' \subseteq V$, also $V = U \oplus U'$, und es gilt:

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

Beweis:
Sei $\mathfrak{a} = (v_1, \ldots, v_n)$ Basis von V und $\mathfrak{b} = (w_1, \ldots, w_k)$ eine Basis von U $(k \leq n)$. Wir ergänzen $\mathfrak{b}$ nach (1.37) mit Mitgliedern $v_{i_{k+1}}, \ldots, v_{i_n}$, paarweise verschieden, zu einer Basis von V und setzen:

$$U' := \langle v_{i_{k+1}}, \ldots, v_{i_n} \rangle \subseteq V \ .$$

Dann ist $U \cap U' = \{0\}, U + U' = V$, also $V = U \oplus U'$. Da $(v_{i_{k+1}}, \ldots, v_{i_n})$ Basis von U' ist gilt:

$$\dim U + \dim U' = k + (n - k) = n = \dim V \ .$$

$\square$

1.47 Korollar: Dimensionsformel

Sei V ein endlich erzeugter K-Vektorraum und $U, U' \subseteq V$ Unterräume. Dann gilt die folgende **Dimensionsformel**:

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

Beweis:
Sei $W \subseteq U$ ein Komplement von $U \cap U'$ in U und $W' \subseteq U'$ Komplement von $U \cap U'$ in U',

$$U = (U \cap U') \oplus W \ ,$$
$$U' = (U \cap U') \oplus W' \ .$$

Es ist dann

$$U + U' = ((U \cap U') \oplus W) + ((U \cap U') \oplus W') = (U \cap U') + W + W'$$

und diese Summe ist sogar direkt (d.h. jeder Summand geschnitten mit der Summe der restlichen beiden ergibt den Nullraum (Übung)). Mit (1.46) ist:

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

$$\dim U' = \dim(U \cap U') + \dim W'$$

und

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

$$\begin{aligned}\Rightarrow \dim(U + U') + \dim(U \cap U') &= 2\dim(U \cap U') + \dim W + \dim W' \\ &= \dim U + \dim U'.\end{aligned}$$

$\square$

2 Lineare Abbildungen

Sei K stets ein Körper.

2.1 Definition: Homomorphismus

Seien V und W K-Vektorräume. Eine Abbildung $f : V \to W$ heißt (K-) linear (oder ein **(K-) Homomorphismus**), wenn gilt:

a) Für alle $v_1, v_2 \in V$ ist:

$$f(v_1 + v_2) = f(v_1) + f(v_2) \ ;$$

b) für alle $\lambda \in K$ und $v \in V$ gilt:

$$f(\lambda v) = \lambda f(v) \ .$$

2.2 Beispiel: Homomorphismen

a) Die Nullabbildung: $f : V \to W$ gegeben durch $f(v) = 0, \ \forall \ v \in V$, ist offenbar K-linear und wird oft mit $0 : V \to W$ bezeichnet.

b) Ist $V = W$ so hat man stets **die Identität** $f : V \to V$, gegeben durch $f(v) = v$, für alle $v \in V$. Sie ist offenbar K-linear.
Bezeichnung: $\mathrm{id}_V = f$.

c) Sei $V = W$ und $\lambda \in K$. Dann ist $f : V \to V$ gegeben durch:

$$f(v) = \lambda v \ ,$$

für alle $v \in V$, auch K-linear und heißt **Streckung** (oder auch Homothetie).

d) Sei $V = \mathbb{R}^2$, $\varphi \in [0, 2\pi)$ und $f : \mathbb{R}^2 \to \mathbb{R}^2$, gegeben druch:

$$f(x_1, x_2) = (\cos(\varphi)x_1 - \sin(\varphi)x_2, \sin(\varphi)x_1 + \cos(\varphi)x_2) \ .$$

Dann ist f $\mathbb{R}$-linear und beschreibt gerade **die Drehung um den Winkel** φ.

e) Die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$

$$f(x_1, x_2) = (x_1, -x_2)$$

ist auch $\mathbb{R}$-linear. Sie beschreibt die **Spiegelung an der x_1-Achse** .

f) Die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$:

$$f(x_1, x_2) = (x_1, 0)$$

ist auch $\mathbb{R}$-linear. Sie heißt **Projektion auf die x_1-Achse** (entlang der x_2-Achse).

g) Die Beispiele (d),(e) & (f) sind offenbar Spezialfälle folgender Situation: Seien $a_{11}, a_{12}, a_{21}, a_{22} \in K$ und

$$f : K^2 \to K^2,\ f(x_1, x_2) = (a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2)\ .$$

Das ist K-linear (Übung).

h) Die Abbildung $D : K[X] \to K[X]$, $D(p) = p'$, d.h für $p = a_0 + a_1X + \ldots + a_nX^n$,

$$D(p) = a_1 + 2a_2X + \ldots + na_nX^{n-1}\ ,$$

ist K-linear ($2 := 1 + 1, \ldots$). Sie heißt **Ableitung**.

i) Sei $\mathfrak{R}([0, 1]) \subseteq \mathrm{Abb}([0, 1)], \mathbb{R})$ der Unterraum der Riemann-integrierbaren Funktionen auf $[0, 1] \subseteq \mathbb{R}$. Dann ist $I : \mathfrak{R}([0, 1]) \to \mathbb{R}$,

$$I(f) = \int_0^1 f(x)dx\ ,$$

auch $\mathbb{R}$-linear.

j) Seien V und W K-Vekorräume. Wir setzen:

$$\mathrm{Hom}_K(V, W) = \{f : V \to W : f \text{ ist } K\text{-linear}\}$$

und nennen dies die **Homomorphismen von V nach W**. Für $f, g \in \mathrm{Hom}_k(V, W)$ und $\lambda \in K$ setzen wir:

$$(f + g)(v) := f(v) + g(v)\ ,$$

$$(\lambda f)(v) := \lambda f(v)\ ,$$

für alle $v \in V$. Dann wird $(\mathrm{Hom}_K(V, W), +, \cdot)$ selbst zu einem K-Vektorraum (Übung).

2.3 Definition: Matrix

Sei K ein Körper und $m, n \in \mathbb{N}$. Seien $a_{ij} \in K$ mit $1 \leq i \leq m$ und $1 \leq j \leq n$. Dann nennen wir das **Zahlentableau**:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & \dots & a_{mn} \end{pmatrix}$$

eine (K-) **Matrix** (mit Einträgen aus K) mit m Zeilen und n Spalten und schreiben dafür:

$$A = (a_{ij})_{1 \leq i \leq m,\ 1 \leq j \leq n} \ .$$

2.4 Kommentar: $(\mathrm{Mat}(m, n, K), +, \cdot)$ als Vektorraum

a) Etwas präziser ist also eine K-Matrix $A = (a_{ij})_{1 \leq i \leq m,\ 1 \leq j \leq n}$ eine Familie in K, die über die Indexmenge $I \times J$ mit $I = \{1, \dots, m\}$ und $J = \{1, \dots, n\}$ indiziert ist.

b) Wir setzen

$$\mathrm{Mat}(m, n; K) := \{A = (a_{ij}) : \ A \text{ ist } K\text{-Matrix mit m Zeilen und n Spalten}\}$$

und weiter für $A = (a_{ij})$ und $B = (b_{ij})$, $\lambda \in K$:

$$(a_{ij}) + (b_{ij}) := (a_{ij} + b_{ij}) \ ,$$

$$\lambda(a_{ij}) := (\lambda a_{ij}) \ .$$

Dann wird auch $(\mathrm{Mat}(m, n, K), +, \cdot)$ zu einem K-Vektorraum (Übung).

2.5 Definition: lineare Abbildung

Sei $A \in \mathrm{Mat}(m, n; K)$. Dann setzen wir $f_A : K^n \to K^m$:

$$f_A(x_1, \dots, x_n) = (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \dots, a_{m1}x_1 + \ldots + a_{mn}x_n) \ ,$$

die **zu A gehörende lineare Abbildung** (von K^n nach K^m).

2.6 Kommentar: Spaltenvektoren

a) f_A ist tatsächlich K-linear (Übung).

b) Schreiben wir die Vektoren $x \in K^n$ und $y \in K^m$ **spaltenförmig**, d.h.:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ und } \; y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

so schreibt sich $f_A : K^n \to K^m$ so:

$$f_A(x) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} =: Ax$$

wenn wir die Multiplikation einer Matrix $A \subseteq \mathrm{Mat}(m, n; K)$ mit einem Vektor $x \in K^n$ so definieren :

$$\mathrm{Mat}(m, n; K) \times K^n \to K^m \; ,$$

$$(A, x) \mapsto Ax$$

mit

$$Ax := \begin{pmatrix} a_{11}x_1 + & \cdots & +a_{1n}x_n \\ & \vdots & \\ a_{m1}x_1 + & \ldots & +a_{mn}x_n \end{pmatrix} .$$

c) Matrizen sind ein sehr effizientes Mittel, um lineare Abbildungen zwischen endlich dimensionalen Vektorräumen zu beschreiben. Z.B. werden wir bald sehen (siehe (2.12)), dass jede lineare Abbildung $f : K^n \to K^m$ von der Form (b) ist, d.h.: es gibt ein $A \in \mathrm{Mat}(m, n; K)$ mit $f = f_A$ (und A ist sogar eindeutig bestimmt).

2.7 Lemma: eindeutige Darstellung

Sei V ein K-Vektorraum. Ein n-Tupel $\mathfrak{a} = (v_1, \ldots, v_n)$ von Vektoren in V ($n \in \mathbb{N}$) ist genau dann eine Basis von V, wenn es für jedes $v \in V$ eine eindeutige Dastellung:

$$(*) \; v = \lambda_1 v_1 + \ldots + \lambda_n x_n$$

(mit $\lambda_1, \ldots, \lambda_n \in K$) gibt.

Beweis:
$\Rightarrow$: Ist $\mathfrak{a} = (v_1, \dots, v_n)$, eine Basis, so insbesondere Erzeugendensystem, $V =< v_1, \dots, v_n >$. Für jedes $v \in V$ gibt es also eine Darstellung $(*)$. Ist nun:

$$\lambda_1 v_1 + \dots + \lambda_n v_n = v = \mu_1 v_1 + \dots + \mu_n v_n$$

mit $\lambda_j,\ \mu_j \in K \quad (j = 1, \dots, n)$, so ist

$$0 = v - v = (\lambda_1 - \mu_1)v_1 + \dots + (\lambda_n - \mu_n)v_n \ .$$

Da $\mathfrak{a}$ auch linear unabhängig ist, gilt also: $\lambda_j - \mu_j = 0 \quad (j = 1, \dots, n)$, also $\lambda_j = \mu_j$. Die Darstellung ist also eindeutig.

$\Leftarrow$: (Übung).

$\square$

2.8 Bemerkung

Seien V und W K-Vektorräume, $\mathfrak{a} = (v_1, \dots, v_n)$ Erzeugendensystem von V und $f : V \to W$ linear. Dann ist f bereist durch seine Werte auf $(v_1, \dots, v_n)$ eindeutig bestimmt.

Beweis:
Für jedes $v \in V$ gibt es $\lambda_1, \dots, \lambda_n \in K$ mit $v = \lambda_1 v_1 + \dots + \lambda_n v_n$. Da f K-linear ist, gilt:

$$\begin{aligned} f(v) &= f(\lambda_1 v_1 + \dots + \lambda_n v_n) \\ &\overset{(2.1a)}{=} f(\lambda_1 v_1) + \dots + f(\lambda_n v_n) \\ &\overset{(2.1b)}{=} \lambda_1 f(v_1) + \dots + \lambda_n f(v_n) \ . \end{aligned}$$

Also ist f durch $f(v_1), \dots, f(v_n) \in W$ bereits festgelegt.

$\square$

2.9 Satz: eindeutige lineare Abbildung

Seien V und W K-Vektorräume, $\mathfrak{a} = (v_1, \dots, v_n)$ eine Basis von V und $w_1, \dots, w_n \in W$ beliebig. Dann gibt es genau eine lineare Abbildung $f : V \to W$ mit

$$f(v_j) = w_j \ \ (j = 1, \dots, n) \ .$$

Beweis:
Eindeutigkeit: Nach (2.8) gibt es höchstens eine solche Abbildung. Existenz: Sei $v \in V$ beliebig und $\lambda_1, \ldots, \lambda_n \in K$ die eindeutig bestimmten Skalare mit

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n$$

(nach 2.7). Wir setzen dann $f : V \to W$,

$$f(v) := \lambda_1 w_1 + \ldots + \lambda_n w_n \ .$$

Zu zeigen: f ist K-linear. Weil $f(v_j) = w_j$ für $1, \ldots, n$ ist, erfüllt f dann das Gewünschte. Seien

$$v = \sum_{j=1}^{n} \lambda_j v_j, \ \ \widetilde{v} = \sum_{j=1}^{n} \widetilde{\lambda}_j v_j$$

$$\Rightarrow v + \widetilde{v} = \sum_{j=1}^{n} (\lambda_j + \widetilde{\lambda}_j) v_j$$

$$\Rightarrow f(v + \widetilde{v}) = \sum_{j=1}^{n} (\lambda_j + \widetilde{\lambda}_j) w_j = \sum_{j=1}^{n} \lambda_j w_j + \sum_{j=1}^{n} \widetilde{\lambda}_j w_j = f(v) + f(\widetilde{v}) \ .$$

Ähnlich sieht man:

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

für $\lambda \in K, v \in V$.

□

2.10 Kommentar

Satz (2.9) ermöglicht es mit Matrizen $A \in \mathrm{Mat}(m, n; K)$ auch lineare Abbildungen zwischen (abstrakten) Vektorräumen V und W zu beschreiben, wenn man Basen $\mathfrak{a} = (v_1, \ldots, v_n)$ von V und $\mathfrak{b} = (w_1, \ldots, w_m)$ von W gegeben hat.

2.11 Definition: lineare Abbildung bezüglich zweier Basen

Sei $\mathfrak{a} = (v_1, \ldots, v_n)$ eine Basis von V, $\mathfrak{b} = (w_1, \ldots, w_m)$ eine Basis von W und $A \in \mathrm{Mat}(m, n; K)$. Wir setzen dann $f_A^{\mathfrak{a},\mathfrak{b}} : V \to W$,

$$f_A^{\mathfrak{a},\mathfrak{b}}(\lambda_1 v_1 + \ldots + \lambda_n v_n) := \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \lambda_j \right) w_i \ .$$

2.12 Kommentar

a) $f_A^{\mathfrak{a},\mathfrak{b}} : V \to W$ ist tatsächlich linear (Übung). Sie ist die lineare Abbildung, die die Mitglieder v_j von $\mathfrak{a}$ gerade auf die Elemente

$$\widetilde{w}_j := \sum_{i=1}^{m} a_{ij} w_i \;\; (j = 1, \ldots, n)$$

abbildet:

$$f_A^{\mathfrak{a},\mathfrak{b}}(v_j) = f_A^{\mathfrak{a},\mathfrak{b}}(\sum_{k=1}^{n} \delta_{jk} v_k)$$

$$= \sum_{i=1}^{m} \underbrace{(\sum_{k=1}^{n} a_{ik}\delta_{jk})}_{=a_{ij}} w_i = \widetilde{w}_j \; .$$

b) Sei $V = K^n$ und $\mathfrak{K} = (e_1, \ldots, e_n)$ die kanonische Basis von K^n (vgl.(1.30 b)). Ist nun $A \in \mathrm{Mat}(m, n; K)$, so ist mit den Bezeichnungen von (2.6) und Teil (a).

$$f_A = f_A^{\mathfrak{K}\mathfrak{K}} \; ,$$

denn f_A bildet offenbar e_j gerade auf $\sum_{i=1}^{m} a_{ij} e_i$ ab.

c) Jede lineare Abbildung $f : V \to W$ ist von der Form in (a): $f = f_A^{\mathfrak{a},\mathfrak{b}}$, für ein eindeutig bestimmtest $A \in \mathrm{Mat}(m, n; K)$! Ist nämlich

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

mit $a_{ij} \in K$ $(1 \leq i \leq m, 1 \leq j \leq n)$, so ist

$$f(v_j) = f_A^{\mathfrak{a},\mathfrak{b}}(v_j)$$

und damit $f(v) = f_A^{\mathfrak{a},\mathfrak{b}}(v)$, für alle $v \in V$ also:

$$f = f_A^{\mathfrak{a},\mathfrak{b}} \; .$$

Wir schreiben für diese Zuordnung.

$$A = M(f; \mathfrak{a}, \mathfrak{b})$$

und nennen dies die **Matrix von f bezüglich der Basen $\mathfrak{a}$ und $\mathfrak{b}$**.

d) Merkregel: Die Bilder (unter f) der Basisvektoren (von $\mathfrak{a}$) stehen (in der Darstellung bezüglich $\mathfrak{b}$) in den Spalten der Matrix ($A = M(f;\mathfrak{a},\mathfrak{b})$).

e) Die Abbildungen:

$$\Phi : \mathrm{Mat}(m,n;K) \to \mathrm{Hom}_K(V,W) \ ,$$

$$\Phi(A) = f_A^{\mathfrak{a},\mathfrak{b}} \ ,$$

$$\Psi : \mathrm{Hom}_K(V,W) \to \mathrm{Mat}(m,n;K) \ ,$$

$$\Psi(f) = M(f;\mathfrak{a},\mathfrak{b})$$

sind bijektiv und invers zueinander (vgl.(2.21)),

$$\Psi \circ \Phi = id, \ \ \Phi \circ \Psi = id \ .$$

2.13 Beispiel: Abbildungsmatrizen

a) Sei $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x_1,x_2) = (x_1,-x_2)$ (vgl.(2.2 e)) und $\mathfrak{K} = (e_1,e_2)$ die kanonische Basis. Weil:

$$f(e_1) = f(1,0) = (1,0) = e_1 = 1 \cdot e_1 + 0 \cdot e_2$$

$$f(e_2) = f(0,1) = (0,-1) = 0e_1 - 1e_2$$

ist, gilt:

$$M(f;\mathfrak{K},\mathfrak{K}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ .$$

b) Sei wieder $f : \mathbb{R}^2 \to \mathbb{R}^2, f(x_1,x_2) = (x_1,-x_2)$, aber diesmal $\mathfrak{a} = (v_1,v_2)$ mit $v_1 = (1,1)$ und $v_2 = (1,-1)$. Weil:

$$f(v_1) = f(1,1) = (1,-1) = v_2 = 0v_1 + 1v_2$$

und

$$f(v_2) = f(1,-1) = (1,1) = v_1 = 1v_2 + 0v_2$$

ist, gilt:

$$M(f;\mathfrak{a},\mathfrak{a}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ .$$

c) Sei $n \in \mathbb{N}$, $V = K[X]^{(n)}$, $\mathfrak{a} = (1, X, \ldots, X^n)$ und $D : K[X]^{(n)} \to K[X]^{(n)}$, $D(p) = p'$ (vgl.(2.2 h)). Wegen

$$D(X^k) = kX^{k-1} \ (k = 0, \ldots, n) \ ,$$

ist

$$\mathrm{Mat}(D; \mathfrak{a}, \mathfrak{a}) = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 0 & & \vdots \\ \vdots & \vdots & 0 & 3 & & 0 \\ 0 & 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & & n \\ 0 & & & \ldots & & 0 \end{pmatrix} \in \mathrm{Mat}(n+1, n+1; K) \ .$$

2.14 Definition: Isomorphismus

Seien V und W Vektorräume. Eine lineare Abbildung $f : V \to W$ heißt ein **Isomorphismus**, wenn f bijektiv ist.

2.15 Kommentar: Isomorphismus

a) Dass $f : V \to W$ bijektiv ist, bedeutet, dass es eine Abbildung $g : W \to V$ gibt, so dass $g \circ f = \mathrm{id}_V$ und $f \circ g = \mathrm{id}_W$ ist (nämlich $g = f^{-1}$). Ist $f : V \to W$ ein Isomorphismus so ist $g = f^{-1}$ auch automatisch linear (also auch ein Isomorphismus), denn sind $w_1, w_2 \in W$ und $v_1 = g(w_1), v_2 = g(w_2)$, so ist:

$$f(v_1 + v_2) = f(v_1) + f(v_2) = f \circ g(w_1) + f \circ g(w_2) = w_1 + w_2 \ ,$$

also

$$g(w_1 + w_2) = g \circ f(v_1 + v_2) = v_1 + v_2 = g(w_1) + g(w_2) \ .$$

Ähnlich sieht man: $g(\lambda w) = \lambda g(w)$, für $\lambda \in K$ und $w \in W$.

b) Man sagt, dass zwei Vektorräume isomorph sind, und schreibt dafür

$$V \cong W \ ,$$

wenn es einen Isomorphismus $f : V \to W$ gibt. Sie sind dann im Wesentlichen gleich, denn f ist nicht nur bijektiv (also eine 1:1-Beziehung), sondern respektiert auch die Strukturen $+$ und $\cdot$ der Vektorräume:

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

für $v, v_1, v_2 \in V, \lambda \in K$.

2.16 Bemerkung

Seien V und W Vektorräume, $f : V \to W$ linear, $\mathfrak{a} = (v_1, \dots, v_n)$ ein n-Tupel von Vektoren in V und $\mathfrak{b} = (f(v_1)), \dots, (f(v_n))$. Dann gilt:

a) Ist f injektiv und $\mathfrak{a}$ linear unabhängig, so ist auch $\mathfrak{b}$ linear unabhängig.

b) Ist f surjektiv und $\mathfrak{a}$ Erzeugendensystem, so ist auch $\mathfrak{b}$ ein Erzeugendensystem.

Beweis a)
Sei $\lambda_1, \dots, \lambda_n \in K$ mit $\lambda_1 f(v_1) + \lambda_n f(v_n) = 0$
$\overset{f\ linear}{\Rightarrow} f(\lambda_1 v_1 + \ldots + \lambda_n v_n) = 0 \Rightarrow$ (f injektiv) $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$
$\Rightarrow$ ($\mathfrak{a}$ linear unabhängig) $\lambda_1 = \ldots = \lambda_n = 0$. Also ist $\mathfrak{b}$ linear unabhängig.

Beweis b) Sei $w \in W$ beliebig $\overset{f\ surjektiv}{\Rightarrow} \exists\ v \in V : f(v) = w \overset{\text{da } \mathfrak{a} \text{ Erz.-S.}}{\Rightarrow}$ $\exists\ \lambda_1, \dots, \lambda_n \in K : v = \lambda_1 v_1 + \ldots + \lambda_n v_n \Rightarrow w = f(v) = f(\lambda_1 v_1 + \ldots + \lambda_n v_n)$ $= \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n)$. Also ist $\mathfrak{b}$ Erzeugendensystem.

$\square$

2.17 Korollar

Seien V und W Vektorräume mit $V \cong W$. dann gilt:

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

Beweis:
Ist $\dim V = n < \infty$ und $\mathfrak{a} = (v_1, \dots, v_n)$ eine Basis von V, so zeigt (2.16), dass für einen Isomorphismus $f : V \to W$ gilt: $(f(v_1), \dots, f(v_n))$ ist Basis von W, also

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

Ist $\dim V = \infty$, so zeigt (2.16), dass auch W unendlich-dimensional ist.

$\square$

2.18 Kommentar: Isomorphierelation

a) (2.17) formuliert man auch so, dass die Dimension eines Vektorraums ein **(Isomorphie-) Invariante** ist. Es ist im Wesentlichen die einzige Invariante, wie wir gleich sehen werden.

b) Sind V, W und U Vektorräume und $f : V \to W$ und $g : W \to U$ linear, so ist auch $g \circ f : V \to W$ linear (Übung). Deshalb gilt auch: Sind f und g Isomorphismus, so auch $g \circ f$. Für die *Isomorphie-Relation* zwischen K-Vektorräumen gilt daher Folgendes:

i) $V \cong V$ (Reflexivität), für alle V, weil id_V ein Isomorphismus ist.

ii) $V \cong W \Rightarrow W \cong V$ (Symmetrie), für alle V, W, denn ist $f : V \to W$ ein Isomorphismus, so auch $f^{-1} : W \to V$.

iii) $V \cong W,\ W \cong U \Rightarrow V \cong U$ (Transitivität), für alle V, W und U.

Es ist deshalb $\cong$ eine **Äquivalenzrelation**.

2.19 Satz

Sei V ein K-Vektorraum der Dimension $n \in \mathbb{N}$. Dann gilt:

$$V \cong K^n \ .$$

Beweis:
Ist $\mathfrak{a} = (v_1, \ldots, v_n)$ eine Basis von V, so wissen wir bereits aus (2.7), dass die Abbildung:

$$T^{\mathfrak{a}} : K^n \to V \ ,$$

$$T^{\mathfrak{a}}(\lambda_1, \ldots, \lambda_n) = \lambda_1 v_1 + \ldots + \lambda_n v_n$$

bijektiv ist. Sie ist offenbar auch K-linear und damit ein Isomorphismus.

□

2.20 Kommentar: Koordinatenisomorphimus

a) Unter den endlich-dimensionalen K-Vektorräumen V gibt es nach (2.19) im Wesentlichen nur K^n ($n \in \mathbb{N}$). Man beachte allerdings, dass der Isomorphismus $T^{\mathfrak{a}}$ nicht kanonisch ist, sondern von der Basiswahl $\mathfrak{a}$ abhängt.

b) Man bezeichnet $T^{\mathfrak{a}}$ auch als **Koordinatenisomorphismus**. Sind V und W Vektorräume gleicher Dimension, $\mathfrak{a} = (v_1, \ldots, v_n)$ und $\mathfrak{b} = (w_1, \ldots, w_n)$ Basen von V und W, und $f : V \to W$ die lineare Abbildung, die durch $f(v_j) = w_j$ ($j = 1, \ldots, n$) bestimmt ist, so ist offenbar folgendes **Diagramm kommutativ**:

Abbildung fehlt

d.h.
$$f \circ T^{\mathfrak{a}} = T^{\mathfrak{b}} \ .$$
Denn $T^{\mathfrak{a}}$ bildet die kanonische Basis $\mathfrak{K}$ von K^n gerade auf $\mathfrak{a}$ ab und damit $f \circ T^{\mathfrak{a}}$ und $T^{\mathfrak{b}}$ beide die kanonische Basis auf $\mathfrak{b}$ ab. Es ist also
$$f = T^{\mathfrak{b}} \circ (T^{\mathfrak{a}})^{-1}$$
Isomorphismus,
$$V \cong W \ .$$

c) Bezeichnet man mit $E_n \in \mathrm{Mat}(n, n; K)$ die **Einheitsmatrix**, d.h.:
$$E_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} = (\delta_{ij}) \ ,$$
so gilt für $T^{\mathfrak{a}}$ und f aus Teil (b) mit (2.11):
$$T^{\mathfrak{a}} = f_E^{\mathfrak{K},\mathfrak{a}}, \qquad f = f_E^{\mathfrak{a},\mathfrak{b}} \ .$$

d) Mit den Bezeichnungen von (2.12) ist deshalb nun für beliebige endlich-dimensionale Vektorräume V und W, jeder linearen Abbildung $f : V \to W$, Basen $\mathfrak{a} = (v_1, \dots, v_n)$ von V und $\mathfrak{b} = (w_1, \dots, w_m)$ von W und der Matrix $A = M(f; \mathfrak{a}, \mathfrak{b}) \in \mathrm{Mat}(m, n; K)$ folgendes Diagramm kommutativ:

Abbildung fehlt

also
$$f \circ T^{\mathfrak{a}} = T^{\mathfrak{b}} \circ f_A \ .$$
Setzt man nämlich $\widetilde{w}_j := \sum_{i=1}^m a_{ij} w_i$ $(j = 1, \dots, n)$, so bilden $f \circ T^{\mathfrak{a}}$ und $T^{\mathfrak{b}} \circ f_A$ beide die kanonische Basis $\mathfrak{K} = (e_1, \dots e_n)$ von K^n gerade auf $(\widetilde{w}_1, \dots, \widetilde{w}_n)$ ab:
$$f \circ T^{\mathfrak{a}}(e_j) = f(v_j) = \sum_{i=1}^m a_{ij} w_i = \widetilde{w}_j \ ,$$
$$T^{\mathfrak{b}} \circ f_A(e_j) = T^{\mathfrak{b}}(Ae_j) = T^{\mathfrak{b}}(\sum_{i=1}^m a_{ij} e_i) = \sum_{i=1}^m a_{ij} T^{\mathfrak{b}}(e_i) = \sum_{i=1}^m a_{ij} w_i = \widetilde{w}_j \ .$$
Man sagt: $A = M(f; \mathfrak{a}, \mathfrak{b})$ beschreibt die lineare Abbildung $f : V \to W$ in den Koordinaten bezüglich $\mathfrak{a}$ und $\mathfrak{b}$.

2.21 Satz

Seien V und W Vektorräume der Dimensionen n und m, $\mathfrak{a}$ eine Basis von V und $\mathfrak{b}$ eine Basis von W. Dann sind die Abbildungen.

$$\Phi : \mathrm{Mat}(m,n) \to \mathrm{Hom}(V,W) \ ,$$

$$\Phi(A) = f_A^{\mathfrak{a},\mathfrak{b}} \ ,$$

$$\Psi : \mathrm{Hom}(V,W) \to \mathrm{Mat}(m,n) \ ,$$

$$\Psi(f) = M(f;\mathfrak{a},\mathfrak{b})$$

Isomorphismen und invers zueinander.

Beweis a)
Ist $\mathfrak{a} = (v_1,\ldots,v_n)$ und $\mathfrak{b} = (w_1,\ldots,w_m)$, so ist $\Phi(A) : V \to W$ die lineare Abbildung, die v_j auf $\widetilde{w}_j := \sum_{i=1}^m a_{ij}w_i$ abbildet. Für $A, B \in \mathrm{Mat}(m,n)$ bildet also $\Phi(A+B)$ den Vektor v_j aus $\mathfrak{a}$ auf

$$\widetilde{\widetilde{w_j}} = \sum_{i=1}^m (a_{ij}+b_{ij})w_i = \sum_{i=1}^m a_{ij}w_i + \sum_{i=1}^m b_{ij}w_i$$

ab, genauso wie $\Phi(A)+\Phi(B)$. Also ist $\Phi(A+B) = \Phi(A)+\Phi(B)$. Ähnlich sieht man: $\Phi(\lambda A) = \lambda\Phi(A), \ \forall\ A \in \mathrm{Mat}(m,n), \ \forall\ \lambda \in K$. Also ist Φ linear.

Beweis b)
Ist $f : V \to W$ linear und $A = M(f;\mathfrak{a},\mathfrak{b})$, so ist $f(v_j) = \sum_{i=1}^m a_{ij}w_i$. Ist $B = M(g;\mathfrak{a},\mathfrak{b})$ für ein $g \in \mathrm{Hom}(V,W)$, so ist also:

$$(f+g)(v_j) = f(v_j)+g(v_j) = \sum_{i=1}^m a_{ij}w_i + \sum_{i=1}^m b_{ij}w_i = \sum_{i=1}^m (a_{ij}+b_{ij})w_i$$

also $A+B = M(f+g;\mathfrak{a},\mathfrak{b})$. d.h. :

$$\Psi(f+g) = \Psi(f)+\Psi(g) \ .$$

Ähnlich:

$$\Psi(\lambda f) = \lambda\Psi(f), \ \forall\ f \in \mathrm{Hom}(V,W), \lambda \in K \ .$$

Damit ist Ψ linear.

Beweis c)
Für $A \in \mathrm{Mat}(m,n)$ ist $f_A^{\mathfrak{a},\mathfrak{b}}(v_j) = \sum_{i=1}^m a_{ij}w_i$, also ist

$$\Psi \circ \Phi(A) = \Psi(f_A^{\mathfrak{a},\mathfrak{b}}) = A \ .$$

Ist $f \in \mathrm{Hom}(V,W)$ und $(a_{ij}) = A = \Psi(f)$, so ist $f(v_j) = \sum_{i=1} a_{ij}w_i$ und daher für jedes $j = 1,\dots,n$:

$$\Phi \circ \Psi(f)(v_j) = \Phi(A)(v_j) = f_A^{\mathfrak{a},\mathfrak{b}}(v_j) = \sum_{i=1}^m a_{ij}w_i = f(v_j) \ ,$$

also ist

$$\Psi \circ \Phi = \mathrm{id},\ \Phi \circ \Psi = \mathrm{id} \ .$$

□

2.22 Kommentar

a) Da $\mathrm{Mat}(m,n;K) \cong K^{mn}$ (mit $(a_{ij}) \mapsto (a_{11}, a_{12}, \dots, a_{1n}, a_{21} \dots . a_{mn})$) ist, wissen wir, dass $\dim_K \mathrm{Mat}(m,n) = m \cdot n$ ist. Es ist also:

$$\dim \mathrm{Hom}(V,W) = \dim V \cdot \dim W$$

b) Satz (2.21) ist der präzise Ausdruck dafür, dass man endlich-dimensionale Probleme über Vektorräume und lineare Abbildungen mit Matrizen behandeln kann.

2.23 Definition: Kern und Bild

Sei $f : V \to W$ eine lineare Abbildung. Dann heißt:

a) $\ker(f) = \{v \in V : f(v) = 0\} \subseteq V$ der **Kern von** f und

b) $\mathrm{im}(f) := \{w \in W : \ \exists\ v \in V : \ f(v) = w\} \subseteq W$ das **Bild von** f.

2.24 Kommentar: Kern und Bild sind Unterräume

$\ker(f) \subseteq V$ und $\mathrm{im}(f) \subseteq W$ sind nicht nur Teilmengen von V und W, sondern sogar Untervektorräume. Sind nämlich $v_1, v_2 \in \ \ker(f),\ \lambda_1, \lambda_2 \in \ K$, so ist:

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) = \lambda_1 0 + \lambda_2 0 = 0 \ ,$$

also (da $0 \in \ker(f)$) ist $\ker(f)$ Untervektorraum. Ähnlich ist offenbar $0 \in \operatorname{im}(f)$ und für $w_1, w_2 \in \operatorname{im}(f)$ (also gibt es $v_1, v_2 \in V$ mit $f(v_1) = w_1$ und $f(v_2) = w_2$), $\lambda_1, \lambda_2 \in K$ ist:

$$f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2) = \lambda_1 w_1 + \lambda_2 w_2 \ ,$$

also $\lambda_1 w_1 + \lambda_2 w_2 \in \operatorname{im}(f)$ und damit $\operatorname{im}(f) \subseteq W$ ein Unterraum von W.

□

2.25 Lemma: Zusammenhang Injektivität und Kern

Sei $f : V \to W$ linear. Dann gilt:

a) Es ist f injektiv genau wenn

$$\ker(f) = \{0\} \ .$$

b) Ist $(v_1, \dots, v_n)$ Erz-System für V, so ist $(f(v_1), \dots, f(v_n))$ Erz-System für $\operatorname{im}(f)$.

