Lineare Algebra für Elektrotechnik

Prof. Dr. Volker Bach

Technische Universität Braunschweig

Wintersemester 2025/26

Diese Vorlesung ist eine Einführung in die Lineare Algebra. Jedes Kapitel enthält am Schluss neben dem in der Vorlesung behandelten und somit prüfungsrelevanten Stoff dargestellt noch einen Abschnitt mit (nicht prüfungsrelevanten) Ergänzungen, d.h. Beweise und Beispiele, die wegen der Fülle des Materials in der Vorlesung nicht behandelt werden können.

Inhaltsverzeichnis

I. Grundlagen, Konventionen und Notationen

Dieses Kapitel stellt eine Übersicht über in der Mathematik häufig gebrauchte Begriffe, Konventionen und Notationen dar. Der Inhalt dieses Kapitels wird den Leser(inne)n größtenteils aus dem Schulunterricht geläufig sein. Die wenigen neu hinzukommenden Begriffe sind so leicht zu lernen, dass es den Leser(inne)n überlassen ist, sich im Laufe der ersten Woche diese anzueignen. In der Vorlesung *Lineare Algebra für Elektrotechnik* wird dieses Kapitel nur schlaglichtartig beleuchtet und wir[1] werden seinen Inhalt ohne weitere Erklärung benutzen.

I.1. Quantoren und Logik

- Eine Aussage im mathematischen Sinne ist ein Wortgebilde, dem man entweder den Wert <u>wahr</u> (w) oder <u>falsch</u> (f) zuordnen kann.
  - „*Es regnet jetzt*“, „*Braunschweig ist die schönste Stadt Deutschlands*“ oder „$1 \cdot 1 = 4$“ sind jeweils Aussagen,
  - „*Grün*“, „*Fünfzehn Mann auf des toten Mannes Kiste -Johoo, johoo, johoo- und 'ne Buddel mit Rum!*“ oder „$15x + 7y$“ sind jeweils keine Aussagen.

  Dabei geht es nur um die prinzipielle Zuordnung von w und f und weder um die praktische Nachprüfbarkeit der Aussage als Fakt noch um deren Objektivität noch um die Frage, ob sie eine Wertung beinhaltet.

- Das Zeichen $=$ bedeutet Gleichheit und ist in seiner Bedeutung evident. Das Zeichen $:=$ bedeutet, dass die linke Seite durch die rechte definiert wird, das Zeichen $=:$ bedeutet, dass die linke Seite durch die rechte (im Sinn einer Namensgebung) abgekürzt wird.

  <u>Beispiel:</u>

$$a \ := \ f(0), \quad f(1) \ =: \ b. \tag{I.1}$$

bedeutet

$$\textit{a wird als Wert der Funktion } f \textit{ bei } 0 \textit{ definiert, der Wert der Funktion } f \textit{ bei } 1 \textit{ wird hingegen mit } b \textit{ bezeichnet.} \tag{I.2}$$

- Das Zeichen $\forall$ bedeutet **für alle** (umgedrehtes A wie <u>A</u>lle).

[1] In mathematischen Texten wird meistens die 1. Person Plural verwendet – selbst wenn es sich nur um einen Autor handelt.

Beispiel:

$$\forall\, x, y > 0 : \quad x \cdot y > 0 \tag{I.3}$$

bedeutet

$$\textit{Für alle } x > 0 \textit{ und } y > 0 \textit{ gilt } x \cdot y > 0. \tag{I.4}$$

- Das Zeichen $\exists$ bedeutet **es existiert** (umgedrehtes E wie Existiert).

  Beispiel:

$$\forall\, x > 0\ \exists\, y < 0 : \quad x + y \;=\; 0 \tag{I.5}$$

bedeutet

Für jedes x größer Null existiert ein y kleiner Null, so dass $x + y = 0$ gilt. (Das gesuchte y ist natürlich $-x$.) (I.6)

- Man beachte, dass Quantoren im Allgemeinen nicht vertauscht werden dürfen. Dazu betrachten wir folgende Beispiele:

$$\forall\, n \in \mathbb{N}\ \exists\, m \in \mathbb{N} : \quad m > n$$

bedeutet

Zu jeder natürlichen Zahl n gibt es eine natürliche Zahl m so, dass m größer als n ist. (Diese Aussage ist wahr). (I.7)

$$\exists\, m \in \mathbb{N}\ \forall\, n \in \mathbb{N} : \quad m > n$$

bedeutet

Es gibt eine natürliche Zahl m, so dass alle natürlichen Zahlen n kleiner sind als m. (Diese Aussage ist falsch). (I.8)

- Das Zeichen $\Rightarrow$ bedeutet **impliziert**.

  Beispiel:

$$A \;\Rightarrow\; B \tag{I.9}$$

bedeutet

Aussage A impliziert Aussage B, d.h.: Ist A wahr, so ist auch B wahr. (I.10)

- Wie oben gesagt, kann eine mathematische Aussage A ein Satz, eine Bedingung oder auch eine Behauptung sein. In jedem Fall ist sie aber wahr oder falsch, $A \in \{w, f\}$.

  Beispiel:

$$A \;:=\; \textit{Es regnet.}\,, \quad B \;:=\; \textit{Die Erde wird nass.} \tag{I.11}$$

Dann gilt die Implikation $A \Rightarrow B$, was gelesen werden muss als

$$(A = w) \quad \Rightarrow \quad (B = w).$$

Dieses Beispiel wirkt etwas künstlich. Darum geben wir ein weiteres Beispiel, in dem die Aussagen von Platzhaltern abhängen:

$$A(x) := \begin{cases} w, & \text{falls } x \geq 5, \\ f, & \text{falls } x < 5, \end{cases} \qquad B(y) := \begin{cases} w, & \text{falls } y \geq 7, \\ f, & \text{falls } y < 7, \end{cases} \qquad C(z) := \begin{cases} w, & \text{falls } z \geq 33, \\ f, & \text{falls } z < 33, \end{cases}$$

dann gilt die folgende Implikation (s. (I.18)–(I.19)):

$$A(x) \wedge B(y) \;\Rightarrow\; C(x \cdot y). \tag{I.12}$$

- Das Zeichen $\Longleftrightarrow$ bedeutet **ist gleichwertig mit** oder **ist äquivalent zu**, d.h.

$$A \iff B \tag{I.13}$$

bedeutet

A ist genau dann wahr, wenn B wahr ist. (I.14)

- Das Zeichen $\vee$ ist ein logisches **oder**, d.h.

$$A \;\vee\; B \tag{I.15}$$

ist wahr, falls A oder B oder beide wahr sind und falsch, falls A und (gleichzeitig auch) B falsch sind. (I.16)

Beispiel:

$$\{\, x \cdot y = 0 \,\} \iff \{\, (x = 0) \;\vee\; (y = 0) \,\}. \tag{I.17}$$

- Das Zeichen $\wedge$ ist ein logisches **und**, d.h.

$$A \;\wedge\; B \tag{I.18}$$

ist wahr, falls A und (gleichzeitig auch) B wahr sind und sonst falsch. (I.19)

Beispiel 1:

$$\{\, (x = 0) \;\wedge\; (y = 0) \,\} \;\Rightarrow\; \{\, x + y = 0 \,\}. \tag{I.20}$$

Beispiel 2:

$$A \iff B \quad \text{ist gleichwertig mit} \quad (A \Rightarrow B) \wedge (B \Rightarrow A) \tag{I.21}$$

d.h. (um die Verwirrung komplett zu machen)

$$\Big(A \iff B\Big) \iff \Big[\big(A \Rightarrow B\big) \wedge \big(B \Rightarrow A\big)\Big]. \tag{I.22}$$

- Die logische **Negation** wird mit $\neg$ bezeichnet, also

$$\neg\, A \tag{I.23}$$

ist wahr, falls A falsch ist und falsch, falls A wahr ist. (I.24)

- Eine wichtige Beobachtung ist die *Kontraposition*, d.h. dass

$$(A \Rightarrow B) \quad \Longleftrightarrow \quad (\neg\, B \Rightarrow \neg\, A). \tag{I.25}$$

  Dies versteht man intuitiv sofort an folgendem Beispiel:

$$\begin{aligned} &\textit{Es regnet.} \Rightarrow \textit{Die Erde wird nass.} \\ &\quad \text{ist gleichwertig mit} \\ &\textit{Die Erde ist nicht nass.} \Rightarrow \textit{Es regnet nicht.} \end{aligned} \tag{I.26}$$

- Es ist nützlich, sich die Werte der Aussagen A, B, $A \Rightarrow B$, $A \Leftrightarrow B$, $A \vee B$, $A \wedge B$ und $\neg A$, in einer Wertetabelle zu verdeutlichen:

A	B	$A \Rightarrow B$	$A \Leftrightarrow B$	$A \vee B$	$A \wedge B$	$\neg A$
w	w	w	w	w	w	f
w	f	f	f	w	f	f
f	w	w	f	w	f	w
f	f	w	w	f	f	w

(I.27)

  Wir sehen, dass $[A \Rightarrow B] \Leftrightarrow [(\neg A) \vee B] \Leftrightarrow [\neg(A \wedge (\neg B))]$ gilt, da die Werte von $A \Rightarrow B$ und $(\neg A) \vee B$ für alle vier möglichen Werte des Paars $(A, B) \in \{(w,w), (w,f), (f,w), (f,f)\}$ übereinstimmen.