Beweis a)
$\Rightarrow$ Sei $f(v) = 0$, also $v \in \ker(f)$. Wegen $f(0) = 0$ und der Injektivität von f folgt: $v = 0$. $\Rightarrow \ker(f) = \{0\}$.
$\Leftarrow$ Seien $v_1, v_2 \in V$ mit $f(v_1) = f(v_2)$

$$\Rightarrow 0 = f(v_2) - f(v_1) \underbrace{=}_{linear} f(v_2 - v_1) \ ,$$

also $v_2 - v_1 \in \ker(f) = \{0\} \Rightarrow v_1 = v_2$. Also ist f injektiv.

Beweis b)
Schränkt man den Wertebreich von f auf $\operatorname{im}(f) \subseteq W$ ein, also $\widetilde{f} : V \to \operatorname{im}(f)$, $\widetilde{f}(v) = f(v)$, so ist $\widetilde{f}$ surjektiv. Nach (2.16.b) ist dann $(\widetilde{f}(v_1), \dots, \widetilde{f}(v_n)) = (f(v_1), \dots, f(v_n))$ Erzeugendensystem von $\operatorname{im}(f)$.

□

2.26 Kommentar: Rang von f

a) Die Dimension von $\operatorname{im}(f)$ kann höchstens $\dim(V)$ sein, denn ist $(v_1, \dots, v_n)$ eine Basis von V, so ist $(f(v_1), \dots, f(v_n))$ immer noch erzeugend für $\operatorname{im}(f)$, also:

$$\dim(\operatorname{im}(f)) \leq n = \dim V \ .$$

b) Es kann daher z.B. keine surjektive, lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^3$ geben. (Nicht lineare surjektive (sogar bijektive) Abbildungen $\mathbb{R}^2 \to \mathbb{R}^3$ gibt es sehr wohl.)

c) Man nennt:

$$\mathrm{rg}(f) := \dim(\mathrm{im}(f))$$

den **Rang von** f. (Beachte: $0 \le \mathrm{rg}(f) \le \min\{\dim V, \dim W\}$.)

2.27 Satz: Dimensionsformel

Seien V und W Vektorräume und $f : V \to W$ linear. Dann gilt:

$$\dim(\ker(f)) + \dim(\mathrm{im}(f)) = \dim V \ .$$

Beweis:
Sei $\dim(\ker(f)) =: k < \infty$ und $\dim(\mathrm{im}(f)) =: l < \infty$. Sei weiter $(v_1, \ldots, v_k)$ Basis von $\ker(f)$ und $(w_{k+1}, \ldots, w_{k+l})$ eine Basis von $\mathrm{im}(f) \subseteq W$, $n := k + l$. Wähle dann:

$$v_{k+1}, \ldots, v_n \in V \text{ mit } f(v_j) = w_j \ (j = k+1, \ldots, n) \ .$$

Behauptung: $(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ ist eine Basis von V.

a) Lineare Unabhängigkeit: Seien $\lambda_1, \ldots, \lambda_n \in K$ mit $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$.

$$\begin{aligned}
\Rightarrow 0 &= f(0) = f(\lambda_1 v_1 + \ldots + \lambda_n v_n) \\
&= \underbrace{\lambda_1 f(v_1)}_{=0} + \ldots + \underbrace{\lambda_k f(v_k)}_{=0} + \lambda_{k+1} f(v_{k+1}) + \ldots + \lambda_n f(v_n) \\
&= \lambda_{k+1} \underbrace{f(v_{k+1})}_{=w_{k+1}} + \ldots + \lambda_n \underbrace{f(v_n)}_{=w_n} \\
&\Rightarrow \lambda_{k+1} w_{k+1} + \ldots + \lambda_n w_n = 0 \\
&\Rightarrow \lambda_{k+1} = \ldots = \lambda_n = 0,
\end{aligned}$$

weil $(w_{k+1}, \ldots, w_n)$ linear unabhängig ist.

$$\begin{aligned}
&\Rightarrow \lambda_1 v_1 + \ldots + \lambda_k v_k = 0 \\
&\Rightarrow \lambda_1 = \ldots = \lambda_k = 0,
\end{aligned}$$

weil $(v_1, \ldots, v_k)$ linear unabhängig ist. Insgesamt: $\lambda_1 = \ldots = \lambda_n = 0$ also: $(v_1, \ldots, v_n)$ ist linear unabhängig.

b) Erzeugendensystem: Sei $v \in V$ beliebig. Da $(w_{k+1}, \ldots, w_n)$ Erz-System von $\mathrm{im}(f)$ ist, existieren $\lambda_{k+1}, \ldots, \lambda_n \in K$ mit $f(v) = \lambda_{k+1} w_{k+1} + \ldots + \lambda_n w_n$

$$\Rightarrow f(v) = \lambda_{k+1} f(v_{k+1}) + \ldots + \lambda_n f(v_n) = f(\lambda_{k+1} v_{k+1} + \ldots + \lambda_n v_n)$$

$$\Rightarrow f(v - \lambda_{k+1} v_{k+1} - \ldots - \lambda_n v_n) = 0 \; .$$

Da $(v_1, \ldots, v_k)$ erzeugend für $\ker(f)$ ist, gibt es $\lambda_1, \ldots, \lambda_k \in K$ mit

$$v - \lambda_{k+1} v_{k+1} - \ldots - \lambda_n v_n = \lambda_1 v_1 + \ldots + \lambda_k v_k$$

$$\Rightarrow v = \lambda_1 v_1 + \ldots + \lambda_k v_k + \lambda_{k+1} v_{k+1} + \ldots + \lambda_n v_n \; .$$

Also ist $(v_1, \ldots, v_n)$ auch erzeugend und es folgt:

$$\dim(\ker(f)) + \dim(\mathrm{im}(f)) = k + l = n = \dim V \; .$$

Teil a) zeigt auch, dass für den Fall, wo $\ker(f)$ oder $\mathrm{im}(f)$ unendlich dimensional sind, es auch V ist. Mit $a + \infty := \infty, \quad \forall\, a \in \mathbb{N}_0 \cup \{\infty\}$, gilt (2.27) also auch im ∞-dimensionalen Fall.

$\square$

2.28 Korollar

Sei $\dim V = \dim W < \infty$ und $f : V \to W$ linear. Dann gilt:

a) Ist f injektiv, so ist f bereits bijektiv.

b) Ist f surjektiv, so ist f bereits bijektiv.

Beweis a)
Aus f injektiv, folgt: $\dim(\ker(f)) = 0$. Also ist $\mathrm{im}(f) \subseteq W$ ein Unterraum mit:

$$\dim(\mathrm{im}(f)) = \dim V - \dim(\ker(f)) = \dim V = \dim W \; .$$

Also ist nach (1.41.b) $\mathrm{im}(f) = W$ und damit f auch surjektiv.

Beweis b)
Ist f surjektiv, so ist also $\mathrm{im}(f) = W$ und damit

$$\dim(\ker(f)) = \dim V - \dim(\mathrm{im}(f)) = \dim V - \dim W = 0 \; ,$$

also $\ker(f) = \{0\}$, und damit f nach (2.25) auch injektiv.

$\square$

3 Matrizen

Sei K (wie immer) ein Körper.

3.1 Definition: Matrizen-Produkt

Das **Matrizen-Produkt** wird wie folgt definiert: Sind $m, n, r \in \mathbb{N}$, so setzt man:

$$\mathrm{Mat}(m, n; K) \times \mathrm{Mat}(n, r; K) \to \mathrm{Mat}(m, r; K)$$

$(A, B) \mapsto AB =: C = (c_{ik})_{i=1,\ldots,m;\ k=1,\ldots,r}$ mit:

$$c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk}\ ,$$

$$\begin{pmatrix} & \vdots & \\ a_{i1} & \cdots & a_{in} \\ & \vdots & \end{pmatrix} \begin{pmatrix} & b_{1k} & \\ \ldots & \vdots & \ldots \\ & b_{nk} & \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots & c_{ik} & \cdots \\ & \vdots & \end{pmatrix}\ .$$

3.2 Kommentar: Matrizen-Produkt

a) Schreibt man die Spalten von B als Vektoren in K^n,

$$B = (b_1, \ldots, b_r)\ ,$$

d.h.

$$b_k = \sum_{j=1}^{n} b_{jk} e_j\ ,$$

so stehen in AB also gerade die Spaltenvektoren $Ab_1, \ldots, Ab_r \in K^n$ (vgl.(2.6 b)).

b) Man beachte, dass man zwei Matrizen $A \in \mathrm{Mat}(m, n)$ und $B \in \mathrm{Mat}(s, r)$ nur dann multiplizieren kann, wenn die Anzahl der Spalten von A mit der Anzahl der Zeilen von B übereinstimmt, $n = s$.

c) Selbst wenn man AB und BA bilden kann, ist dies i.A. verschieden. Ist z.B. $A \in \mathrm{Mat}(1, n), A = (a_1, \ldots, a_n)$, und $B \in \mathrm{Mat}(n, 1)$,

$$B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} ,$$

so ist $AB \in Mat(1, 1)$ die 1×1-Matrix mit dem Eintrag

$$a_1 b_1 + \cdots + a_n b_n \ .$$

Dagegen ist $BA \in Mat(n, n)$ die $n \times n$-Matrix

$$C = (c_{ik})_{1 \leq i,\, k \leq n}$$

mit Einträgen

$$c_{ik} = a_i b_k \ .$$

d) Selbst im Fall, dass $A, B \in \mathrm{Mat}(n, n)$ sind, ist i.A. (für $n \geq 2$): $AB \neq BA$ (Übung).

3.3 Bemerkung: Matrizen und deren linearen Abbildungen

Seien $A \in \mathrm{Mat}(m, n; K)$ und $B \in \mathrm{Mat}(n, r; K)$ sowei $f_A : K^n \to K^m$ und $f_B : K^r \to K^n$ die zugehörigen linearen Abbildungen. Dann gilt:

$$f_{AB} = f_A \circ f_B$$

Beweis:
Für alle $z \in K^r$ ist $f_A \circ f_B(z) = f_A(Bz) = A(Bz)$. Da

$$Ax := \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) e_i \ ,$$

$$Bz := \sum_{j=1}^{n} \left(\sum_{k=1}^{r} b_{jk} z_k \right) e_j \ ,$$

ist, folgt:

$$
\begin{aligned}
A(Bz) &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \left(\sum_{k=1}^{r} b_{jk} z_k \right) \right) e_i \\
&= \sum_{i=1}^{m} \left(\sum_{k=1}^{r} \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right) z_k \right) e_i \\
&= (AB)z \ .
\end{aligned}
$$

Also ist:

$$f_A \circ f_B(z) = f_{AB}(z), \quad \forall z \in K^r \ ,$$

und damit

$$f_A \circ f_B = f_{AB} \ .$$

□

3.4 Kommentar: allgemeines Assoziativgesetz

a) Die Anwendung einer Matrix $A \in \mathrm{Mat}(m, n)$ auf einen Vektor $x \in K^n$ (siehe (2.6)) kann man als Spezialfall von (3.1) auffassen, wenn man K^n mit $\mathrm{Mat}(n, 1)$ identifiziert, die Vektoren $x \in K^n$ also als Spaltenvektoren auffasst. Die Regel $A(Bz) = (AB)z$ ist dann ein Spezialfall des allgemeinen Assoziativgesetzes

$$A(BC) = (AB)C$$

für $A \in \mathrm{Mat}(m, n)$, $B \in \mathrm{Mat}(n, r)$ und $C \in \mathrm{Mat}(r, s)$ (Übung).

b) Das Matrizenprodukt entspricht also dem Hintereinanderschalten von linearen Abbildungen (zwischen Vektorräumen K^n, K^m und K^r). Das gilt auch für (abstrakte) Vektorräume in folgendem Sinn:

3.5 Satz

Seien V, W, U endlich dimensionale Vektorräume und $\mathfrak{a}, \mathfrak{b}$ und $\mathfrak{c}$ Basen von V, W und U. Seien $f : V \to W$ und $g : W \to U$ linear. Dann gilt:

$$M(g \circ f; \mathfrak{a}, \mathfrak{c}) = M(g; \mathfrak{b}, \mathfrak{c}) \cdot M(f; \mathfrak{a}, \mathfrak{b})$$

Beweis:
Seien $A = M(f; \mathfrak{a}, \mathfrak{b})$, $B = M(g; \mathfrak{b}, \mathfrak{c})$ und $C = M(g \circ f; \mathfrak{a}, \mathfrak{c})$ sowie $\mathfrak{a} = (v_1, \ldots, v_n)$, $\mathfrak{b} = (w_1, \ldots, w_m)$ und $\mathfrak{c} = (u_1, \ldots, u_r)$. Es ist dann

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i \ ,$$

$$g(w_i) = \sum_{k=1}^{r} b_{ki} u_k \ ,$$

$$g \circ f(v_j) = \sum_{k=1}^{r} c_{kj} u_k \ .$$

Es folgt:

$$\begin{aligned} \sum_{k=1}^{r} c_{kj} u_k &= g \circ f(v_j) = g(f(v_j)) = g\left(\sum_{i=1}^{m} a_{ij} w_i\right) = \sum_{i=1}^{m} a_{ij} g(w_i) \\ &= \sum_{i=1}^{m} a_{ij} \left(\sum_{k=1}^{r} b_{ki} u_k\right) = \sum_{k=1}^{r} \left(\sum_{i=1}^{m} b_{ki} a_{ij}\right) u_k \ , \end{aligned}$$

für alle $j = 1, \ldots, n$. Koeffizientenvergleich liefert ($(u_1, \ldots, u_r)$ ist linear unabhängig!) für alle $1 \leq k \leq r$ und $1 \leq j \leq n$:

$$c_{kj} = \sum_{i=1}^{m} b_{ki} a_{ij} \ ,$$

also

$$C = BA \ .$$

□

3.6 Definition: Spaltenrang

Sei $A \in \mathrm{Mat}(m, n; K)$ bestehend aus den Spaltenvektoren $a_1, \ldots, a_n \in K^m$,

$$A = (a_1, \ldots, a_n) \ .$$

Dann heißt die Dimension des Untervektorraums $\langle a_1, \ldots, a_n \rangle \subseteq K^m$ der **Spaltenrang von A**:

$$\mathrm{rg}_s(A) := \dim\langle a_1, \ldots, a_n \rangle \ .$$

3.7 Kommentar: Zeilenrang

Schreibt man $A \in \mathrm{Mat}(n, m; K)$ als eine Spalte von Zeilenvektoren $\widetilde{a}_1, \ldots, \widetilde{a}_m \in K^n$, so definiert man analog zu (3.6) den **Zeilenrang von** A als Dimension von $\langle \widetilde{a}_1, \ldots, \widetilde{a}_m \rangle \subseteq K^n$,

$$\mathrm{rg}_z(A) := \dim \langle \widetilde{a}_1, \ldots, \widetilde{a}_m \rangle \ .$$

Später werden wir sehen (siehe (3.36)), dass stets der Zeilenrang gleich dem Spaltenrang ist, also $\mathrm{rg}_s(A) = \mathrm{rg}_z(A)$, $\forall A \in \mathrm{Mat}(m, n)$, und diese Zahl (zwischen 0 und $\min(m, n)$) heißt dann einfach **der Rang von** A,

$$\mathrm{rg}(A) := \mathrm{rg}_s(A) \ \ (= \mathrm{rg}_z(A)) \ .$$

Z.B. ist für

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & 9 \end{pmatrix} \in \mathrm{Mat}(2, 3; \mathbb{R})$$

$\mathrm{rg}_z(A) = 1$, da $((1, 2, 3), (3, 6, 9))$ linear abhängig in $\mathbb{R}^3$ sind, aber $\mathrm{rg}_s(A) = 1$, weil $((1, 3)(2, 6))$ und $((1, 3), (3, 9))$ jeweils linear abhängig in $\mathbb{R}^2$ sind.

3.8 Bemerkung: Rang von f

Sei $f : V \mapsto W$ linear, $\mathfrak{a}$ und $\mathfrak{b}$ endliche Basen von V und W. Dann gilt für den Rang von f (vgl. (2.26.b)):

$$\mathrm{rg}(f) = \mathrm{rg}_s(M(f; \mathfrak{a}, \mathfrak{b})) \ .$$

Beweis:
Der Rang von f ist die Dimension von $\mathrm{im}(f) \subseteq W$. Ist $\mathfrak{a} = (v_1, \ldots, v_n)$, so wird das Bild $\mathrm{im}(f)$ von f von $f(v_1), \ldots, f(v_n) \subseteq W$ aufgespannt. Ist nun $f(v_{i_1}), \ldots, f(v_{i_r})$ (mit $1 \leq i_1 < \ldots < i_r \leq n$) ein minimales Erzeugendensystem, so ist dies eine Basis von $\mathrm{im}(f)$, also $\mathrm{rg}(f) = r$.
Ist andererseits $f(v_j) = \sum_{i=1}^m a_{ij} w_i$, so ist $M(f; \mathfrak{a}, \mathfrak{b}) = (a_{ij}) = A$, also wird $f(v_j)$ unter dem Koordinatenisomorphismus $T^{\mathfrak{b}} : K^m \mapsto W$ gerade durch den Spaltenvektor $a_j = \sum_{i=1}^m a_{ij} e_i \in K^m$ beschreiben. Es ist daher auch $(a_{i_1}, \ldots a_{i_r})$ ein minimales Erzeugendensystem von $\langle a_1, \ldots, a_n \rangle \subseteq K^m$, also auch $r = \mathrm{rg}_s(A)$.

3.9 Kommentar

Sind V und W Vektorräume der Dimension n und m und $f : V \to W$ linear, so kann nicht jede Matrix $A \in \mathrm{Mat}(m, n)$ als beschreibende Matrix für f auftauchen, wenn man die Basen $\mathfrak{a}$ von V und $\mathfrak{b}$ von W variieren lässt. Es muss nämlich stets $\mathrm{rg}_s(A) = \mathrm{rg}(f)$ sein.

3.10 Satz

Seien V und W endlich dimensionale Vektorräume der Dimension n und m und $f : V \to W$ linear vom Rang r. Dann gibt es Basen $\mathfrak{a}$ und $\mathfrak{b}$ von V bzw. W, so dass gilt:

$$M(f;\mathfrak{a},\mathfrak{b}) = \begin{array}{c} \left(\begin{array}{cc} E_r & 0 \\ \underbrace{0}_{r} & \underbrace{0}_{n-r} \end{array} \right) \end{array} \begin{array}{l} \}r \\ \}m-r \end{array} .$$

Beweis:
Ähnlich wie im Beweis von (2.27) wählen wir eine Basis $\mathfrak{a} = (v_1, \dots, v_n)$ von V so, dass $(f(v_1), \dots, f(v_r))$ eine Basis von $\mathrm{im}(f)$ und $(v_{r+1}, \dots, v_n)$ eine Basis von $\ker(A)$ ist. Dann ergänzen wir $(f(v_1), \dots, f(v_r))$ (irgendwie) zu einer Basis $\mathfrak{b}$ von W. Es ist dann:

$$\begin{aligned} f(v_j) &= \sum_{k=1}^{r} \delta_{kj} f(v_k) \quad \text{für } j = 1, \dots, r \\ f(v_j) &= 0 \quad \text{für } j = r+1, \dots, n , \end{aligned}$$

also

$$M(f;\mathfrak{a},\mathfrak{b}) = \begin{pmatrix} (\delta_{kj}) & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} .$$

3.11 Motivation: verschiedene Umformungen

Zum Bestimmen des Zeilenrangs (und damit auch des Spaltenrangs) einer Matrix (und zu einer ganzen Reihe von Anwendungsproblemen) führen wir folgende Umformungen von einer $m \times n$-Matrix A mit Zeilenvektoren $a_1, \dots, a_m \in K^n$ ein:

Typ I: Sei $i \in \{1, \dots, m\}$ und $\lambda \in \underbrace{K^*}_{=K \setminus \{0\}}$. Dann multipliziere a_i mit λ:

$$\begin{pmatrix} a_1 \\ \vdots \\ \vdots \\ \vdots \\ a_m \end{pmatrix} \to \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i \\ \vdots \\ a_m \end{pmatrix} .$$

Typ II: Seien $i, j \in \{1, \ldots, m\}$, $i \neq j$ und $\lambda \in K$. Dann multipliziere a_i mit λ und addiere dies zu a_j:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_m \end{pmatrix} \to \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ \lambda a_i + a_j \\ \vdots \\ a_m \end{pmatrix} .$$

Typ III: Seien $i, j \in \{1, \ldots, m\}, i \neq j$. Dann vertausche a_i mit a_j:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_m \end{pmatrix} \to \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_m \end{pmatrix} .$$

3.12 Kommentar: Elementarmatrizen

a) Wendet man auf eine Matrix $A \in \mathrm{Mat}(m, n; K)$ Umformungen vom Typ I, II oder III an, so hat die neue Matrix $B \in \mathrm{Mat}(m, n; K)$ den gleichen Zeilenrang wie A, denn:

i) $\langle a_1, \ldots, a_i, \ldots, a_m \rangle = \langle a_i, \ldots, \lambda a_i, \ldots, a_m \rangle$ für $\lambda \neq 0$

ii) $\langle a_1, \ldots, a_m \rangle = \langle a_1, \ldots, a_i, \ldots, \lambda a_i + a_j, \ldots, a_m \rangle$ für $\lambda \in K$

iii) $\langle a_1, \ldots, a_m \rangle = \langle a_1, \ldots, a_j, \ldots, a_i, \ldots, a_m \rangle$

weil $a_i = \frac{1}{\lambda}(\lambda a_i)$ im Fall I und $a_j = -\lambda a_i + (\lambda a_i + a_j)$ im Fall II ist. Die Zeilen von B spannen sogar den gleichen Unterraum von K^n auf wie die von A.

b) Die Umformung von A nach B kann man als Multiplikation von links mit so genannten **Elementarmatrizen** $F_i(\lambda)$ $(\lambda \in K^*)$, $G_{ij}(\lambda)$ $(\lambda \in K)$, H_{ij} bekommen. Sei dazu für $1 \leq i \leq m$ und $1 \leq j \leq n$

$$E_{ij} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

also $E_{ij} = (a_{kl})_{1\leq k\leq n,\ 1\leq l\leq m}$, mit $a_{kl} := \delta_{ik}\delta_{jl}$. Wir nennen dann $(E_{11},\ldots,E_{mn})$ **die kanonische Basis** von $\mathrm{Mat}(m,n;K)$. Für die Einheitsmatrix $E_m \in \mathrm{Mat}(m,m;K)$ gilt dann offenbar:

$$E_m = \sum_{i=1}^{m} E_{ii} \ .$$

Nun setzen wir:

i)

$$F_i(\lambda) = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix} \leftarrow i$$

$$= E_m + (\lambda - 1)E_{ii} \text{ für } \lambda \in K^* \text{ und } i \in \{1,\ldots,m\} \ .$$

ii)

$$G_{ij}(\lambda) = \begin{pmatrix} 1 & & & 0 \\ & & \lambda & \\ & \ddots & & \\ & & & \\ 0 & & & 1 \end{pmatrix}$$

$$= E_m + \lambda E_{ij} \text{ für } \lambda \in K \text{ und } i \neq j \ .$$

iii)

$$H_{ij} = \begin{pmatrix} 1 & & & 0 \\ & 0 & 1 & \\ & 1 & 0 & \\ 0 & & & 1 \end{pmatrix}$$

$$= E_m - E_{ii} - E_{jj} + E_{ij} + E_{ji} \text{ für } i \neq j \ .$$

Es ist dann (Übung):

i) Typ I: $B = F_i(\lambda)A,\ \lambda \in K^*$;

ii) Typ II : $B = G_{ij}(\lambda)A,\ \lambda \in K$;

iii) Typ III: $B = H_{ij}(\lambda)A$.

3.13 Satz: Gaußsches Eliminationsverfahren

Sei $A \in \mathrm{Mat}(m, n; K)$. Dann kann man A durch endlich viele Umformungen von Typ I, II und III in eine **Zeilenstufenform** B überführen, d.h.:
B hat folgende Gestalt:

$$B = \begin{pmatrix} 0 & \cdots & 0 & |\underline{\beta_1 *} & \cdots & \cdots & \cdots & * \\ 0 & & & 0\cdots 0 & |\underline{\beta_2 *} & \cdots & \cdots & * \\ \cdot & & & & 0\cdots 0 & |\underline{\beta_3 *} & \cdots & * \\ \cdot & & & & \cdot & 0\cdots 0 & \ddots & \vdots \\ \cdot & & & & \cdot & \cdot & & |\underline{\beta_r\ * \cdots *} \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

mit $0 \leq r \leq min(m, n)$ und Einträgen $\beta_1, \ldots, \beta_r \in K^*$. Die ersten r Zeilenvektoren $b_1, \ldots, b_r \in K^n$ von B bilden dann eine Basis von $\langle a_1, \ldots, a_m \rangle \subseteq K^n$, wo $a_1, \ldots, a_m \in K^n$ die Zeilen von A sind. Insbesondere ist: $\mathrm{rg}_Z(A) = r$.

Beweis a) O.E. $A \neq 0$. Suche zunächst ein minimales $j_1 \in \{1, \ldots, n\}$, so dass es ein $i \in \{1, \ldots, m\}$ mit $a_{ij} \neq 0$ gibt und vertausche dann die 1. mit der i-ten Zeile (falls $i \neq 1$ ist). Setze dann $\beta_1 = a_{ij_1} \neq 0$. Benutze dann Typ II-Umformungen (mit $i = 1, j = 2, \ldots, m$), um in allen Zeilen $i = 2, \ldots, m$ unter β_1 eine Null zu produzieren. Erhalte also:

$$\left(\begin{array}{ccc|c} 0 & \ldots & \beta_1 & * \quad \ldots \quad * \\ \hline \vdots & & 0 & \\ 0 & \ldots & 0 & A^{(1)} \end{array}\right)$$

Verfahre nun mit $A^{(1)}$ so wie vorher mit A. Nach endlich vielen Schritten erhalte B.

Beweis b) Nach (3.12 a) erzeugen die Zeilenvektoren $b_1, \ldots, b_m \in K^n$ von B den gleichen Unterraum von K^n wie die Zeilenvektoren $a_1, \ldots, a_m$ von A. Es ist dann aber $b_{r+1} = \cdots = b_m = 0$ und $(b_1, \ldots, b_r)$ linear unabhängig, denn ist $\lambda_1 b_1 + \cdots + \lambda_r b_r = 0$ für $\lambda_1, \ldots, \lambda_r \in K$ so ist:

$$\begin{aligned} \lambda_1 \beta_1 &= 0 \\ \lambda_1 b_{1j_2} + \lambda_2 \beta_2 &= 0 \\ &\vdots \\ \lambda_1 b_{1j_r} + \cdots + \lambda_{r-1} b_{r-1,j_r} + \lambda_r \beta_r &= 0 \end{aligned}$$

Wegen $\beta_1 \neq 0$ ist dann zunächst $\lambda_1 = 0$ und sukzessive $\lambda_2 = \lambda_3 = \cdots = \lambda_r = 0$. Also ist $(b_1, \ldots, b_r)$ eine Basis von $\langle a_1, \ldots, a_m \rangle$, insbesondere ist $\mathrm{rg}_Z(A) = r$.

3.14 Beispiel: Zeilenrang

$$A = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 4 & 4 & -1 \\ -1 & -2 & 1 & 1 \end{pmatrix} \overset{\substack{G_{13}(1) \\ G_{12}(-2)}}{\longrightarrow} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 3 & 2 \end{pmatrix}$$

$$\overset{H_{23}}{\longrightarrow} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & -3 \end{pmatrix}$$

Es folgt:

$$\mathrm{rg}_Z(A) = 3$$

3.15 Definition: Ring

Ein Tripel $(R, +, \cdot)$ mit einer Menge R und zwei (inneren) Verknüpfungen $+, \cdot : R \times R \to R$ heißt **ein Ring**, wenn gilt:

a) $(R, +)$ ist eine abelsche Gruppe

b) Für alle $a, b, c \in R$ gilt

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c$$

(Assoziativgesetz);

c) Für alle $a, b, c \in R$ gilt

$$(a + b) \cdot c = a \cdot c + b \cdot c$$

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

(Distributivgesetze).

3.16 Kommentar

a) Wie in einem Körper bezeichnet man das neutrale Element **der Addition** $+$ als **Nullelement** des Rings. Ein Ring hat also stets mindestens ein Element, seine Null. $R = \{0\}$ (mit $0+0 := 0, 0 \cdot 0 := 0$) heißt Nullring.

b) Hat ein Ring $(R, +, \cdot)$ auch ein neutrales Element **der Multiplikation** so wird dies als **Einselement** des Rings bezeichnet. Das Nullelement und das Einselement (falls vorhanden) eines Rings sind eindeutig bestimmt (Übung).

c) I.A. ist die Multiplikation in einem Ring nicht kommutativ. Falls doch, so heißt $(R, +, \cdot)$ ein **kommutativer Ring**.

d) Ein Körper $(K, +, \cdot)$ ist also ein kommutativer Ring mit Eins, in dem jedes Element $a \neq 0$ ein Inverses hat, also ein Element $b \in K$ existiert mit $ab = ba = 1$.

3.17 Beispiele: Ring

a) Natürlich ist jeder Körper ein Ring.

b) $R = \mathbb{Z}$ mit den natürlichen Verknüpfungen $+$ und $\cdot$ ist ein kommutativer Ring mit Eins. Man beachte, dass $a \in \mathbb{Z}$ mit $a \neq \pm 1$ kein Inverses (in $\mathbb{Z}$) hat.

c) Ist K ein Körper, so definieren wir auf $R := K[X]$ eine Multiplikation durch

$$\left(\sum_{i=0}^{d} a_i X^i\right) \cdot \left(\sum_{j=0}^{e} b_j X^j\right) := \sum_{k=0}^{d+e} \left(\sum_{i+j=k} a_i b_j\right) X^k$$

Dann wird $(K[X], +, \cdot)$ zu einem kommutativen Ring mit Eins (Übung).

d) Sei K ein Körper, $n \in \mathbb{N}$ und

$$R = \mathrm{Mat}_n(K) := \mathrm{Mat}(n, n, K) \ .$$

Auf R hatten wir bereits die Matritzenmultiplikation $(A, B) \mapsto AB$ definiert. Es ist dann $(R, +, \cdot)$ ein Ring mit Eins, denn die Einheitsmatrix

$$E_n = (\delta_{ij})$$

ist offenbar ein Einselement für R. R ist für $n \geq 2$ nicht kommutativ.

e) Sei V ein K-Vektorraum. Wir nennen eine K-lineare Abbildung $f : V \to V$ einen **K-Endomorphismus** und notieren

$$\mathrm{End}_K(V) := \mathrm{Hom}_K(V, V) \ .$$

Auf $\mathrm{End}_K(V)$ hatten wir bereits eine Addition (und eine skalare Multiplikation (siehe (2.2 j))) erklärt. Eine innere Multiplikation wird nun durch die Komposition von Abbildungen gegeben,

$$f \cdot g := f \circ g \ .$$

Es ist dann $R := \mathrm{End}(V)$ (zusammen mit $+$ und $\cdot$) ein Ring mit Eins (der für $\dim_K V \geq 2$ nicht kommutativ ist). Das Einselement ist dabei gerade die Identität id_V von V.

3.18 Definition: Ring-Homomorphismus

Seien R und S Ringe und $f : R \to S$ eine Abbildung.

a) Es heißt f ein **Ring-Homomorphismus**, wenn für alle $a, b \in R$ gilt:

$$f(a + b) = f(a) + f(b) \ ,$$
$$f(ab) = f(a)f(b) \ .$$

b) Ein Ringhomomorphismus $f : R \to S$ heißt ein **Ringisomorphismus**, wenn f zudem bijektiv ist.

3.19 Satz

Sei V ein endlich-dimensionaler Vektorraum, $n = \dim(V)$ und $\mathfrak{a}$ eine Basis von V. Dann ist die Abbildung

$$\Psi : \mathrm{End}_K(V) \to \mathrm{Mat}_n(K) \ ,$$
$$\Psi(f) = M(f; \mathfrak{a}, \mathfrak{a}) \ ,$$

ein Ringisomorphismus.

Beweis:
Wir wissen bereits, dass Ψ ein Vektorraum-Isomorphismus ist, also bijektiv ist und $\Psi(f + g) = \Psi(f) + \Psi(g)$, $\forall f, g \in \mathrm{End}(V)$ (vgl. (2.21)) erfüllt. Aber nach (3.5) mit $\mathfrak{a} = \mathfrak{b}$ gilt auch

$$\Psi(fg) = \Psi(f)\Psi(g) \ ,$$

$\forall f, g \in \mathrm{End}(V)$. Also ist Ψ ein Ring-Isomorphismus.

$\square$

3.20 Definition: Einheit

Sei R ein Ring mit Eins. Ein Element $a \in R$ heißt **eine Einheit**, wenn a ein Inverses besitzt, also ein $b \in R$ mit $ab = ba = 1$ (Bezeichnung für $b : a^{-1}$).

3.21 Kommentar: Nullteiler, Einheitsgruppe

a) In einem Körper K ist also jedes Element $a \neq 0$ eine Einheit.

b) Ein Element $a \in R$ heißt ein **Nullteiler**, wenn es ein $b \neq 0$ gibt, so dass $ab = 0$ ist. In einem Ring kann es Nullteiler verschieden von Null geben. Ist $a \in R$ eine Einheit, so ist allerdings a sicher kein Nullteiler, denn aus $ab = 0$ für ein $b \in R$ würde folgen:

$$0 = a^{-1} \cdot 0 = a^{-1}(ab) = \underbrace{(a^{-1}a)}_{=1} b = b$$

Ein Körper hat also keine von Null verschiedenen Nullteiler.

c) Setzt man bei einem Ring R mit Eins

$$R^* = \{a \in R : \text{a ist Einheit}\}$$

so ist $(R^*, +, \cdot)$ eine Gruppe (Übung), **die Einheitengruppe von** R.