- Für logische Verknüpfungen gelten

$$\textbf{das Kommutativgesetz:} \tag{I.28}$$
$$\begin{aligned} A \vee B &= B \vee A, \\ A \wedge B &= B \wedge A, \end{aligned}$$

$$\textbf{das Assoziativgesetz:} \tag{I.29}$$
$$\begin{aligned} A \vee \big(B \vee C\big) &= \big(A \vee B\big) \vee C, \\ A \wedge \big(B \wedge C\big) &= \big(A \wedge B\big) \wedge C, \end{aligned}$$

$$\textbf{das Distributivgesetz:} \tag{I.30}$$
$$\begin{aligned} A \vee \big(B \wedge C\big) &= \big(A \vee B\big) \wedge \big(A \vee C\big), \\ A \wedge \big(B \vee C\big) &= \big(A \wedge B\big) \vee \big(A \wedge C\big). \end{aligned}$$

$$\textbf{sowie:} \tag{I.31}$$
$$\begin{aligned} \neg\,(A \vee B) &= (\neg A) \wedge (\neg B), \\ \neg\,(A \wedge B) &= (\neg A) \vee (\neg B). \end{aligned}$$

I.2. Mengen

Mengen sind (endliche, abzählbare oder sogar überabzählbare) Sammlungen mathematischer Objekte.

- $\{1, 5, 9\}$ ist die Menge, die die Zahlen 1,5 und 9 enthält,
- $\{x \in \mathbb{N} \mid x < 5\}$ enthält alle natürlichen Zahlen, die kleiner als 5 sind (also $\{x \in \mathbb{N} \mid x < 5\} = \{1, 2, 3, 4\}$).
- Jedes Element einer Menge wird nur einmal aufgeführt, beispielsweise ist $\{1, 5, 9, 9, 5\} = \{1, 5, 9\}$.
- $x \in M$ heißt x **ist Element der Menge** M.
- $x \notin M$ heißt x **ist nicht in der Menge** M **enthalten**.
- Die **Anzahl der Elemente** einer Menge M wird mit $|M|$ oder auch $\#[M]$ bezeichnet. Beispielsweise ist $|\{1, 5, 9\}| = 3$.
- $A \subseteq B$ bedeutet, dass die Menge A in der Menge B enthalten ist, also dass A **Teilmenge** von B ist. Umgekehrt heißt $A \supseteq B$, dass die Menge A die Menge B enthält:

$$\big(A \subseteq B\big) \iff \big(B \supseteq A\big) \iff (x \in A \Rightarrow x \in B). \tag{I.32}$$

- $\emptyset = \{\ \}$ ist die leere Menge, die kein Element enthält.
- Gleichheit von Mengen bedeutet elementweise Übereinstimmung,

$$\big(A = B\big) \iff \big(x \in A \iff x \in B\big). \tag{I.33}$$

- $\{x \mid E(x)\}$ und $\{x\}_{E(x)}$ bezeichnen die *Menge aller x, die die Eigenschaft $E(x)$ besitzen.*

  Beispiel:

$$U := \{n \mid \exists\, k \in \mathbb{N} :\ n = 2k - 1\} \tag{I.34}$$

  ist

$$\begin{array}{l}\textit{die Menge aller } n\textit{, für die es eine natürliche Zahl } k \textit{ gibt,}\\ \textit{so dass } n = 2k-1 \textit{ gilt}\end{array}, \tag{I.35}$$

  d.h.

$$U = \{2k - 1 \mid k \in \mathbb{N}\} = 2\mathbb{N} - 1 \tag{I.36}$$

  ist die Menge aller ungeraden Zahlen.

  Die Eigenschaft $E(x)$ kann auch durch eine **Indexmenge** $\mathcal{I}$ charakterisiert sein. Beispiel: Mit $\mathcal{I} := \{1, 3, 5, 7\}$ ist

$$\{x_i \mid i \in \mathcal{I}\} = \{x_i\}_{i \in \mathcal{I}} = \{x_1, x_3, x_5, x_7\}. \tag{I.37}$$

- Haben wir mehrere Mengen, etwa A_1, A_2 und A_3, so bildet $M = \{A_1, A_2, A_3\}$ wieder eine Menge – eine Menge von Mengen. Dies kann man so fortsetzen und kommt zu Mengen von Mengen von Mengen u.s.w. Der Übersichtlichkeit halber hat sich deshalb im Sprachgebrauch bewährt, die übergeordnete Menge M als **Familie, System, Kollektion** oder auch **Klasse** zu bezeichnen. Somit ist M *die Familie der Mengen A_1, A_2 und A_3.*

- Die Familie aller Teilmengen einer Menge M bezeichnet man als ihre **Potenzmenge** $\mathfrak{P}(M)$. Dabei zählen auch die leere Menge $\emptyset$ und M selbst als Teilmenge von M. Für $|M| < \infty$ ist $|\mathfrak{P}(M)| = 2^{|\{M\}|}$. (Warum?)
  Beispiel:
  $$M := \{1, 2, 3\} \quad \Rightarrow \quad \mathfrak{P}(M) = \big\{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\big\}. \tag{I.38}$$

- Die **Vereinigung**, der **Durchschnitt** und die **Differenz** zweier Mengen A, B werden wie folgt bezeichnet:
  $$\text{Vereinigung:} \quad A \cup B \;=\; \{x \mid (x \in A) \vee (x \in B)\}, \tag{I.39}$$
  $$\text{Durchschnitt:} \quad A \cap B \;=\; \{x \mid (x \in A) \wedge (x \in B)\}, \tag{I.40}$$
  $$\text{Differenz:} \quad A \setminus B \;=\; \{x \mid (x \in A) \wedge (x \notin B)\}. \tag{I.41}$$

- Ist A eine Teilmenge einer **Obermenge** M, d.h. $A \subseteq M$, so bezeichnet
  $$A^c \;:=\; M \setminus A \tag{I.42}$$
  das Komplement von A bezüglich M. (Vorsicht, der Notation A^c für das Komplement sieht man die Grundmenge M, auf die sie sich bezieht nicht mehr an!)

  Beispiel: Seien $A = \{1, 2, 3\}$, $B = \{3, 4, 5\}$, $M = \{1, 2, \ldots, 10\}$. Dann sind
  $$A \cup B \;=\; \{1, 2, 3, 4, 5\}, \qquad A \cap B = \{3\}, \tag{I.43}$$
  $$A \setminus B \;=\; \{1, 2\}, \qquad A^c = \{4, 5, 6, 7, 8, 9, 10\}. \tag{I.44}$$

- Vereinigungen und Durchschnitte können auch über Familien $\{A_i\}_{i \in \mathcal{I}}$ von Mengen A_i gebildet werden, wobei i eine Indexmenge $\mathcal{I}$ durchläuft.

  Beispiel:
  $$\bigcup_{i \in \mathcal{I}} A_i \;:=\; \big\{x \;\big|\; \exists\, i \in \mathcal{I} :\; x \in A_i\big\}, \tag{I.45}$$
  ist

  die Vereinigung der Mengen A_i, d.h. die x, die in (mindestens) einer Menge A_i mit $i \in \mathcal{I}$ enthalten sind; (I.46)

  $$\bigcap_{i \in \mathcal{I}} A_i \;:=\; \big\{x \;\big|\; \forall\, i \in \mathcal{I} :\; x \in A_i\big\}, \tag{I.47}$$
  ist

  der Durchschnitt der Mengen A_i, d.h. die x, die in allen Mengen A_i mit $i \in \mathcal{I}$ enthalten sind. (I.48)

- Für Vereinigung, Durchschnitt und Komplementbildung von Mengen gelten Kommutativ-, Assoziativ-, und Distributionsgesetze, analog zu den entsprechenden Gesetzen für die logischen Verknüpfungen $\vee$, $\wedge$ und $\neg$.

- Sind A und B zwei nichtleere Mengen, so bezeichnet

$$A \times B \;=\; \big\{(a,b) \;\big|\; a \in A\,,\; b \in B\big\} \tag{I.49}$$

das kartesische Produkt von A **und** B, d.h. die Menge aller Paare (a,b), die sich mit Elementen a aus A und b aus B bilden lässt. (I.50)

Allgemeiner ist

$$A_1 \times A_2 \times \cdots \times A_n = \big\{(a_1, a_2, \ldots, a_n) \,\big|\, a_1 \in A_1,\, \ldots,\, a_n \in A_n\big\} \tag{I.51}$$

das (n-fache) kartesische Produkt der Mengen $A_1, A_2, \ldots, A_n$, d.h. die Menge aller n*-Tupel* $(a_1, a_2, \ldots, a_n)$, die sich mit Elementen a_i aus A_i, für $i \in \{1, 2, \ldots, n\}$ bilden lässt. (I.52)

Vorsicht! Oft wird $A_1 \times A_2 \times \cdots \times A_n$ mit $A_1 \cup A_2 \cup \cdots \cup A_n$ verwechselt. Der Unterschied wird aber schon deutlich, wenn man die Zahl der Elemente für $A := A_1 = A_2$ betrachtet:

$$|A \times A| \;=\; \big(|A|\big)^2\,, \tag{I.53}$$

$$|A \cup A| \;=\; |A|\,. \tag{I.54}$$

- Sind $a_1, a_2, a_3, \ldots$ Zahlen, so bezeichnen wir $(a_n)_{n\in\mathbb{N}}$ als **Zahlenfolge**. Man beachte auch hier den Unterschied zwischen dem Tupel $(a_n)_{n\in\mathbb{N}}$ und der Menge $\{a_n\}_{n\in\mathbb{N}}$. So ist beispielsweise für die Zahlenfolge $a_1 = a_2 = a_3 = \ldots = 1$, die konstant gleich eins ist,

$$(a_n)_{n\in\mathbb{N}} \;=\; (1, 1, 1, \ldots), \tag{I.55}$$

aber

$$\{a_n\}_{n\in\mathbb{N}} \;=\; \{1, 1, 1, \ldots\} \;=\; \{1\}. \tag{I.56}$$

- Häufig wiederkehrende Mengen haben in der Mathematik eine eigene Bezeichnung bekommen. Wir listen die Symbole für die wichtigsten Zahlenmengen auf:

die **natürlichen Zahlen:** $\mathbb{N} \;:=\; \{1, 2, 3, \ldots\},$ (I.57)

die **natürlichen Zahlen mit Null:** $\mathbb{N}_0 \;:=\; \{0, 1, 2, 3, \ldots\},$ (I.58)

die **ganzen Zahlen:** $\mathbb{Z} \;:=\; \{0, 1, -1, 2, -2, 3, -3, \ldots\},$ (I.59)

die **rationalen Zahlen:** $\mathbb{Q} \;:=\; \left\{ \frac{p}{q} \;\middle|\; p \in \mathbb{Z},\; q \in \mathbb{N}\right\},$ (I.60)

die **reellen Zahlen:** $\mathbb{R}$ (I.61)

die **komplexen Zahlen:** $\mathbb{C}$ (I.62)

Die präzise Definition der reellen oder gar der komplexen Zahlen geht über den üblichen Schulstoff hinaus. Wir werden dies in den kommenden Wochen in der Vorlesung behandeln.