3.22 Beispiel: Nullteiler

a) Für einen Körper K ist also tatsächlich

$$K^* = K \backslash \{0\} \ .$$

b) Für $R = \mathbb{Z}$ hat man

$$\mathbb{Z}^* = \{\pm 1\} \ ,$$

und $\mathbb{Z}$ hat keine **nicht-trivialen Nullteiler**.

c) Für $R = K[X]$ ist $R^* = K^* \subseteq K[X]$, wobei wir den Körper K als Teilmenge von $K[X]$ als die Polynome vom Grad 0 auffassen, $K \subseteq K[X]$ (Übung). Auch K[X] ist **nullteilerfrei**.

d) Für $R = \mathrm{Mat}_n(K)$ und $n \geq 2$ gibt es nicht-triviale Nullteiler. Z.B. ist

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0 \ .$$

3.23 Definition: invertierbar

Sei K ein Körper und $n \in \mathbb{N}$. Eine Matrix $S \in \mathrm{Mat}_n(K)$ heißt **invertierbar** (oder **regulär**), wenn sie eine Einheit im Matritzenring $\mathrm{Mat}_n(K)$ ist, es also ein $T \in \mathrm{Mat}_n(K)$ gibt mit

$$ST = TS = E_n \ .$$

Die Einheitsgruppe von $\mathrm{Mat}_n(K)$ wird mit

$$GL_n(K) = \{S \in \mathrm{Mat}_n(K) : S \text{ ist invertierbar}\}(= \mathrm{Mat}_n(K)^*)$$

(General Linear Group) bezeichnet.

3.24 Satz

Es ist $A \in \mathrm{Mat}_n(K)$ genau dann invertierbar, wenn $\mathrm{rg}_s(A) = n$ ist.

Beweis:
$\Rightarrow$ Sei $S \in \mathrm{Mat}_n(K)$ invertierbar und $T = S^{-1}$. Es ist dann nach (3.3) für die linearen Abbildungen $f = f_S : K^n \to K^n$ und $g = f_T : K^n \to K^n$.

$$f \circ g = f_S \circ f_T = f_{ST} = f_E = id \ ,$$

$$g \circ f = f_T \circ f_S = f_{TS} = f_E = id$$

(mit $E := E_n$).
Deshalb ist f bijektiv, insbesondere surjektiv, also $\mathrm{im}(f) = K^n$ und damit

$$\mathrm{rg}_s(S) = \mathrm{rg}(f_S) = \dim(\mathrm{im}(f)) = n \ .$$

$\Leftarrow$ Ist $A \in \mathrm{Mat}_n(K)$ mit $\mathrm{rg}_S(A) = n$, so ist also (vgl.(3.8)) $n = \mathrm{rg}(f_A)$, also f_A surjektiv (vgl.(1.39)). Dann ist $f_A : K^n \to K^n$ nach (2.28 b) ein Isomorphismus. Es gibt also ein lineares $g : K^n \to K^n$ mit:

$$f_A \circ g = \mathrm{id}, \ g \circ f_A = \mathrm{id}$$

(nämlich $g = f_A^{-1}$). Ist $\mathfrak{K}$ die kanonische Basis von K^n, so ist $g = f_B$ mit $B := M(g; \mathfrak{K}, \mathfrak{K})$ (siehe (2.12 b)). Es folgt:

$$f_{AB} = f_A \circ f_B = f_A \circ g = \mathrm{id} = f_E$$

$$f_{BA} = f_B \circ f_A = g \circ f_A = \mathrm{id} = f_E$$

und daher nach (2.21)

$$AB = E, \ BA = E \ ,$$

d.h.: A ist invertierbar.

$\square$

3.25 Beispiel: invertieren

a) Die Elementarmatrizen $F_i(\lambda)$ $(1 \leq i \leq n,\ \lambda \in K^*), G_{ij}(\lambda)$ $(1 \leq i \neq j \leq n,\ \lambda \in K)$ und H_{ij} $(1 \leq i \neq j \leq n)$ sind allesamt invertierbar und es gilt (Übung):

$$F_i(\lambda)^{-1} = F_i\left(\frac{1}{\lambda}\right) ,$$

$$G_{ij}(\lambda)^{-1} = G_{ij}(-\lambda) ,$$

$$H_{ij}^{-1} = H_{ij} .$$

b) Jedes $A \in GL_n(K)$ kann man nach (3.13) auf die Zeilenstufenform E_n bringen (wegen (3.24)): Deshalb ist A ein Produkt von Elementarmatrizen:

$$S_1 \cdot \ldots \cdot S_r A = E_n$$

$$A = S_r^{-1} \ \ldots \ S_1^{-1} E_n = S_r^{-1} \ldots \ S_1^{-1}$$

(für Elementarmatrizen $S_1, \ldots, S_r$).

c) Es folgt dann auch

$$A^{-1} = S_1 \cdots S_r ,$$

was ein praktisches Verfahren zur Bestimmung der **inversen Matrix** liefert. Führe einfach alle elementaren Zeilenumformungen nicht nur an A sondern simultan auch an E_n durch. Hat man nämlich A auf E_n gebracht, so ist E_n nach A^{-1} übergegangen,

$$S_1 \cdots S_r A = E_n$$

$$\Rightarrow S_1 \cdots S_r E_n = S_1 \cdots S_r = A^{-1} .$$

3.26 Beispiel: Matrixinversion

Sei
$$A := \begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & -1 \\ 3 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{Q}) \ .$$

Bestimmung von A^{-1} (wobei man erst während der Umformungen merkt, ob A überhaupt invertierbar, als $\mathrm{rg}_Z A = \mathrm{rg}_S A = 3$, ist).

$$\left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\overset{\substack{G_{12}(-2)\\G_{13}(-3)}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 2 & -1 & -2 & 1 & 0 \\ 0 & 4 & 1 & -3 & 0 & 1 \end{array}\right)$$

$$\overset{F_1(2)}{\longrightarrow} \left(\begin{array}{ccc|ccc} 2 & -2 & 0 & 2 & 0 & 0 \\ 0 & 2 & -1 & -2 & 1 & 0 \\ 0 & 4 & 1 & -3 & 0 & 1 \end{array}\right)$$

$$\overset{\substack{G_{21}(1)\\G_{23}(-2)}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 2 & 0 & -1 & 0 & 1 & 0 \\ 0 & 2 & -1 & -2 & 1 & 0 \\ 0 & 0 & 3 & 1 & -2 & 1 \end{array}\right)$$

$$\overset{\substack{F_1(3)\\F_2(3)}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 6 & 0 & -3 & 0 & 3 & 0 \\ 0 & 6 & -3 & -6 & 3 & 0 \\ 0 & 0 & 3 & 1 & -2 & 1 \end{array}\right)$$

$$\overset{\substack{G_{31}(1)\\G_{32}(1)}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 6 & 0 & 0 & 1 & 1 & 1 \\ 0 & 6 & 0 & -5 & 1 & 1 \\ 0 & 0 & 3 & 1 & -2 & 1 \end{array}\right)$$

$$\overset{F_3(2)}{\longrightarrow} \left(\begin{array}{ccc|ccc} 6 & 0 & 0 & 1 & 1 & 1 \\ 0 & 6 & 0 & -5 & 1 & 1 \\ 0 & 0 & 6 & 2 & -4 & 2 \end{array}\right)$$

$$\Rightarrow A^{-1} = \frac{1}{6} \begin{pmatrix} 1 & 1 & 1 \\ -5 & 1 & 1 \\ 2 & -4 & 2 \end{pmatrix}$$

3.27 Definition: Automorphismus

Sei V ein K-Vektorraum. Ein Endomorphismus $f : V \to V$ heißt ein **(K-)Automorphismus**, wenn f ein Isomorphismus ist.

3.28 Kommentar

a) Bezeichnet

$$\mathrm{Aut}_K(V) := \{f \in \mathrm{End}_K V : f \text{ ist Automorphismus}\}$$

so ist $\mathrm{Aut}(V) \subseteq \mathrm{End}(V)$ gerade die Einheitengruppe im **Endomorphismenring**. Denn f ist genau dann ein Automorphismus, wenn es ein $g \in \mathrm{End}(V)$ gibt mit

$$g \circ f = \mathrm{id}_V, f \circ g = \mathrm{id}_V$$

(siehe(2.15)).

b) Ist V endlich dimensional (der Dimension n), so bildet $\Psi : \mathrm{End}_K(V) \to \mathrm{Mat}_n(K)$ aus (2.21) gerade $\mathrm{Aut}(V)$ (isomorph) nach $GL_n(K)$ ab.

3.29 Bemerkung

Sei V ein endlich dimensionaler Vektorraum. Ein Endomorphismus $f : V \to V$ ist genau dann ein Automorphismus, wenn er eine Basis $\mathfrak{a} = (v_1, \ldots, v_n)$ in eine Basis $(f(v_1), \ldots, f(v_n))$ abbildet.

Beweis:
$\Rightarrow$ Ist $\mathfrak{a} = (v_1, \ldots, v_n)$ eine Basis und $f : V \to V$ ein Automorphismus, so ist $(f(v_1), \ldots, f(v_n))$ erzeugend, weil $(v_1, \ldots, v_n)$ es ist und f surjektiv (siehe (2.16.b)). Ist $g = f^{-1}$, so folgt aus (2.16.a), dass $(f(v_1), \ldots, f(v_n))$ linear unabhängig ist, denn wäre $(f(v_1), \ldots, f(v_n))$ linear abhängig, so wäre auch $(g \circ f(v_1), \ldots, g \circ f(v_n)) = (v_1, \ldots, v_n)$ linear abhängig, was es aber nicht ist.

$\Leftarrow$ Sei $\mathfrak{a} = (v_1, \ldots, v_n)$ und $f(v_j) = \sum_i a_{ij} v_i$ also $A = (a_{ij}) = M(f; \mathfrak{a}, \mathfrak{a})$. Ist $(f(v_1), \ldots, f(v_n))$ eine Basis von V, inbesondere erzeugend und damit f surjektiv, so ist $\mathrm{rg}_S(A) = \mathrm{rg}(f) = n$, also nach (3.24) $A \in GL_n(K)$. Mit (3.26 a) ist dann $f \in \mathrm{Aut}(V)$.

$\square$

3.30 Motivation

Seinen V und W endlich-dimensionale Vektorräume (der Dimensionen n und m) und $f : V \to W$ linear. Wir wollen nun untersuchen, wie sich zwei Matrizen $A, B \in \mathrm{Mat}(m, n, ; K)$ zueinander verhalten, wenn A die Abbildung f bezüglich zweier Basen $\mathfrak{a}$ und $\mathfrak{b}$ sowie B die Beschreibung von f bezüglich zweier Basen $\mathfrak{a}'$ und $\mathfrak{b}'$ ist.

3.31 Lemma

Sei V eine K-Vektorraum und $\mathfrak{a} = (v_1, \ldots, v_n)$ eine Basis. Sei weiter $\mathfrak{b} = (w_1, \ldots, w_n)$ n-Tupel in V mit:

$$w_j = \sum_{i=1}^{n} a_{ij} v_i \ \ (j = 1, \ldots, n)$$

Dann ist $\mathfrak{b}$ genau dann eine Basis, wenn $A \in GL_n(K)$ ist.

Beweis:
Wir betrachten den Endomorphismus $f : V \to V$, der durch $f(v_j) = w_j$ $(j = 1, \ldots, n)$ gegeben ist (vgl.(2.9)). Dann ist nach (3.29) $\mathfrak{b}$ genau dann eine Basis, wenn f ein Automorphismen ist. Und dies ist nach (3.28.b) genau dann der Fall, wenn $A = M(f; \mathfrak{a}, \mathfrak{a}) = \Psi(f)$ in $GL_n(K)$ ist.

$\square$

3.32 Kommentar: Basiswechselmatrix

a) Invertierbare Matrizen sind daher geeignet, den Basiswechsel zwischen zwei Basen $\mathfrak{a}$ und $\mathfrak{a}'$ eines endlich-dimensionalen Vektorraums zu beschreiben. Wir nennen für zwei Basen $\mathfrak{a}$ und $\mathfrak{a}'$ von V

$$S = M(\mathrm{id}; \mathfrak{a}', \mathfrak{a}) \ \in GL_n(K)$$

die Basiswechselmatrix von $\mathfrak{a}'$ auf $\mathfrak{a}$, weil für $S = (a_{ij})$ gilt:

$$v'_j = \mathrm{id}(v'_j) = \sum_{i=1}^{n} a_{ij} v_i \quad (j = 1, \ldots, n)$$

b) Das Inverse $S^{-1} \in GL_n(K)$ beschreibt dann offensichtlich den Wechsel von $\mathfrak{a}$ nach $\mathfrak{a}'$, denn $M(\mathrm{id}; \mathfrak{a}, \mathfrak{a}) = E_n$ für jede Basis $\mathfrak{a}$ und deshalb gilt:

$$M(\mathrm{id}; \mathfrak{a}', \mathfrak{a}) \cdot M(\mathrm{id}; \mathfrak{a}, \mathfrak{a}') = M(\mathrm{id}; \mathfrak{a}, \mathfrak{a}) = E_n$$

und

$$M(\mathrm{id}, \mathfrak{a}, \mathfrak{a}') \cdot M(\mathrm{id}; \mathfrak{a}', \mathfrak{a}) = M(\mathrm{id}; \mathfrak{a}', \mathfrak{a}') = E_n$$

nach (3.5).

3.33 Satz: Transformationsformel

Seien V und W Vektorräume der Dimension n und m und $\mathfrak{a}, \mathfrak{a}'$ Basen von V sowie $\mathfrak{b}, \mathfrak{b}'$ Basen von W. Sei weiter $T \in GL_n(K)$ die Basiswechselmatrix zwischen $\mathfrak{a}'$ und $\mathfrak{a}$, $S \in GL_m(K)$ die Basiswechselmatrix zwischen $\mathfrak{b}$ und $\mathfrak{b}'$. Sei schließlich $f : V \to W$ linear und $A, B \in \mathrm{Mat}(m, n; K)$ die Beschreibungen von f bezüglich $\mathfrak{a}$ und $\mathfrak{b}$ bzw. $\mathfrak{a}'$ und $\mathfrak{b}'$. Dann gilt:

$$B = S \cdot A \cdot T \ .$$

Beweis:
Es ist also

$$T = M(\mathrm{id}; \mathfrak{a}', \mathfrak{a}) \ \in GL_n(K) \ ,$$
$$S = M(\mathrm{id}; \mathfrak{b}, \mathfrak{b}') \ \in GL_m(K) \ ,$$

sowie

$$A = M(f; \mathfrak{a}, \mathfrak{b}) \ ,$$
$$B = M(f; \mathfrak{a}', \mathfrak{b}') \ .$$

Weil nun

$$f = \mathrm{id}_W \circ f \circ \mathrm{id}_V$$

ist, gilt nach (3.5)

$$B = M(f; \mathfrak{a}', \mathfrak{b}') = M(\mathrm{id}_W; \mathfrak{b}, \mathfrak{b}') \cdot M(f; \mathfrak{a}, \mathfrak{b}) \cdot M(\mathrm{id}_V; \mathfrak{a}', \mathfrak{a}) = S \cdot A \cdot T \ .$$

□

3.34 Kommentar: Basiswechselmatrix der Koordinatenisomorphismen

Beachtet man, dass für die Basiswechselmatrix $T = M(\mathrm{id}; \mathfrak{a}', \mathfrak{a})$ gilt, dass $f_T = (T^{\mathfrak{a}'})^{-1} \cdot T^{\mathfrak{a}}$ (Übung) für die Koordinatenisomorphismen $T^{\mathfrak{a}}, T^{\mathfrak{a}'} : K^n \to V$ ist, so drückt die Transmormationsformel gerade die Kommutativität des Basiswechseldiagramms aus.

Abbildung fehlt

3.35 Kommentar: elementare Spaltenumformungen

Genauso wie man elementare Umformungen an Zeilen vornimmt, kann man dies auch an Spalten tun. Aus der Zeilenstufenform kann man dann eine Matrix $A \in \mathrm{Mat}(m,n;K)$ vom Rang r in die Matrix $\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$ überführen. Da elementare Spaltenumformungen durch Multiplikation mit Elementarmatrizen von rechts beschrieben werden, sieht man, dass es zu jedem $A \in \mathrm{Mat}(m,n;K)$ stets Matrizen $S \in Gl_m(K)$ und $T \in Gl_n(K)$ gibt, so dass

$$SAT = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

für $0 \leq r \leq \min(m,n)$ ist. (Dieses Resultat haben wir auch schon in (3.10) zusammen mit (3.33) gesehen.)

3.36 Motivation

Wir wollen nun noch nachtragen, dass der Zeilenrang $\mathrm{rg}_z(A)$ einer Matrix $A \in \mathrm{Mat}(m,n;K)$ stets gleich seinem Spaltenrang ist. Sei dazu:

$$^t = \mathrm{Trans} : \mathrm{Mat}(m,n) \to \mathrm{Mat}(n,m) \ ,$$

$$A \mapsto A^t \ ,$$

wobei die Transponierte von $A = (a_{ij})_{1\leq i\leq m,\ 1\leq j\leq n}$ gegeben ist durch
$A^t = (b_{ji})_{1\leq j\leq m,\ 1\leq i\leq n}$ mit

$$b_{ji} := a_{ij} \ .$$

Es ist dann Trans linear, denn

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

und bijektiv (also ein Isomorphismus), denn Trans $\circ$ Trans = id.

3.37 Bemerkung

Ist $A \in \mathrm{Mat}(m,n)$ und $B \in \mathrm{Mat}(n,r)$, so gilt

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

Beweis:
Setzen wir $C := AB$ und weiter $A^t = (\widetilde{a}_{ji})$, $B^t = (\widetilde{b}_{ji})$, $C^t = (\widetilde{c}_{ki})$ mit $1 \le i \le m, 1 \le j \le n$ und $1 \le k \le r$, so ist

$$\widetilde{c}_{ki} = c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} = \sum_{j=1}^{n} \widetilde{b}_{kj} \widetilde{a}_{ji} \ ,$$

also

$$C^t = B^t \cdot A^t \ .$$

□

3.38 Kommentar

Trans : $\mathrm{Mat}_n(K) \to \mathrm{Mat}_n(K)$ ist deshalb zwar kein Ringhomomorphismus, überführt aber Einheiten in Einheiten, denn ist $S \in Gl_n(K)$ und $T = S^{-1}$, also $ST = E_n = TS$, so ist

$$E_n = E_n^t = T^t \cdot S^t = S^t \cdot T^t$$

also $S^t \in Gl_n(K)$ mit $(S^t)^{-1} = (S^{-1})^t$.

3.39 Satz: Zeilenrang gleich Spaltenrang

Seien $m, n \in \mathbb{N}$ beliebig, K ein beliebiger Körper. Dann gilt für alle $A \in \mathrm{Mat}(m, n; K)$:

$$\mathrm{rg}_z(A) = \mathrm{rg}_s(A) \ .$$

Beweis:
Man beachte zunächst, dass natürlich $\mathrm{rg}_z(A) = \mathrm{rg}_s(A^t)$ ist, weil die Zeilen von A gerade die Spalten A^t sind. Weiter beobachten wir, dass für alle $S \in Gl_m(K)$ und $T \in Gl_n(K)$

$$\mathrm{rg}_s(SAT) = \mathrm{rg}_s(A)$$

ist, weil man SAT als beschreibende Matrix von $f_A : K^n \to K^m$ bezüglich der Basen $\mathfrak{a}$ und $\mathfrak{b}$ auffassen darf, die aus der kanonischen Basis $\mathfrak{K}$ durch die Matrizen S bzw. T hervorgehen, $S = M(\mathrm{id}; \mathfrak{K}, \mathfrak{b}), T = M(\mathrm{id}; \mathfrak{a}, \mathfrak{K})$ (siehe (3.8)).

$$\mathrm{rg}_s(A) = \mathrm{rg}(f_A) \overset{(3.8)}{=} \mathrm{rg}_s(M(f; \mathfrak{a}, \mathfrak{a})) \overset{(3.33)}{=} \mathrm{rg}_s(SAT) \ .$$

Wählen wir aber nun $S \in Gl_m(K)$ und $T \in Gl_n(K)$ so, dass

$$SAT = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

ist und benutzen noch die offensichtliche Tatsache, dass

$$\mathrm{rg}_s \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = r = \mathrm{rg}_z \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

ist, so können wir schließen:

$$\mathrm{rg}_s(A) = \mathrm{rg}_s(SAT) = \mathrm{rg}_s \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = r$$

und

$$\begin{aligned} \mathrm{rg}_z(A) &= \mathrm{rg}_s(A^t) \overset{(*)}{=} \mathrm{rg}_s(T^t A^t S^t) = \mathrm{rg}_s((SAT)^t) \\ &= \mathrm{rg}_z(SAT) = \mathrm{rg}_z \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = r \ , \end{aligned}$$

wobei $(*)$ ausnutzt, dass auch $T^t \in Gl_n(K)$ und $S^t \in Gl_m(K)$ ist. Also ist

$$\mathrm{rg}_s(A) = r = \mathrm{rg}_z(A) \ .$$

□

3.40 Definition: Lineares Gleichungssystem

Seien $m, n \in \mathbb{N}, a_{ij} \in K$ für $1 \leq i \leq m$ und $1 \leq j \leq n$ sowie $b_i \in K$ für $1 \leq i \leq m$. Man nennt das System von m Gleichungen $(*)$

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

ein lineares Gleichungssystem (LGS).

3.41 Kommentar: Lineares Gleichungssystem

a) Gegeben sind dabei also die Zahlen $a_{ij}, b_i \in K$, gesucht sind die Zahlen $x_1, \ldots x_n \in K$, so dass $(*)$ erfüllt ist.

b) Setzt man $A := (a_{ij}) \in \mathrm{Mat}(m,n;K)$ und $b = (b_1,\dots,b_m) \in K^m$, so schreibt sich $(*)$ mit Hilfe der Matrizenmultiplikation als eine Gleichung in K^n (mit $x = (x_1,\dots,x_n) \in K^n$) einfacher als

$$Ax = b\ . \quad (*)$$

c) Für gegebenes $A \in \mathrm{Mat}(m,n;K)$ und $b \in K^m$ nennen wir die Teilmenge

$$L_{A,b} := \{x \in K^n : Ax = b\} \subseteq K^n$$

den Lösungsraum des LGS $(*)$ (mit $L_{A,0} =: L_A$).

d) Ist $b = 0$, so heißt das LGS **homogen**, für $b \neq 0$ heißt es **inhomogen**.

3.42 Satz

Sei $A \in \mathrm{Mat}(m,n;K)$ und $L_A \subseteq K^n$ der Lösungsraum des homogenen Gleichungssystems $Ax = 0$. Dann gilt: L_A ist ein Unterraum von K^n der Dimension

$$\dim(L_A) = n - \mathrm{rg}(A)\ .$$

Beweis:
Betrachte die zu A gehörende lineare Abbildung $f_A : K^n \to K^m$, $f_A(x) = Ax$. Dann ist offenbar $L_A = \ker(f_A)$ und damit nach (2.24) ein Unterraum von K^n der Dimension (nach (2.27))

$$\dim(L_A) = n - \mathrm{rg}(f_A) = n - \mathrm{rg}(A)\ .$$

3.43 Kommentar

Zur Berechnung von L_A beachten wir, dass für $x \in K^n$ und $S \in Gl_m(K)$ gilt:

$$Ax = 0 \Leftrightarrow SAx = 0\ .$$

Deshalb verändern elementare Zeilenumformungen an A den Lösungsraum $L = L_A$ nicht. Nach eventueller Spaltenvertauschung (d.h. der Austausch der Komponenten von $x \in K^n$) können wir A auf folgende Gestalt

$$B = \left(\begin{array}{c|c} E_r & (b_{kl}) \\ \hline 0 & 0 \end{array}\right)$$

mit $r := \mathrm{rg}(A)$ und $b_{kl} \in \mathrm{Mat}(r, n-r)$ (mit $1 \le k \le r,\ r+1 \le l \le n$) bringen.

Dann können wir alle Lösungen $x \in L$ einfach ablesen, denn aus

$$x_k + b_{k,r+1} \cdot x_{r+1} + \cdots + b_{k,n} \cdot x_n = 0$$

für $k = 1, \ldots, r$ sieht man, dass man $x_{r+1}, \ldots, x_n \in K$ frei wählen kann und sich $x_1, \ldots, x_r$, also die abhängigen Variablen, daraus ergeben. Eine Basis von L erhält man daher z.B. so, dass man für jedes $s \in \{r+1, \ldots, n\}$ $x_l := \delta_{ls}$ $(l = r+1, \ldots, n)$ setzt und dann offenbar $x_k = -b_{ks}$ erhält für $k = 1, \ldots, r$. Es ist also $(v_1, \ldots, v_{n-r})$ mit

$$\begin{aligned} v_1 &= (-b_{1,r+1}, \ldots, -b_{r,r+1}, 1, 0, \ldots, 0, 0) \\ &\vdots \\ v_{n-r} &= (-b_{1,n}, \ldots, -b_{r,n}, 0, 0, \ldots, 0, 1) \end{aligned}$$

eine Basis von L_A.

3.44 Beispiel: LGS lösen

Löse das LGS (im $\mathbb{R}^3$)

$$\begin{aligned} 4x_1 + 4x_2 - 5x_3 &= 0 \\ 2x_1 + 0x_2 - x_3 &= 0 \\ -x_1 + 2x_2 - x_3 &= 0 \ . \end{aligned}$$

Also:

$$\begin{pmatrix} 4 & 4 & -5 \\ 2 & 0 & -1 \\ -1 & 2 & -1 \end{pmatrix} \overset{H_{13}}{\longrightarrow} \begin{pmatrix} -1 & 2 & -1 \\ 2 & 0 & -1 \\ 4 & 4 & -5 \end{pmatrix} \overset{\substack{G_{12}(2) \\ G_{13}(4)}}{\longrightarrow} \begin{pmatrix} -1 & 2 & -1 \\ 0 & 4 & -3 \\ 0 & 12 & -9 \end{pmatrix}$$

$$\overset{F_2(2)}{\longrightarrow} \begin{pmatrix} -2 & 4 & -2 \\ 0 & 4 & -3 \\ 0 & 12 & -9 \end{pmatrix} \overset{\substack{G_{21}(-1) \\ G_{23}(-3)}}{\longrightarrow} \begin{pmatrix} -2 & 0 & 1 \\ 0 & 4 & -3 \\ 0 & 0 & 0 \end{pmatrix} \overset{\substack{F_1\left(-\frac{1}{2}\right) \\ F_2\left(\frac{1}{4}\right)}}{\longrightarrow} \begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{3}{4} \\ 0 & 0 & 0 \end{pmatrix}$$

Also ist $v = \left(\frac{1}{2}, \frac{3}{4}, 1\right)$ eine Basis von L_A.

3.45 Satz

Sei $A \in \mathrm{Mat}(m, n; K)$ und $b \in K^m$. Dann ist das LGS $Ax = b$ genau dann lösbar, wenn gilt:

$$\mathrm{rg}(A, b) = \mathrm{rg}(A) \ .$$

3.46 Kommentar

Hierbei ist $(A, b) \in \mathrm{Mat}(m, n+1; K)$ die Matrix, die entsteht, wenn man A um eine Spalte mit dem Eintrag b ergänzt. Es gibt nur folgende zwei Möglichkeiten für $\mathrm{rg}(A, b)$:

i) $\mathrm{rg}(A, b) = \mathrm{rg}(A)$

ii) $\mathrm{rg}(A, b) = \mathrm{rg}(A) + 1$.

Genau im 1. Fall, so sagt Satz (3.45), ist das Sytem $(*)$ lösbar.

Beweis von (3.45):
Man betrachte die zu A gehörende lineare Abbildung $f_A : K^n \to K^m$, $f_A(x) = Ax$. Das Bild von f_A wird von den Spaltenvektoren $a_1, \dots, a_n \in K^m$ von A aufgespannt, denn $f(e_j) = a_j$ für $j = 1, \dots, n$,

$$\mathrm{im}(f_A) = \langle a, \dots, a_n \rangle \subseteq K^m \ .$$

Offenbar ist nun $(*)$ genau dann lösbar, wenn es $x_1, \dots, x_n \in K$ gibt mit

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

also wenn $b \in \langle a_1, \dots, a_n \rangle$. Das ist genau dann der Fall, wenn

$$\langle a_1, \dots, a_n \rangle = \langle a_1, \dots, a_n, b \rangle$$

ist, also

$$\mathrm{rg}(A) = \mathrm{rg}(A, b) \ .$$

$\square$

3.47 Satz: Lösungsraum

Sei $A \in \mathrm{Mat}(m, n; K)$, $b \in K^m$ und y eine (spezielle) Lösung von $Ax = b$. Sei weiter $L_A \subseteq K^n$ der Lösungsraum **des zugehörigen homogenen Systems** $Ax = 0$. Dann gilt für den Lösungsraum $L_{A,b} \subseteq K^n$ von $Ax = b$:

$$L_{A,b} = y + L_A := \{y + x \in K^n : x \in L_A\} \ .$$

3.48 Kommentar: affiner Unterraum

a) Eine Teilmenge $L \subseteq V$ (in einem Vektorraum V) heißt **ein affiner Unterraum** von V wenn es ein $v \in V$ und ein Untervektorraum $U \subseteq V$ gibt, so dass folgendes gilt:

$$L = v + U$$

Man setzt dann:

$$\dim_K L := \dim_K U$$

L ist sozusagen „ein verschobener Unterraum“.

b) Lösungsräume von linearen Gleichungssystemen $Ax = b$ sind also affine Unterräume von K^n der Dimension $n - \mathrm{rg}(A)$.

Beweis von (3.47):
Sei also $y \in K^n$ eine Lösung von $Ax = b$, also: $Ay = b$. Ist $z \in L_{A,b}$ beliebig, also $Az = b$, so gilt:

$$A(z - y) = Az - Ay = b - b = 0$$

Also ist $x := z - y \in L_A$ und damit $z = y + x \in y + L_A$, also $L_{A,b} \subseteq y + L_A$. Ist umgekehrt $x \in L_A$, so ist

$$A(y + x) = Ay + Ax = b + 0 = b\ ,$$

also ist $z := y + x \in L_{A,b}$ und damit $y + L_A \subseteq L_{A,b}$.

□

3.49 Kommentar

a) Zur Bestimmung einer speziellen Lösung nimmt man alle elementaren Zeilenumformungen an A auch an b – also an (A, b) – vor. Es ist nämlich für $S \in Gl_n K$

$$Ax = b \Leftrightarrow SAx = Sb\ .$$

b) Hat man (A, b) durch elementare Zeilenumformungen (und evtl. Spaltenumformungen von A) dann nach $(B, \widetilde{c})$ mit

$$B = \begin{pmatrix} E_r & * \\ 0 & 0 \end{pmatrix} \text{ und } \widetilde{c} = \begin{pmatrix} c \\ 0 \end{pmatrix}$$

mit $c \in K^r$ überführt (falls $\mathrm{rg}(A) = \mathrm{rg}(A, b)$ ist), so ist offenbar $y := (c, 0) \in K^n$ eine spezielle Lösung. Also ist

$$L_{A,b} = y + L_A\ .$$

3.50 Beispiel: LGS lösen

Löse das LGS $Ax = b$ in $\mathbb{R}^3$ mit

$$A = \begin{pmatrix} 4 & 4 & 5 \\ 2 & 0 & -1 \\ -1 & 2 & -1 \end{pmatrix}, \; b = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} .$$

Die Zeilenumformungen vom (3.44) überführen (A,b) nach (Übung)

$$\begin{pmatrix} 1 & 0 & -\frac{1}{2} & | & 0 \\ 0 & 1 & -\frac{3}{4} & | & \frac{1}{2} \\ 0 & 0 & 0 & | & 0 \end{pmatrix} .$$

Nach (3.49) und (3.44) ist deshalb

$$L_{A,b} = \{\underbrace{\left(0, \frac{1}{2}, 0\right) + \lambda \left(\frac{1}{2}, \frac{3}{4}, 1\right)}_{\text{allgemeine Lösung von } (*)} : \lambda \in K\} .$$

4 Determinanten

4.1 Erinnerung: symmetrische Gruppe

Sei $n \in \mathbb{N}_0$ und $[[1, n]] := \{1, 2, \ldots, n\}$ ($[[1, 0]] = \varnothing$), dann hatten wir in (1.5) mit

$$\mathcal{S}_n := \mathrm{Bij}\,([[1, n]]) = \{\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}\text{: bijektiv}\}$$

die **symmetrische Gruppe** mit n Einträgen. Es ist $\mathcal{S}_n$ eine Gruppe mit der Komposition von Abbildungen,

$$\sigma_1\sigma_2 = \sigma_1 \circ \sigma_2 \ .$$

Jedes Element $\sigma \in \mathcal{S}_n$ heißt eine **Permutation**.

4.2 Kommentar: Permutationen

a) Eine Permutation von $\sigma \in \mathcal{S}_n$ wird oft so notiert:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

z.B. ist also $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ die Permutationen, die $1 \to 2, 2 \to 3$ und $3 \to 1$ abbildet.

b) In Teil I hatten wir gesehen, dass $\mathcal{S}_n$ genau $n!$ Elemente hat. (Beachte: $\mathcal{S}_0 = \{id_\varnothing\}$, ist also 1-elementig, $0! = 1$.)

c) Eine Permutation τ heißt **Transposition**, wenn es $1 \leq i < j \leq n$ gibt, so dass gilt:

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

Beachte, dass $\tau^2 = \tau\tau = \mathrm{id}$ ist.

d) Jede Permutation σ ist Produkt von Transpositionen, d.h.: es gibt ein $s \in \mathbb{N}_0$ und Transposition $\tau_1, \ldots, \tau_s$ mit:

$$\sigma = \tau_1 \ldots \tau_s$$

($s = 0$ bedeutet : $\sigma = \mathrm{id}$) (Übung).

4.3 Definition: Signum

Sei $n \in \mathbb{N}_0$ und $\sigma \in \mathcal{S}_n$. Wir definieren das **Signum von** σ durch

$$\operatorname{sgn}(\sigma) := \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} \ .$$

4.4 Kommentar: Signum

a) Die Menge der 2-elementigen Teilmengen von $[[1, \ldots, n]]$ bezeichnen wir mit $\mathfrak{P}_2([[1, n]])$,

$$\mathfrak{P}_2([[1, n]]) := \{\{i, j\} \subseteq \mathfrak{P}([[1, n]]) : 1 \leq i < j \leq n\} \ .$$

($\mathfrak{P}(X)$ sei die Menge aller Teilmengen einer Menge X.) Sie hat $\binom{n}{2} = \frac{1}{2}(n-1)n$ Elemente. Das Produkt in (4.3) hat also $\binom{n}{2}$ Faktoren.

b) Wir werden gleich sehen (siehe (4.5)), dass sgn nur die Werte ± 1 annehmen kann. Ist σ Produkt einer geraden Anzahl von Transpositionen, so wird $\operatorname{sgn}(\sigma) = +1$ sein, im Falle einer ungeraden Anzahl ist $\operatorname{sgn}(\sigma) = -1$ (siehe(4.10)).

c) Offenbar ist $\operatorname{sgn}(\mathrm{id}) = +1$, denn dann ist:

$$\operatorname{sgn}(\mathrm{id}) = 1^{\binom{n}{2}} = 1 \ .$$

d) Sei $\sigma \in \mathcal{S}_n$ und $1 \leq i < j \leq n$ derart, dass $\sigma(i) > \sigma(j)$ ist. Wir nennen dann (i, j) **einen Fehlstand von** σ. Z.B. hat

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \text{ die Fehlstände } (1,3) \text{ und } (2,3) \ .$$

4.5 Bemerkung: Fehlstände

Sei $n \in \mathbb{N}_0$, $\sigma \in \mathcal{S}_n$ und $0 \leq s \leq \binom{n}{2}$, die Anzahl der Fehlstände von σ. Dann gilt:

$$\operatorname{sgn}(\sigma) = (-1)^s$$

Beweis:
Ist $\sigma \in \mathcal{S}_n$, so bildet σ auch die 2-elementigen Teilmenge von $[[1, n]]$ bijektiv auf sich ab,

$$\mathfrak{P}_2([[1, n]]) \to \mathfrak{P}_2([[1, n]]) \ ,$$
$$\{i, j\} \to \{\sigma(i), \sigma(j)\} \ .$$

Es ist deshalb

$$\prod_{1\le i<j\le n}(j-i)=\pm\prod_{1\le i<j\le n}(\sigma(j)-\sigma(i))\ .$$

Auf der rechten Seite sind dabei genau s Faktoren negativ, die anderen $\binom{n}{2}-s$ Faktoren sind positiv (links sind alle positiv). Es folgt:

$$\mathrm{sgn}(\sigma)=\prod_{i<j}\frac{\sigma(j)-\sigma(i)}{j-i}=\frac{\prod_{i<j}(\sigma(j)-\sigma(i))}{\prod_{i<j}(j-i)}=(-1)^s$$

4.6 Korollar: Signum der Transposition

Für eine Transposition $\tau\in\mathcal{S}_n$ gilt:

$$\mathrm{sgn}(\tau)=-1\ .$$

Beweis:
Seien also $1\le i<j\le n$ derart, dass τ gerade i und j vertauscht, die anderen Elemente von $[[1,n]]$ festhält,

$$\tau=\begin{pmatrix}1\dots(i-1)i(i+1)\dots(j-1)j(j+1)\dots n\\ 1\dots(i-1)j(i+1)\dots(j-1)i(j+1)\dots n\end{pmatrix}\ .$$

Es gibt dann die Fehlstände (k,j) für $k=i,\dots,j-1$ (also j-i Stück) und die Fehlstände (i,k) für $(k=i+1,\dots,j)$ (also j-i Stück), wobei nur der Fehlstand (i,j) zweimal gezählt wurde. τ hat also $2(j-i)-1$ Fehlstände und das ist ungerade.

4.7 Vorbreitung: Gruppenhomomorphismus

a) Sind G,G' Gruppen, so heißt eine Abbildung $\phi\ :\ G\ \to\ G'$ ein **Gruppenhomomorphismus**, wenn für alle $g,h\in G$ gilt:

$$\phi(gh)=\phi(g)\phi(h)\ .$$

b) Die Menge $G=\{-1,+1\}\subseteq\mathbb{Z}$ besteht gerade aus den Einheiten des Rings $\mathbb{Z}$ (siehe (3.22 b)). Sie bildet damit eine (abelsche) Gruppe.

4.8 Satz: Signum ist Gruppenhomomorphismus

Sei $n\in\mathbb{N}_0$. Das Signum

$$\mathrm{sgn}:\mathcal{S}_n\to\{\pm1\},\ \ \sigma\mapsto\mathrm{sgn}(\sigma)$$

ist ein Gruppenhomomorphismus, also: Für alle $\sigma_1,\sigma_2\in\mathcal{S}_n$ gilt:

$$\mathrm{sgn}(\sigma_1\sigma_2)=\mathrm{sgn}(\sigma_1)\mathrm{sgn}(\sigma_2)$$

4.9 Lemma

Sei $n \in \mathbb{N}_0$ und $M \subseteq [[1,n]]^2$ derart, dass die Abbildung $M \to \mathfrak{P}_2([[1,n]]),\ (i,j) \to \{i,j\}$ definiert und bijektiv ist. Dann gilt für alle $\sigma \in \mathcal{S}_n$:

$$\operatorname{sgn}(\sigma) = \prod_{(i,j)\in M} \frac{\sigma(j)-\sigma(i)}{j-i} \ .$$

Beweis:
Wegen

$$\frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\sigma(i)-\sigma(j)}{i-j}$$

für alle (i,j) mit $i \neq j$ folgt

$$\prod_{(i,j)\in M} \frac{\sigma(j)-\sigma(i)}{j-i} = \prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i} = \operatorname{sgn}(\sigma)$$

Beweis von (4.6):
Für $\sigma_1, \sigma_2 \in \mathcal{S}_n$ ist nun

$$\operatorname{sgn}(\sigma_1\sigma_2) = \prod_{i<j} \frac{\sigma_1\sigma_2(j)-\sigma_1\sigma_2(i)}{j-i} = \prod_{i<j} \frac{\sigma_1\sigma_2(j)-\sigma_1\sigma_2(i)}{\sigma_2(j)-\sigma_2(i)} \frac{\sigma_2(j)-\sigma_2(i)}{j-i}$$

Setzt man nun:

$$M = \{(\sigma_2(i), \sigma_2(j)) \in [[1,n]]^2 : 1 \leq i < j \leq n\} \ ,$$

so ist $M \to \mathfrak{P}_2([[1,n]]),\ (k,l) \mapsto \{k,l\}$ bijektiv, also ist nach (4.9):

$$\begin{aligned}
\operatorname{sgn}(\sigma_1\sigma_2) &= \prod_{i<j} \frac{\sigma_1(\sigma_2(j))-\sigma_1(\sigma_2(i))}{\sigma_2(j)-\sigma_2(i)} \prod_{i<j} \frac{\sigma_2(j)-\sigma_2(i)}{j-i} \\
&= \prod_{(k,l)\ \in\ M} \frac{\sigma_1(l)-\sigma_1(k)}{l-k} \prod_{i<j} \frac{\sigma_2(j)-\sigma_2(i)}{j-1} \\
&\overset{(4.9)}{=} \operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2)
\end{aligned}$$

4.10 Korollar: Signum einer Permutation

Sei $n \in \mathbb{N}_0$ und $\sigma \in \mathcal{S}_n$ Produkt von s Transpositionen ($s \in \mathbb{N}_0$), $\sigma = \tau_1 \dots \tau_s$. Dann gilt:

$$\operatorname{sgn}(\sigma) = -1^s$$

Beweis:
Ist $\sigma = \tau_1 \dots \tau_s$ mit Transpositionen τ_k (k=1,...,s), so ist wegen (4.6) und (4.8):

$$\operatorname{sgn}(\sigma) = \prod_{k=1}^{s} \underbrace{\operatorname{sgn}(\tau_k)}_{-1} = (-1)^s \; .$$

4.11 Kommentar: alternierende Gruppe

Wegen (4.8) ist

$$\mathfrak{a}_n := \{\sigma \in \mathcal{S}_n : \operatorname{sgn}(\sigma) = +1\}$$

eine Untergruppe von $\mathcal{S}_n$. Sie heißt **die alternierende Grupe** (in n Einträgen). Für $n \geq 2$ hat $\mathfrak{a}_n$ hat genau $\frac{1}{2}n! = 3 \cdot 4 \dots n$ Elemente, denn ist $\tau \in \mathcal{S}_n$ eine Transposition, so ist

$$\mathfrak{a}_n \longrightarrow \mathcal{S}_n \backslash \mathfrak{a}_n, \; \sigma \mapsto \sigma \cdot \tau$$

bijektiv, denn wegen $\tau^2 = \mathrm{id}$ ist $\mathcal{S}_n \backslash \mathfrak{a}_n \longrightarrow \mathfrak{a}, \; \sigma \mapsto \sigma \cdot \tau$ die Umkehrung. Es ist also $\mathcal{S}_n \backslash \mathfrak{a}_n = \mathfrak{a}_n \cdot \tau = \{\sigma\tau \in \mathcal{S}_n : \sigma \in \mathfrak{a}_n\}$. Die Elemente von $\mathfrak{a}_n$ heißen **gerade Permutationen** die von $\mathcal{S}_n \backslash \mathfrak{a}_n$ **ungerade**.

4.12 Definition: multilinear, alternierend

Seien $V_1, \dots, V_r$ und W Vektorräume.

a) Eine Abbildung $s : V_1 \times \dots \times V_r \to W$ heißt **multilinear**, wenn s linear in jedem Argument ist, d.h. für alle $v_1 \in V_1, \dots, v_r \in V_r, \; i = 1, \dots, r, \; v_i' \in V_i$ und $\lambda \in K$ gilt:

$$\begin{aligned} s(v_1, \dots, v_{i-1}, v_i + v_i', v_{i+1}, \dots, v_r) &= s(v_1, \dots, v_i, \dots, v_r) \\ &\quad + s(v_1, \dots, v_i', \dots, v_r) \; , \\ s(v_1, \dots, \lambda v_i, \dots, v_r) &= \lambda s(v_1, \dots v_i, \dots, v_r) \; . \end{aligned}$$

b) Ist $V_1 = \dots = V_r = V$ für einen Vektorraum V, so heißt eine multilineare Abbildung $s : V^r \to W$ **alternierend**, wenn folgendes gilt:
Ist $1 \leq i < j \leq r, \; v_1, \dots, v_r \in V$ und $v_i = v_j$, so ist $s(v_1, \dots, v_r) = 0$.

4.13 Kommentar

a) So wie die Homomorphismen zwischen zwei Vektorräumen V und W (Bez.: $\mathrm{Hom}(V,W)$) selbst einen Vektorraum bilden, so ist dies auch unter punktweiser Addition und skalarer Multiplikation für

$$\mathrm{Mult}(V_1,\ldots,V_r;W) := \{s : V_1 \times \cdots \times V_r \to W : \text{ s ist multilinear}\}$$

der Fall: $\mathrm{Mult}(V_1,\ldots,V_r;W)$ ist selbst ein K-Vektorraum.

b) Im Falle $V_1 = \cdots = V_r = V$ bilden die alternierenden multilinearen Abbildungen offenbar einen Unterraum von $\mathrm{Mult}(V,\ldots,V;W)$, den wir mit

$$\mathrm{Alt}_r(V,W) := \{s \in \mathrm{Mult}_r(V,\ldots,V;W) : \text{ s in alternierend}\}$$

bezeichnen. Im Fall W=K schreiben wir nur noch

$$\mathrm{Alt}_r(V) := \mathrm{Alt}_r(V;K)\ .$$

4.14 Definition: Determinantenform

Sei V ein K-Vektorraum der Dimenstion n. Wir nennen

$$\Delta : V^n \to K$$

eine Determinantenform (kurz: DF) auf V, wenn Δ multilinear und alternierend ist.

4.15 Kommentar: nicht-triviale DF

a) Man beachte, das hier noch keineswegs klar ist, ob es überhaupt nicht triviale DF'en Δ auf V gibt. Natürlich ist $\Delta = 0$ eine DF auf V und $\Delta \in \mathrm{Alt}_n(V)$ heißt **nicht-trivial**, wenn $\Delta \neq 0$ ist.

b) Es wird sich aber bald zeigen (siehe (4.24)), dass, sobald wir eine nicht trivilale DF auf V gefunden haben, sagen wir Δ_1, jede andere DF Δ nur ein skalares Vielfaches von Δ_1 ist, d.h.: Es gibt ein $\lambda \in K$ mit $\Delta = \lambda\Delta_1$ also:

$$\dim(\mathrm{Alt}_n(V)) = 1$$

(unabhängig von n).

4.16 Bemerkung: Regeln für DFen

Sei V ein Vektorraum der Dimension n und $\Delta : V^n \to K$ eine DF. Dann gilt:

a) Ist $(v_1, \ldots, v_n)$ ein linear abhäniges n-Tupel in V, so gilt:

$$\Delta(v_1, \ldots, v_n) = 0 \ ;$$

b) für alle $v_1, \ldots, v_n \in V$ und $1 \leq i < j \leq n$ gilt:

$$\Delta(\ldots, v_j, \ldots, v_i, \ldots) = -\Delta(\ldots, v_i, \ldots, v_j, \ldots) \ ;$$

c) für alle $v_1, \ldots, v_n \in V$ und jedem $\sigma \in \mathcal{S}_n$ gilt:

$$\Delta(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = sgn(\sigma)\Delta(v_1, \ldots, v_n) \ ;$$

d) für alle $v_1, \ldots, v_n \in V,\ 1 \leq i < j \leq n$ und $\lambda \in K$ ist:

$$\Delta(\ldots, v_i, \ldots, v_j + \lambda v_i, \ldots) = \Delta(\ldots, v_i, \ldots, v_j, \ldots) \ .$$

Beweis a) Ist $(v_1, \ldots, v_n)$ l.a., so ist eines der Mitglieder des n-Tupels Linearkombination des anderen (siehe (1.20)), sagen wir v_n.
Es gibt also $\lambda_1, \ldots, \lambda_{n-1} \in K$ mit

$$v_n = \lambda_1 v_1 + \ldots . + \lambda_{n-1} v_{n-1} \ .$$

Wegen der Multilinearität von Δ ist dann:

$$\Delta(v_1, \ldots, v_n) = \Delta(v_1, \ldots, \sum_{i=1}^{n-1})\lambda_i v_i = \sum_{i=1}^{n-1} \lambda_i \Delta(v_1, \ldots, v_{n-1}, v_i) \ ,$$

und jeder dieser $n-1$ Summanden ist Null, weil Δ alternierend ist.

Beweis b) Wir notieren nur die i. und j. Stelle. Da Δ multilinear und alternierend ist, gilt:

$$\begin{aligned} 0 = \ & \Delta \ (\ldots, v_i + v_j, \ldots, v_i + v_j, \ldots) \\ = \ & \Delta \ \underbrace{(\ldots, v_i, \ldots, v_i, \ldots)}_{=0} + \Delta(\ldots, v_i, \ldots, v_j, \ldots) \\ & + \ \Delta(\ldots, v_j, \ldots, v_i, \ldots) + \Delta \underbrace{(\ldots, v_j, \ldots, v_j, \ldots)}_{=0} \ , \end{aligned}$$

also:

$$\Delta(\dots, v_j, \dots, v_i, \dots) = -\Delta(\dots, v_i, \dots, v_j, \dots)\ .$$

Beweis c) Ist $\sigma = \tau$ eine Transposition, so ist

$$\Delta(v_{\tau(1)}, \dots, v_{\tau(n)}) \overset{(b)}{=} -\Delta(v_1, \dots, v_n) \overset{(4.6)}{=} \operatorname{sgn}(\tau)\Delta(v_1, \dots, v_n)\ .$$

Ist $\sigma \in \mathcal{S}_n$ beliebig, so wähle man Transposition $\tau_1, \dots, \tau_s$ (vgl. (4.2 d)) mit

$$\sigma = \tau_1 \dots \tau_s$$

und benutze (4.8):

$$\begin{aligned}
\Delta(v_{\sigma(1)}, \dots, v_{\sigma(n)}) &= \Delta(v_{\tau_1(\tau_2 \dots \tau_s)(1)}, \dots, v_{\tau_1(\tau_2 \dots \tau_s)(n)}) \\
&= \operatorname{sgn}(\tau_1)\Delta(v_{\tau_2 \dots \tau_s(1)}, \dots, v_{\tau_2 \dots \tau_s(n)}) \\
&= \operatorname{sgn}(\tau_1) \dots \operatorname{sgn}(\tau_s)\Delta(v_1, \dots, v_n) \\
&= \operatorname{sgn}(\tau_1 \dots \tau_s)\Delta(v_1, \dots, v_n) \\
&= \operatorname{sgn}(\sigma)\Delta(v_1, \dots, v_n)\ .
\end{aligned}$$

Beweis d) Für $v_1, \dots, v_n \in V,\ \lambda \in K$ und $1 \le i \ne j \le n$ ist.

$$\begin{aligned}
\Delta(v_1, \dots, v_i, \dots, v_j + \lambda v_i, \dots, v_n) = \ & \Delta\ (v_1, \dots, v_i, \dots, v_j, \dots, v_n) \\
& + \lambda\Delta\, \underbrace{(v_1, \dots, v_i, \dots, v_i, \dots, v_n)}_{=0} \\
= \ & \Delta\ (v_1, \dots, v_n)\ .
\end{aligned}$$

4.17 Definition: Determinante

Sei K ein Körper und $n \in \mathbb{N}_0$.

a) Wir definieren die **Determinate** einer $n \times n$-Matrix $A = (a_{ij})_{1 \le i,j \le n}$ durch

$$(*) \qquad \det(A) := \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}$$

b) Für $V = K^n$ definieren wir die **kanonische Determinantenform** $\Delta_1 : (K^n)^n \to K$ durch

$$\Delta(x_1, \dots, x_n) = \det(x_1 \ \dots \ x_n) \ ,$$

wo $A = (x_1, \dots, x_n) \in \mathrm{Mat}_n(K)$ die Matrix bezeichnet, in deren Spalten $x_1, \dots, x_n \in K^n$ stehen.

4.18 Kommentar: Regel von Sarrus

a) Es gibt also in der Leibniz-Formel $(*)$ $n!$ Summenden. Jeder Summand ist (bis auf das Vorzeichen) ein Produkt $c_1 \dots c_n$ aus n Faktoren, wobei aus jeder Spalte (und jeder Zeile) von A genau ein Faktor stammt.

i) $n = 0 : \mathcal{S}_0 = \{\mathrm{id}\} \Rightarrow \det() = 1$

ii) $n = 1 : \mathcal{S}_1 = \{\mathrm{id}\} \Rightarrow \det(a_{11}) = a_{11}$

iii) $n = 2 : \mathcal{S}_2 = \{\mathrm{id}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\} \Rightarrow \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12}$

iv) $n = 3 \Rightarrow \mathcal{S}_3$ hat bereits $3! = 6$ Elemente. Es gilt dann die **Regel von Sarrus** (Übung).

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \ .$$

b) Der Summand, der zum Element $\mathrm{id} \subseteq \mathcal{S}_n$ gehört, ist offenbar das Produkt der Diagonalelemente $a_{11}a_{22} \dots a_{nn}$. Ist $A \in \mathrm{Mat}_n(K)$ eine Diagonalmatrix, d. h.

$$A = \mathrm{diag}(\alpha_1, \dots, \alpha_n) := \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix} \ ,$$

so sind offenbar alle anderen Summanden in $\det(A)$ gleich Null, also:

$$\det(\mathrm{diag}(\alpha_1, \dots, \alpha_n)) = \alpha_1 \cdot \dots \cdot \alpha_n \ .$$

Insbesondere gilt für die Einheitsmatrix $E_n = \mathrm{diag}(1,\dots,1)$ bzw. für die kanonische Basis $\mathfrak{K} = (e_1,\dots,e_n)$ von K^n:

$$\det(E_n) = 1 \ ,$$

$$\Delta_1(e_1,\dots,e_n) = 1 \ .$$

4.19 Satz: kanonische DF

Sei $n \in \mathbb{N}_0$ und $\Delta_1 : (K_n)^n \to K$ die kanonische DF auf K^n. Dann ist Δ_1 tatsächlich multilinear und alternierend.

Beweis:
Jeder der Summanden in $\det(A)$ ist von der Form $\pm c_1 \dots c_n$, wobei c_j ein Eintrag aus der j-ten Spalte von A ist $(j = 1,\dots,n)$. Da $K^n \to K$,
$(c_1,\dots,c_n) \mapsto c_1 \dots c_n$ multilinear ist, ist dies auch $\Delta_1 : (K^n)^n \to K$.
Sei nun $n \geq 2$ und die i. und j. Spalte von A gleich, $i \neq j$, sagen wir $A = (a_1 \dots a_n)$ mit
$a_k \in K^n (k = 1,\dots,n)$ und $a_i = a_j$. D. h.:

$$(*) \qquad\qquad a_{ki} = a_{kj}, \qquad \text{für } k = 1,\dots,n \ .$$

Wir zerlegen nun die symmetrische Gruppe disjunkt in die geraden und ungeraden Permutationen,

$$\mathcal{S}_n = \mathfrak{a}_n \ \dot\cup \ (\mathcal{S}_n \backslash \mathfrak{a}_n) = \mathfrak{a}_n \ \dot\cup \ \mathfrak{a}_n\tau \ ,$$

wobei wir für τ die Transposition wählen, die gerade i und j vertauscht (vgl. (4.11)). Dann gilt:

$$\det(a_1 \dots a_n) = \sum_{\sigma \in \mathfrak{a}_n} a_{\sigma(1)1} \dots a_{\sigma(n)n} - \sum_{\sigma \in \mathfrak{a}_n} a_{\sigma\circ\tau(1),1} \dots a_{\sigma\circ\tau(n),n} \ .$$

Nun ist für $k \neq i, j$ natürlich $a_{\sigma\circ\tau(k),k} = a_{\sigma(k),k}$ und wegen $(*)$ ist

$$a_{\sigma\circ\tau(i),i} = a_{\sigma(j),i} \overset{(*)}{=} a_{\sigma(j),j} \ ,$$

$$a_{\sigma\circ\tau(j),j} = a_{\sigma(i),j} \overset{(*)}{=} a_{\sigma(i),i} \ .$$

Deshalb ist

$$\Delta_1(a_1,\dots,a_n) = \det(a_1 \dots a_n) = 0 \ .$$

Es ist also Δ_1 eine Determinatenform.

4.20 Kommentar: nicht-triviale DF

a) Für $V = K^n$ haben wir also nun eine nicht-triviale DF Δ_1 gefunden, denn Δ_1 verschwindet z.B. auf der kanonischen Basis von K^n nicht,

$$\Delta_1(e_1, \ldots, e_n) = 1 \ .$$

Ist nun V ein beliebiger K-Vektorraum der Dimension n, so hat V zwar i.A. keine kanonische Determinantenform (DF), mit Hilfe einer Basis $\mathfrak{a}$ können wir aber vermöge des Koordinaten-Isomorphismus' $T^{\mathfrak{a}}$: $K^n \to V$ eine nicht-triviale DF so definieren: Wir setzen $\Delta_{\mathfrak{a}} : V^n \to K$,

$$\Delta_{\mathfrak{a}}(w_1, \ldots, w_n) := \Delta_1((T^{\mathfrak{a}})^{-1}(w_1), \ldots, (T^{\mathfrak{a}})^{-1}(w_n)) \ .$$

Multilinearität (weil $(T^{\mathfrak{a}})^{-1}$ linear ist) und Alterniertheit folgen unmittelbar. Ist $\mathfrak{a} = (v_1, \ldots, v_n)$, so gilt $(T^{\mathfrak{a}})^{-1}(v_j) = e_j$ $(j = 1, \ldots, n)$ und daher

$$\Delta_{\mathfrak{a}}(v_1, \ldots, v_n) = 1 \ .$$

Also ist $\Delta_{\mathfrak{a}}$ tatsächlich nicht trivial.

b) Die Determinantenformel von Leibniz fiel in der Definition (4.17) scheinbar vom Himmel. Dass sie sich in natürlicher Weise ergibt, zeigt folgendes Lemma.

4.21 Lemma

Sei V ein K-Vektorraum der Dimension n und $\mathfrak{a} = (v_1, \ldots, v_n)$ eine Basis von V. Sei weiter $\Delta : V^n \to K$ eine beliebige Determinantenform und seien $w_1, \ldots, w_n \in V$ beliebig. Ist dann $w_j = \sum_{i=1}^n a_{ij} v_i$ und $A = (a_{ij})$, so gilt:

$$\Delta(w_1, \ldots, w_n) = \det(A)\Delta(v_1, \ldots, v_n) \ .$$

Beweis:
Direkt aus der Multilinearität von Δ ergibt sich:

$$\begin{aligned}
\Delta(w_1, \ldots, w_n) &= \Delta\left(\sum_{i_1=1}^n a_{i_1 1} v_{i_1}, \ldots, \sum_{i_n=1}^n a_{i_n n} v_n\right) \\
&= \sum_{i_1, \ldots, i_n = 1}^n a_{i_1 1} \ldots a_{i_n n} \Delta(v_1, \ldots, v_n)
\end{aligned}$$

Wegen (4.16 a) sind in dieser Summe (von n^n Summmanden) höchstens die von Null verschieden, wo durch $i_k =: \sigma(k)$ $(k = 1, \ldots, n)$ eine Permutation gegeben ist, und wegen (4.16 c) ist deshalb:

$$\begin{aligned}\Delta(w_1, \ldots, w_n) &= \sum_{\sigma \in \mathcal{S}_n} a_{\sigma(1)1} \ldots a_{\sigma(n)n} \Delta(v_{\sigma(1),1}, \ldots, v_{\sigma(n),n}) \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \ldots a_{\sigma(n)n} \Delta(v_1, \ldots, v_n) \\ &= \det(A) \Delta(v_1, \ldots, v_n)\ .\end{aligned}$$

4.22 Satz:

Sei V ein K-Vektorraum der Dimension n und $\Delta : V^n \to K$ eine nicht triviale DF, dann ist $\mathfrak{a} = (v_1, \ldots, v_n)$ genau dann eine Basis von V, wenn $\Delta(v_1, \ldots, v_n) \neq 0$ ist.

Beweis:
$\Leftarrow$ Ist $\mathfrak{a} = (v_1, \ldots, v_n)$ keine Basis, so muss $(v_1, \ldots, v_n)$ linear abhängig sein (wegen $\dim V = n$) und daher wissen wir nach (4.16 a), dass $\Delta(v_1, \ldots, v_n) = 0$ ist.
$\Rightarrow$ Ist nun $\mathfrak{a}$ eine Basis von V und wäre $\Delta(v_1, \ldots, v_n) = 0$, so zeigt (4.21), dass $\Delta(w_1, \ldots, w_n) = 0$ ist, für alle $w_1, \ldots, w_n \in V$, also $\Delta = 0$. Da Δ nicht trivial ist, folgt $\Delta(v_1, \ldots, v_n) \neq 0$.

4.23 Kommentar: DF erkennen Basen

a) Determinantenformen sind also geeignet um zu erkennen ob ein vorgelegtes n-Tupel $(v_1, \ldots, v_n)$ in einem n-dimensionalen Vektorraum eine Basis ist oder nicht.

b) Lemma (4.21) zeigt auch, dass eine DF komplett festgelegt ist, wenn man sie auf einer einzigen Basis $\mathfrak{a} = (v_1, \ldots, v_n)$ von V kennt. Es gibt also auf V zum Beispiel genau eine DF Δ mit der Eigenschaft:

$$\Delta(v_1, \ldots, v_n) = 1\ ,$$

nämlich $\Delta = \Delta_{\mathfrak{a}}$.

4.24 Korollar

Sei V ein K-Vektorraum der Dimension n und Δ_1 eine nicht-triviale DF. Ist nun Δ eine beliebige DF, so gibt es genau ein $\lambda \in K$ mit $\Delta = \lambda\Delta_1$, also

$$\dim(\mathrm{Alt}_n(V)) = 1 \ .$$

Beweis:
Sei $\mathfrak{a} = (v_1, \ldots, v_n)$ Basis von V. Dann ist $d_1 := \Delta_1(v_1, \ldots, v_n) \neq 0$, sonst wäre nach (4.22) $\Delta_1 = 0$ Ist nun $\Delta(v_1, \ldots, v_n) =: d \in K$, so ist also mit $\lambda := \frac{d}{d_1} \in K$:

$$\Delta(v_1, \ldots, v_n) = d = \frac{d}{d_1} d_1 = \lambda\Delta_1(v_1, \ldots, v_n)$$

und deshalb wegen (4.21) für alle $(w_1, \ldots, w_n) \in V^n$ mit $A = (a_{ij})$, wenn $w_j = \sum a_{ij} v_i$ ist:

$$\begin{aligned} \Delta(w_1, \ldots, w_n) &= \det(A)\Delta(v_1, \ldots, v_n) \\ &= \det(A)\lambda\Delta_1(v_1, \ldots, v_n) \\ &= \lambda\Delta(w_1, \ldots, w_n) \ , \end{aligned}$$

also

$$\Delta = \lambda\Delta_1 \ .$$

4.25 Kommentar: wohldefinierte DFen

a) Sei V ein K-Vektorraum der Dimension n und $f : V \to V$ ein Endomorphismus. Ist $\Delta_1 : V^n \to K$ eine nicht triviale DF auf V, so wird durch $\Delta_f : V^n \to K$ mit:

$$\Delta_f(v_1, \ldots, v_n) := \Delta_1(f(v_1), \ldots, f(v_n))$$

wieder eine DF auf V definiert (evtl. ist Δ_f trivial, vgl. auch (4.20)). Nach (4.24) gibt es deshalb eine eindeutig bestimmte Zahl $\alpha_f \in K$, so dass gilt:

$$\Delta_f = \alpha_f \Delta_1 \ .$$

b) Wegen (4.24) sieht man nun auch, dass die Zuordnung $\mathrm{End}(V) \to K$, $f \mapsto \alpha_f$, nicht von der Wahl der nicht trivialen DF Δ_1 abhängt.
Ist nämlich Δ_1' eine andere, so ist $\Delta_1' = \lambda\Delta_1$ für ein $\lambda \in K^*$ und daher gilt für $\Delta_f' : V^n \to K^n, \Delta_f'(v_1, \ldots, v_n) := \Delta_1'(f(v_1), \ldots, f(v_n))$:

$$\begin{aligned}\Delta'_f(v_1,\dots,v_n) &= \lambda\Delta_1(f(v_1),\dots,f(v_n)) = \lambda\Delta_f(v_1,\dots,v_n)\\ &= \lambda\alpha_f\Delta_1(v_1,\dots,v_n) = \alpha_f\Delta'_1(v_1,\dots,v_n)\end{aligned}$$

für alle $v_1,\dots,v_n \in V$. Es ist also α_f **wohldefiniert** (d.h. unabhängig von der Auswahl von $\Delta_1 \neq 0$).

4.26 Definition: Determinante eines Endomorphismus

Sei V ein K-Vektorraum der Dimension n und Δ_1 eine nicht-triviale DF auf V. Sei weiter $f : V \to V$ ein Endomorphismus. Dann definieren wir die **Determinante von f**, $\det(f) \in K$, durch die Bedingung:

$$\Delta_1(f(v_1),\dots,f(v_n)) = \det(f)\Delta_1(v_1,\dots,v_n)$$

für alle $(v_1,\dots,v_n) \in V^n$.

4.27 Bemerkung

Sei V ein K-Vektorraum und $\mathfrak{a} = (v_1,\dots,v_n)$ eine Basis von V. Dann gilt für alle Endomorphismen $f : V \to V$:

$$\det(f) = \det(M(f;\mathfrak{a},\mathfrak{a}))\ .$$

Beweis:
Sei $\Delta_1 = \Delta_{\mathfrak{a}}$ die DF auf V mit $\Delta_1(v_1,\dots,v_n) = 1$. Ist $f(v_j) = \sum_{i=1}^n a_{ij}v_i$ $(j = 1,\dots,n)$, also $(a_{ij}) = A = M(f;\mathfrak{a},\mathfrak{a})$, so ist nach (4.21)

$$\begin{aligned}\det(f) &= \det(f)\Delta_1(v_1,\dots,v_n) = \Delta_1(f(v_1),\dots,f(v_n))\\ &\overset{4.20}{=} \det(A)\Delta_1(v_1,\dots,v_n) = \det(A)\ .\end{aligned}$$

4.28 Vorbemerkung: Matrizen vs. Endomorphismen

Sei V ein K-Vektorraum der Dimension n und $\mathfrak{a}$ eine Basis von V. Weil

$$\Psi : \mathrm{End}_K(V) \to \mathrm{Mat}_n(K) \ ,$$

$$\Psi(f) = M(f; \mathfrak{a}, \mathfrak{a}) \ ,$$

ein Ringisomorphismus ist (vgl.(3.19)), kann man in folgenden beiden Sätzen (4.29) und (4.29') wahlweise die Endomorphismenversion oder die Matrizenversion benutzen, z.B. gilt für den **Determinanten-Multiplikationssatz** (c'), wenn man (c) bewiesen hat: Für $A, B \in \mathrm{Mat}_n(K)$ seien $f = \Psi^{-1}(A),\ g = \Psi^{-1}(B)$. Dann ist:

$$\begin{aligned}
\det(AB) &= \det(\Psi(f)\Psi(g)) \overset{3.19}{=} \det(\Psi(fg)) \\
&\overset{4.27}{=} \det(fg) \overset{c)}{=} \det(f)det(g) \\
&= \det(\Psi(f))\det(\Psi(g)) = \det(A)\det(B) \ .
\end{aligned}$$

4.29 Satz: Endomorphismenversion

Sei V ein K-Vektorraum der Dimension n. Dann gilt für alle $f, g, \in \mathrm{End}(V)$ und für alle $\lambda \in K$:

a) $\det(\mathrm{id}) = 1$

b) $\det(\lambda f) = \lambda^n \det(f)$

c) $\det(f \circ g) = \det(f)\det(g)$

d) $\det(f) \neq 0,$ genau dann, wenn f ein Automorphismus ist und dann ist

$$\det(f^{-1}) = \frac{1}{\det(f)}$$

4.29' Satz: Matrizenversion

Sei K ein Körper und $n \in \mathbb{N}_0$. Dann gilt für alle $A, B \in \mathrm{Mat}_n(K)$ und $\lambda \in K$:

a') $\det(E_n) = 1$

b') $\det(\lambda A) = \lambda^n \det(A)$

c') $\det(AB) = \det(A)\det(B)$

d') Es ist $\det(A) \neq 0$, genau dann, wenn $A \in GL_n(K)$ ist, und dann gilt:

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

Beweis a') Haben wir schon bewiesen (4.18 b).