- Weiterhin führen wir für $m, n \in \mathbb{N}_0$ mit $m \leq n$ noch die Bezeichnung

$$\mathbb{Z}_m^n \;:=\; \{m, m+1, m+2, \ldots, n\} \tag{I.63}$$

für die natürlichen Zahlen zwischen m und n ein.

I.3. Ordnungsrelationen

Die Zeichen $<$, $>$, $\leq$, $\geq$ haben wir in verschiedenen Beispielen in den vorigen Abschnitten wie selbstverständlich benutzt.

- $a \leq b$ heißt a **ist kleiner als oder gleich** b.
- $a \geq b$ heißt a **ist größer als oder gleich** b.
- $a < b$ heißt a **ist kleiner als** b. Zur Unterscheidung dieser Relation von $a \leq b$ sagt man auch a **ist echt kleiner als** b oder a **ist strikt kleiner als** b.
- $a > b$ heißt a **ist (echt, strikt) größer als** b.
- Offenbar gilt

$$a < b \iff b > a, \tag{I.64}$$

$$a \leq b \iff (a < b) \vee (a = b), \tag{I.65}$$

$$a \geq b \iff (a > b) \vee (a = b). \tag{I.66}$$

Die Ordnungsrelation $<$ lässt sich aber auch auf andere Mengen, als den uns vertrauten Zahlen übertragen. Deshalb ist es zweckmäßig, den Begriff einer geordneten Menge präzise zu definieren.

Definition I.1. Eine Menge $S \neq \emptyset$ heißt **(total) geordnet** bezüglich "$<$" $\;:\Leftrightarrow$

(i) Sind $a, b \in S$, so gilt genau eine der drei Relationen $a < b$, $a = b$ oder $a > b$. (I.67)

(ii) Sind $a, b, c \in S$, und gilt $a < b$ und $b < c$, dann gilt auch $a < c$. (I.68)

Das Symbol "$<$" heißt **Ordnungsrelation** auf S.

Beispiele für total geordnete Mengen sind $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$. Auf den komplexen Zahlen gibt es keine (mit den Verknüpfungen verträgliche) Ordnungsrelation. Ebenso gibt es keine (mit den Verknüpfungen verträgliche) Ordnungsrelation auf den Vektoren in $\mathbb{R}^3$.

Mit Hilfe der Ordnungsrelation kann man **Intervalle** in $\mathbb{R}$ definieren. Seien $a, b \in \mathbb{R}$ und

$a \leq b$. Dann heißen

$$(a\,,\ b) \ := \ \{x \in \mathbb{R} \mid a < x < b\} \tag{I.69}$$

das **offene Intervall** von a nach b,

$$[a\,,\ b] \ := \ \{x \in \mathbb{R} \mid a \leq x \leq b\} \tag{I.70}$$

das **abgeschlossene Intervall** von a nach b,

$$[a\,,\ b) \ := \ \{x \in \mathbb{R} \mid a \leq x < b\} \tag{I.71}$$

das **rechts halboffene Intervall** von a nach b,

$$(a\,,\ b] \ := \ \{x \in \mathbb{R} \mid a < x \leq b\} \tag{I.72}$$

das **links halboffene Intervall** von a nach b

und insbesondere

$$\mathbb{R}^+ \ := \ \{x \in \mathbb{R} \mid x > 0\}, \qquad \mathbb{R}^- \ := \ \{x \in \mathbb{R} \mid x < 0\}, \tag{I.73}$$

$$\mathbb{R}_0^+ \ := \ \{x \in \mathbb{R} \mid x \geq 0\}, \qquad \mathbb{R}_0^- \ := \ \{x \in \mathbb{R} \mid x \leq 0\}. \tag{I.74}$$

I.4. Funktionen

Funktionen, auch **Abbildungen** genannt, sind die wichtigsten Objekte der Mathematik. Eine Funktion f ordnet jedem Element x seiner **Definitionsmenge** $\mathcal{D}$ genau ein Element $f(x)$ seiner **Wertemenge** $\mathcal{W}$ zu. Die symbolische Schreibweise dafür ist

$$f \ : \ \mathcal{D} \to \mathcal{W}, \quad x \mapsto f(x). \tag{I.75}$$

Dabei ist die Definitionsmenge zwar voll ausgeschöpft, denn $f(x)$ ist für jedes $x \in \mathcal{D}$ definiert. Für die Wertemenge muss das aber nicht der Fall sein. Sind $\mathcal{D}' \subseteq \mathcal{D}$ und $\mathcal{W}' \subseteq \mathcal{W}$ Teilmengen von $\mathcal{D}$ bzw. $\mathcal{W}$, so bezeichnen wir mit

$$f(\mathcal{D}') \ := \ \big\{f(x) \in \mathcal{W} \ \big| \ x \in \mathcal{D}'\big\} \tag{I.76}$$

die **Bildmenge** (oder das Bild) von $\mathcal{D}'$ und

$$f^{-1}(\mathcal{W}') \ := \ \big\{x \in \mathcal{D} \ \big| \ f(x) \in \mathcal{W}'\big\} \tag{I.77}$$

die **Urbildmenge** (oder das Urbild) von $\mathcal{W}'$.

Es kann also $f(\mathcal{D}) \subset \mathcal{W}$ durchaus eine *echte* Teilmenge des Wertebereichs sein. (Wir sollten hier aber erwähnen, dass diese Konvention nicht einheitlich akzeptiert ist. Manche Autoren verlangen, dass für $f : \mathcal{D} \to \mathcal{W}$ auch stets $f(\mathcal{D}) = \mathcal{W}$ gilt, andere fordern noch nicht einmal, dass $f^{-1}(\mathcal{W}) = \mathcal{D}$.)

Definition I.2. Seien $\mathcal{D}, \mathcal{W} \neq \emptyset$ und $f : \mathcal{D} \to \mathcal{W}$ eine Abbildung.

$$f \text{ heißt } \textbf{surjektiv} \quad :\Leftrightarrow \quad f(\mathcal{D}) \ = \ \mathcal{W}, \tag{I.78}$$

$$f \text{ heißt } \textbf{injektiv} \quad :\Leftrightarrow \quad \forall\, x, x' \in \mathcal{D} : \ \big(f(x) = f(x')\big) \ \Rightarrow \ \big(x = x'\big), \tag{I.79}$$

$$f \text{ heißt } \textbf{bijektiv} \quad :\Leftrightarrow \quad f \text{ ist surjektiv und injektiv.} \tag{I.80}$$

Bemerkungen und Beispiele.

- $\exp : \mathbb{R} \to \mathbb{R},\ x \mapsto e^x$ ist nicht surjektiv, (wegen $e^x > 0$) aber injektiv (wegen der strengen Monotonie der Exponentialfunktion).
- $\sin : \mathbb{R} \to [-1, 1],\ x \mapsto \sin x$ ist surjektiv aber nicht injektiv.
- $\tan : (-\pi/2\,,\ \pi/2) \to \mathbb{R},\ x \mapsto \tan x$ ist bijektiv.

Definition I.3. Für $g : A \to B$ und $f : g(A) \to C$ ist die **Verkettung** oder **Komposition** oder auch **Hintereinanderschaltung** $f \circ g$ von g und f wie folgt definiert:

$$f \circ g \ : \ A \to C, \quad x \mapsto f\big(g(x)\big). \tag{I.81}$$

Satz I.4. *Seien $g : A \to B$ und $f : B \to C$.*

(i) Sind f und g surjektiv, so ist auch $f \circ g : A \to C$ surjektiv.

(ii) Sind f und g injektiv, so ist auch $f \circ g : A \to C$ injektiv.

(iii) Sind f und g bijektiv, so ist auch $f \circ g : A \to C$ bijektiv.

Beweis.
Zu (i): Sei $c \in C$. Weil f surjektiv ist, gibt es ein $b \in B$ mit $f(b) = c$. Weil g surjektiv ist, gibt es ein $a \in A$ mit $g(a) = b$. Also gilt $(f \circ g)(a) = c$.

Zu (ii): Seien $a, a' \in A$ mit $(f \circ g)(a) = (f \circ g)(a')$. Weil f injektiv ist, folgt $g(a) = g(a')$. Weil g injektiv ist, folgt dann auch $a = a'$.

Zu (iii): Folgt aus (i) und (ii). □

Wichtig ist also zu beachten, dass der Definitionsbereich von f mit dem Bildbereich von g übereinstimmt. Man beachte auch die Reihenfolge: obwohl die Komposition $f \circ g$ heißt, wird erst g auf $x \in A$ angewandt und danach f auf das Ergebnis $g(x) \in B$.

Die Bedeutung der Bijektivität liegt darin, dass sie die Existenz und Eindeutigkeit der Umkehrfunktion sichert, wie der folgende Satz zeigt.

Satz I.5. *Seien A, B zwei nichtleere Mengen und $f : A \to B$ eine bijektive Abbildung. Dann gibt es eine eindeutige Abbildung $g : B \to A$ so, dass $g \circ f = \mathbb{1}_A$ und $f \circ g = \mathbb{1}_B$ gelten, d.h. dass*

$$\forall\, x \in A : \quad g[f(x)] = x \quad \textit{und} \quad \forall\, y \in B : \quad f[g(y)] = y \tag{I.82}$$

*gelten. In diesem Fall heißt $g : B \to A$ **Umkehrabbildung zu $\boldsymbol{f}$**, und wir schreiben $g =: f^{-1}$.*

Beweis. Sei $y \in B$. Weil f surjektiv ist, gibt es ein $x \in A$ so, dass $y = f(x)$, und weil f injektiv ist, ist x das einzige Element in A, für das $y = f(x)$ ist. Also definiert $g(y) := x$ eine Abbildung $g : B \to A$. Diese Abbildung hat die Eigenschaft, dass $f[g(y)] = f(x) = y$ für alle $y \in B$ gilt. Ist umgekehrt $x \in A$ beliebig, so setzen wir $y := f(x)$ und beobachten, dass $f(x) = y = f[g(y)] = f\big(g[f(x)]\big)$ gilt. Aus der Injektivität von f folgt nun jedoch $x = g[f(x)]$. □

Eine wichtige Klasse von Funktionen ist die der charakteristischen Funktionen, auch Indikatorfunktionen genannt. Ist $\mathcal{D}$ eine nichtleere Menge und $A \subseteq \mathcal{D}$ eine Teilmenge, so ist die **charakteristische Funktion** von A gegeben als

$$\mathbb{1}_A : \mathcal{D} \to \{0,1\}, \quad x \mapsto \begin{cases} 1 & \text{falls } x \in A, \\ 0 & \text{falls } x \notin A. \end{cases} \tag{I.83}$$

Mit anderen Worten: $\mathbb{1}_A[x]$ ist genau dann gleich 1, wenn x in A liegt und anderenfalls gleich 0.

I.5. Beweistechniken

I.5.1. Vollständige Induktion

Eine häufig verwendete Beweistechnik ist die **vollständige Induktion**. Zunächst stellen wir das Verfahren abstrakt vor. Nehmen wir an, wir wollten Aussagen $A(1), A(2), A(3), \ldots$ beweisen. Dann können wir folgende Tatsache verwenden.

Satz I.6. *Gibt es ein $n_0 \in \mathbb{N}$, so dass $A(n_0)$ wahr ist, und gilt die Implikation*

$$A(n) \quad \Rightarrow \quad A(n+1), \tag{I.84}$$

für jedes $n \geq n_0$, $n \in \mathbb{N}$, so ist $A(m)$ wahr, für jedes $m \geq n_0$, $m \in \mathbb{N}$.

Beweis. Wendet man (I.84) $(m - n_0)$-mal an, so erhält man

$$A(n_0) \Rightarrow A(n_0+1) \Rightarrow A(n_0+2) \Rightarrow \cdots \Rightarrow A(m-1) \Rightarrow A(m). \tag{I.85}$$

□

Der Beweis durch vollständige Induktion wird an einem Beispiel am deutlichsten. Wir wollen für $n \in \mathbb{N}$ die Summe

$$F(n) := 1 + 2 + \cdots + n \tag{I.86}$$

berechnen, und wir haben die Vermutung, dass $F(n) = G(n)$, wobei

$$G(n) := \frac{n(n+1)}{2}. \tag{I.87}$$

Nun gilt es, die Aussage

$$A(n) \;=\; w \quad :\Leftrightarrow \quad F(n) = G(n) \tag{I.88}$$

für alle $n \in \mathbb{N}$ zu beweisen.

- **Induktionsanfang**: Wähle $n_0 := 1$. Dann ist

  $$F(n_0) \;=\; F(1) \;=\; 1 \;=\; \frac{1(1+1)}{2} \;=\; G(1) \;=\; G(n_0), \tag{I.89}$$

  und $A(n_0) = A(1) = w$.

- **Induktionsannahme**: Seien $n \geq 1$ und gelte $A(n) = w$, also $F(n) = G(n)$.
- **Induktionsschritt**: Wir zeigen, dass aus $A(n) = w$ auch $A(n+1) = w$ folgt. Dazu beobachten wir, dass unter Verwendung von $F(n) = G(n)$ auch

$$\begin{aligned} F(n+1) &= n+1+F(n) \;=\; n+1+G(n) \;=\; n+1+\frac{n(n+1)}{2} \\ &= \frac{(n+1)(n+2)}{2} \;=\; G(n+1) \end{aligned} \tag{I.90}$$

  gilt, dass somit also $A(n+1) = w$ richtig ist.

Nach Satz I.6 ist damit $A(n) = w$ für alle $n \in \mathbb{N}$ bewiesen.

I.5.2. Beweis durch Kontraposition

Neben der vollständigen Induktion ist auch der Beweis durch Kontraposition eine häufig verwendete Methode, die auf (I.25) beruht,

$$(A \Rightarrow B) \;\Leftrightarrow\; (\neg B \Rightarrow \neg A). \tag{I.91}$$

Wir illustrieren dies wieder mit einem Beispiel: Seien die Aussagen A und B definiert durch

$$A \;=\; w \;:\Leftrightarrow\; \sqrt{2} \in \mathbb{Q}\,, \tag{I.92}$$

$$B \;=\; w \;:\Leftrightarrow\; \exists p, q \in \mathbb{N},\ \mathrm{ggT}(p,q) = 1: \quad 2q^2 \;=\; p^2 \tag{I.93}$$

wobei $\mathrm{ggT}(a,b) \in \mathbb{N}$ den größten gemeinsamen Teiler zweier natürlichen Zahlen $a, b \in \mathbb{N}$ bezeichne.

Ist nun $A = w$, so gibt es Zahlen $p', q' \in \mathbb{N}$ so, dass $\sqrt{2} = \frac{p'}{q'}$ gilt. Enthalten p' und q' einen gemeinsamen ganzzahligen Faktor $r \in \mathbb{N}$, sodass also $p' = pr$ und $q' = qr$ gelten, so können wir r herauskürzen und erhalten $\sqrt{2} = \frac{p}{q}$ mit teilerfremden p und q, d.h. $\mathrm{ggT}(p,q) = 1$. Damit ist auch $B = w$, und wir erhalten

$$A \Rightarrow B\,. \tag{I.94}$$

Die Aussage B ist jedoch stets falsch, weil dann der Primfaktor 2 in $2q^2$ in ungerader Anzahl und in p^2 in gerader Anzahl auftreten müsste.

$$\{B = f\} \Rightarrow \{\neg B = w\} \Rightarrow \{\neg A = w\} \Rightarrow \{\sqrt{2} \;\notin\; \mathbb{Q}\}\,. \tag{I.95}$$

Wir bemerken, dass der Beweis durch Kontraposition sehr ähnlich zur Methode des *Widerspruchsbeweises* ist, letzterer beruht auf $[A \Rightarrow B] = (\neg A) \vee B = \neg[A \wedge (\neg B)]$.

I.6. Notationen

Seien $m, n \in \mathbb{Z}$ mit $m \leq n$ und

$$A \;=\; \{a_m, a_{m+1}, a_{m+2}, \ldots, a_n\} \tag{I.96}$$

eine Menge von Zahlen. Dann ist das Summenzeichen wie folgt definiert,

$$\sum_{i=m}^{n} a_i \;:=\; a_m + a_{m+1} + a_{m+2} + \ldots + a_n. \tag{I.97}$$

Wir bemerken, dass der *Summationsindex* i durch irgend einen anderen Buchstaben außer m oder n ersetzt werden kann,

$$\sum_{i=m}^{n} a_i \;=\; \sum_{k=m}^{n} a_k \;=\; \sum_{j=m}^{n} a_j. \tag{I.98}$$

Für $m > n$ wird $\sum_{i=m}^{n} a_i \;:=\; 0$ definiert.
Mit $\mathcal{I} := \{m, m+1, \ldots, n\}$ und A wird die Summe auch oft noch anders geschrieben:

$$\sum_{i=m}^{n} a_i \;=\; \sum_{i\in\mathcal{I}} a_i \;=\; \sum_{a\in A} a. \tag{I.99}$$

I.7. Ergänzungen

I.7.1. Äquivalenzrelationen

Häufig lässt sich eine Menge in eine Familie disjunkter Teilmengen zerlegen, deren Elemente jeweils ähnliche Eigenschaften haben. Beispiel:
Wir zerlegen $\mathbb{Z}$ in $\mathbb{Z} = A_0 \cup A_1 \cup A_2$, wobei

$$A_0 := 3\mathbb{Z} \;=\; \{3k \mid k \in \mathbb{Z}\}, \tag{I.100}$$

$$A_1 := 3\mathbb{Z} + 1 \;=\; \{3k + 1 \mid k \in \mathbb{Z}\}, \tag{I.101}$$

$$A_2 := 3\mathbb{Z} + 2 \;=\; \{3k + 2 \mid k \in \mathbb{Z}\}. \tag{I.102}$$

Offensichtlich sind $A_0 \cap A_1 = A_1 \cap A_2 = A_0 \cap A_2 = \emptyset$. Die Elemente in A_j $(j = 0, 1, 2)$ lassen sich dadurch charakterisieren, dass sie einen Rest j beim Teilen durch 3 ergeben.