Beweis b') Ist $A = (a_1, \ldots, a_n)$ mit $a_j \in K^n$ $(j = 1, \ldots, n)$, so ist wegen der Multilinearität der kanonischen DF Δ_1 auf K^n:

$$\det(\lambda A) = \Delta_1(\lambda a_1, \ldots, \lambda a_n) = \lambda^n \Delta_1(a_1, \ldots, a_n) = \lambda^n \det(A) \ .$$

Beweis c) Sei $\mathfrak{a} = (v_1, \ldots, v_n)$ eine Basis von V und Δ_1 die DF auf V mit $\Delta_1(v_1, \ldots, v_n) = 1$. Dann ist:

$$\begin{aligned}
\det(f \circ g) &= \det(f \circ g)\Delta_1(v_1, \ldots, v_n) = \Delta_1(f \circ g(v_1), \ldots, f \circ g(v_n)) \\
&= \Delta_1(f(g(v_1)), \ldots, f(g(v_n))) = \det(f)\Delta_1(g(v_1), \ldots, g(v_n)) \\
&= \det(f)\det(g)\Delta_1(v_1, \ldots, v_n) = \det(f)\det(g) \ .
\end{aligned}$$

Beweis d) Sei $\mathfrak{a} = (v_1, \ldots, v_n)$ eine Basis und Δ_1 wie unter (c).
Nach (4.22) ist ein n-Tupel $(w_1, \ldots, w_n) \in V^n$ denau dann eine Basis von V, wenn $\Delta_1(w_1, \ldots, w_n) \neq 0$ ist. Nach (3.27) ist $f \in \mathrm{End}(V)$ genau dann Automorphismus, wenn $(f(v_1), \ldots, f(v_n))$ Basis ist. Wegen

$$\Delta_1(f(v_1), \ldots, f(v_n)) = \det(f)$$

folgt der erste Teil der Behauptung. Ist $f \in \mathrm{Aut}(V)$, so ist wegen (a) und (c):

$$1 = \det(\mathrm{id}) = \det(f \circ f^{-1}) \overset{(c)}{=} \det(f)\det(f^{-1}) \ ,$$

also:

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

4.30 Satz: Determinante der transponierten Matrix

Für alle $A \in \mathrm{Mat}_n(K)$ gilt:

$$\det(A^t) = \det(A) \ .$$

Beweis:
Da die Abbildung $\mathcal{S}_n \to \mathcal{S}_n, \sigma \mapsto \sigma^{-1}$, bijektiv ist (ihr Inverses ist sie selbst), kann man in der Leibnizschen Formel statt über $\sigma \in \mathcal{S}_n$ über σ^{-1} summieren (d.h. man nur ändert die Reihenfolge der Summanden):

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma^{-1}) a_{\sigma^{-1}(1),1}, \ldots, a_{\sigma^{-1}(n),n} \ .$$

Nennt man bei festem $\sigma \in \mathcal{S}_n$ noch $k := \sigma(i)$ $(i = 1, \ldots, n)$ so ist $i = \sigma^{-1}(k)$ $(k = 1, \ldots, n)$, also

$$a_{\sigma^{-1}(i),i} = a_{k,\sigma(k)}$$

und damit

$$\prod_{i=1}^{n} a_{\sigma^{-1}(i),i} = \prod_{k=1}^{n} a_{k,\sigma(k)} \ .$$

(Man vertauscht nur die Faktoren.) Schließlich ist wegen

$$1 = \operatorname{sgn}(\mathrm{id}) = \operatorname{sgn}(\sigma\sigma^{-1}) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\sigma^{-1}) \ ,$$

also

$$\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma) \ ,$$

da $\operatorname{sgn}(\sigma) = \pm 1$ ist. Setzen wir $A^t = (\widetilde{a}_{ij})$, also $\widetilde{a}_{ij} = a_{ji}$ so sieht man:

$$\begin{aligned} \det(A) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{k=1}^{n} a_{k,\sigma(k)} \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \widetilde{a}_{\sigma(1),1}, \ldots, \widetilde{a}_{\sigma(n),n} \\ &= \det(A^t) \end{aligned}$$

4.31 Bemerkung

Sei $A \in \mathrm{Mat}_n(K), 1 \leq i \neq j \leq n$ und $\lambda \in K$. Dann gilt:

a) Multipliziert man die i-te Zeile (bzw. Spalte) mit λ, so multipliziert sich $\det(A)$ mit λ.

b) Addiert man zur j-ten Zeile (bzw. Spalte) das λ-fache der i-ten Zeile (bzw. Spalte), so ändert sich die Determinante von A nicht.

c) Vertauscht man i-te mit der j-ten Zeile (bzw. Spalte) so ändert $\det(A)$ sein Vorzeichen.

Beweis:
Wegen $\det(A) = \det(A^t)$ reicht es die Aussagen für elementare Spaltenumforungen zu beweisen. Aber dort folgen sie unmittelbar daraus, dass die kanonische Determinantenform $\Delta_1 : (K^n)^n \to K$ multilinear und alternierend ist (vgl.(4.16)).

4.32 Kommentar: Verfahren zur Determinantenberechnung

Mit Hilfe des Gaußchen Eliminationsverfahrens ergibt das eine effektive Methode Determinanten zu berechnen:

a) Ist z.B. $A \in \mathrm{Mat}_n(K)$ eine obere Dreieckmetrix, d.h. für $A = (a_{ij})$ gilt, dass $a_{ij} = 0$ für $i > j$,

$$A = \begin{pmatrix} a_{11} & \cdots & a_{in} \\ 0 & \ddots & \vdots \\ 0 & 0 & a_{nn} \end{pmatrix} ,$$

so sind im Fall, dass $a_{ii} = 0$ ist, für ein $i \in \{1, \ldots, n\}$, die Spalten 1 bis i linear abhängig und daher $\det(A) = 0$. Ist $a_{ii} \neq 0$, für alle $i = 1, \ldots, n$, so multipliziere man die i-te Zeile mit a_{ii}^{-1} (und verändere damit $\det(A)$ um $a_{11}^{-1} \ldots a_{nn}^{-1}$) und erhalte

$$B = \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} .$$

Dann kann man aber B mit elementaren Zeilenumforungen vom Typ II (ohne Änderung der Determinante) nach E_n überführen. Es ist also:

$$1 = \det(E_n) = \det(B) = a_{11}^{-1} \ldots a_{nn}^{-1} \det(A) ,$$

also:

$$\det(A) = a_{11} \ldots a_{nn} ,$$

und diese Formel beinhaltet offenbar auch den Fall, dass $a_{ii} = 0$ ist, für ein $i \in \{1, \ldots, n\}$.

b) Bei einer beliebigen Matrix $A \in \mathrm{Mat}_n(K)$ reicht es daher, sie mit dem Eliminationsverfahren auf Zeilenstufenform and damit auf obere Dreiecksgestalt zu bringen. Man beachte, dass man die Umformungen vom Typ I und III im Auge behalten muss.

4.33 Beispiel: Berechnung der Determinante

Sei $A \in \mathrm{Mat}_4(\mathbb{Q})$,

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

$$\overset{F_4(\frac{1}{2}),\ H_{14}}{\Longrightarrow} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 2 \\ 2 & 2 & -1 & 1 \\ -3 & 0 & 2 & 1 \end{pmatrix}$$

$$\overset{G_{13}(-2),\ G_{14}(3)}{\Longrightarrow} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 2 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 2 & -2 \end{pmatrix}$$

$$\overset{G_{23}(2),\ F_4(\frac{1}{2})}{\Longrightarrow} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 2 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

$$\overset{G_{34}(-1),\ F_4(\frac{1}{2})}{\Longrightarrow} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 2 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & -8 \end{pmatrix}$$

Es ist deshalb

$$\det(A) = (-1)\left(\frac{1}{2}\right)^{-1}\left(\frac{1}{2}\right)^{-1} 1(-1)1(-8) = -32\ .$$

4.34 Definition: Streichungsmatrix

Sei $A \in \mathrm{Mat}_n(K)$ und $1 \leq i, j \leq n$. Wir definieren die **Streichungsmatrix** $A^{ij} \in \mathrm{Mat}_{n-1}(K)$ durch Streichen der i-ten Zeile und der j-ten Spalte.

$$A^{ij} = \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \\ a_{i1} & & a_{ij} & & a_{in} \\ \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix}$$

4.35 Lemma

Sei $A \in \mathrm{Mat}_n(K)$, $a_1, \ldots, a_n \in K^n$ die Spalten von A und $1 \leq i, j \leq n$. Sei weiter $B^{ij} \in \mathrm{Mat}_n(K)$ die Matrix, die aus A entsteht, wenn man die j-te Spalte durch $e_i := (\delta_{1i}, \ldots, \delta_{ni}) \in K^n$ ersetzt,

$$B^{ij} = (a_1 \ldots . a_{j-1} e_i a_{j+1} \ldots . a_n)$$

$$= \begin{pmatrix} a_{1,1} & \ldots & a_{1,j-1} & 0 & a_{1,j+1} & \ldots & a_{1,n} \\ \vdots & & & \vdots & & & \vdots \\ a_{i-1,1} & \ldots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \ldots & a_{i-1,n} \\ a_{i,1} & \ldots & a_{i,j-1} & 1 & a_{i,j+1} & \ldots & a_{i,n} \\ a_{i+1,1} & \ldots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \ldots & a_{i+1,n} \\ \vdots & & & \vdots & & & \vdots \\ a_{n,1} & \ldots & a_{n,j-1} & 0 & a_{n,j+1} & \ldots & a_{n,n} \end{pmatrix} .$$

Dann gilt:

$$\det(B^{ij}) = (-1)^{i+j} \det(A^{ij}) .$$

Beweis:
Weil wir elementare Spaltenumformungen tom TYP II ohne Veränderung der Determinate von B^{ij} vornehmen können, ist:

$$\det(B^{ij}) = \det(C^{ij})$$

mit

$$C^{ij} = \begin{pmatrix} a_{1,1} & \ldots & a_{1,j-1} & 0 & a_{1,j+1} & \ldots & a_{1,n} \\ \vdots & & & \vdots & & & \vdots \\ a_{i-1,1} & \ldots & & 0 & & \ldots & a_{i-1,n} \\ 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ a_{i+1,1} & \ldots & & 0 & & \ldots & a_{i+1,n} \\ \vdots & & & \vdots & & & \vdots \\ a_{n,1} & \ldots & a_{n,j-1} & 0 & a_{n,j+1} & \ldots & a_{n,n} \end{pmatrix} ,$$

denn Hinzuaddieren des (a_{ik})-fachen der j-ten Spalte zur k-ten Spalte von C^{ij} überführt diese nach B^{ij} $(k \neq j)$. Indem man an C^{ij} die i-te Zeile mit $i-1$ Zeilenvertauschungen in die erste Zeile bringt und ebenso die j-te Spalte mit $j-1$ Vertauschungen in die erste Spalte bringt, erhält man:

$$\det(C^{ij}) = (-1)^{i-1+j-1} \det \left(\begin{array}{c|c} 1 & 0 \quad - \quad 0 \\ \hline 0 & \\ | & A^{ij} \\ 0 & \end{array} \right) .$$

Für jede Matrix $D \in \mathrm{Mat}_{n-1}(K)$ ist

$$\det\left(\begin{array}{c|ccc} 1 & * & \cdots & * \\ \hline 0 & & & \\ | & & D & \\ 0 & & & \end{array}\right) = \det(D)$$

denn die elementaren Zeilenumformungen, die man an D vornimmt, um diese in obere Dreiecksform zu bringen, bringen auch $\left(\begin{array}{c|c} 1 & * \\ \hline 0 & D \end{array}\right)$ in obere Dreieckgestalt. Für

$$D = \begin{pmatrix} d_{22} & & * \\ & \ddots & \\ 0 & & d_{nn} \end{pmatrix}$$

gilt aber

$$\det(D) = d_{22} \cdots d_{nn} = 1 \cdot d_{22} \cdots d_{nn} = \det\begin{pmatrix} 1 & * \\ 0 & D \end{pmatrix} .$$

Insgesamt gilt also:

$$\det(B^{ij}) = \det(C^{ij}) = (-1)^{i+j}\det(A^{ij}) .$$

4.36 Definition: adjungierte Matrix

Sei K ein Körper, $n \in \mathbb{N}$ und $A \in \mathrm{Mat}_n(K)$. Wir setzen für alle $1 \leq i, j \leq n$:

$$\widetilde{a}_{ij} := (-1)^{i+j}\det(A^{ij}) ,$$

$\widetilde{A} = (\widetilde{a}_{ij})_{i \leq i,j \leq n}$, und nennen $A^{ad} := \widetilde{A}^t$ die **zu A adjungierte Matrix**.

4.37 Satz

Sei K ein Körper, $n \in \mathbb{N}$ und $A \in \mathrm{Mat}_n(K)$. Dann gilt:

$$A^{ad} \cdot A = A \cdot A^{ad} = (\det A) \cdot E_n .$$

Beweis:
Sei $1 \leq i, j \leq n$. Es ist dann der (i, j)-Eintrag von $A^{ad}A$ gerade

$$\sum_{k=1}^{n} \widetilde{a}_{ki} a_{kj} = \sum_{k=1}^{n} a_{kj}(-1)^{k+i}\det(A^{ki}) .$$

Sind $a_1, \ldots, a_n \in K^n$ die Spaltenvektoren von A, ist
$B^{ki} = (a_1 \ldots a_{i-1}\, e_k\, a_{i+1} \ldots a_n)$, so ist wegen $a_j = \sum_{k=1}^n a_{kj} e_k$ und (4.35):

$$\begin{aligned}
\sum_{k=1}^n \widetilde{a}_{ki} a_{kj} &= \sum_{k=1}^n a_{kj} \det(B^{ki}) \\
&= \sum_{k=1}^n a_{kj} \det(a_1 \ldots a_{i-1}\, e_k\, a_{i+1} \ldots a_n) \\
&= \det(a_1 \ldots . a_{i-1} a_j a_{i+1} \ldots . a_n) = \delta_{ij} \det(A) \ ,
\end{aligned}$$

und das ist der (i,j)-Eintrag von $\det(A)E_n$. Also ist $A^{ad}A = (\det A)E_n$ und ähnlich sieht man, dass auch $AA^{ad} = (\det A)E_n$ ist.

4.38 Korollar: Laplacescher Entwicklungssatz

Sei $A \in \mathrm{Mat}_n(K)$. Dann gilt:

a) Für alle $j = 1, \ldots, n$ ist

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A^{ij}) \ .$$

b) Für alle $i = 1, \ldots, n$ ist

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A^{ij}) \ .$$

Beweis a) Der Eintrag von $A^{ad}A$ an der Stelle (j,j) ist offenbar gerade

$$\sum_{i=1}^n \widetilde{a}_{ij} a_{ij} = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A^{ij})$$

und das ist nach (4.37) gerade $(\det A)\delta_{jj} = \det A$.

Beweis b) Das ergibt sich ebenso aus $AA^{ad} = (\det A)E_n$.

4.39 Kommentar: Vorzeichenregel mit Hilfe des Schachbretts

a) Die **Vorzeichenregel** im Laplaceschen Entwicklungssatz kann man sich als **Schachbrett** vorstellen, z.B. bei $n = 4$:

$$\begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix} .$$

b) Im Grunde ordnet der Laplacesche Entwicklungsatz die Permutationen in der Leibniz-Formel nur in einer bestimmten Weise an (und klammert dann aus), ist also i.A. zur Berechnung der Determinanten ungeeignet. Sind allerdings in einer Zeile (oder Spalte) besonders viele Nullen enthalten, so ist (4.38) durchaus nützlich. Wie im Beispiel (4.33) mit der Bezeichnung $|A| := \det(A)$:

$$\begin{aligned}
\begin{vmatrix} -3 & 0 & 2 & 1 \\ 0 & -1 & 1 & 2 \\ 2 & 2 & -1 & 1 \\ 2 & 0 & 0 & -2 \end{vmatrix} &= (-2) \cdot \begin{vmatrix} 0 & 2 & 1 \\ -1 & 1 & 2 \\ 2 & -1 & 1 \end{vmatrix} + (-2) \cdot \begin{vmatrix} -3 & 0 & 2 \\ 0 & -1 & 1 \\ 2 & 2 & 1 \end{vmatrix} \\
&= (-2) \cdot \left(-(-1) \cdot \begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} \right) \\
&\quad -2 \cdot \left((-3) \cdot \begin{vmatrix} -1 & 1 \\ 2 & -1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 0 & 2 \\ -1 & 1 \end{vmatrix} \right) \\
&= -2 \cdot (2 + 1 + 2 \cdot (4 - 1)) \\
&\quad -2 \cdot ((-3) \cdot (1 - 2) + 2 \cdot (0 + 2)) \\
&= -32 .
\end{aligned}$$

4.40 Korollar: Berechnung der inversen Matrix

Für alle $A \in GL_n(K)$ gilt:

$$A^{-1} = \frac{1}{\det(A)} \cdot A^{ad}$$

Beweis:
Dies folgt aus $A^{ad} A = \det(A) \cdot E_n$ durch Multiplikation mit A^{-1} von rechts.

4.41 Kommentar

a) Auch diese Formel für das Inverse einer Matrix $A \in GL_n(K)$ ist i.A. ungeeignet für praktische Zwecke. Man müsste dazu nämlich die Determinante einer $(n \times n)$-Matrix und von weiteren $n^2\,[(n-1)\times(n-1)]$-Matrizen berechnen.

b) Für theoretische Zwecke ist (4.40) aber häufig sehr nützlich. Zum Beispiel sieht man an ihr, dass die Abbildung

$$F : GL_n(\mathbb{R}) \to \mathrm{Mat}_n(\mathbb{R}) \cong \mathbb{R}^{n^2}\ ,\ \ F(A) = A^{-1}$$

differenzierbar (sogar rational) ist.

c) Ähnliches gilt für die folgende Regel, die zeigt, dass die (nach (3.42) und (3.44)) eindeutige Lösung eines LGS $Ax = b$ mit $A \in GL_n(K)$ und $b \in K^n$ im Fall von $K = \mathbb{R}$ oder $K = \mathbb{C}$ differenzierbar von A und b abhängig ist:

4.42 Korollar: Cramersche Regel

Sei $A \in GL_n(K)$ mit Spaltenvektoren $a_1, \dots, a_n \in K^n$ und $b \in K^n$. Ist $x = (x_1, \dots, x_n) \in K^n$ die Lösung von $Ax = b$, so gilt für alle $i = 1, \dots, n$:

$$x_i = \frac{1}{\det(A)}\det(a_1 \dots a_{i-1} b a_{i+1} \dots a_n)\ .$$

Beweis:
Die eindeutig bestimmte Lösung von $Ax = b$ für $A \in GL_n(K)$ und $b \in K^n$ ist natürlich

$$x = A^{-1}b\ .$$

Benutzt man nun (4.40) für die inverse Matrix A^{-1} und erneut (4.35), so ergibt sich mit

$$B^{ij} = (a_1, \dots, a_{i-1}, e_i, a_{i+1}, \dots, a_n)$$

für die i-te Komponente von $x = A^{-1}b = \frac{1}{\det(A)} \cdot A^{ad} \cdot b$

$$\begin{aligned}
\det(A)^{-1}\sum_{j=1}^{n}\widetilde{a}_{ji}b_j &= \det(A)^{-1}\sum_{j=1}^{n} b_j \det(B^{ji}) \\
&= \frac{1}{\det(A)}\sum_{j=1}^{n} b_j \det(a_1 \dots a_{i-1}\, e_j\, a_{i+1} \dots a_n) \\
&= \frac{1}{\det(A)}\det(a_1 \dots a_{i-1}\, b\, a_{i+1} \dots a_n)\ .
\end{aligned}$$

5 Eigenwerte

5.1 Motivation

a) Sind V und W Vektorräume der Dimension n bzw. m und ist $f : V \to W$ linear vom Rang r, so hatten wir in (3.10) gesehen, dass es Basen $\mathfrak{a}$ von V und $\mathfrak{b}$ von W gibt, so dass gilt:

$$M(f; \mathfrak{a}, \mathfrak{b}) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} .$$

Ist $V = W$, also $f \in \mathrm{End}(V)$, so ist es (z.B. wegen (3.19)) sinnvoll, nur eine Basis von V zu suchen, so dass $M(f; \mathfrak{a}, \mathfrak{a})$ möglichst einfach wird.

b) Ist $f \in \mathrm{End}(V)$, $\mathfrak{a}$ zunächst beliebig und $A = M(f; \mathfrak{a}, \mathfrak{a})$, so wissen wir aus (3.31), dass für eine andere Basis $\mathfrak{b}$ von V und $B = M(f; \mathfrak{b}, \mathfrak{b})$ gilt:

$$B = SAS^{-1} .$$

Die invertierbare Matrix $S \in Gl_n(K)$ beschreibt dabei gerade den Basiswechsel von $\mathfrak{a}$ nach $\mathfrak{b}$, $S = M(\mathrm{id}; \mathfrak{a}, \mathfrak{b})$. Ist umgekehrt $S \in Gl_n(K)$ beliebig, so setzt man für $\mathfrak{a} = (v_1, \ldots, v_n)$ nun $\mathfrak{b} := (w_1, \ldots, w_n)$ mit $T := S^{-1}, T = (t_{ij})$ und

$$w_j = \sum_{i=1}^{n} t_{ij} v_i \quad (j = 1, \ldots, n) .$$

Dann ist $M(f; \mathfrak{b}, \mathfrak{b})$ gerade SAS^{-1}.

c) Man sagt, dass zwei Matrizen $A, B \in \mathrm{Mat}_n(K)$ **ähnlich** sind, wenn es ein $S \in GL_n(K)$ gibt mit $B = SAS^{-1}$. Ist also für ein $f \in \mathrm{End}(V)$ und einer Basis $\mathfrak{a}$ von V zunächst $A = M(f; \mathfrak{a}, \mathfrak{a})$, so muss man unter allen zu A ähnlichen Matrizen versuchen, eine „beste“ zu finden.

d) Beachte, dass man z.B. für $f \neq \mathrm{id_V}$ nie erreichen kann, dass $M(f; \mathfrak{a}, \mathfrak{a}) = E_n$ ist für eine Basis $\mathfrak{a} = (v_1, \ldots, v_n)$, denn dann wäre

$$f(v_j) = \sum_{i=1}^{n} \delta_{ij} v_i = v_j \quad (j = 1, \ldots n) ,$$

also $f = \mathrm{id}_V$.

e) Schließlich beachte man, dass die Determinante (und die Spur) von ähnlichen Matrizen gleich bleibt (Übung):

$$\det(SAS^{-1}) = \det(A) \ ,$$

$$\mathrm{spur}(SAS^{-1}) = \mathrm{spur}(A) \ .$$

e) Vorerst das Beste, was man evtl. erreichen kann, ist eine Basis $\mathfrak{a} = (v_1, \ldots, v_n)$ von V zu finden, so dass $M(f; \mathfrak{a}, \mathfrak{a})$ eine Diagonalmatrix ist, also

$$M(f; \mathfrak{a}, \mathfrak{a}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \ .$$

Es ist dann offenbar

$$f(v_j) = \lambda_j v_j \quad (j = 1, \ldots, n) \ .$$

Das motiviert:

5.2 Definition: Eigenwert von f

Sei V ein K-Vektorraum und $f : V \to V$ ein Endomorphismus. Ein Skalar $\lambda \in K$ heißt **ein Eigenwert von f** (EW), wenn es ein $v \in V$ mit $v \neq 0$ gibt, so dass folgendes gilt:

$$f(v) = \lambda v \quad (*) \ .$$

5.3 Kommentar: Eigenvektor

a) Ist $\lambda \in K$ ein Eigenwert von f, so heißt jeder Vektor $v \in V$ mit $v \neq 0$, der $(*)$ erfüllt, ein **Eigenvektor** (EV) von f (zum Eigenwert λ).

b) Ist etwa $K = \mathbb{R}$, so wird also ein Eigenvektor $v \in V \backslash \{0\}$ von f unter f gestreckt (für $\lambda > 1$), fixiert (für $\lambda = 1$), gestaucht (für $0 < \lambda < 1$), annuliert (für $\lambda = 0$) oder umgedreht und dann gestreckt, fixiert oder gestaucht (für $\lambda < 0$).

c) Ist $A \in \mathrm{Mat}_n(K)$ eine Matrix, so heißt ein $\lambda \in K$ ein **Eigenwert von A**, wenn λ Eigenwert der zu A gehörenden linearen Abbildung $f_A : K^n \to K^n, x \mapsto Ax$, ist. D.h.: Es gibt ein $x \in K^n$, $x \neq 0$ mit

$$Ax = \lambda x \ .$$

5.4 Bemerkung

Sei V ein K-Vektorraum der Dimension n und $\mathfrak{a}$ eine Basis von V. Dann ist $\lambda \in K$ genau dann ein Eigenwert eines Endomorphismus' $f : V \to V$, wenn λ Eigenwert von $A = M(f; \mathfrak{a}, \mathfrak{a})$ ist.

Beweis:
Sei $T^{\mathfrak{a}} : K^n \to V$ der Koordinatenisomorphismus zu $\mathfrak{a}$, also $T^{\mathfrak{a}}(e_j) = v_j$ $(j = 1, \dots, n)$, wenn $\mathfrak{a} = (v_1, \dots, v_n)$ ist. Dann ist das folgende Diagramm kommutativ (siehe (2.20 d)):
Zeichnung fehlt

mit $f_A(x) = Ax$,
also:

$$f_A = (T^{\mathfrak{a}})^{-1} \circ f \circ T^{\mathfrak{a}} \ .$$

Ist nun λ Eigenwert von f und $v \neq 0$ ein Eigenvektor zum Eigenwert λ, so ist also $f(v) = \lambda v$. Setzt man dann $x := (T^{\mathfrak{a}})^{-1}(v)$, so ist $x \neq 0$ und

$$\begin{aligned} f_A(x) &= (T^{\mathfrak{a}})^{-1} \circ f \circ T^{\mathfrak{a}}(x) = (T^{\mathfrak{a}})^{-1}(f(v)) \\ &= (T^{\mathfrak{a}})^{-1}(\lambda v) = \lambda (T^{\mathfrak{a}})^{-1}(v) = \lambda x \ . \end{aligned}$$

Also ist λ auch Eigenwert von A (und x ist Eigenvektor dazu).
Genauso sieht man: λ EW von $A \Rightarrow \lambda$ EW von f.

□

5.5 Kommentar

a) Das zeigt auch, dass ähnliche Matrizen $A, B \in \mathrm{Mat}_n(K)$ die gleichen Eigenwerte haben, denn A und B können nach (5.1) als beschreibende Matrizen für ein und denselben Endomorphismus $f : K^n \to K^n$ aufgefasst werden (z.B. $f = f_A$).

b) Ein Endomorphismus $f : V \to V$ darf als vollständig verstanden gelten, wenn es eine Basis $\mathfrak{a} = (v_1, \dots, v_n)$ aus Eigenvektoren gibt. Ist v_j dabei Eigenvektor zum Eigenwert $\lambda_j \in K$, so ist also $f(v_j) = \lambda_j v_j$ $(j = 1, \dots, n)$ und die beschreibende Matrix $D = M(f; \mathfrak{a}, \mathfrak{a})$ diagonal,

$$D = M(f; \mathfrak{a}, \mathfrak{a}) = \mathrm{diag}(\lambda_1, \dots, \lambda_n) \ .$$

5.6 Definition: Diagonalisierbarkeit

a) Sei V ein endlich-dimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus. Es heißt f **diagonalisierbar**, wenn es eine Basis $\mathfrak{a}$ von V gibt, sodass $M(f;\mathfrak{a},\mathfrak{a})$ diagonal ist.

b) Eine Matrix $A \in \mathrm{Mat}_n(K)$ heißt **diagonalisierbar**, wenn es eine invertierbare Matrix $S \in GL_n(K)$ gibt, so dass gilt:

$$D = SAS^{-1}$$

ist diagonal, $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$.

5.7 Kommentar: Diagonalisierbarkeit

a) Es ist also f genau dann diagonalisierbar, wenn es eine Basis $\mathfrak{a} = (v_1, \dots, v_n)$ aus Eigenvektoren v_k gibt,

$$f(v_k) = \lambda_k v_k \ \ (k = 1, \dots, n)$$

für gewisse $\lambda_1, \dots, \lambda_n \in K$.

b) Eine Matrix $A \in \mathrm{Mat}_n(K)$ ist also nach (5.1 b) genau dann diagonalisierbar, wenn $f_A : K^n \to K^n,\ x \mapsto Ax$, diagonalisierbar ist.

c) Ein Endomorphismus $f : V \to V$ ist genau dann diagonalisierbar, wenn für eine (und dann für jede) Basis $\mathfrak{a}$ von V die Matrix $M(f;\mathfrak{a},\mathfrak{a})$ diagonalisierbar ist.

5.8 Beispiel

a) Ist $\dim V = 1$, so ist jeder Endomorphismus $f : V \to V$ diagonalisierbar. Ist $v \in V \backslash \{0\}$, also $\mathfrak{a} = (v)$ eine Basis von V, so muss $f(v) = \lambda v$ für ein $\lambda \in K$ sein. Es ist dann also

$$M(f;\mathfrak{a},\mathfrak{a}) = (\lambda)$$

diagonal.

b) Sei $f : \mathbb{R}^2 \to \mathbb{R}^2$ die 90°-Drehung um Null in positiver Richtung, also $f(e_1) = e_2,\ f(e_2) = -e_1$ und damit (für die kanonische Basis $\mathfrak{K} = (e_1, e_2)$ von $\mathbb{R}^2$):

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

Dann hat f offenbar keine Eigenwerte, da kein Vektor $x \neq 0$ auf ein Vielfaches von sich abgebildet wird.

c) Betrachtet man aber A als komplexe 2×2-Matrix, $A \in \mathrm{Mat}_2(\mathbb{C})$, so hat A sehr wohl einen Eigenwert $\lambda \in \mathbb{C}$ (und damit auch $f_A : \mathbb{C}^2 \to \mathbb{C}^2$), z.B. $\lambda = i$, denn $x = (1, -i) \in \mathbb{C}^2 \backslash \{0\}$ ist nun ein Eigenvektor zu λ von $f_A : \mathbb{C}^2 \to \mathbb{C}^2$:

$$Ax = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} i \\ 1 \end{pmatrix} = i \begin{pmatrix} 1 \\ -i \end{pmatrix} = ix \ .$$

Ebenso ist $-i \in \mathbb{C}$ ein Eigenwert und $(1, i)$ ein Eigenvektor dazu. Setzt man

$$T = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \quad \text{und} \quad S := T^{-1} = \frac{1}{2} \cdot \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$$

so ist also

$$SAS^{-1} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = D$$

diagonal.

5.9 Satz: linear unabhängige Eigenvektoren

Sei $f : V \to V$ ein Endomorphismus und $\lambda_1, \ldots, \lambda_r \in K$ paarweise verschiedene Eigenwerte von f. Seien weiter $v_1, \ldots, v_r \in V \backslash \{0\}$ jeweils Eigenvektoren zum Eigenwert $\lambda_1, \ldots, \lambda_r$. Dann gilt $(v_1, \ldots, v_r)$ ist linear unabhängig.

Beweis: Induktion über r:

- $r = 1$: Ist v Eigenvektor zum Eigenwert $\lambda \in K$, so ist insbesondere $v \neq 0$, also (v) linear unabhängig.
- $r \to r + 1$: Gegeben seien $\lambda_1, \ldots, \lambda_{r+1} \in K$ und $v_1, \ldots, v_{r+1} \in V \backslash \{0\}$ mit $f(v_k) = \lambda_k v_k$ mit $(k = 1, \ldots, r + 1)$. Seien $\mu_1, \ldots, \mu_{r+1} \in K$ mit $\mu_1 v_1 + \cdots + \mu_{r+1} v_{r+1} = 0$. Dann gilt einerseits

$$\lambda_{r+1}\mu_1 v_1 + \cdots + \lambda_{r+1}\mu_r v_r + \lambda_{r+1}\mu_{r+1} v_{r+1} = 0 \quad (*)$$

und andererseits:

$$\begin{aligned} 0 &= f(0) = \mu_1 f(v_1) + \cdots + \mu_{r+1} f(v_{r+1}) \\ &= \mu_1 \lambda_1 v_1 + \cdots + \mu_r \lambda_r v_r + \mu_{r+1} \lambda_{r+1} v_{r+1} \ . \end{aligned}$$

Aus diesen beiden Gleichungen ergibt sich dann:

$$0 = (\lambda_{r+1} - \lambda_1)\mu_1 v_1 + \cdots + (\lambda_{r+1} - \lambda_r)\mu_r v_r \ .$$

Nach Induktionsvoraussetzung gilt dann:

$$(\lambda_{r+1} - \lambda_k)\mu_k = 0 \ \ (k = 1, \ldots, r) \ ,$$

und weil nun $\lambda_k \neq \lambda_{r+1}$ ist, für $k = 1, \ldots, r$, folgt daraus

$$\mu_1 = \cdots = \mu_r = 0 \ .$$

Setzt man dies in $(*)$ ein, so muss auch $\mu_{r+1} = 0$ sein (weil $v_{r+1} \neq 0$ ist) und damit ist $(v_1, \ldots, v_{r+1})$ linear unabhängig.

□

5.10 Korollar

Ist V ein K-Vektorraum der Dimension $n \in \mathbb{N}$, so gibt es höchstens n paarweise verschiedene Eigenwerte von f.

Beweis:
Klar, weil V höchstens n Vektoren $v_1, \ldots, v_n$ hat, so dass $(v_1, \ldots, v_n)$ linear unabhängig ist.

□

5.11 Korollar

Ist V ein K-Vektorraum der Dimension $n \in \mathbb{N}$ und hat $f \in \mathrm{End}(V)$ n paarweise verschiedene Eigenwerte, so ist f diagonalisierbar.

Beweis:
Sind $\lambda_1, \ldots, \lambda_n \in K$ die paarweise verschiedenen Eigenwerte von f, so wähle für jedes $k \in \{1, \ldots, n\}$ einen Eigenvektor $v_k \in V \backslash \{0\}$ zum Eigenwert λ_k. Nach (5.9) ist $\mathfrak{a} = (v_1, \ldots, v_n)$ dann linear unabhängig, also wegen $\dim(V) = n$ eine Basis (aus Eigenvektoren) von V, also ist f diagonalisierbar (vgl (5.7 a)).

□

5.12 Kommentar

Es muss $f : V \to V$ aber keineswegs $n = \dim(V)$ paarweise verschiedene Eigenwerte haben, damit f diagonalisierbar ist, z.B. hat die Nullabbildung $0 : V \to V$ offenbar nur einen Eigenwert $\lambda \in K$, nämlich $\lambda = 0$, ist aber offensichtlich diagonalisierbar, da $M(f; \mathfrak{a}, \mathfrak{a}) = 0$ ist (für alle Basen $\mathfrak{a}$ von V). Ebenso hat die Identität $f = \mathrm{id}_V : V \to V$ nur einen Eigenwert, nämlich $\lambda = 1$, ist aber offensichtlich ebenso diagonalisierbar, weil $M(\mathrm{id}; \mathfrak{a}, \mathfrak{a}) = E_n$ ist (wieder sogar für alle Basen $\mathfrak{a}$ von V).