Wir formalisieren nun diese Überlegungen.

Definition I.7. Sei A eine Menge. Eine Abbildung $R : A \times A \to \{w, f\}$ heißt **Relation auf** A. Für $R(a, b) = w$ schreiben wir auch $a \sim b$.

Definition I.8. Eine Relation $R : A \times A \to \{w, f\}$ auf einer Menge A, mit $R(a, b) = w \Leftrightarrow: a \sim b$ heißt **Äquivalenzrelation**, falls folgende drei Eigenschaften gelten:

$$\textbf{Reflexivität} \quad \forall\, a \in A: \qquad a \sim a, \tag{I.103}$$

$$\textbf{Symmetrie} \quad \forall\, a, b \in A: \qquad a \sim b \Leftrightarrow b \sim a, \tag{I.104}$$

$$\textbf{Transitivität} \quad \forall\, a, b, c \in A: \quad (a \sim b \wedge b \sim c) \Rightarrow (a \sim c). \tag{I.105}$$

Satz I.9. *Eine Äquivalenzrelation auf einer Menge A bewirkt eine Zerlegung von A in disjunkte Teilmengen. Dabei sind zwei Elemente aus A genau dann äquivalent, wenn sie derselben Teilmenge angehören.*

Beweis. Zu $a \in A$ definieren wir

$$[a]_\sim := \{x \in A \mid a \sim x\}. \tag{I.106}$$

Wegen $a \in [a]_\sim$ ist $[a]_\sim$ nicht leer. Wir zeigen nun für $a, b \in A$, dass

$$\text{entweder} \quad [a]_\sim \;\cap\; [b]_\sim \;=\; \emptyset \tag{I.107}$$

$$\text{oder} \quad [a]_\sim \;=\; [b]_\sim \tag{I.108}$$

gilt. (Wegen $[a]_\sim \neq \emptyset$ können (I.107) und (I.108) nicht gleichzeitig gelten.)
Sei $[a]_\sim \cap [b]_\sim \neq \emptyset$. Dann gibt es also ein gemeinsames Element $c \in [a]_\sim$, $c \in [b]_\sim$. Damit gelten $a \sim c$ und $c \sim b$, also auch $a \sim b$. Ist nun $x \in [a]_\sim$, dann gilt $x \sim a$ und mit $a \sim b$ auch $x \sim b$, also $x \in [b]_\sim$. Es folgt, dass $[a]_\sim \subseteq [b]_\sim$. Genauso erhält man $[b]_\sim \subseteq [a]_\sim$, also $[a]_\sim = [b]_\sim$. Damit ist (I.107)–(I.108) gezeigt.
Schreiben wir jetzt

$$A \;=\; \bigcup_{a \in A} [a]_\sim, \tag{I.109}$$

folgt die Aussage unmittelbar durch Zusammenfassen gleicher $[a]_\sim$. □

Definition I.10. Die Teilmengen $[a]_\sim$ heißen **Äquivalenzklassen**. Die Familie der Äquivalenzklassen bezeichnet man mit

$$A/\sim \quad (\text{sprich: }''A \text{ modulo } \sim ''). \tag{I.110}$$

Liegt a in einer Äquivalenzklasse, so heißt a **Repräsentant der Klasse**.

Bemerkungen und Beispiele.
Sind $A := \mathbb{Z}$ die ganzen Zahlen und $p \in \mathbb{N}$ eine natürliche Zahl, so sind

$$m \sim n \;:\Leftrightarrow\; \exists\, k \in \mathbb{Z} : \quad m - n \;=\; kp. \tag{I.111}$$

$$\Rightarrow\; \mathbb{Z} \;=\; [0]_\sim \cup [1]_\sim \cup \ldots \cup [p-1]_\sim, \tag{I.112}$$

$$[j]_\sim \;=\; \{kp + j \mid k \in \mathbb{Z}\}. \tag{I.113}$$

$\mathbb{Z}/\sim$ bezeichnet man auch mit $\mathbb{Z}/p\mathbb{Z}$ oder $\mathbb{Z}_p$ und $[j]_\sim =: [j]_{\bmod p}$.

Definition I.11. Sei $\sim$ eine Äquivalenzrelation auf einer Menge A. Eine Teilmenge $S \subseteq A$ heißt ein **vollständiges Repräsentantensystem zu** $\sim$, falls folgende zwei Eigenschaften gelten:

(i) Jedes Element aus A ist zu einem Element aus S äquivalent.

(ii) Die Elemente aus S sind paarweise nicht äquivalent.

Bemerkungen und Beispiele.

$$A \;:=\; \big\{ g \subseteq \mathbb{R}^2 \;\big|\; g \text{ ist eine Gerade} \big\}, \tag{I.114}$$

$$g_1 \sim g_2 :\Leftrightarrow\; g_1 \text{ und } g_2 \text{ sind parallel.}$$

$$\Rightarrow\; S \;=\; \big\{ g \in A \;\big|\; g \cap \{\vec{0}\} = \{\vec{0}\} \big\} \tag{I.115}$$

$$\text{ist ein vollständiges Repräsentantensystem.} \tag{I.116}$$

I.7.2. Das griechische Alphabet

Das griechische Alphabet wird in der Mathematik häufig verwendet. Zum Abschluss geben wir noch eine Liste der gebräuchlichsten griechischen Buchstaben:

					KLEIN				
α	alpha	β	beta	γ	gamma	δ	delta	ϵ	epsilon
ε	epsilon	ζ	zeta	η	eta	θ	theta	ϑ	theta
ι	jota	κ	kappa	λ	lambda	μ	mü	ν	nü
ξ	xi	o	o	π	pi	φ	phi	ρ	rho
ϱ	rho	σ	sigma	ς	sigma	τ	tau	υ	upsilon
ϕ	phi	φ	phi	χ	chi	ψ	psi	ω	omega
					GROSS				
Γ	Gamma	Δ	Delta	Θ	Theta	Λ	Lambda	Ξ	Xi
Π	Pi	Σ	Sigma	Υ	Upsilon	Φ	Phi	Ψ	Psi
Ω	Omega								

II. Gruppen, Ringe und Körper

Im vorigen Kapitel I wurden die wichtigsten Zahlenmengen bereits genannt:
die *natürlichen* Zahlen,

$$\mathbb{N} \ := \ \{1, 2, 3, \ldots\}, \tag{II.1}$$

die *ganzen* Zahlen,

$$\mathbb{Z} \ := \ \{0, +1, -1, +2, -2, \ldots\}, \tag{II.2}$$

die *rationalen* Zahlen,

$$\mathbb{Q} \ := \ \left\{ \frac{p}{q} \ \middle| \ p \in \mathbb{Z}, q \in \mathbb{N} \right\}, \tag{II.3}$$

sowie die *reellen* und die *komplexen* Zahlen,

$$\mathbb{R} \quad \text{und} \quad \mathbb{C}. \tag{II.4}$$

Wir wenden uns zunächst $\mathbb{N}$, $\mathbb{Z}$ und $\mathbb{Q}$ zu. Für $a, b \in \mathbb{N}$ ist auch $a+b \in \mathbb{N}$. Diese Tatsache bezeichnet man als *Abgeschlossenheit* oder *Stabilität* von $\mathbb{N}$ bezüglich Addition.

I.A. gilt $a - b \in \mathbb{N}$ jedoch nicht. Dafür geht man von $\mathbb{N}$ zu $\mathbb{Z}$ über; für $a, b \in \mathbb{Z}$ sind $a + b$ und $a - b \in \mathbb{Z}$. Insbesondere ist 0 das *neutrale Element* bezüglich Addition in $\mathbb{Z}$: $a + 0 = 0 + a = a$. Man sagt, dass $\mathbb{Z}$ bezüglich der Addition $+$ eine *Gruppe* bildet.

Weiterhin ist $\mathbb{Z}$ auch bezüglich Multiplikation abgeschlossen, d.h. für $a, b \in \mathbb{Z}$ ist auch $a \cdot b \in \mathbb{Z}$, und es gilt das Distributivgesetz, $a(b + c) = ab + bc$. Somit ist $\mathbb{Z}$ bezüglich der Addition $+$ und der Multiplikation $(\cdot)$ ein *Ring*.

Schließlich gelangt man von $\mathbb{Z}$ zu $\mathbb{Q}$ durch die Forderung, dass auch Abgeschlossenheit bezüglich Division gelten soll: Für $a, b \in \mathbb{Q}$ sind $a + b, a - b, a \cdot b \in \mathbb{Q}$ und $\frac{a}{b} \in \mathbb{Q}$, falls $b \neq 0$. Diese Eigenschaften von $\mathbb{Q}$ stehen auch exemplarisch für die allgemeine Definition eines *Körpers*.

II.1. Gruppen

Definition II.1. Eine Menge G heißt **Gruppe** $:\Leftrightarrow$
Auf G ist eine Verknüpfung $\circ : G \times G \to G$ definiert, die die folgenden Eigenschaften besitzt:

$$(G_1) \quad \forall\, a, b, c \in G : \qquad (a \circ b) \circ c \ = \ a \circ (b \circ c), \tag{II.5}$$

$$(G_2) \quad \exists\, e \in G\ \forall\, a \in G : \qquad a \circ e \ = \ e \circ a \ = \ a, \tag{II.6}$$

$$(G_3) \quad \forall\, a \in G\ \exists\, a^{-1} \in G : \quad a \circ a^{-1} \ = \ a^{-1} \circ a \ = \ e. \tag{II.7}$$

Dabei bezeichnet man (G_1) als **Assoziativität**, e als das **neutrale Element** und a^{-1} als das **zu a inverse Element**. Die Anzahl $|G|$ der Elemente in G bezeichnet man als **Ordnung von G**.
Gilt außerdem noch

$$(G_4) \quad \forall\, a, b \in G : \quad a \circ b \;=\; b \circ a \quad \text{(Kommutativität)}, \tag{II.8}$$

so nennt man G **kommutativ** oder **abelsch**.

Bemerkungen und Beispiele.

- Häufig wird das Verknüpfungszeichen weg gelassen, und man schreibt

$$a \circ b \;=:\; a\, b. \tag{II.9}$$

- Die Assoziativität erlaubt es uns, Klammern bei der Gruppenverknüpfung einfach wegzulassen oder bei Bedarf einzufügen,

$$(a\, b)\, c \;=\; a\, (b\, c) \;=:\; a\, b\, c. \tag{II.10}$$

- Das neutrale Element e einer Gruppe ist eindeutig. Ist nämlich e' irgendein (möglicherweise von e verschiedenes) Element von G, das die Eigenschaft (G_2) besitzt, so folgt $e' = e\, e' = e$.
- Sind G eine Gruppe, $a \in G$ und $b \in G$ ein (möglicherweise von a^{-1} verschiedenes) zu a inverses Element, also $ab = ba = e$, so folgt dass $b = (a^{-1}\, a)b = a^{-1}(ab) = a^{-1}$, und das zu a inverse Element, a^{-1}, ist eindeutig.
- Aus der Eindeutigkeit des inversen Elements folgen dann auch

$$(a^{-1})^{-1} \;=\; a \qquad \text{und} \qquad (ab)^{-1} \;=\; b^{-1}a^{-1}, \tag{II.11}$$

  letzteres wegen $a\, b\, b^{-1}\, a^{-1} \;=\; a\, e\, a^{-1} \;=\; a\, a^{-1} \;=\; e$.
- Für abelsche Gruppen G schreibt man die Verknüpfung "$\circ$" häufig als Addition $+ : G \times G \to G$, und die Eigenschaften (G_1)–(G_4) nehmen folgende Gestalt an:

$$(\widetilde{G}_1) \quad \forall\, a, b, c \in G : \qquad (a + b) + c \;=\; a + (b + c), \tag{II.12}$$

$$(\widetilde{G}_2) \quad \exists\, 0 \in G\, \forall\, a \in G : \qquad a + 0 \;=\; 0 + a \;=\; a, \tag{II.13}$$

$$(\widetilde{G}_3) \quad \forall\, a \in G\, \exists\, -a \in G : \quad a + (-a) \;=\; (-a) + a \;=\; 0, \tag{II.14}$$

$$(\widetilde{G}_4) \quad \forall\, a, b \in G : \qquad a + b \;=\; b + a. \tag{II.15}$$

- Die Menge $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ der ganzen Zahlen ist bezüglich der Addition eine abelsche Gruppe mit 0 als neutralem Element und $k^{-1} = -k$ als zu $k \in \mathbb{Z}$ inversem Element.
- Die Menge $\mathbb{Q} \setminus \{0\}$ der rationalen Zahlen ohne Null ist bezüglich der Multiplikation eine abelsche Gruppe mit 1 als neutralem Element und $q^{-1} = 1/q$ als zu $q \in \mathbb{Q} \setminus \{0\}$ inversem Element.
- Die Menge $G := \mathbb{R} \setminus \{1\}$ bildet bezüglich $a \circ b := a + b - ab$ eine Gruppe.

II.1.1. Permutationen

Definition II.2. Zu vorgegebenem $n \in \mathbb{N}$ sei $\mathbb{Z}_1^n := \{1, 2, \ldots, n\}$. Die Menge der **Permutationen** von n Elementen ist durch

$$\mathcal{S}_n \ := \ \left\{ \pi : \mathbb{Z}_1^n \to \mathbb{Z}_1^n \;\middle|\; \pi \text{ ist bijektiv} \right\} \tag{II.16}$$

gegeben. Das **Signum** $(-1)^\pi \in \{-1, 1\}$ von π ist definiert durch

$$(-1)^\pi \ := \ \frac{\prod_{1 \le i < j \le n} \pi(i) - \pi(j)}{\prod_{1 \le i < j \le n} i - j}. \tag{II.17}$$

Die Permutationen schreibt man auch häufig als Schema

$$\pi \ \equiv \ \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}. \tag{II.18}$$

Die Komposition von Bijektionen ist nach Satz I.4 (iii) stets selbst bijektiv. Somit bildet die Komposition zweier Permutationen eine Verknüpfung $\circ : \mathcal{S}_n \times \mathcal{S}_n \to \mathcal{S}_n$. Die Komposition von Abbildungen ist stets assoziativ, deshalb gilt auch (G_1). Weiterhin agiert, wie in (G_2) gefordert,

$$\mathbb{1}_{\mathbb{Z}_1^n} \ \equiv \ \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix} \ \in \ \mathcal{S}_n \tag{II.19}$$

als neutrales Element bezüglich $\circ$. Schließlich ist mit $\pi \in \mathcal{S}_n$ auch die Umkehrabbildung $\pi^{-1} \in \mathcal{S}_n$ eine Permutation, und es gilt $\pi^{-1} \circ \pi = \pi \circ \pi^{-1} = \mathbb{1}_{\mathbb{Z}_1^n}$, also (G_3). Zusammenfassend stellen wir fest, dass die Permutationen $\mathcal{S}_n$ bezüglich der Komposition $\circ$ eine Gruppe bilden. Dabei ist die Ordnung gleich $|\mathcal{S}_n| = n!$.

Bemerkungen und Beispiele.

- Beachte, dass

$$\prod_{1 \le i < j \le n} F(i, j) \ = \ \prod_{i=1}^{n} \left(\prod_{j=i+1}^{n} F(i, j) \right). \tag{II.20}$$

- Für $n = 2$ sind $|\mathcal{S}_2| = 2! = 2$ und

$$\mathcal{S}_2 \ = \ \left\{ \mathbb{1} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \ \ \sigma = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\}, \tag{II.21}$$

  mit $(-1)^{\mathbb{1}} = +1$ und $(-1)^\sigma = -1$.

- Für $n = 3$ sind $|\mathcal{S}_3| = 3! = 6$ und

$$\mathcal{S}_3 \ = \ \left\{ \pi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \ \ \pi_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \ \ \pi_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \right. \tag{II.22}$$

$$\left. \pi_4 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \ \ \pi_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \ \ \pi_6 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \right\}, \tag{II.23}$$

  mit $(-1)^{\pi_1} = (-1)^{\pi_2} = (-1)^{\pi_3} = +1$ und $(-1)^{\pi_4} = (-1)^{\pi_5} = (-1)^{\pi_6} = -1$.

- Für $n = 3$ sind $\pi_5(1) = 3$, $\pi_5(2) = 2$, $\pi_5(3) = 1$ und somit

$$\prod_{1 \le i < j \le n} (i - j) \;=\; (1-2) \cdot (1-3) \cdot (2-3) \;=\; (-1) \cdot (-2) \cdot (-1) \;=\; (-2) \tag{II.24}$$

und

$$\begin{aligned} \prod_{1 \le i < j \le n} \big(\pi_5(i) - \pi_5(j)\big) &= \big(\pi_5(1) - \pi_5(2)\big) \cdot \big(\pi_5(1) - \pi_5(3)\big) \cdot \big(\pi_5(2) - \pi_5(3)\big) \\ &= (3-2) \cdot (3-1) \cdot (2-1) \;=\; 1 \cdot 2 \cdot 1 \;=\; 2. \end{aligned} \tag{II.25}$$

Also ist in der Tat

$$(-1)^{\pi_5} \;:=\; \frac{\prod_{1 \le i < j \le n} \big(\pi_5(i) - \pi_5(j)\big)}{\prod_{1 \le i < j \le n} (i - j)} \;=\; \frac{2}{-2} \;=\; -1. \tag{II.26}$$

- Für $n = 3$ sind $\pi_6(1) = 2$, $\pi_6(2) = 1$, $\pi_6(3) = 3$ sowie $\pi_2(1) = 3$, $\pi_2(2) = 1$, $\pi_2(3) = 2$. Also sind

$$[\pi_6 \circ \pi_2](1) \;= \pi_6[\pi_2(1)] \;=\; \pi_6[3] \;=\; 3, \tag{II.27}$$
$$[\pi_6 \circ \pi_2](2) \;= \pi_6[\pi_2(2)] \;=\; \pi_6[1] \;=\; 2, \tag{II.28}$$
$$[\pi_6 \circ \pi_2](3) \;= \pi_6[\pi_2(3)] \;=\; \pi_6[2] \;=\; 1, \tag{II.29}$$

und dementsprechend ist

$$\pi_6 \circ \pi_2 \;=\; \begin{pmatrix} 1 & 2 & 3 \\ [\pi_6 \circ \pi_2](1) & [\pi_6 \circ \pi_2](2) & [\pi_6 \circ \pi_2](3) \end{pmatrix} \;=\; \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \;=\; \pi_5. \tag{II.30}$$

- Analog erhält man für $n = 3$

$$\pi_2 \circ \pi_6 \;=\; \pi_4 \;\neq\; \pi_5 \;=\; \pi_6 \circ \pi_2, \tag{II.31}$$

was zeigt, dass die Gruppe $\mathcal{S}_3$ der Permutationen nicht kommutativ ist. Dies gilt in der Tat ganz allgemein für alle $\mathcal{S}_n$ mit $n \ge 3$. (Trivialerweise ist $\mathcal{S}_2$ kommutativ.)