5.13 Definition: Eigenraum

Sei f ein Endomorphismus eines K-Vektorraums V und $\lambda \in K$ ein Eigenwert von f. Es heißt dann

$$\mathrm{Eig}(f; \lambda) := \{v \in V : f(v) = \lambda v\}$$

der **Eigenraum** von f zum Eigenwert λ.

5.14 Kommentar

a) Es besteht also $\mathrm{Eig}(f; \lambda)$ genauso aus den Eigenvektoren zum Eigenwert λ von f und zusätzlich dem Nullvektor. Wegen

$$f(v) = \lambda v \Leftrightarrow f(v) = (\lambda \mathrm{id})(v) \Leftrightarrow 0 = (\lambda \mathrm{id} - f)(v)$$

ist $\mathrm{Eig}(f; \lambda) = \ker(\lambda \mathrm{id} - f)$ und damit ist klar, dass $\mathrm{Eig}(f; \lambda) \subseteq V$ ein Unterraum von V ist.

b) Man könnte natürlich für jedes $\lambda \in K$ den Unterraum $V_\lambda := \ker(\lambda \mathrm{id}_V - f)$ betrachten. Offenbar ist dann $V_\lambda \neq 0$ genau dann, wenn λ ein Eigenwert von f ist.

c) Satz (5.9) zeigt, dass der Durchschnitt eines Eigenraums $\mathrm{Eig}(f; \lambda_1)$ mit der Summe der Eigenräume zu anderen Eigenwerten $\lambda_2, \ldots, \lambda_r \in K$ von f nur aus dem Nullvektor besteht:

$$\mathrm{Eig}(f; \lambda_1) \cap (\mathrm{Eig}(f; \lambda_2) + \cdots + \mathrm{Eig}(f; \lambda_r)) = (0) \ .$$

Die Summe der verschiedenen Eigenräume von f ist also direkt,

$$\mathrm{Eig}(f; \lambda_1) + \cdots + \mathrm{Eig}(f; \lambda_r) = \mathrm{Eig}(f; \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(f; \lambda_r) \ .$$

5.15 Lemma

Sei $f : V \to V$ linear und $\mathfrak{a}$ eine Basis aus Eigenvektoren. Sei $D = M(f; \mathfrak{a}, \mathfrak{a}) = \operatorname{diag}(\lambda_1, \ldots, \lambda_1, \ldots, \lambda_r, \ldots, \lambda_r)$, $\lambda_1, \ldots, \lambda_r \in K$ seien dabei die verschiedenen Eigenwerte und λ_k tauche dabei ν_k-mal auf ($\nu_k \in \mathbb{N}$). Ist dann $\mathfrak{a} = (v_1^{(1)}, \ldots, v_{\nu_1}^{(1)}, \ldots, v_1^{(r)}, \ldots, v_{\nu_r}^{(r)})$, wobei $v_i^{(k)}$ Eigenvektor zu λ_k ist, $(i = 1, \ldots, \nu_k;\ k = 1, \ldots, r)$, so gilt:

$$\operatorname{Eig}(f; \lambda_k) = \langle v_1^{(k)}, \ldots, v_{\nu_k}^{(k)} \rangle \ .$$

Beweis:
Sei o.E. $k = 1$. Da $v_i^{(1)}$ Eigenvektor zum Eigenwert λ_1 ist, ist die Inklusion „ $\supseteq$ " klar. Sei andererseits $v \in \operatorname{Eig}(f; \lambda_1)$. Da $\mathfrak{a}$ eine Basis von V ist, gibt es $\mu_1^{(1)}, \ldots, \mu_{\nu_r}^{(r)} \in K$ mit:

$$v = \sum_{l=1}^{r} \sum_{i=1}^{\nu_l} \mu_i^{(l)} \nu_i^{(l)} = \sum_{l=1}^{r} w^{(l)} \ ,$$

wenn wir $w^{(l)} := \sum_{i=1}^{\nu_l} \mu_i^{(l)} \nu_i^{(l)}$ setzen. Weil $w^{(l)} \in \operatorname{Eig}(f; \lambda_l)$ und $\lambda_1, \ldots, \lambda_r$ paarweise verschieden sind, folgt nun aus (5.9) und $v - w^{(1)} = \sum_{l=2}^{r} w^{(l)}$, dass $v - w^{(1)} = 0$ ist, also

$$v = w^{(1)} \in \langle v_1^{(1)}, \ldots, v_{\nu_k}^{(1)} \rangle$$

sein muss. Es ist also auch

$$\operatorname{Eig}(f; \lambda_1) \subseteq \langle v_1^{(1)}, \ldots, v_{\nu_k}^{(1)} \rangle \ .$$

□

5.16 Satz

Sei $f : V \to V$ ein Endomorphismus eines endlich-dimensionalen Vektorraumes V und seien $\lambda_1, \ldots, \lambda_r \in K$ die paarweise verschiedenen Eigenwerte von f. Dann gilt: Es ist f genau dann diagonalisierbar, wenn V in die direkte Summe der Eigenräume von f zerfällt, d.h.:

$$V = \operatorname{Eig}(f; \lambda_1) \oplus \cdots \oplus \operatorname{Eig}(f; \lambda_r) \ .$$

Beweis:
$\Leftarrow$: Seien $(v_1^{(k)}, \dots, v_{\nu_k}^{(k)})$ Basen von $\mathrm{Eig}(f;\lambda_k)$ (also $\nu_k = \dim(\mathrm{Eig}(f;\lambda_k)))$ für $k = 1, \dots, r$. Wegen $V = \bigoplus_{k=1}^r \mathrm{Eig}(f;\lambda_k)$ und (1.43) ist dann $n = \nu_1 + \dots + \nu_r$ und damit $(v_1^{(1)}, \dots, v_{\nu_r}^{(r)}) = \mathfrak{a}$ eine (**Eigen-**)**Basis** von V aus Eigenvektoren zu f. Es folgt:

$$M(f;\mathfrak{a},\mathfrak{a}) = \left(\begin{array}{ccc|c|ccc}
\lambda_1 & & & & & & \\
& \ddots & & & & & \\
& & \lambda_1 & & & & \\
\hline
& & & \ddots & & & \\
\hline
& & & & \lambda_r & & \\
& & & & & \ddots & \\
& & & & & & \lambda_r
\end{array}\right) .$$

$$\nu_1 - mal \qquad \cdots \qquad \nu_r - mal$$

Also ist f diagonalisierbar.
$\Rightarrow$: Sei f diagonalisierbar. Es gibt dann also eine Basis $\mathfrak{a} = (v_1^{(1)}, \dots, v_{\nu_r}^{(r)})$ von V, so dass $D = M(f;\mathfrak{a},\mathfrak{a})$ diagonal ist mit

$$D = \mathrm{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{\nu_1 - mal}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{\nu_r - mal})$$

mit $\lambda_1, \dots, \lambda_r$ paarweise verschieden. Nach (5.15) ist dann

$$V = \bigoplus_{k=1}^r \langle v_1^{(k)}, \dots, v_{\nu_k}^{(k)} \rangle = \bigoplus_{k=1}^r \mathrm{Eig}(f;\lambda_k) \ .$$

$\lambda_1, \dots, \lambda_r$ sind Eigenwerte von f und es gibt auch keine weiteren Eigenwerte, denn: Ist $\lambda \in K$ und $v \in V \backslash \{0\}$ mit $f(v) = \lambda v$ sowie $\lambda \neq \lambda_k$ für
$k = 1, \dots, r$, so ist v nach (5.9) keine Linearkombination von $\mathfrak{a} = (v_1^{(1)}, \dots, v_{\nu_r}^{(r)})$. Das wäre ein Widerspruch, denn $\mathfrak{a}$ ist ein Erzeugendensystem von V. Also ist tatsächlich $V = \bigoplus_{k=1}^r \mathrm{Eig}(f;\lambda_k)$, wo $\lambda_k,\ k = 1, \dots, r$, alle Eigenwerte von f durchläuft.

$\square$

5.17 Definition: geometrische Vielfachheit

Ist $f \in \operatorname{End}(V)$ und $\lambda \in K$ ein Eigenwert von f, so nennen wir

$$\nu_\lambda := \dim(\operatorname{Eig}(f;\lambda))$$

die **geometrische Vielfachheit** des Eigenwertes λ.

5.18 Kommentar: geometrische Vielfachheiten

a) Sind $\lambda_1, \dots, \lambda_r \in K$ die verschiedenen Eigenwerte eines Endomorphismus $f : V \to V$, $n = \dim(V)$ und $\nu_1, \dots, \nu_r \in \mathbb{N}$ die geometrischen Vielfachheiten von $\lambda_1, \dots, \lambda_r$, so ist also f genau dann diagonalisierbar, wenn gilt:

$$n = \nu_1 + \cdots + \nu_r \ .$$

b) Wenn $f \in \operatorname{End}(V)$ diagonalisierbar ist, so zeigt (5.16) auch, dass die Diagonalmatrix $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \in \operatorname{Mat}_n(K)$, die f bezüglich einer Eigenbasis darstellt, bis auf die Reihenfolge der Diagonalelemente eindeutig bestimmt ist. Es stehen dort nämlich genau die Eigenwerte $\lambda_1, \dots, \lambda_r \in K$ von f und jedes λ_k genau ν_k-mal.

c) Für eine Matrix $A \in \operatorname{Mat}_n(K)$ und einen Eigenwert $\lambda \in K$ definiert man dessen geometrische Vielfachheit $\nu_k \in \mathbb{N}$ als die Dimension des Eigenraums $\operatorname{Eig}(f_A;\lambda)$. So wie die Eigenwerte $\lambda_1, \dots, \lambda_r \in K$ von A sich bei Konjugation mit einem Element $S \in GL_n(K)$ nicht ändern (siehe (5.5 a)), ändern sich auch die geometrischen Vielfachheiten $\nu_1, \dots, \nu_r \in \mathbb{N}$ nicht, denn ist $B = SAS^{-1}$ so gilt (ähnlich wie in (5.4)):

$$\operatorname{Eig}(f_B;\lambda) = f_S(\operatorname{Eig}(f_A;\lambda)) \ .$$

Insbesondere ist also $\nu(f_B;\lambda) = \nu(f_A;\lambda)$. Man sagt: Die Eigenwerte und ihre geometrischen Vielfachheiten sind **Invarianten von A** (unter Konjugation).

d) Ist $f \in \operatorname{End}(V)$ mit Eigenwerten $\lambda_1, \dots, \lambda_r \in K$ und geometrischen Vielfachheiten $\nu_1, \dots, \nu_r \in \mathbb{N}$, so ist, weil die Summe der Eigenräume direkt ist (vgl. (5.14 c)), stets:

$$\nu_1 + \cdots + \nu_r \leq n \ .$$

Um nun die Frage zu klären, ob ein gegebenes $f \in \mathrm{End}(V)$ (oder ein $A \in \mathrm{Mat}_n(K)$) diagonalisierbar ist, muss man folgende Aufgaben lösen:

i) Bestimme alle Eigenwerte $\lambda \in K$ von f.

ii) Bestimme die geometrischen Vielfachheiten $\nu_\lambda \in \mathbb{N}$ von λ.

e) Aufgabe (*ii*) in (d) ist einfach. Ist $A = M(f; \mathfrak{a}, \mathfrak{a})$ die Matrix von f bezüglich irgendeiner Basis von V, so ist

$$\begin{aligned} \nu_\lambda &= \dim(\mathrm{Eig}(f; \lambda)) = \dim(\ker(\lambda \mathrm{id} - f)) \\ &= \dim(\ker(\lambda E_n - A)) = n - \mathrm{rg}(\lambda E_n - A) \; . \end{aligned}$$

Will man zudem eine Basis aus Eigenvektoren und damit eine Matrix $S \in GL_n(K)$ haben, so dass $B = SAS^{-1}$ diagonal ist, muss man zusätzlich das LGS

$$(\lambda E_n - A)x = 0$$

lösen.

f) Aufgabe (*i*) ist entschieden schwerer. Sie wird uns zum ersten Mal aus der linearen Algebra heraus in die **Algebra** führen. Beachte zunächst:

5.19 Satz

Sei V ein K-Vektorraum endlicher Dimension, $\mathfrak{a}$ eine Basis von V, $f : V \to V$ ein Endomorphismus und $A = M(f; \mathfrak{a}, \mathfrak{a})$. $\lambda \in K$ ist genau dann ein Eigenwert von f, wenn gilt:

$$\det(\lambda E_n - A) = 0 \; .$$

Beweis:
Nach Definition ist $\lambda \in K$ Eigenwert von f, wenn $\lambda \mathrm{id}_V - f$ nicht injektiv ist (vgl. (2.25 a)). Nach (2.28) ist dies genau dann der Fall, wenn $\lambda \mathrm{id}_V - f$ kein Automorphismus ist. Da $\lambda E_n - A = M(\lambda \mathrm{id}_V - f; \mathfrak{a}, \mathfrak{a})$, ist dies genau dann so, wenn $\lambda E_n - A$ nicht invertierbar ist (3.28 b). Nach (4.29 d') schließlich merkt dies genau die Determinante:

$$\lambda \text{ Eigenwert von } f \;\Leftrightarrow\; \det(\lambda E_n - A) = 0 \; .$$

5.20 Kommentar

Man ersetzt deshalb nun $\lambda \in K$ in $\det(\lambda E_n - A)$ durch eine Unbestimmte T im Polynomring $K[T]$ (vgl. (1.15 e) und (3.17 c)) und setzt:

$$P_A(T) = \det(TE_n - A) \ .$$

Dabei muss man ein wenig aufpassen, weil nun die Matrix $TE_n - A$ ihre Einträge in $R := K[T]$ hat, was kein Körper mehr ist, sondern nur noch ein kommutativer Ring. Da aber in der Definition der Determinanten (siehe (4.17)) nur multipliziert und addiert (bzw. subtrahiert), nicht aber dividiert wird, macht folgende Definition tatsächlich Sinn:

5.21 Definition: charakteristisches Polynom von A

Sei K ein Körper, $n \in \mathbb{N}_0$ und $A = (a_{ij}) \in \mathrm{Mat}_n(K)$. Wir nennen dann $P_A \in K[T]$,

$$P_A(T) := \det(TE_n - A) := \sum_{\sigma \in \mathcal{S}_n} \mathrm{sgn}(\sigma) \prod_{k=1}^{n} (T \cdot \delta_{\sigma(k),k} - a_{\sigma(k),k})$$

das **charakteristische Polynom von A**.

5.22 Kommentar

a) In jedem Summanden von diesem Polynom P_A steht ein Produkt aus n Faktoren aus $K[T]$ vom Grad höchstens 1. So ist also P_A höchstens vom Grad n.

b) Es gibt nur einen Summanden aus P_A, wo jeder Faktor die Unbestimmte T enthält, nämlich der zu $\sigma = \mathrm{id}$, und dieser ist

$$\prod_{k=1}^{n} (T - a_{kk}) = T^n - (a_{11} + \cdots + a_{nn})T^{n-1} + \widetilde{Q}(T)$$

wo $\deg(\widetilde{Q}) \leq n - 2$ ist. Da jeder andere Summand von P_A höchstens $n - 2$ Faktoren hat, in denen T vorkommt, gilt:

$$P_A(T) = T^n - \mathrm{spur}(A)T^{n-1} + Q(T), \quad \text{mit } \deg(Q) \leq n - 2 \ .$$

Hierbei ist $\mathrm{spur}(A) := a_{11} + \ldots . a_{nn}$ die **Spur von** A. Es ist also P_A ein normiertes Polynom vom Grad n, d.h. $\deg P_A = n$ und der Koeffizient vor T^n ist 1.

c) Setzt man in P_A für T nun $\lambda = 0$ ein, so bekommt man gerade den Koeffizienten vor $T^0 = 1$. Wegen $P_A(0) = \det(-A) = (-1)^n \det(A)$ ist also:

$$P_A(T) = T^n - \operatorname{spur}(A)T^{n-1} + \cdots + (-1)^n \det(A) \ .$$

d) Ist $A \in \operatorname{Mat}_n(K)$ und $B \in \operatorname{Mat}_n(K)$ ähnlich zu A, also $B = SAS^{-1}$ für ein $S \in GL_n(K)$, so ist zunächst einmal $S(TE_n) = (TE_n)S$ in $\operatorname{Mat}_n(K[T])$ (rechne wie in $\operatorname{Mat}_n(K)$) und deshalb

$$S(TE_n)S^{-1} = (TE_n)SS^{-1} = TE_n \ .$$

Daraus sieht man nun:

$$\begin{aligned} P_B(T) &= \det(TE_n - B) = \det(STE_nS^{-1} - SAS^{-1}) \\ &= \det(S(TE_n - A)S^{-1}) = \det(S) \cdot \det(TE_n - A) \cdot \det(S^{-1}) \\ &= \det(TE_n - A) = P_A(T) \end{aligned}$$

(weil der Determinanten-Multiplikationssatz auch in $\operatorname{Mat}_n(R)$ für kommutative Ringe R gilt). Deshalb hängt folgende Definition nicht von der Basiswahl ab.

5.23 Definition: charakteristisches Polynom von f

Sei V ein K-Vektorraum der Dimension n, $\mathfrak{a}$ eine Basis von V und $f : V \to V$ ein Endomorphismus. Ist $A = M(f; \mathfrak{a}, \mathfrak{a}) \in \operatorname{Mat}_n(K)$, so setzt man

$$P_f := P_A \in K[T]$$

und nennt dies das **charakteristische Polynom von f**.

5.24 Satz: Nullstellen des charakteristischen Polynoms

Sei V ein endlich-dimensionaler K-Vektorraum, $f : V \to V$ ein Endomorphismus und $P_f \in K[T]$ sein charakteristisches Polynon, dann gilt: Die Eigenwerte von f sind genau die Nullstellen von P_f (in K).

Beweis:
Klar, weil nach (5.19) für $A = M(f; \mathfrak{a}, \mathfrak{a})$ und einer Basis $\mathfrak{a}$ von V gilt:

$$\begin{aligned} \lambda \text{ ist EW von } f &\Leftrightarrow \det(\lambda E_n - A) = 0 \Leftrightarrow P_A(\lambda) = 0 \\ &\Leftrightarrow P_f(\lambda) = 0 \ . \end{aligned}$$

□

5.25 Kommentar: Division mit Rest, Zerfall in Linearfaktoren

a) Nullstellen von Polynomen $p \in K[T]$ höheren Grades zu bestimmen ist schwer und ein zentrales Problem der Algebra. (Nicht immer wird p überhaupt Nullstellen haben, z.B. $p(T) = T^2 + 1 \in \mathbb{R}[T]$.) Wir wollen hier benutzen, dass man im Polynomring $K[T]$ (ähnlich wie in $K = \mathbb{Z}$) **Division mit Rest** durchführen kann. Insbesondere gilt folgendes: Ist $p \in K[T]$ vom Grad n und $\lambda \in K$ eine Nullstelle von p, d.h. $p(\lambda) = 0$, so existiert genau ein Polynom q vom Grad $n-1$, so dass gilt:

$$p(T) = (T - \lambda)q(T) \ .$$

b) Weil ein Polynom (verschieden von Null) vom Grad 0 gar keine Nullstellen hat, folgt durch iterierte Division mit Rest (in diesem Fall ohne Rest), dass ein Polynom p vom Grad n, $p \neq 0$, höchstens n Nullstellen hat. Hat es n Nullstellen $\lambda_1, \ldots . \lambda_n \in K$, so gilt also

$$p(T) = c(T - \lambda_1) \ldots (T - \lambda_n)$$

mit einem $c \in K^*$ und wir sagen, dass p in **Linearfaktoren** zerfällt.

c) Aus der Division mit Rest folgt auch, dass es für jede Nullstelle $\lambda \in K$ eines Polynoms $p \neq 0$ ein maximales $\mu \in \mathbb{N}$ gibt, so dass gilt:

$$p(T) = (T - \lambda)^{\mu} q(T)$$

mit einem $q(T) \in K[T]$. Es ist dann $q(\lambda) \neq 0$ und man nennt $\mu = \mu_\lambda$ die **Vielfachheit der Nullstelle λ in p**.

d) Ist daher $p \in K[T] \backslash \{0\}$ und sind $\lambda_1, \ldots, \lambda_r \in K$ die verschiedenen Nullstellen von p mit Vielfachheiten $\mu_1, \ldots, \mu_r \in \mathbb{N}$, so ist offenbar $\mu_1 + \cdots + \mu_r \leq n$ und

$$\mu_1 + \cdots + \mu_r = n$$

genau dann, wenn p in Linearfaktoren zerfällt,

$$p(T) = c(T - \lambda_1)^{\mu_1} \ldots (T - \lambda_r)^{\mu_r} \ .$$

e) Ist nun wieder f ein Endomorphismus von V für einen endlich-dimensionalen Vektorraum V, so sagen wir, dass ein Eigenwert $\lambda \in K$ von f die **algebraische Vielfachheit** $\mu_\lambda \in \mathbb{N}$ hat, wenn μ_λ die Vielfachheit der Nullstelle λ im charakteristischen Polynom P_f von f ist.

5.26 Satz

Sei V ein endlich-dimensionaler K-Vektorraum, $f : V \to V$ ein Endomorphismus und $\lambda \in K$ ein Eigenwert von f mit geometrischer Vielfachheit $\nu \in \mathbb{N}$ und algebraischer Vielfachheit $\mu \in \mathbb{N}$. Dann gilt:

$$\nu \le \mu \ .$$

Beweis:
Sei $(v_1, \dots v_\nu)$ eine Basis von $\mathrm{Eig}(f; \lambda)$. Wir ergänzen diese nach (1.37) zu einer Basis $\mathfrak{a} = (v_1, \dots, v_\nu, v_{\nu+1}, \dots, v_n)$ von V. Die Matrix $A = M(f; \mathfrak{a}, \mathfrak{a})$ sieht dann so aus:

$$A = \left(\begin{array}{c|c} \lambda E_\nu & * \\ \hline 0 & C \end{array}\right) ,$$

mit einem $C \in \mathrm{Mat}_{n-\nu}(K)$.
Es ist dann

$$TE_n - A = \left(\begin{array}{c|c} (T-\lambda)E_\nu & * \\ \hline 0 & TE_{n-\nu} - C \end{array}\right)$$

und daher (vgl. A2, Blatt 10):

$$\begin{aligned} P_f(T) &= P_A(T) = \det(TE_n - A) = \det((T-\lambda)E_\nu)\ \det(TE_{n-\nu} - C) \\ &= (T-\lambda)^\nu \cdot P_C(T) \end{aligned}$$

Deshalb ist $\nu \le \mu$ (und $\nu = \mu$ genau dann, wenn λ kein Eigenwert von C mehr ist).

$\square$

5.27 Theorem

Sei V ein endlich-dimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus. Dann gilt: f ist genau dann diagonalisierbar, wenn folgendes richtig ist:

i) Das charakteristische Polynom $P_f \in K[T]$ zerfällt in Linearfaktoren.

ii) Für jeden Eigenwert $\lambda \in K$ ist die geometrische Vielfachheit $\nu_\lambda \in \mathbb{N}$ von λ gleich der algebraischen Vielfachheit $\mu_\lambda \in \mathbb{N}$ von λ, also $\nu_\lambda = \mu_\lambda$.

Beweis:
Sind $\lambda_1, \ldots, \lambda_r \in K$ die verschiedenen Eigenwerte von f mit geometrischen Vielfachheiten $\nu_1, \ldots, \nu_r \in \mathbb{N}$ und algebraischen Vielfachheiten $\mu_1, \ldots, \mu_r \in \mathbb{N}$, so gilt wegen (5.26) und (5.25 c):

$$\nu_1 + \cdots + \nu_r \leq \mu_1 + \cdots + \mu_r \leq n \ .$$

Die 2.Ungleichung wird genau dann zur Gleichung, wenn P_f in Linearfaktoren zerfällt (5.25 c) und die erste genau dann, wenn $\nu_k = \mu_k$ ist, für alle $k = 1, \ldots, r$ (wegen (5.26)). Da f genau dann diagonalisierbar ist, wenn $\nu_1 + \cdots + \nu_r = n$ ist (siehe (5.18 a)), folgt die Behauptung.

$\square$

5.28 Beispiel

Sei $A \in \mathrm{Mat}_3(\mathbb{Q})$ gegeben durch

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

Dann ist das charakteristische Polynom

$$\begin{aligned} P_A(T) &= \begin{vmatrix} T+3 & 1 & 1 \\ -8 & T-3 & -2 \\ 2 & 1 & T \end{vmatrix} \\ &= (T+3)(T-3)T - 4 - 8 - 2(T-3) + 2(T+3) + 8T \\ &= T^3 - 9T - 12 - 2T + 6 + 2T + 8T + 6 \\ &= T^3 - T = T(T-1)(T+1) \ . \end{aligned}$$

Also hat A die Eigenwerte $\lambda_1 = 1,\ \lambda_2 = -1,\ \lambda_3 = 0$. Wegen (5.11) ist A damit diagonalisierbar. Alle Vielfachheiten sind hier 1, also $\nu_1 = \mu_1 = \nu_2 = \mu_2 = \nu_3 = \mu_3 = 1$. Will man $S \in GL_3(\mathbb{Q})$, so dass $SAS^{-1} = \mathrm{diag}(1, -1, 0)$ ist, so muss man eine Eigenbasis $\mathfrak{a} = (v_1, v_2, v_3)$ von v bestimmen, also $0 \neq v_k \in \mathrm{Eig}(\lambda_k) = \ker(\lambda_k E_3 - A)$, $(k = 1, 2, 3)$:

$$\begin{aligned}
\lambda_1 &= 1 : \begin{pmatrix} 4 & 1 & 1 \\ -8 & -2 & -2 \\ 2 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{pmatrix} \to \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\
&\Rightarrow \quad v_1 = (0, -1, 1) \text{ ist EV zu } \lambda_1 = 1;
\end{aligned}$$

$$\begin{aligned}
\lambda_2 &= -1 : \begin{pmatrix} 2 & 1 & 1 \\ -8 & -4 & -2 \\ 2 & 1 & -1 \end{pmatrix} \to \begin{pmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\
&\Rightarrow \quad v_2 = (-\frac{1}{2}, 1, 0) \text{ ist EV zu } \lambda_2 = -1.
\end{aligned}$$

Auch $\quad 2v_2 \ = (-1, 2, 0)$ ist dann Eigenvektor zum Eigenwert λ_2.

$$\begin{aligned}
\lambda_3 &= 0 : \begin{pmatrix} 3 & 1 & 1 \\ -8 & -3 & -2 \\ 2 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \\
&\Rightarrow \quad v_3 = (-1, 2, 1) \text{ ist EV zu } \lambda_3 = 0.
\end{aligned}$$

Wir setzen nun

$$T := \begin{pmatrix} 0 & -1 & -1 \\ -1 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix} ,$$

berechnen noch $S := T^{-1}$ zu

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

und weil nun $Te_j = v_j$, also $Sv_j = e_j$ für $j = 1, 2, 3$ ist, gilt:

$$SAS^{-1}(e_j) = SAv_j = S(\lambda_j v_j) = \lambda_j S v_j = \lambda_j e_j \ ,$$

also ist wie gewünscht:

$$B = SAS^{-1} = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix} = \mathrm{diag}(1, -1, 0) \ .$$

5.29 Kommentar: Fundamentalsatz

a) Ein Körper K heißt **algebraisch abgeschlossen**, wenn jedes Polynom $p \in K[T]$ vom Grad $n \geq 1$ mindestens eine Nullstelle in K hat. Es zerfällt dann in Linearfaktoren. Der so genannte **Fundamentalsatz der Algebra** (Gauss 1796) besagt: $\mathbb{C}$ ist algebraisch abgeschlossen.

b) Man kann jeden Körper K in einen Körper $\widetilde{K}$ einbetten (d. h. es gibt einen injektiven Körperhomomorphismus $K \to \widetilde{K}$), der algebraisch abgeschlossen ist (z. B. $\mathbb{R} \to \mathbb{C}$). Man kann dann auch jeden K-Vektorraum V zu einem Vektorraum $\widetilde{V}$ über $\widetilde{K}$ erweitern (mit $\dim_{\widetilde{K}} \widetilde{V} = \dim_K V$, z.B. $K^n \to \widetilde{K}^n$) und jeden Endomorphismus $f \in \mathrm{End}(V)$ zu einem Endomorphismus $\widetilde{f} \in \mathrm{End}(\widetilde{V})$, so dass folgendes Diagramm kommutiert:

Zeichnung fehlt

Es ist dann

$$P_{\widetilde{f}} = P_f \in K[T] \subseteq \widetilde{K}[T] \ .$$

Nach Übergang zu einem größeren Körper zerfällt also das charakteristische Polynom von f. (Aber Achtung: Die Eigenwerte von f liegen nun in $\widetilde{K}$ und die Eigenvektoren in $\widetilde{V}$ (vgl. (5.8))).

c) Ist $f \in \mathrm{End}(V)$ derart, dass ihr charakteristisches Polynom zerfällt, so kann man f immerhin **trigonalisieren**, d.h. es gibt eine Basis $\mathfrak{a}$ von V, so dass $A = M(f; \mathfrak{a}, \mathfrak{a})$ **trigonal** ist, also:

$$A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \ .$$

(Auf der Diagonalen von A stehen dabei die Eigenwerte von f und zwar so oft, wie die algebraischen Vielfachen es vorgeben, vgl. z.B. G. Fischer, § 4.4).

d) Die beste Form (so genannte **Normalform**) erreicht man bei Endomorphismen $f : V \to V$, deren charakteristisches Polynom zerfällt, durch eine Basiswahl $\mathfrak{a}$, so dass $A = M(f; \mathfrak{a}, \mathfrak{a})$ **Jordan-Form** hat, d.h.:

$$A = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_s \end{pmatrix}$$

wobei, jede Matrix J_l eine **Jordan-Matrix** ist, d.h.:

$$J_l = \begin{pmatrix} \lambda_l & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_l \end{pmatrix},$$

wobei λ_l wieder Eigenwert von f ist, $l = 1, \ldots, s$, (siehe z.B. G. Fischer, § 4.6).

e) Zerfällt das charakteristische Polynom von f nicht (und möchte man keine Körpererweiterung machen (vgl. (a))), so ist der Normalformensatz schwierig (vgl. z.B. S. Bosch, Kapitel 6).

6 Euklidische und unitäre Vektorräume

6.1 Motivation: Länge eines Vektors

Um nun in Vektorräumen V über $\mathbb{R}$ oder $\mathbb{C}$ auch Geometrie treiben zu können, d.h. hier insbesondere den Abstand zwischen zwei Vektoren erklären zu können, braucht man neben den rein algebraischen Vektorraumstrukturen $+$ und $\cdot$ eine weitere Struktur. (In einem „blanken" $\mathbb{R}$-Vektorraum ist die Länge eines Vektors nicht erklärt.) Man nimmt hierbei üblicherweise den Satz des Pythagoras als Motivation, der die Länge des Vektors $x = (x_1, x_2) \in \mathbb{R}^2$ als

$$|| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} || = \sqrt{x_1^2 + x_2^2}$$

vorschlägt und damit von einem „Skalarprodukt" $\langle -, - \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ kommt, d.h.:

$$||x|| = \sqrt{\langle x, x \rangle}$$

mit

$$\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \rangle := x_1 y_1 + x_2 y_2 \ .$$

Die Eigenschaften dieses „Skalarprodukts" werden nun axiomatisiert.

6.2 Vereinbarung: komplexe Zahlen

Im Folgenden steht $\mathbb{K}$ stets für die reellen Zahlen $\mathbb{R}$ oder die komplexen Zahlen $\mathbb{C}$. Ist $z \in \mathbb{C}$ mit $z = x + iy$ $(x, y \in \mathbb{R})$, so schreiben wir

$$x = \mathrm{Re}(z) \quad (\textbf{Realteil von z}) \ ,$$

$$y = \mathrm{Im}(z) \quad (\textbf{Imaginärteil von z}) \ ,$$

sowie

$$\overline{z} := x - iy \ ,$$

das **Komplex-Konjugierte** von z. Es ist dann

$$\mathrm{Re}(z) = \frac{1}{2}(z + \overline{z}) \ ,$$

$$\mathrm{Im}(z) = \frac{1}{2i}(z - \overline{z}) \ .$$

Weiter heißt für jedes $z \in \mathbb{C}$

$$|z| = \sqrt{z\overline{z}} = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2}$$

der **Betrag von** z. Und $z \in \mathbb{C}$ ist genau dann reell (d.h. $\mathrm{Im}(z) = 0$), wenn $z = \overline{z}$ ist.

6.3 Definition: bilinear, sesquilinear

Sei V ein Vektorraum über $\mathbb{K}$. Eine Abbildung $s : V \times V \mapsto \mathbb{K}$ heißt

a) im Falle $\mathbb{K} = \mathbb{R}$ **bilinear**, wenn für alle $v, v_1, v_2, w, w_1, w_2 \in V$ und $\lambda, \mu \in \mathbb{R}$ gilt:

i) $s(v_1 + v_2, w) = s(v_1, w) + s(v_2, w)$ und $s(\lambda v, w) = \lambda s(v, w)$

ii) $s(v, w_1 + w_2) = s(v, w_1) + s(v, w_2)$ und $s(v, \mu w) = \mu s(v, w)$ (vgl. (4.12))

b) im Falle $\mathbb{K} = \mathbb{C}$ **sesquilinear**, wenn für alle $v, v_1, v_2, w, w_1, w_2 \in V$ und $\lambda, \mu \in \mathbb{C}$ gilt:

i) $s(v_1 + v_2, w) = s(v_1, w) + s(v_2, w)$ und $s(\lambda v, w) = \overline{\lambda} s(v, w)$

ii) $s(v, w_1 + w_2) = s(v, w_1) + s(v, w_2)$ und $s(v, \mu w) = \mu s(v, w)$

6.4 Kommentar: Bilinear- und Sesquilinearform

a) Im Fall $\mathbb{K} = \mathbb{R}$ nennt man s **eine Bilinearform auf V**, weil s im 1. und 2. Argument linear ist.

b) Im Fall $\mathbb{K} = \mathbb{C}$ heißt s **eine Sesquilinearform auf V** (sesqui = eineinhalb), da s im 1. Argument semilinear und im 2. Argument linear ist.

6.5 Definition: symmetrisch, Hermitesch

Sei V ein Vektorraum über $\mathbb{K}$.

a) Eine Bilinearform (im Fall von $\mathbb{K} = \mathbb{R}$) $s : V \times V \mapsto \mathbb{R}$ heißt **symmetrisch**, wenn für alle $v, w \in V$ gilt:

$$s(v, w) = s(w, v) \ .$$

b) Eine Sesquilinearform (im Fall $\mathbb{K} = \mathbb{C}$) $s : V \times V \mapsto \mathbb{C}$ heißt **Hermitesch**, wenn für alle $v, w \in V$ gilt:

$$s(v, w) = \overline{s(w, v)} \ .$$

Sie heißt dann **eine Hermitesche Form (HF)**.