- Das Signum von Permutationen ist multiplikativ, d.h. für $n \in \mathbb{N}$ und $\pi, \kappa \in \mathcal{S}_n$ gilt

$$(-1)^{\pi \circ \kappa} \;=\; (-1)^{\pi} \cdot (-1)^{\kappa}. \tag{II.32}$$

Dies wird in Lemma II.12 bewiesen.

- Die Permutationen der Form

$$\sigma \;=\; \begin{pmatrix} 1 & \ldots & i-1 & \mathbf{i} & i+1 & \ldots & j-1 & \mathbf{j} & j+1 & \ldots & n \\ 1 & \ldots & i-1 & \mathbf{j} & i+1 & \ldots & j-1 & \mathbf{i} & j+1 & \ldots & n \end{pmatrix} \;\in\; \mathcal{S}_n, \tag{II.33}$$

die nur zwei Elemente (hier: $i \leftrightarrow j$) gegeneinander austauschen, heißen **Transpositionen**. Gemäß Lemma II.11 gilt stets $(-1)^{\sigma} = -1$.

- Jede Permutation $\pi \in \mathcal{S}_n$ lässt sich als Komposition von Transpositionen schreiben, d. h. es gibt Transpositionen $\sigma_1, \ldots, \sigma_m \in \mathcal{S}_n$, so dass
$$\pi \;=\; \sigma_1 \circ \sigma_2 \circ \ldots \circ \sigma_m. \tag{II.34}$$
Gemäß (II.32) und (II.62) gilt also
$$(-1)^\pi \;=\; (-1)^{\sigma_1} \cdot (-1)^{\sigma_2} \cdots (-1)^{\sigma_m} \;=\; (-1)^m, \tag{II.35}$$
d.h. das Signum der Permutation π ist -1 hoch die Anzahl der Transpositionen, deren Komposition π ergibt.

Weitere Details zu Permutationen findet man in Abschnitt II.4.2.

II.2. Ringe

Definition II.3. Eine Menge R heißt **Ring** :⇔
Auf R sind zwei Verknüpfungen *Addition* $+ : R \times R \to R$ und *Multiplikation* $(\cdot) : R \times R \to R$ definiert, die die folgenden Eigenschaften besitzen:

$$(R_1) \qquad R \text{ ist bezüglich der Addition } + \text{ eine abelsche Gruppe,} \tag{II.36}$$
$$(R_2) \quad \forall a, b, c \in R: \qquad (a \cdot b) \cdot c \;=\; a \cdot (b \cdot c), \tag{II.37}$$
$$(R_3) \quad \forall a, b, c \in R: \quad a \cdot (b + c) \;=\; a \cdot b + a \cdot c, \quad (b + c) \cdot a \;=\; b \cdot a + c \cdot a. \tag{II.38}$$

Dabei bezeichnet man (R_3) als **Distributivität** und vereinbart, dass Multiplikation vor Addition ausgeführt wird (Punktrechnung vor Strichrechnung).

Bemerkungen und Beispiele.

- Die Menge $\mathbb{N} := \{1, 2, 3, \ldots\}$ der natürlichen Zahlen ist kein Ring.
- Die Menge $\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ der ganzen Zahlen bildet einen Ring.
- Die Menge $2\mathbb{Z} := \{\ldots, -4, -2, 0, 2, 4, \ldots\}$ der geraden Zahlen bildet einen Ring.
- Die Menge $2\mathbb{Z} + 1 := \{\ldots, -5, -3, -1, 1, 3, 5, \ldots\}$ der ungeraden Zahlen ist gegenüber Addition nicht abgeschlossen und deswegen auch kein Ring.
- $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ sind Ringe.

II.2.1. Die Restklassenringe $\mathbb{Z}_p$ von $\mathbb{Z}$ modulo p

Für $p \in \mathbb{N}$ definieren wir die **Restklassen modulo p** durch
$$\forall\, k \in \mathbb{Z}: \qquad [k]_p \;:=\; k + p\mathbb{Z} \;=\; \big\{ k + pn \;\big|\; n \in \mathbb{Z} \big\}. \tag{II.39}$$
Wir beobachten, dass $[k]_p = [\ell]_p$ gleichwertig mit $(k - \ell) \in p\mathbb{Z} = \{pn | n \in \mathbb{Z}\}$ ist. Offensichtlich gibt es genau p solcher Restklassen modulo p. Ihre Menge bezeichnen wir mit
$$\mathbb{Z}_p \;:=\; \Big\{ [0]_p,\; [1]_p,\; [2]_p,\; \ldots,\; [p-1]_p \Big\}, \tag{II.40}$$

sie bilden eine paarweise disjunkte Zerlegung der ganzen Zahlen, $\mathbb{Z} = \bigcup_{K \in \mathbb{Z}_p} K$. Definieren wir Addition und Multiplikation auf $\mathbb{Z}_p$ durch

$$[k]_p + [\ell]_p \;:=\; [k+\ell]_p \qquad \text{und} \qquad [k]_p \cdot [\ell]_p \;:=\; [k \cdot \ell]_p, \tag{II.41}$$

so bilden die Restklassen $\mathbb{Z}_p$ modulo p mit den Verknüpfungen in (II.41) einen Ring, den **Restklassenring modulo p**.

Um einzusehen, dass $\mathbb{Z}_p$ einen Ring bildet, braucht man natürlich nur die Ringaxiome nachzuprüfen, was eine reine Fleißaufgabe ist. Eine Subtilität liegt allerdings im Beweis der *Wohldefiniertheit* der Verknüpfungen in (II.41): Damit (II.41) überhaupt Verknüpfungen $+ : \mathbb{Z}_p \times \mathbb{Z}_p \to \mathbb{Z}_p$ und $(\cdot) : \mathbb{Z}_p \times \mathbb{Z}_p \to \mathbb{Z}_p$ definiert, ist zu zeigen, dass die Gleichungen in (II.41) unabhängig von den gewählten Repräsentanten $k, \ell \in \mathbb{Z}$ sind. Seien dazu $k, k', \ell, \ell' \in \mathbb{Z}$ und $[k]_p = [k']_p$ sowie $[\ell]_p = [\ell']_p$, es gibt also $m, n \in \mathbb{Z}$, sodass

$$k' \;=\; k + mp \quad \text{und} \quad \ell' \;=\; \ell + np. \tag{II.42}$$

Dann sind aber auch

$$[k+\ell]_p \;=\; [k'+\ell']_p \quad \text{und} \quad [k \cdot \ell]_p \;=\; [k' \cdot \ell']_p, \tag{II.43}$$

denn

$$k' + \ell' \;=\; (k+\ell) + (m+n)p \quad \text{und} \quad k' \cdot \ell' \;=\; k \cdot \ell + (kn + \ell m + mn)p, \tag{II.44}$$

also sind die Verknüpfungen in (II.41) unabhängig von den gewählten Repräsentanten.

Eine Anwendung der Restklassenringe ist die Regel, dass eine Zahl $n \in \mathbb{N}$ genau dann durch 9 teilbar ist, wenn ihre Quersumme durch 9 teilbar ist, denn es gilt

$$\begin{aligned}
&\big[a_0 + a_1 \cdot 10^1 \;+\; a_2 \cdot 10^2 \;+\; \ldots \;+\; a_m \cdot 10^m\big]_9 \\
&\quad = \; [a_0]_9 \;+\; [a_1]_9 \cdot \big[10\big]_9 \;+\; \ldots \;+\; [a_m]_9 \cdot \big[10\big]_9^m \\
&\quad = \; [a_0]_9 \;+\; [a_1]_9 \;+\; [a_2]_9 \;+\; \ldots \;+\; [a_m]_9 \\
&\quad = \; [a_0 + a_1 + a_2 + \ldots + a_m]_9.
\end{aligned} \tag{II.45}$$

II.3. Körper

Definition II.4. Ein Ring $\mathbb{F}$ heißt **Körper**[1] $\Leftrightarrow$
$\mathbb{F}$ besitzt zu (R_1)–(R_3) zusätzlich die folgenden Eigenschaften:

$$(K_1) \quad \forall a, b \in \mathbb{F} : \qquad a \cdot b \;=\; b \cdot a, \tag{II.46}$$

$$(K_2) \quad \exists 1 \in \mathbb{F} \setminus \{0\} \; \forall a \in \mathbb{F} : \qquad 1 \cdot a \;=\; a \cdot 1 \;=\; a, \tag{II.47}$$

$$(K_3) \quad \forall a \in \mathbb{F} \setminus \{0\} \; \exists \frac{1}{a} \in \mathbb{F} \setminus \{0\} : \qquad a \cdot \frac{1}{a} \;=\; 1. \tag{II.48}$$

Bemerkungen und Beispiele.