6.6 Kommentar

Man beachte, dass für eine HF s und jedes $v \in V$ gilt:

$$s(v,v) = \overline{s(v,v)} \ ,$$

also $s(v,v) \in \mathbb{R} \subseteq \mathbb{C}$. Wenn wir im Folgenden für eine komplexe Zahl z schreiben „$z > 0$“, so ist gemeint: $z \in \mathbb{R} \subseteq \mathbb{C}$ und $z > 0$ (als reelle Zahl). (Erinnere, dass $\mathbb{C}$ keine Anordnung besitzt.)

6.7 Definition: Skalarprodukt

Sei V ein Vektorraum über $\mathbb{K}$. Eine symmetrische Bilinearform (sBF) bzw. eine HF $s : V \times V \mapsto \mathbb{K}$ heißt **positiv definit** (oder ein **Skalarprodukt** auf V), wenn für alle $v \in V \backslash \{0\}$ gilt:

$$s(v,v) > 0 \ .$$

6.8 Kommentar: euklidische/unitäre Vektorräume

a) Ein Skalarprodukt wird auch häufig mit $\langle -,- \rangle : V \times V \to \mathbb{K}$ notiert.

b) Ist V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\langle -,- \rangle$ ein Skalarprodukt auf V, so heißt das Paar $(V, \langle -,- \rangle)$ im Fall $\mathbb{K} = \mathbb{R}$ ein **euklidscher Vektorraum** und im Fall $\mathbb{K} = \mathbb{C}$ ein **unitärer Vektorraum**.

6.9 Beispiel: kanonische Skalarprodukte

a) Sei $V = \mathbb{R}^n$ $(n \in \mathbb{N})$ und $\langle -,- \rangle : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ gegeben durch

$$\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n$$

(= $x^t y$, wenn $x, y \in \mathbb{R}^n$ Spaltenvektoren sind). Dann ist $\langle -,- \rangle$ ein Skalarprodukt auf $\mathbb{R}^n$. Es heißt das **kanonische Skalarprodukt**. $(\mathbb{R}^n, \langle -,- \rangle)$ heißt der (standarde) euklidische Raum.

b) Sei $V = \mathbb{C}^n$ und $\langle -,- \rangle : \mathbb{C}^n \times \mathbb{C}^n \mapsto \mathbb{C}$ gegeben durch

$$\langle x, y \rangle = \overline{x_1} y_1 + \cdots + \overline{x_n} y_n = \overline{x}^t y.$$

Dann ist $\langle -,- \rangle$ ein Skalarprodukt auf $\mathbb{C}^n$. Es heißt das **kanonische Skalarprodukt auf $\mathbb{C}^n$**.

c) Sei $V = \mathbb{R}^n$ und $s : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ gegeben durch

$$s(x, y) = x_1 y_1 + \cdots + x_{n-1} y_{n-1} - x_n y_n \ .$$

Damit ist s zwar bilinear und symmetrisch, nicht aber positiv definit.

d) Sei $V = \mathfrak{C}([0,1]) = \{f : [0,1] \to \mathbb{K} : f \text{ stetig}\}$ und

$$\langle f, g \rangle := \int_0^1 \overline{f}(x) g(x) dx$$

Dann ist $\langle -, - \rangle$ ein Skalarprodukt auf V (Übung).

6.10 Definition: Norm

Ist $\langle -, - \rangle$ ein Skalarprodukt auf einem $\mathbb{K}$-Vektorraum V, so erklärt man **die Norm auf** V durch $||\_|| : V \to [0, \infty)$,

$$||v|| := \sqrt{\langle v, v \rangle} \ .$$

6.11 Kommentar: Metrik

a) Im Falle des kanonischen Skalarproduktes $\langle -, - \rangle$ auf $\mathbb{K}^n$ wird das offensichtlich zu

$$||x|| = \sqrt{x_1^2 + \cdots + x_n^2} \quad \text{(im Fall } \mathbb{K} = \mathbb{R})$$

bzw.

$$||x|| = \sqrt{|x_1|^2 + \cdots + |x_n|^2} \quad \text{(im Fall } \mathbb{K} = \mathbb{C}) \ .$$

b) Offenbar gilt tatsächlich $||v|| \geq 0$, für alle $v \in V$ und wegen der positiven Definitheit auch: $||v|| = 0$, genau dann wenn $v = 0$ ist.

c) Es gilt auch folgende **Homogenität**:

$$||\lambda v|| = |\lambda| \cdot ||v||$$

für alle $v \in V$, $\lambda \in \mathbb{K}$, denn:

$$||\lambda v||^2 = \langle \lambda v, \lambda v \rangle = \overline{\lambda} \lambda \langle v, v \rangle = |\lambda|^2 \cdot ||v||^2$$

also

$$||\lambda v|| = |\lambda| \cdot ||v|| \ .$$

Im Fall $\mathbb{K} = \mathbb{R}$ ist natürlich $\overline{\lambda} = \lambda$ und $|\lambda|^2 = \lambda^2$.

d) **Der Abstand** (oder **die Metrik**) $d : V \times V \to [0, \infty)$ zwischen zwei Vektoren $v, w \in V$ wird dann definiert durch:

$$d(v, w) := ||w - v|| \ .$$

6.12 Satz: Cauchy-Schwarz

Sei $\langle -, - \rangle$ ein Skalarprodukt auf einem $\mathbb{K}$-Vektorraum. Dann gilt für alle $v, w \in V$:

a)

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

b) (v, w) ist genau dann linear abhängig, wenn gilt:

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

Beweis:
Für $v = 0$ ist die Aussage offenbar richtig. Sei deshalb o.E. $v \neq 0$.

a) Für alle $v, w \in V$ und $\lambda, \mu \in \mathbb{K}$ ist:

$$\begin{aligned} 0 &\leq ||\lambda v - \mu w||^2 = \langle \lambda v - \mu w,\ \lambda v - \mu w\rangle \\ &= \overline{\lambda}\lambda\langle v, v\rangle - \overline{\lambda}\mu\langle v, w\rangle - \overline{\mu}\lambda\langle w, v\rangle + \overline{\mu}\mu\langle w, w\rangle \\ &= |\lambda|^2||v||^2 - (\overline{\lambda}\mu\langle v, w\rangle + \overline{\overline{\lambda}\mu\langle v, w\rangle}) + |\mu|^2 \cdot ||w||^2 \end{aligned}$$

denn $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$, $\overline{\overline{z}} = z$ und $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$ für alle $z, z_1, z_2 \in \mathbb{C}$.
Es ist also

$$0 \leq |\lambda|^2 \cdot ||v||^2 - 2\mathrm{Re}(\overline{\lambda}\mu\langle v, w\rangle) + |\mu|^2 \cdot ||w||^2 \ .$$

Setzt man speziell

$$\lambda := \langle v, w\rangle \text{ und } \mu = \langle v, v\rangle = ||v||^2 > 0 \ ,$$

so ist:

$$\begin{aligned} 0 &\leq |\langle v, w\rangle|^2 \cdot ||v||^2 - 2\mathrm{Re}(|\langle v, w\rangle|^2||v||^2) + ||v||^4||w||^2 \\ &= ||v||^2(|\langle v, w\rangle|^2 - 2|\langle v, w\rangle|^2 + ||v||^2||w||^2) \\ &= ||v||^2(||v||^2||w||^2 - |\langle v, w\rangle|^2) \ . \end{aligned}$$

Wegen $||v||^2 > 0$ folgt

$$||v||^2||w||^2 - |\langle v, w\rangle|^2 \geq 0$$

und damit (a).

b) Ist (v, w) linear abhängig, so $\exists \lambda \in \mathbb{K}$ mit $w = \lambda v$, also

$$|\langle v, w\rangle|^2 = |\langle v, \lambda v\rangle|^2 = |\lambda\langle v, v\rangle|^2 = |\lambda|^2 \cdot ||v||^4 \ .$$

Aber ebenso ist

$$||v||^2 \cdot ||w||^2 = ||v||^2 \cdot |\lambda|^2 \cdot ||v||^2 = |\lambda|^2 \cdot ||v||^4 \ .$$

Ist umgekehrt $|\langle v, w\rangle| = ||v|| \cdot ||w||$, so zeigt die Rechnung unter (a), dass $\lambda v - \mu w = 0$ sein muss mit $\mu = ||v||^2 > 0$ (und $\lambda = \langle v, w\rangle$). Also ist (v, w) linear abhängig.

□

6.13 Kommentar: Dreiecks-Ungleichung

a) Wegen der Ungleichung von Cauchy-Schwarz ist für alle $v, w \in V \backslash \{0\}$ in einem euklidischen Vektorraum $(V, \langle -, -\rangle)$

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

Mit $\arccos : [-1, 1] \to [0, \pi]$ (vgl. Teil 1), kann man daher den Winkel zwischen v und w durch

$$\varphi := \angle(v, w) := arccos \left(\frac{\langle v, w\rangle}{||v|| \cdot ||w||} \right)$$

erklären. Es gilt dann also (nach Definition):

$$\langle v, w\rangle = ||v|| \cdot ||w|| \cdot \cos(\varphi)$$

(mit $\varphi \in [0, \pi]$) für alle $v, w \in V$.

b) Es gilt nun auch für $||\_||$ die so gennante **Dreiecks-Ungleichung**:

$$||v + w|| \leq ||v|| + ||w||$$

für alle $v, w \in V$, denn

$$\begin{aligned} ||v + w||^2 &= \langle v + w, v + w\rangle = ||v||^2 + 2\mathrm{Re}(\langle v, w\rangle) + ||w||^2 \\ &\leq ||v||^2 + 2|\langle v, w\rangle| + ||w||^2 \quad (\text{weil } \mathrm{Re}(z) \leq |z| \quad \forall z \in \mathbb{C}) \\ &\overset{(C.S.)}{\leq} ||v||^2 + 2||v|| \cdot ||w|| + ||w||^2 = (||v|| + ||w||)^2 \ . \end{aligned}$$

c) Für die induzierte Metrik $d : V \times V \to [0, \infty)$ bedeutet dies für alle $v, w, u \in V$ (Übung):

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

6.14 Vorbereitung: komplex-konjugierte Matrix

a) Ist $A \in \mathrm{Mat}_n(\mathbb{C})$ eine komplexe $n \times n$-Matrix, $A = (a_{ij})$, so heißt

$$\overline{A} := (\overline{a}_{ij})$$

die komlex-konjugierte Matrix zu A. Es gilt dann:

$$\overline{A+B} = \overline{A} + \overline{B}, \quad \overline{\lambda A} = \overline{\lambda}\,\overline{A}\ ,$$

$$\overline{AB} = \overline{A} \cdot \overline{B}, \quad \det(\overline{A}) = \overline{\det(A)}\ ,$$

für alle $A, B \in \mathrm{Mat}_n(\mathbb{C})$ und $\lambda \in \mathbb{C}$. Ist $S \in GL_n(\mathbb{C})$, so ist deshalb auch

$$\overline{S^{-1}} = \overline{S}^{-1}\ .$$

b) Erinnere auch die Rechenregeln für das Transponieren:

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

$$(AB)^t = B^t \cdot A^t, \quad \det(A^t) = \det(A)\ ,$$

sowie

$$(S^{-1})^t = (S^t)^{-1}\ ,$$

falls $S \in GL_n(\mathbb{C})$.

6.15 Definition: symmetische und Hermitesche Matrizen

Sei $n \in \mathbb{N}$ und $A \in \mathrm{Mat}_n(\mathbb{K})$.

a) Im Fall $\mathbb{K} = \mathbb{R}$ heißt A **symmetrisch**, wenn gilt:

$$A^t = A\ .$$

b) Im Fall $\mathbb{K} = \mathbb{C}$ heißt A **Hermitesch**, wenn gilt:

$$\overline{A}^t = A\ .$$

6.16 Kommentar

a) Im Fall $\mathbb{K} = \mathbb{R}$ notieren wir

$$\mathrm{Sym}_n(\mathbb{R}) := \{A \in \mathrm{Mat}_n(\mathbb{R}) : \ A^t = A\} \ .$$

Es ist dann $\mathrm{Sym}_n(\mathbb{R}) \subseteq \mathrm{Mat}_n(\mathbb{R})$ ein Unterraum der Dimension $\frac{1}{2}n(n+1)$.

b) Im Fall $\mathbb{K} = \mathbb{C}$ setzen wir

$$\mathrm{Herm}_n(\mathbb{C}) := \{A \in \mathrm{Mat}_n(\mathbb{C}) : \ \overline{A}^t = A\} \ .$$

Hier ist $\mathrm{Herm}_n(\mathbb{C}) \subseteq \mathrm{Mat}_n(\mathbb{C})$ nur ein reeller Unterraum (im komplexen $\mathrm{Mat}_n(\mathbb{C})$) der (reellen) Dimension n^2 (Übung).

6.17 Beispiele

Sei $n \in \mathbb{N}$, $V = \mathbb{K}^n$ und $A \in \mathrm{Mat}_n(\mathbb{K})$.

a) Im Fall $\mathbb{K} = \mathbb{R}$ gilt, dass

$$s_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad s(x,y) = x^t A y$$

bilinear ist und (genau dann) symmetrisch, wenn A symmetrisch ist, denn

$$\begin{aligned} s_A(y,x) \quad &= \quad y^t \cdot A \cdot x = (x^t \cdot A^t \cdot y)^t = x^t \cdot A^t \cdot y \\ &\overset{(A=A^t)}{=} \quad x^t \cdot A \cdot y = s_A(x,y) \end{aligned}$$

für alle $x, y \in \mathbb{R}^n$. Der Fall $A = E_n$ liefert das kanonische Skalarprodukt.

b) Im Fall $\mathbb{K} = \mathbb{C}$ sei $A \in \mathrm{Herm}_n(\mathbb{C})$. Dann ist

$$s_A : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}^n, \quad s(x,y) = \overline{x}^t A y$$

eine HF (ähnlich wie unter (a)).

6.18 Definition: positiv definite Matrix

Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ symmetrisch bzw. Hermitsch. Es heißt A **positiv definit**, wenn für alle $x \in \mathbb{K}^n \backslash \{0\}$ gilt:

$$\overline{x}^t A x > 0 \quad (\text{kurz } A > 0) \ .$$

6.19 Kommentar

a) Wir bezeichnen für $n \in \mathbb{N}$ mit

$$\mathrm{Pos}_n(\mathbb{R}) := \{A \in \mathrm{Sym}_n(\mathbb{R}) : A > 0\}$$

und

$$\mathrm{Pos}_n(\mathbb{C}) := \{A \in \mathrm{Herm}_n(\mathbb{C}) : A > 0\}$$

die positiv definiten Matrizen. Sie sind keine Unterräume mehr, denn z.B. ist für $A \in \mathrm{Pos}_n(\mathbb{K})$ sicher $-A \notin \mathrm{Pos}_n(\mathbb{K})$.

b) Man beachte, dass $A > 0$ nicht etwa bedeutet, dass $a_{ij} > 0$ ist, für alle $1 \leq i, j \leq n$. Z.B. ist

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

positiv definit, während $B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ es nicht ist (Übung).

c) Ist nun $A \in \mathrm{Pos}_n(\mathbb{K})$, so ist klar, dass durch

$$s_A : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}, \quad s_A(x, y) = \overline{x}^t A y$$

ein Skalarprodukt auf $\mathbb{K}^n$ gegeben wird. (Wir werden bald sehen (siehe (6.23 c)), dass man so alle Skalarprodukte auf $\mathbb{K}^n$ bekommt.)

d) Ähnlich wie man durch Matrizen $A \in \mathrm{Mat}_n(\mathbb{K})$ lineare Abbildungen $f_A : \mathbb{K}^n \to \mathbb{K}^n$, durch $f_A(x) = Ax$ bekommt, bekommt man also auch durch $s_A : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$, $s_A(x, y) = \overline{x}^t Ay$ Sesquilinearformen, insbesondere durch $A \in \mathrm{Pos}_n(\mathbb{K})$ Skalarprodukte auf $\mathbb{K}^n$.

e) Umgekehrt haben wir in §2 gesehen, wie man Endomorphismen $f : V \to V$ (bei $n = \dim V < \infty$) durch Matrizen $A \in \mathrm{Mat}_n(\mathbb{K})$ durch eine Basiswahl $\mathfrak{a}$ von V beschreiben kann, $A = M(f; \mathfrak{a}, \mathfrak{a})$. Ähnliches gilt nun auch für Sesquilinearformen, insbesondere für Skalarprodukte:

6.20 Definition: Matrix einer Sesquilinearform

Sei V ein $\mathbb{K}$-Vektorraum der Dimension $n \in \mathbb{N}$ und $s : V \times V \to \mathbb{K}$ eine Sesquilinearform. Ist nun $\mathfrak{a} = (v_1, \ldots, v_n)$ eine Basis von V, so setzen wir $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$,

$$a_{ij} = s(v_i, v_j) \ ,$$

und nennen A **die Matrix von** s **bzgl.** $\mathfrak{a}$ und notieren:

$$A =: M(s; \mathfrak{a}) \ .$$

6.21 Kommentar

a) Wie in (6.20) fassen wir im Folgenden den reellen Fall $\mathbb{K} = \mathbb{R}$ häufig als einen Spezialfall des komplexen Falls auf. (Es ist dabei natürlich eine Sesquilinearform auf einem reellen Vektorraum V bilinear, weil $\lambda = \overline{\lambda}$ ist, für alle $\lambda \in \mathbb{R}$.)

b) Es folgt nun leicht (z.B. aus dem folgenden Satz), dass s durch $A = M(s;\mathfrak{a})$ vollständig bestimmt ist und s genau dann Hermitesch ist, wenn A es ist (und positiv definit, genau wenn A es ist).

6.22 Satz

Ist V ein $\mathbb{K}$-Vektorraum der Dimension n und $\mathfrak{a}$ eine Basis von V, $s : V \times V \to \mathbb{K}$ eine Sesquilinearform und $T^{\mathfrak{a}} : \mathbb{K}^n \to V$ der Koordinaten-Isomorphismus zu $\mathfrak{a}$, so gilt für die Matrix $A = M(s;\mathfrak{a})$ und für alle $x, y \in \mathbb{K}^n$:

$$s(T^{\mathfrak{a}}(x), T^{\mathfrak{a}}(y)) = s_A(x, y)$$

d.h. das folgende Diagramm kommutiert:
$\boxed{\text{Zeichnung fehlt}}$

Beweis:
Es ist für $x, y \in \mathbb{K}^n$ bei $\mathfrak{a} = (v_1, \ldots, v_n)$ wegen $T^{\mathfrak{a}}(x) = \sum_{i=1}^n x_i v_i$:

$$\begin{aligned} s(T^{\mathfrak{a}}(x), T^{\mathfrak{a}}(y)) &= s(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j) \\ &= \sum_{i,j} \overline{x}_i y_j \underbrace{s(v_i, v_j)}_{=a_{ij}} = \sum_{i,j} \overline{x}_i a_{ij} y_j \\ &= \overline{x}^t A y = s_A(x, y) \ . \end{aligned}$$

□

6.23 Beispiele

a) Ist $\langle -, - \rangle : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ das kanonische Skalarprodukt und $\mathfrak{K} = (e_1, \ldots, e_n)$ die kanonische Basis von $\mathbb{K}^n$, so ist $M(\langle -, - \rangle; \mathfrak{K}) = E_n$, denn:

$$\langle e_i, e_j \rangle = \delta_{ij} \ ,$$

für alle $1 \leq i, j \leq n$.

b) Ist $A \in \mathrm{Mat}_n(\mathbb{K})$ und $s_A : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ die zugehörige Sesquilinearform, so gilt für die kanonische Basis $\mathfrak{K} = (e_1, \ldots, e_n)$ von $\mathbb{K}^n$: $M(s_A; \mathfrak{K}) = A$, denn

$$s_A(e_i, e_j) = \overline{e}_i^t A e_j = a_{ij} \ ,$$

für $1 \leq i, j \leq n$.

c) Ist $s : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ eine beliebige Sesquilinearform, $\mathfrak{K} = (e_1, \ldots, e_n)$ kanonisch und $A = M(s; \mathfrak{K})$ so gilt: $s = s_A$, denn zunächst ist für alle $1 \leq i, j \leq n$:

$$s(e_i, e_j) = a_{ij} = s_A(e_i, e_j)$$

und daher auch für alle $x, y \in \mathbb{K}^n$ wegen der Sesquilinearität:

$$s(x, y) = s_A(x, y)$$

also $s = s_A$.

6.24 Kommentar

a) Ähnlich wie bei linearen Abbildungen zwischen Vektorräumen kann man also auch bei Bilinearformen bzw. Sesquilinearformen Matrizen sehr effektiv als Hilfsmittel einsetzen.

b) Ist $\mathfrak{a}$ Basis von V und $A \in \mathrm{Mat}_n(\mathbb{K})$ beliebig, so gibt es genau ein $s : V \times V \to \mathbb{K}$ sesquilinear mit $M(s; \mathfrak{a}) = A$. Die Zuordnung

$$\mathrm{Sesqui}_n(\mathbb{K}) \to \mathrm{Mat}_n(\mathbb{K}) \ ,$$

$$s \mapsto M(s; \mathfrak{a})$$

ist ein $\mathbb{K}$-Isomorphismus.

c) Wir wollen nun das entsprechende Transformationsproblem lösen. Sei dazu V ein $\mathbb{K}$-Vektorraum der Dimension n und $s : V \times V \to \mathbb{K}$ eine Sesquilinearform. Seien $\mathfrak{a}$ und $\mathfrak{b}$ Basen von V und $A = M(s; \mathfrak{a})$ und $B = M(s; \mathfrak{b})$ die s beschreibenden Matrizen. Sei schließlich $S \in GL_n(\mathbb{K})$ die Basiswechselmatrix von $\mathfrak{b}$ nach $\mathfrak{a}$, $S = M(\mathrm{id}; \mathfrak{b}, \mathfrak{a})$. Wie transformieren sich nun A und B?

6.25 Lemma

Seien $A, B \in \mathrm{Mat}_n(\mathbb{K})$, so dass für alle $x, y \in \mathbb{K}^n$ gilt: $\overline{x}^t Ay = \overline{x}^t By$, dann gilt:

$$A = B \ .$$

Beweis:
Speziell für $x = e_i$ und $y = e_j$ $(1 \leq i, j \leq n)$ erhält man:

$$a_{ij} = \overline{e}_i^t A e_j = \overline{e}_i^t B e_j = b_{ij} \ .$$

□

6.26 Satz: Transformationsformel

Ist $s : V \times V \to \mathbb{K}$ eine Sesquilinearform auf einem Vektorraum V der Dimension n, sind $\mathfrak{a}$ und $\mathfrak{b}$ zwei Basen von V mit Basiswechselmatrix $S \in GL_n(\mathbb{K})$ und ist $A = M(s; \mathfrak{a})$ und $B = M(s; \mathfrak{b})$, so gilt:

$$B = \overline{S}^t AS \ .$$

Beweis:
Sind $T^{\mathfrak{a}}, T^{\mathfrak{b}} : \mathbb{K}^n \to V$ die Koordinaten-Isomorphismen und ist $f_S : \mathbb{K}^n \to \mathbb{K}^n, f_S(x) = Sx$, so gilt: $T^{\mathfrak{a}} \cdot f_S = T^{\mathfrak{b}}$, d.h. es kommutiert:
Zeichnung fehlt

denn mit $\mathfrak{a} = (v_1, \ldots, v_n)$ und $\mathfrak{b} = (w_1, \ldots, w_n)$ ist

$$\begin{aligned} T^{\mathfrak{b}}(e_j) &= w_j = \sum_i s_{ij} v_i = \sum_i s_{ij} T^{\mathfrak{a}}(e_i) \\ &= T^{\mathfrak{a}}(\sum_i s_{ij} e_i) = T^{\mathfrak{a}}(f_S(e_j)) = T^{\mathfrak{a}} \circ f_S(e_j) \end{aligned}$$

für $j = 1, \ldots, n$ und $(e_1, \ldots, e_n)$ kanonisch für $\mathbb{K}^n$. Deshalb kommutiert nun auch das folgende Basiswechseldiagramm (vgl. (3.31)):
Zeichnung fehlt

Deshalb ist für alle $x, y \in \mathbb{K}^n$:

$$\begin{aligned}\overline{x}^t By &= s_B(x,y) = s(T^{\mathfrak{b}}(x), T^{\mathfrak{b}}(y)) = s(T^{\mathfrak{a}} \circ f_S(x), T^{\mathfrak{a}} \circ f_S(y)) \\ &= s_A(Sx, Sy) = \overline{(Sx)}^t A(Sy) = \overline{x}^t \overline{S}^t ASy \ ,\end{aligned}$$

also ist nach (6.25):

$$B = \overline{S}^t AS \ .$$

□

6.27 Kommentar

a) Ähnlich wie bei den Endomorphismen eines $\mathbb{K}$-Vektorraumes drängt sich daher nun auch bei den Bilinear- bzw. Sesquilinearformen das Normalformenproblem auf: Welches ist „die beste Matrix“ für ein gegebenes $s : V \times V \to \mathbb{K}$?

b) Wir wollen dies hier zunächst nur für den Fall von Skalarprodukten behandeln und werden gleich feststellen (siehe (6.31)), dass der Normalenformensatz dann viel einfacher ist als im Endomorphismenfall. Dazu zunächst:

6.28 Definition: Orthonomalbasis

Sei $(V, \langle -,- \rangle)$ ein euklidischer bzw. unitärer Vektorraum. Eine Basis $\mathfrak{a} = (e_1, \ldots, e_n)$ heißt **Orthonormalbasis** (ON-Basis) von V, wenn für alle $1 \leq i, j \leq n$ gilt:

$$\langle e_i, e_j \rangle = \delta_{ij} \ .$$

6.29 Kommentar

a) Man beachte, dass wir bisher mit $(e_1, \ldots, e_n)$ stets die kanonische Basis von $\mathbb{K}^n$ bezeichnet haben. Nun benutzen wir diese Notation für jede ON-Basis eines euklidischen bzw. unitären Vektorraums.

b) Zum Glück ist die kanonische Basis $\boldsymbol{\mathfrak{K}} = (e_1, \ldots, e_n)$ von $\mathbb{K}^n$ aber orthonormal bzgl. des kanonischen Skalarprodukts $\langle -,- \rangle$ auf $\mathbb{K}^n$, denn

$$\langle e_i, e_j \rangle = \overline{e}_i^t \cdot e_j = \delta_{ij}$$

für $1 \leq i, j \leq n$.

6.30 Beispiele

a) Sei $\langle -,- \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ das kanonische Skalarprodukt und $e_1 = \frac{1}{\sqrt{2}}(1,1)$, $e_2 = \frac{1}{\sqrt{2}}(-1,1)$. Dann ist (e_1, e_2) eine ON-Basis von $(\mathbb{R}^2, \langle -,- \rangle)$.

b) Sei $V = \{f : \mathbb{R} \to \mathbb{R}$ stetig: f ist 2π periodisch$\}$ und $U := \langle \sin, \cos \rangle \subseteq V$ der von sin und cos aufgespannte Unterraum. Sei $\langle -,- \rangle : U \times U \to \mathbb{R}$,

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x)g(x)dx \ .$$

Dann ist (e_1, e_2) eine ON-Basis von U (Übung),

$$e_1 = \sqrt{2} \sin \ , \quad e_2 = \sqrt{2} \cos \ .$$

6.31 Satz

Sei $(V, \langle -,- \rangle)$ ein euklidischer bzw. unitärer Vektorraum. Dann besitzt V eine ON-Basis.

6.32 Kommentar

Im Fall, dass eine Bilinear- bzw. Sesquilinearform $s : V \times V \to \mathbb{K}$ symmetrisch bzw. Hermitesch und positiv definit ist, ist das Normalformenproblem also denkbar einfach gelöst: Es gibt eine Basis $\mathfrak{a}$ von V, so dass

$$M(s; \mathfrak{a}) = E_n$$

ist. Man wähle nämlich eine ON-Basis.

6.33 Korollar

Sei $A \in \mathrm{Pos}_n(\mathbb{K})$ positiv definit. Dann gibt es ein invertierbares $S \in GL_n(\mathbb{K})$ mit

$$\overline{S}^t A S = E_n \ .$$

Beweis:
Betrachte das Skalarprodukt $s_A : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$, $(x, y) \to \overline{x}^t A y$ und wähle dafür eine ON-Basis $\mathfrak{a}$. Ist dann $S \in GL_n(\mathbb{K})$ die Basiswechselmatrix von $\mathfrak{a}$ auf die kanonische Basis $\mathfrak{K}$, so ist wegen $M(s_A; \mathfrak{K}) = A$ nach (6.25).

$$E_n = M(s_A, \mathfrak{a}) = \overline{S}^t M(s_A; \mathfrak{K}) S = \bar{S}^t A S \ .$$

Beweis von (6.31):
Induktion nach $n = \dim V$.

- $n = 1$: Ist $v \in V\backslash\{0\}$ beliebig, so setze $e := \frac{v}{||v||}$ Dann ist (e) eine ON-Basis von V.

- $n \to n+1$: Wähle $v \in V\backslash\{0\}$ und setzte zunächst (wieder) $e_{n+1} := \frac{v}{||v||}$. Nun betrachte

  $$U := e_{n+1}^{\perp} := \{v \in V : \langle e_{n+1}, v\rangle = 0\}$$

  (das zu e_{n+1} **orthogonale Komplement**). Dann ist $U \subseteq V$ ein Unterraum der Dimension n, denn $U = \ker(f)$ für $f : V \to \mathbb{K},\ f(v) := \langle e_{n+1}, v\rangle$. Da $f(e_{n+1}) = 1 \neq 0$ ist, ist rg(f)=1, also

  $$\dim(U) = n + 1 - \mathrm{rg}(f) = n\ .$$

  Schränkt man nun $\langle -, -\rangle$ auf $U \times U$ ein, so erhält man einen euklidischen bzw. unitären Vektorraum der Dimension n, auf den man die Induktionsvoraussetzung anwenden kann. Sei also $(e_1, \dots, e_n)$ eine ON-Basis von U. Weil $||e_{n+1}|| = 1$ und $\langle e_{n+1}, e_i\rangle = 0$ für $i = 1, \dots, n$, ist dann $(e_1, \dots, e_n, e_{n+1})$ eine ON-Basis von V.

$\square$

6.34 Kommentar

a) Ist $s : V \times V \to \mathbb{K}$ eine nicht mehr notwendig positiv definite, aber immer noch symmetrische BF bzw. Hermitesche Form, so kann man ganz ähnlich wie im definiten Fall zeigen, dass es eine Basis $\mathfrak{a} = (e_1, \dots, e_n)$ von V gibt, so dass

$$M(s;\mathfrak{a}) = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}$$

$$k-\text{mal} \qquad l-\text{mal} \qquad m-\text{mal}$$

mit $0 \leq k, l, m \leq n$ und $k+l+m = n$ ist. Das Tripel (k, l, m) ist durch s eindeutig bestimmt (d.h. hängt nicht von $\mathfrak{a}$ ab, ist also eine Invariante), siehe z.B G. Fischer, §5.7 (**Sylvesterscher Trägheitssatz**).

b) Man nennt dann

$$\begin{aligned} \mathrm{rg}(s) &:= k+l \\ \mathrm{ind}(s) &:= k \\ \mathrm{sign}(s) &:= k-l \end{aligned}$$

den **Rang**, den **Index** und die **Signatur** von s.

c) Speziell der Fall $n = 4$ mit $k = 3$ und $l = 1$ (also $m = 0$) ist in der Elektrodynamik und in der Relativitätstheorie von Bedeutung ($\mathbb{K} = \mathbb{R}$). Das Tripel (V, s) heißt dann ein **Lorentz-Raum**. Der Standard-Lorentzraum ist $(\mathbb{R}^4, s)$ mit

$$s(x, y) = x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4 \ .$$

d) Man kann Satz (6.31) wie folgt verfeinern, indem man aus einer beliebigen Basis $\mathfrak{a} = (v_1, \ldots, v_n)$ eines euklidischen bzw. unitären Vektorraumen eine ON-Basis konstruiert:

6.35 Satz: ON-Verfahren von E. Schmidt

Sei $(V, \langle -, - \rangle)$ ein euklidischer bzw. unitärer Vektorraum und $\mathfrak{a} = (v_1, \ldots, v_n)$ eine Basis von V. Definiert man dann rekursiv:

$$\begin{aligned} e_1 &= \frac{v_1}{||v_1||} \ , \\ \tilde{e}_{k+1} &:= v_{k+1} - \sum_{i=1}^{k} \langle e_i, v_{k+1} \rangle e_i \ , \\ e_{k+1} &:= \frac{\tilde{e}_{k+1}}{||\tilde{e}_{k+1}||} \quad (k = 1, \ldots, n-1) \ , \end{aligned}$$

so ist $(e_1, \ldots, e_n)$ (wohldefiniert und) eine ON-Basis von V.

6.36 Kommentar

Man normiert also zunächst v_1 ähnlich wie in (6.31). Dann zerlegt man v_2 in eine Komponente in Richtung e_1 und dazu,

$$v_2 = \underbrace{\langle e_1, v_2 \rangle e_1}_{\in \mathbb{K} e_1} + \underbrace{(v_2 - \langle e_1, v_2 \rangle e_1)}_{\in e_1^{\perp}} \ .$$

Den senkrecht stehenden Anteil $\tilde{e_2}$ normiert man dann wieder zu e_2 und verfährt mit v_3 analog.

6.37 Beispiel

Wir betrachten die Basis $((1,1),(0,1))$ des euklidischen Raums $(\mathbb{R}^2, \langle -, - \rangle)$. Es ist dann:

$$||(1,1)|| = \sqrt{1+1} = \sqrt{2} \ ,$$

also

$$e_1 = \frac{1}{\sqrt{2}}(1,1) \ .$$

Weiter ist

$$\begin{aligned} \tilde{e}_2 &= v_2 - \langle e_1, v_2 \rangle e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \langle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \end{aligned}$$

und dann:

$$||\tilde{e}_2|| = \frac{1}{2}\sqrt{(-1)^2+1} = \frac{\sqrt{2}}{2} \, .$$

Also ist:

$$e_2 = \frac{\sqrt{2}}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix} \, .$$

Beweis von (6.35):
Induktion über $n = \dim V$.

- $n = 1$: Ein normierter Vektor ist eine ON-Basis.
- $n \to n+1$: Ist $U := \langle v_1, \dots v_n \rangle \subseteq V$, so dürfen wir nach Induktionsvoraussetzung für $(e_1, \dots, e_n)$ (aus $(v_1, \dots v_n)$ konstruiert) annehmen, dass sie ON-Basis für U ist. Für

$$\widetilde{e}_{n+1} = v_{n+1} - \sum_{k=1}^{n} \langle e_k, v_{n+1} \rangle e_k$$

ist nun $\widetilde{e}_{n+1} \notin U$, sonst wäre auch $v_{n+1} \in U$ und $\mathfrak{a} = (v_1, \dots, v_{n+1})$ keine Basis von V. Es ist damit $(e_1, \dots, e_n, \widetilde{e}_{n+1})$ eine Basis von V. Insbesondere ist $\widetilde{e}_{n+1} \neq 0$ und damit dann $e_{n+1} := \frac{\widetilde{e}_{n+1}}{||\widetilde{e}_{n+1}||}$ wohldefiniert. Und nach Konstruktion ist $||e_{n+1}|| = 1$. Für alle $1 \leq i, j \leq n$ ist nach Induktionsvoraussetzung $\langle e_i, e_j \rangle = \delta_{ij}$. Es bleibt also zu zeigen, dass $\langle e_i, e_{n+1} \rangle = 0$ für $i = 1, \dots, n$:

$$\begin{aligned}
\langle e_i, e_{n+1} \rangle &= \frac{1}{||\widetilde{e}_{n+1}||} \langle e_i, \widetilde{e}_{n+1} \rangle \\
&= \frac{1}{||\widetilde{e}_{n+1}||} \left(\langle e_i, v_{n+1} \rangle - \sum_{k=1}^{n} \langle e_i, \langle e_k, v_{n+1} \rangle e_k \rangle \right) \\
&= \frac{1}{||\widetilde{e}_{n+1}||} \left(\langle e_i, v_{n+1} \rangle - \sum_{k=1}^{n} \langle e_k, v_{n+1} \rangle \underbrace{\langle e_i, e_k \rangle}_{=\delta_{ik}} \right) \\
&= 0 \, .
\end{aligned}$$

6.38 Motivation

Bisher haben wir (alle) symmetrischen Bilinearformen bzw. HF'en auf endlich-dimensionalen $\mathbb{K}$-Vektorräumen studiert, insbesondere Skalarprodukte. Ab nun fixieren wir ein Skalarprodukt $\langle -, - \rangle$ auf V und studieren nun (spezielle) Endomorphismen von $(V, \langle -, - \rangle)$.

6.39 Vorbereitung

a) In (2.21) hatten wir gesehen, dass bei Vorgabe einer Basis $\mathfrak{a}$ eines $\mathbb{K}$-Vektorraums V die Zuordnung

$$\operatorname{End}_{\mathbb{K}}(V) \to \operatorname{Mat}_n(\mathbb{K}),$$
$$f \mapsto M(f;\mathfrak{a}) := M(f;\mathfrak{a},\mathfrak{a})$$

ein Vektorraumisomorphismus ist. Bezeichnet $T^{\mathfrak{a}} : \mathbb{K}^n \to V$ den zu $\mathfrak{a}$ gehörenden Koordinatenisomorphismus, so ist nämlich $A \mapsto T^{\mathfrak{a}} \circ f_A \circ (T^{\mathfrak{a}})^{-1}$ die Umkehrabbildung, weil nämlich das folgende Diagramm kommutiert:

Zeichnung fehlt

Nun bilden auch die Sesqui-Linearformen

$$\operatorname{Sesqui}_{\mathbb{K}}(V) = \{s : V \times V \to \mathbb{K} \text{ sesqui-linear}\}$$

einen $\mathbb{K}$-Vektorraum (in natürlicher Weise) und wiederum ist die Zuordnung

$$\operatorname{Sesqui}_{\mathbb{K}}(V) \to \operatorname{Mat}_n(\mathbb{K}) \ ,$$
$$s \mapsto M(s;\mathfrak{a}) \ ,$$

ein Isomorphismus mit Umkehrung $A \mapsto s_A \circ (T^{\mathfrak{a}} \times T^{\mathfrak{a}})^{-1}$, denn nun kommutiert das folgende Diagramm:

Zeichnung fehlt

Insbesondere ist also

$$\dim(\operatorname{End}(V)) = \dim \operatorname{Mat}_n(\mathbb{K}) = n^2 = \dim \operatorname{Sesqui}(V) \ .$$

b) Ist nun $(V, \langle -,- \rangle)$ ein euklidischer bzw. unitärer Vektorraum, so können wir jedem Endomorphismus $f : V \to V$ eine Sesqui-Linearform $s_f : V \times V \to \mathbb{K}$ durch $s_f : V \times V \to \mathbb{K}$,

$$s_f(v,w) := \langle v, f(w) \rangle \ ,$$

und ebenso $\widetilde{s}_f : V \times V \to \mathbb{K}$ durch

$$\widetilde{s}_f(v,w) := \langle f(w), v \rangle$$

zuordnen. Die Zuordnungen

$$\Phi : \operatorname{End}(V) \to \operatorname{Sesqui}(V),\ f \mapsto s_f$$
$$\Psi : \operatorname{End}(V) \to \operatorname{Sesqui}(V),\ f \mapsto \widetilde{s}_f$$

sind dann linear bzw. semilinear und noch mehr:

6.40 Lemma

Φ ist ein Isomorphismus, Ψ ist ein Semiisomorphismus.

Beweis:
Wir müssen die Bijektivität von Φ und Ψ zeigen. Wegen (2.28) reicht dazu die Injektivität, denn beide Vektorräume $\mathrm{End}(V)$ und $\mathrm{Sesqui}(V)$ haben die gleiche (endliche) Dimension. Sei also etwa $s_f = \Phi(f) = 0$. Für alle $v, w \in V$ ist dann

$$0 = s_f(v, w) = \langle v, f(w) \rangle$$

Insbesondere für $v = f(w)$ (bei zunächst festem $w \in V$), ist also

$$0 = \langle f(w), f(w) \rangle \;\Rightarrow\; f(w) = 0$$

für alle $w \in V$, also $f = 0$. Φ ist also injektiv. Ähnlich geht man bei Ψ vor.

$\square$

6.41 Satz

Sei $(V, \langle -, - \rangle)$ ein euklidischer bzw. unitärer Vektorraum und $f : V \to V$ linear. Dann gibt es genau eine lineare Abbildung $f^* : V \to V$, so dass für alle $v, w \in V$ gilt:

$$\langle v, f(w) \rangle = \langle f^*(v), w \rangle \ .$$

Beweis:
Die Bedingung an f^* lautet mit den Beziehungen von (6.34) gerade:

$$\Phi(f) = s_f = \widetilde{s}_{f^*} = \Psi(f^*) \ .$$

Da Ψ bijektiv ist, ist dies äquivalent zu

$$f^* = \Psi^{-1} \circ \Phi(f) \ .$$

Das zeigt die Existenz und die Eindeutigkeit von f^*.

$\square$

6.42 Definition: adjungierte Abbildung

Sei $(V, \langle -, - \rangle)$ ein euklidischer bzw. unitärer Vektorraum und $f \in \mathrm{End}(V)$. Der nach (6.41) eindeutig bestimmte Endomorphismus $f^* : V \to V$ mit

$$\langle v, f(w) \rangle = \langle f^*(v), w \rangle$$

für alle $v, w \in V$, heißt die zu f **adjungierte Abbildung**.

6.43 Beispiel: adjungierte Abbildung

Sei $(\mathbb{K}^n, \langle -,- \rangle)$ der (standarde) euklidische bzw. unitäre Raum und $A \in \mathrm{Mat}_n(\mathbb{K})$. Dann ist

$$f_A^* = f_{A^*}$$

wenn man

$$A^* := \overline{A}^t$$

setzt, denn für alle $x, y \in \mathbb{K}^n$ gilt:

$$\begin{aligned} \langle x, f_A(y) \rangle &= \langle x, Ay \rangle = \overline{x}^t Ay = \overline{(\overline{A}^t x)}^t y \\ &= \langle \overline{A}^t x, y \rangle = \langle A^* x, y \rangle = \langle f_{A^*}(x), y \rangle \end{aligned}$$

also $f_A^* = f_{A^*}$, weil $\overline{(\overline{A}^t x)}^t = (A^t \overline{x})^t = \overline{x}^t A$ ist.

6.44 Kommentar: adjungierte Matrix

a) Man nennt deshalb für $A \in \mathrm{Mat}_n(\mathbb{K})$

$$A^* := \overline{A}^t$$

die zu A **adjungierte Matrix**. Beachte, dass im reellen Fall $A^* = A^t$ ist.

b) Vorsicht! Diese „Adjungierte“ hat nichts mit der adjungierten Matrix A^{ad} aus (4.36) zu tun.

c) Wegen den Rechenregeln für das Konjugieren und Transponieren gilt:

$$\begin{aligned} (A+B)^* &= A^* + B^* \ , \\ (\lambda A)^* &= \overline{\lambda} A^* \ , \\ (AB)^* &= B^* A^* \ , \\ \det(A^*) &= \overline{\det(A)} \ , \end{aligned}$$

und falls $A \in GL_n(\mathbb{K})$ ist:

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

d) Wir wollen nun sehen, wie sich das Adjungieren von Abbildungen, $f \mapsto f^*$, durch die sie beschreibende Matrizen ausdrückt, wenn man eine Basis von V wählt. Da V nun eine Zusatzstruktur $\langle -,- \rangle$ trägt, ist es konsequent, nicht beliebige Basen zu betrachten, sondern nur ON-Basen, die dieser Zusatzstruktur angepasst sind.

6.45 Bemerkung

Sei $(V, \langle -, - \rangle)$ ein euklidischer bzw. unitärer Vektorraum und $\mathfrak{a}$ eine ON-Basis von V. Seit $f : V \to V$ linear und $A = M(f; \mathfrak{a})$. Ist dann $B = M(f^*; \mathfrak{a})$, so gilt:

$$B = A^* \ .$$

Beweis:
Ist $\mathfrak{a} = (e_1, \ldots, e_n)$, so ist also mit $A = (a_{ij})$ und $B = (b_{ij})$:

$$f(e_j) = \sum_{i=1}^{n} a_{ij} e_i \ ,$$

$$f^*(e_j) = \sum_{i=1}^{n} b_{ij} e_i \ ,$$

$$j = 1, \ldots, n \ .$$

Nach Definition von f^* folgt für $1 \leq j, k \leq n$:

$$\langle f^*(e_j), e_k \rangle = \sum_{i=1}^{n} \overline{b_{ij}} \langle e_i, e_k \rangle = \overline{b_{kj}} \ ,$$

während

$$\langle e_j, f(e_k) \rangle = \sum_{i=1}^{n} a_{ik} \langle e_j, e_k \rangle = a_{jk}$$

ist, also $b_{kj} = \overline{a_{jk}}$, d.h. $B = \overline{A}^t = A^*$.

□

6.46 Definition: Isometrie von V

Sei $(V, \langle -, - \rangle)$ ein euklidischer bzw. unitärer Vektorraum. Eine lineare Abbildung $f : V \to V$ heißt eine **Isometrie von** V, wenn für alle $v, w \in V$ gilt:

$$\langle f(v), f(w) \rangle = \langle v, w \rangle \ .$$

6.47 Kommentar

a) Isometrien sind insbesondere Automorphismen von V, denn ist für eine Isometrie $f : V \to V$ nun $f(v) = 0$, so ist auch

$$\langle v, v \rangle = \langle f(v), f(v) \rangle = 0 \ ,$$

also $v = 0$ und damit f injektiv, also bijektiv. Eine Isometrie erhält aber nicht nur die Vektorraumstruktur, sondern auch Längen (und damit Abstände) und im reellen Fall auch Winkel, denn

$$||f(v)||^2 = \langle f(v), f(v)\rangle = \langle v, v\rangle = ||v||^2 \ ,$$

für alle $v \in V$.

b) Da $\langle f(v), f(w)\rangle = \langle f^* \circ f(v), w\rangle$, für alle $v, w \in V$ ist und weil $\Psi = \mathrm{End}(V) \to \mathrm{Sesqui}(V),\ f \mapsto \widetilde{s}_f$ aus (6.39) injektiv ist, ist f genau dann Isometrie, wenn gilt:

$$f^* \circ f = \mathrm{id}$$

also $f^* = f^{-1}$ und damit auch $f \circ f^* = \mathrm{id}$.

6.48 Definition: orthogonal und unitär

Sei $n \in \mathbb{N}$.

a) Eine Matrix $S \in \mathrm{Mat}_n(\mathbb{R})$ heißt **orthogonal**, wenn gilt: $S^t \cdot S = E_n$.

b) Eine Matrix $S \in \mathrm{Mat}_n(\mathbb{C})$ heißt **unitär**, wenn gilt: $S^* \cdot S = E_n$.

6.49 Kommentar: orthogonale und unitäre Gruppe, Einheitskreislinie

a) Orthogonale bzw. unitäre Matrizen beschreiben gerade die Isometrien des standarden euklidischen bzw. unitären Raumes $(\mathbb{K}^n, \langle -, -\rangle)$, denn

$$\begin{aligned} \langle f_S(x), f_S(y)\rangle &= \langle Sx, Sy\rangle = (\overline{Sx})^t(Sy) \\ &= \overline{x}^t \cdot \overline{S}^t \cdot S \cdot y = \langle x, y\rangle \end{aligned}$$

für alle $x, y \in \mathbb{K}^n$, genau wenn $S^* \cdot S = E_n$ ist.

b) Jedes orthogonale bzw. unitäre $S \in \mathrm{Mat}_n(\mathbb{K})$ ist insbesondere invertierbar (da $f_S \in \mathrm{Aut}(\mathbb{K}^n)$). Man nennt

$$O(n) := \{S \in GL_n(\mathbb{R}) : S^t \cdot S = E_n\}$$

die **orthogonale Gruppe** (zum Index n) und

$$U(n) := \{S \in GL_n(\mathbb{C}) : S^* \cdot S = E_n\}$$

die **unitäre Gruppe** (zum Index n). Wegen den Rechenregeln von $A \mapsto A^*$ sind sie tatsächlich Untergruppen von $GL_n(\mathbb{K})$ (Übung).

c) Ist $S \in O(n)$, so ist $\det(S) \in \{-1, 1\}$, denn:

$$\begin{aligned} 1 &= \det(E_n) = \det(S^t S) = \det(S^t)\det(S) \\ &= \det(S)\det(S) = \det(S)^2 \ , \end{aligned}$$

während für $S \in U(n)$ die Determinante auf der **Einheitskreislinie**

$$S^1 := \{z \in \mathbb{C} : |z| = 1\}$$

liegt, denn

$$\begin{aligned} 1 &= \det(E_n) = \det(S^* S) = \det(S^*)\det(S) \\ &= \overline{\det(S)}\det(S) = |\det(S)|^2 \ . \end{aligned}$$

d) Auch die Eigenwerte einer orthogonalen bzw. unitären Matrix können nur ganz spezielle Werte annehmen:

i) Ist $\lambda \in \mathbb{R}$ ein Eigenwert von $S \in O(n)$, so ist $\lambda \in \{\pm 1\}$, denn ist $x \in \mathbb{R}^n \backslash \{0\}$ ein Eigenvektor zu λ, also $Sx = \lambda x$, so gilt:

$$\lambda^2 \cdot ||x||^2 = \langle \lambda x, \lambda x \rangle = \langle Sx, Sx \rangle = \langle x, x \rangle = ||x||^2$$

also $\lambda^2 = 1$.

ii) Ist $\lambda \in \mathbb{C}$ ein Eigenwert von $S \in U(n)$, so ist $\lambda \in S^1$, denn ist $z \in \mathbb{C}^n \backslash \{0\}$ ein Eigenvektor zu λ, also $Sz = \lambda z$, so gilt:

$$|\lambda|^2 \cdot ||z||^2 = \langle \lambda z, \lambda z \rangle = \langle Sz, Sz \rangle = \langle z, z \rangle = ||z||^2$$

also $|\lambda|^2 = 1$.

e) Nach Definition ist eine Matrix $S \in GL_n(\mathbb{K})$ genau dann orthogonal bzw. unitär, wenn ihre Spalten- (bzw. Zeilen-) Vektoren eine ON-Basis von $(\mathbb{K}^n, \langle -, - \rangle)$ bilden,

$$\langle Se_i, Se_j \rangle = \langle e_i, e_j \rangle = \delta_{ij}$$

(für alle $1 \leq i, j \leq n$).

6.50 Beispiele

a) Für jedes $\varphi \in \mathbb{R}$ ist

$$S(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \in O(2) \ .$$

b) Ist $z = (z_1, z_2) \in \mathbb{C}^2$ mit $||z||^2 = |z_1|^2 + |z_2|^2 = 1$, so ist

$$S(z) = \begin{pmatrix} z_1 & -\overline{z_2} \\ z_2 & \overline{z_1} \end{pmatrix} \in U(2) \ .$$

6.51 Bemerkung

Sei $(V, \langle -, - \rangle)$ euklidisch bzw. unitär, $\mathfrak{a}$ eine ON-Basis und $f \in \mathrm{End}(V)$. Es ist f genau dann eine Isometrie von V, wenn $S = M(f; \mathfrak{a})$ orthogonal bzw. unitär ist.

Beweis:
Ist $\mathfrak{a} = (e_1, \ldots e_n)$, so ist mit $S = (a_{ij})$ für alle $1 \leq i, j \leq n$:

$$\begin{aligned} \langle f(e_i), f(e_j) \rangle &= \langle \sum_k a_{ki} e_k, \sum_l a_{lj} e_l \rangle \\ &= \sum_{k,l} \overline{a}_{ki} a_{lj} \langle e_k, e_l \rangle \\ &= \sum_{k=1}^{n} \overline{a}_{ki} a_{kj} = \langle e_i, e_j \rangle \\ &\Leftrightarrow \sum_k \overline{a}_{ki} a_{kj} = \delta_{ij} \quad \Leftrightarrow \quad S^* \cdot S = E_n \ . \end{aligned}$$

□

6.52 Definition: selbstadjungiert

Sei $(V, \langle -, - \rangle)$ ein euklidischer bzw. unitärer Vektorraum. Man nennt einen Endomorphismus $f : V \to V$ **selbstadjungiert** (in der Physik auch symmetrisch), wenn gilt $f^* = f$, also:

$$\langle f(v), w \rangle = \langle v, f(w) \rangle$$

für alle $v, w \in V$.

6.53 Bemerkung

Sei $(V, \langle -, - \rangle)$ euklidischer bzw. unitärer Vektorraum und $\mathfrak{a} = (e_1, \dots e_n)$ eine ON-Basis. Ein Endomorphismus $f : V \to V$ ist genau dann selbstadjungiert, wenn für seine beschreibende Matrix $A = M(f; \mathfrak{a})$ gilt:

$$A = A^* \ .$$

Beweis:
Für $1 \leq i, j \leq n$ ist mit $f(e_j) = \sum_i a_{ij} e_i$ und $A = (a_{ij})$:

$$\langle f(e_i), e_j \rangle = \langle \sum_k a_{ki} e_k, e_j \rangle = \sum_k \overline{a}_{ki} \langle e_k, e_j \rangle = \overline{a}_{ji}$$

und

$$\langle e_i, f(e_j) \rangle = \sum_k a_{kj} \langle e_i, e_k \rangle = a_{ij} \ .$$

Also ist $f = f^*$, genau wenn $A = A^*$ ist.

6.54 Kommentar: Spektrum

a) Selbstadjungierte Abbildungen tauchen insbesondere in der Quantenmechanik als so genannte Observablen auf (dort allerdings als selbstadjungierte Abbildungen auf unendlich dimensionalen, komplexen Vektorräumen V mit Skalarprodukt). Ihre Eigenwerte sind dann die möglichen Messwerte.

b) Ist $f \in \mathrm{End}(V)$ für einen (endlich-dimensionalen) $\mathbb{K}$-Vektorraum V, so nennt man

$$\sigma(f) := \{\lambda \in \mathbb{K} : \lambda \, \text{ist EW von } f\}$$

auch das **Spektrum von** f. So sind die Eigenwerte von Observablen wie unter (a) z.B. die Wellenlängen von Spektrallinien für eine selbstadjungierte Abbildung H, die ein Atom beschreibt.

c) So wie das Spektrum $\sigma(f) \subseteq \mathbb{C}$ einer unitären Abbildung $f : V \to V$ (d.h. f ist Isometrie) stets auf der Einheitskreislinie $S^1 \subseteq \mathbb{C}$ liegt (siehe (6.49 b)), unterliegt auch das Spektrum einer selbstadjungierten Abbildung einer starken Einschränkung:

6.55 Bemerkung: Eigenwerte selbstadjungierter Abbildungen

Ist $f : V \to V$ eine selbstadjungierte Abbildung eines unitären Vektorraums, so ist jeder Eigenwert von f reell.

Beweis:
Sei $\lambda \in \mathbb{C}$ Eigenwert und $v \in V \backslash \{0\}$ ein Eigenvektor von f bezüglich λ, dann gilt:

$$\lambda ||v||^2 = \langle v, \lambda v\rangle = \langle v, f(v)\rangle = \langle f(v), v\rangle = \langle \lambda v, v\rangle = \overline{\lambda}||v||^2$$

also wegen $||v||^2 \neq 0$ tatsächlich $\lambda = \overline{\lambda}$, d.h. $\lambda \in \mathbb{R}$.

□

6.56 Kommentar

Dass das charakteristische Polynom $p_f \in \mathbb{C}[T]$ eines Endomorphismus' $f \in \mathrm{End}(V)$ eines komplexen Vektorraumes V zerfällt, folgt aus dem Fundamentalsatz der Algebra. Dagegen ist die folgende Aussage durchaus überraschend:

6.57 Satz

Sei $(V, \langle -, -\rangle)$ ein euklidischer Vektorraum und $f : V \to V$ selbstadjungiert. Dann zerfällt das charakteristische Polynom $p_f \in \mathbb{R}[T]$ von f in Linearfaktoren.

Beweis:
Wir wählen eine ON-Basis $\mathfrak{a}$ von V und erhalten ein symmetrisches $A \in \mathrm{Sym}_n(\mathbb{R})$, $A = M(f; \mathfrak{a})$. Nun betrachten wir die komplex-lineare Abbildung $f_A : \mathbb{C}^n \to \mathbb{C}^n$, $z \mapsto Az$. Da $A = A^t = \overline{A}^t = A^*$ ist, ist $f_A^* = f_A$ bezüglich des kanonischen Skalarprodukts $\langle -, -\rangle$ auf $\mathbb{C}^n$. Da $\mathbb{C}$ algebraisch abgeschlossen ist (vgl. (5.29)), zerfällt nun $p_A \in \mathbb{C}[T]$ in Linearfaktoren:

$$p_A(T) = (T - \lambda_1) \cdot (T - \lambda_2) \ldots (T - \lambda_n) \ ,$$

mit $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. Wegen (6.55) ist aber jedes λ_k $(k = 1, \ldots, n)$ reell. Da schließlich für das charakteristische Polynom $p_f \in \mathbb{R}[T]$ von f gilt $p_f = p_A$, zerfällt p_f (über $\mathbb{R}$) tatsächlich in Linearfaktoren.

□

6.58 Kommentar

Auch für die Eigenvektoren selbstadjungierter Abbildungen $f : V \to V$ gelten Besonderheiten. Waren für einen beliebigen Endomorphismus die Eigenvektoren paarweise verschiedener Eigenwerte lediglich linear unabhängig (siehe (5.9)), so stehen sie im selbstadjungierten Fall sogar senkrecht aufeinander.

6.59 Bemerkung

Sei $(V, \langle -, - \rangle)$ euklidisch bzw. unitär und $f : V \to V$ selbstadjungiert. Sind dann $\lambda, \mu \in \mathbb{R}$ verschiedene Eigenwerte von f, $\lambda \neq \mu$, und $v, w \in V \backslash \{0\}$ Eigenvektoren zu λ bzw. μ, so gilt:

$$v \perp w \ , \ \text{d.h.} \ \langle v, w \rangle = 0 \ .$$

Beweis:
Es ist:

$$\begin{aligned} 0 &= \langle f(v), w \rangle - \langle v, f(w) \rangle = \langle \lambda v, w \rangle - \langle v, \mu w \rangle \\ &= (\overline{\lambda} - \mu) \langle v, w \rangle = (\lambda - \mu) \langle v, w \rangle \ , \end{aligned}$$

also $\langle v, w \rangle = 0$.

□

6.60 Theorem

Sei $(V, \langle -, - \rangle)$ euklidisch bzw. unitär und $f : V \to V$ selbstadjungiert. Dann existiert eine ON-Basis von V bestehend aus Eigenvektoren von f.

6.61 Kommentar: orthogonale Summe

a) f ist also diagonalisierbar! Der Vektorraum V zerfällt damit in die direkte Summe der Eigenräume (sogar **orthogonale Summe**) (vgl. (5.16)), d.h. sind $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ die paarweise verschiedenen Eigenwerte von f, so gilt:

$$V = \mathrm{Eig}(f; \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(f; \lambda_r) = \mathrm{Eig}(f; \lambda_1) \ominus \ldots \ominus \mathrm{Eig}(f; \lambda_r) \ .$$

b) Wir schreiben „$\ominus$“, wenn die Summe der Unterräume nicht nur direkt, sondern sogar orthogonal ist, d.h. $U_1 \ominus U_2 \subseteq V$ bedeutet, dass nicht nur $U_1 \oplus U_2$, sondern auch $U_1 \perp U_2$, also

$$\langle u_1, u_2 \rangle = 0 \ ,$$

für alle $u_1 \in U_1$, $u_2 \in U_2$.

6.62 Korollar

a) Ist $A \in \mathrm{Sym}_n(\mathbb{R})$, so existiert eine Diagonalmatrix $D \in \mathrm{Mat}_n(\mathbb{R})$ und eine orthogonale Matrix $S \in O(n)$ mit

$$S^{-1}AS = D\ .$$

b) Ist $A \in \mathrm{Herm}_n(\mathbb{C})$, so existiert eine reelle Diagonalmatrix $D \in \mathrm{Mat}_n(\mathbb{C})$ und eine unitäre Matrix $S \in U(n)$ mit

$$S^{-1}AS = D\ .$$

Beweis:
Da $f := f_A : \mathbb{K}^n \to \mathbb{K}^n$ (bzgl. des kanonischen Skalarprodukts auf $\mathbb{K}^n$) selbstadjungiert ist, existiert eine ON-Basis $(\widetilde{e}_1, \ldots, \widetilde{e}_n) = \mathfrak{a}$ von $(\mathbb{K}^n, \langle -, - \rangle)$ aus Eigenvektoren von f, also:

$$M(f; \mathfrak{a}) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = D$$

mit den Eigenwerten $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ (evtl. mehrfach aufgezählt). Ist $S \in GL_n(\mathbb{K})$ der Basiswechsel von $\mathfrak{a}$ nach $\mathfrak{K} = (e_1, \ldots, e_n)$ von $\mathbb{K}^n$, so ist S sogar in $O(n)$ bzw. $U(n)$, weil in den Spalten von S gerade die Vektoren $\widetilde{e}_1, \ldots, \widetilde{e}_n \in \mathbb{K}^n$ stehen (vgl. (6.50 e)). Nach dem Transformationssatz (3.31) ist also:

$$D = M(f; \mathfrak{a}, \mathfrak{a}) = S^{-1} \cdot M(f; \mathfrak{K}, \mathfrak{K}) \cdot S = S^{-1} \cdot A \cdot S\ .$$

6.63 Kommentar

Dass im reellen Fall das charakteristische Polynom einer symmetrischen Matrix A zerfällt, ist schon eine Überraschung. Dass A sogar diagonalisierbar ist umso mehr und dass diese Diagonalisierung sogar mit einer orthogonalen Matrix möglich ist, ist eine kleine Sensation.

Beweis von (6.60):
Induktion über $n = \dim V$:

- $n = 1$: Nehme irgendeinen Vektor $e \in V$ mit $||e|| = 1$.
- $n \to n+1$: Sei $\lambda \in \mathbb{R}$ ein Eigenwert von f (existiert nach (6.57) auch im reellen Fall) und $e_{n+1} \in V \backslash \{0\}$ ein Eigenvektor und sei ohne Einschränkung $||e_{n+1}|| = 1$ (über ein geeignetes Vielfaches). Wir betrachten nun

$$U := e_{n+1}^{\perp} = \{v \in V : \langle v, e_{n+1} \rangle = 0\}$$

(vgl. auch den Beweis von (6.31)). Entscheidend ist nun die Aussage $f(U) \subseteq U$, denn dann kann man die Induktionsvoraussetzung auf $f|U : U \to U$ anwenden (denn $\dim(U) = n$). Ist nämlich dann $(e_1, \ldots, e_n)$ eine ON-Basis von U bestehend aus Eigenvektoren von $f|U$, so ist $(e_1, \ldots, e_n, e_{n+1})$ eine ON-Basis von V bestehend aus Eigenvektoren von f. Sei also $u \in U$, also $\langle u, e_{n+1} \rangle = 0$. Da $f^* = f$ ist, ist dann auch

$$\langle f(u), e_{n+1} \rangle = \langle u, f(e_{n+1}) \rangle = \langle u, \lambda e_{n+1} \rangle = \lambda \langle u, e_{n+1} \rangle = 0 \ .$$

also auch $f(u) \in U$.

$\square$

6.64 Kommentar: Spektralsatz für selbstadjungierte Abbildungen

a) Theorem (6.60) zusammen mit (6.55) wird manchmal als **Spektralsatz für selbstadjungierte Abbildungen** bezeichnet, weil er Auskunft über das Spektrum σ (hier $\sigma \subseteq \mathbb{R}$) gibt, wie auch eine Aussage über die Lage der Eigenräume macht.

b) Auch für **unitäre Endomorphismen** (d.h. Isometrien von unitären Vektorräumen) gilt ein Spektralsatz: $\sigma(f) \subseteq S^1$ und es existiert eine ON-Basis aus Eigenvektoren von f (mit einem ganz ähnlichen Beweis wie für (6.60); Übung).

c) Eine entsprechende Aussage für **orthogonale Endomorphismen** (d.h. Endomorphismen euklidischer Vektorräume) ist im Allgemeinen falsch, weil im Allgemeinen bereits das charakteristische Polynom nicht zerfällt (z.B. das von $S(\varphi)$ für $\varphi \in (0, \pi)$ aus (6.50 a)). Der Normalformensatz für orthogonale Endomorphismen ist entsprechend komplizierter (siehe G. Fischer 5.5).

6.65 Korollar: Satz über Hauptachsentransformation

Sei $(V, \langle -, - \rangle)$ ein euklidischer bzw. unitärer Vektorraum und $s : V \times V \to \mathbb{K}$ eine (weitere) symmetrische Bilinearform bzw. Hermitesche Form, dann gibt es eine ON-Basis $\mathfrak{a}$ von V und $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, so dass gilt:

$$M(s; \mathfrak{a}) = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \ .$$

Beweis:
Da $\Phi : \operatorname{End}(V) \to \operatorname{Sesqui}(V)$, $f \mapsto s_f$, bijektiv ist (siehe (6.40)) gibt es also ein (eindeutig bestimmtes) $f \in \operatorname{End}(V)$ mit $s = s_f$, d.h.:

$$s(v, w) = \langle v, f(w) \rangle \ ,$$

für alle $v, w \in V$. Da s symmetrisch ist (bzw. Hermitesch), ist f selbstadjungiert, denn:

$$\langle f(v), w \rangle = \overline{\langle w, f(v) \rangle} = \overline{s(w, v)} = s(v, w) = \langle v, f(w) \rangle \ ,$$

für alle $v, w \in V$. Nach (6.60) gibt es $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ (die Eigenwerte von f mit Vielfachheiten) und eine ON-Basis $\mathfrak{a} = (e_1, \ldots, e_n)$ von V mit $f(e_j) = \lambda_j \cdot e_j$ $(j = 1, \ldots, n)$. Deshalb ist mit $A = (a_{ij}) := M(s; \mathfrak{a})$:

$$a_{ij} = s(e_i, e_j) = \langle e_i, f(e_j) \rangle = \langle e_i, \lambda_j e_j \rangle = \lambda_j \langle e_i, e_j \rangle = \lambda_j \cdot \delta_{ij}$$

also $M(s; \mathfrak{a}) = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

□

6.66 Kommentar

Die Zahlen $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ (und auch ihre Vielfachheiten) sind durch s eindeutig bestimmt, denn ist $\widetilde{\mathfrak{a}} = (\widetilde{e}_1, \ldots, \widetilde{e}_n)$ eine (weitere) ON-Basis für $(V, \langle -, - \rangle)$ mit $M(s; \mathfrak{a}) = \operatorname{diag}(\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n) =: \widetilde{D}$ für $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n \in \mathbb{R}$, so gilt für die Basiswechselmatrix $S \in GL_n(\mathbb{K})$ von $\widetilde{\mathfrak{a}}$ nach $\mathfrak{a}$:

$$\widetilde{D} = S^* \cdot D \cdot S$$

(wo $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, siehe (6.26)). Weil aber S zwischen zwei ON-Basen vermittelt, ist $S \in O(n)$ bzw. $S \in U(n)$, also $S^* = S^{-1}$. Deshalb ist auch

$$\widetilde{D} = S^{-1} \cdot D \cdot S$$

also sind dann $\widetilde{D}$ und D ähnlich, haben deshalb die selben Eigenwerte (und deren Vielfachheiten), welches gerade ihre Diagonaleinträge sind.

6.67 Beispiel: Hauptachsentransformation

Der Trägheitstensor eines starren Körpers $K \subseteq \mathbb{R}^3$ (mit Schwerpunkt in 0) ist gegeben durch eine symmetrische Bilinearform $s : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$. Nach (6.65) existiert eine ON-Basis $\mathfrak{a} = (\widetilde{e}_1, \widetilde{e}_2, \widetilde{e}_3)$ von $(\mathbb{R}^3, \langle -, - \rangle)$, so dass $M(s; \mathfrak{a}) = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$ mit $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ ist. $\mathbb{R}\widetilde{e}_1, \mathbb{R}\widetilde{e}_2, \mathbb{R}\widetilde{e}_3 \subseteq \mathbb{R}^3$ werden **Hauptachsen** von K genannt und $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ die **Hauptträgheitsmomente** von K. Die Transformation $S : \mathbb{R}^3 \to \mathbb{R}^3$, die die kanonische Basis $\mathfrak{K} = (e_1, e_2, e_3)$ nach $\mathfrak{a} = (\widetilde{e}_1, \widetilde{e}_2, \widetilde{e}_3)$ überführt, $Se_j = \widetilde{e}_j$ $(j = 1, 2, 3)$, ist eine **Hauptachsentransformation**. $\boxed{\text{Zeichnung fehlt}}$