[1] engl.: Field

- Ist $\mathbb{F}$ ein Ring, der die Eigenschaften (K_2) und (K_3), aber nicht (K_1) besitzt, so bezeichnet man $\mathbb{F}$ als Schiefkörper.
- Die Eigenschaften (R_1)–(R_3) sowie (K_1)–(K_3) eines Körpers $\mathbb{F}$ implizieren, dass $\mathbb{F} \setminus \{0\}$ bezüglich der Multiplikation eine abelsche Gruppe mit neutralem Element 1 ist. Man bezeichnet $\mathbb{F}^\times := \mathbb{F} \setminus \{0\}$ als multiplikative Gruppe von $\mathbb{F}$.
- Der Ring $\mathbb{Z}$ der ganzen Zahlen ist kein Körper.
- Die Menge $\mathbb{Q} := \{p/q \mid p \in \mathbb{Z}, q \in \mathbb{N}\}$ der *rationalen Zahlen* bilden einen Körper.
- Weitere Körper sind die Menge der *reellen Zahlen* $\mathbb{R}$ und die Menge der *komplexen Zahlen* $\mathbb{C}$. Diese Körper sind für diese Vorlesung am wichtigsten. Deshalb schreiben wir $\mathbb{K}$ statt $\mathbb{F}$, falls $\mathbb{F} = \mathbb{R}$ oder $\mathbb{F} = \mathbb{C}$.
- Ist p eine Primzahl, so bilden die Restklassen $\mathbb{Z}_p$ modulo p einen Körper (s. Abschnitt II.2.1).

II.4. Ergänzungen

II.4.1. Untergruppen

Definition II.5. Sei $(G, \circ)$ eine Gruppe. Eine Teilmenge $U \subseteq G$ heißt **Untergruppe**

$$:\Leftrightarrow \quad U \text{ ist bezüglich der Verknüpfung } \circ \text{ in } G \text{ selbst eine Gruppe.} \tag{II.49}$$

Lemma II.6 (Untergruppenkriterium)**.** *Sei $(G, \circ)$ eine Gruppe. Eine Teilmenge $U \subseteq G$ ist eine Untergruppe, falls folgende drei Kriterien erfüllt sind:*

$$(i) \qquad e \in U, \tag{II.50}$$

$$(ii) \quad \forall a, b \in U: \qquad a \circ b \in U, \tag{II.51}$$

$$(iii) \quad \forall a \in U: \qquad a^{-1} \in U. \tag{II.52}$$

Beweis. Wegen (i) gilt $\circ : U \times U \to U$, d.h. U ist bezüglich $\circ$ abgeschlossen. Da die Verknüpfung auf G assoziativ ist, ist sie (erst recht) auch auf $U \subseteq G$ assoziativ. (i) sichert (G_2) und (iii) sichert (G_3). □

Bemerkungen und Beispiele.

- $\{e\}, G \subseteq G$ sind stets Untergruppen – die trivialen Untergruppen.
- Die geraden Zahlen $2\mathbb{Z} \subseteq \mathbb{Z}$ bilden bezüglich Addition eine Untergruppe.
- Die ungeraden Zahlen $2\mathbb{Z} + 1 \subseteq \mathbb{Z}$ sind bezüglich Addition keine Untergruppe, denn $0 \notin 2\mathbb{Z} + 1$.

Definition II.7. Sei $(G, \circ)$ eine Gruppe. Das **Zentrum** $Z(G) \subseteq G$ von G ist definiert durch

$$Z(G) := \{a \in G \mid \forall\, x \in G: \; ax \; = \; xa\}. \tag{II.53}$$

Lemma II.8. *$Z(G)$ ist eine abelsche Untergruppe von G.*

Beweis. Zunächst weisen wir mit Hilfe des Untergruppenkriteriums, Lemma II.6, nach, dass $Z(G)$ eine Untergruppe von G ist.

(i) Wegen $ex = x = xe$ ist $e \in Z(G)$.

(ii) Für $a, b \in Z(G)$ und $x \in G$ ist $abx = axb = xab$, also ist $ab \in Z(G)$.

(iii) Für $a \in Z(G)$ und $x \in G$ ist $x^{-1}a = ax^{-1}$ und daher $a^{-1}x = (x^{-1}a)^{-1} = (ax^{-1})^{-1} = xa^{-1}$. Also ist mit a auch $a^{-1} \in Z(G)$.

Es folgt, dass $Z(G)$ eine Untergruppe ist. Weiterhin ist für $a, b \in Z(G)$ insbesondere $b \in G$ und deshalb ist $ab = ba$. Also ist $Z(G)$ abelsch. □

II.4.2. Permutationen, Transpositionen, Zyklen und Signum

Definition II.9. Sei $\pi \in \mathcal{S}_n$ eine Permutation. Sind $k_1, k_2, \ldots, k_r \in \mathbb{Z}_1^n$ mit $\pi(k_1) = k_2$, $\pi(k_2) = k_3, \ldots, \pi(k_r) = k_1$, also

$$k_1 \stackrel{\pi}{\longmapsto} k_2 \stackrel{\pi}{\longmapsto} k_3 \stackrel{\pi}{\longmapsto} \cdots \stackrel{\pi}{\longmapsto} k_r \stackrel{\pi}{\longmapsto} k_1, \tag{II.54}$$

so heißt $(k_1, k_2, \ldots, k_r)$ **Zyklus** von π der Länge r.

Bemerkungen und Beispiele.

- Die Permutation $\pi \in \mathcal{S}_9$,

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 4 & 1 & 9 & 3 & 8 & 7 & 6 & 2 \end{pmatrix}, \tag{II.55}$$

  besitzt die Zyklen

$$(1,5,3), \quad (2,4,9), \quad (6,8), \quad (7) \tag{II.56}$$

  (etwa $1 \longmapsto 5 \longmapsto 3 \longmapsto 1$). Jede Permutation ist offensichtlich durch ihre Zyklen eindeutig bestimmt, und man schreibt

$$\pi := (1,5,3)\,(2,4,9)\,(6,8). \tag{II.57}$$

  (Zur Vereinfachung lässt man Zyklen der Länge 1 weg.)

- Umgekehrt kann für $n \geq r$ jeder Zyklus $(k_1, k_2, \ldots, k_r)$ als Permutation in $\mathcal{S}_n$ gelesen werden: $(k_1, k_2, \ldots, k_r)$ lässt alle Elemente in $\mathbb{Z}_1^n \setminus \{k_1, k_2, \ldots, k_r\}$ invariant. Mit dieser Lesart wird (II.57) zu

$$\pi = (1,5,3) \circ (2,4,9) \circ (6,8). \tag{II.58}$$

- Außerdem kommutieren disjunkte Zyklen miteinander, deshalb ist $\pi = (1,5,3) \circ (2,4,9) \circ (6,8) = (2,4,9) \circ (6,8) \circ (1,5,3)$.

- Ein Vergleich mit (II.33) zeigt, dass Transpositionen genau die Zyklen der Länge 2 sind, nämlich

$$\sigma = \begin{pmatrix} 1 & \ldots & i-1 & \mathbf{i} & i+1 & \ldots & j-1 & \mathbf{j} & j+1 & \ldots & n \\ 1 & \ldots & i-1 & \mathbf{j} & i+1 & \ldots & j-1 & \mathbf{i} & j+1 & \ldots & n \end{pmatrix} = (i,j). \tag{II.59}$$

Satz II.10. *Sei $n \geq 2$. Jede Permutation $\pi \in \mathcal{S}_n$ kann als Komposition von Transpositionen geschrieben werden, d.h. es gibt $a_1, b_1,\ a_2, b_2,\ \ldots,\ a_m, b_m \in \mathbb{Z}_1^n$, mit $a_i \neq b_i$, so dass*

$$\pi = (a_1, b_1) \circ (a_2, b_2) \circ \cdots \circ (a_m, b_m). \tag{II.60}$$

Beweis. Zunächst stellen wir fest, dass es genügt, für einen beliebigen Zyklus $(a_1, \ldots, a_r)$ der Länge $2 \leq r \leq n$ Glg. (II.60) zu zeigen. Es ist aber

$$(a_1, a_2, \ldots, a_r) = (a_1, a_r) \circ (a_1, a_{r-1}) \circ \cdots \circ (a_1, a_2). \tag{II.61}$$

$\square$

Lemma II.11. *Sind $n \in \mathbb{N}$, $n \geq 2$, $1 \leq i < j \leq n$ und $\sigma = (i,j) \in \mathcal{S}_n$ eine Transposition, so ist*

$$(-1)^\sigma \;=\; -1. \tag{II.62}$$

Beweis. Der Beweis ist eine kleine Rechnung,

$$\begin{aligned}
(-1)^\sigma \;&=\; \prod_{p<q} \frac{\sigma(q)-\sigma(p)}{q-p} \;=\; \prod_{p<q;\ \{p=i \,\vee\, q=j\}} \frac{\sigma(q)-\sigma(p)}{q-p} \\
&= \Big(\prod_{p=i;\ i<q<j} \frac{\sigma(q)-\sigma(p)}{q-p} \Big) \cdot \Big(\prod_{p=i;\ q=j} \frac{\sigma(q)-\sigma(p)}{q-p} \Big) \cdot \Big(\prod_{p=i;\ q>j} \frac{\sigma(q)-\sigma(p)}{q-p} \Big) \\
&\quad \cdot \Big(\prod_{p=j;\ q>j} \frac{\sigma(q)-\sigma(p)}{q-p} \Big) \cdot \Big(\prod_{p<i;\ q=i} \frac{\sigma(q)-\sigma(p)}{q-p} \Big) \cdot \Big(\prod_{p<i;\ q=j} \frac{\sigma(q)-\sigma(p)}{q-p} \Big) \\
&\quad \cdot \Big(\prod_{i<p<j;\ q=j} \frac{\sigma(q)-\sigma(p)}{q-p} \Big) \\
&= \Big(\prod_{p=i;\ i<q<j} \frac{q-j}{q-i} \Big) \cdot \Big(\frac{i-j}{j-i} \Big) \cdot \Big(\prod_{p=i;\ q>j} \frac{q-j}{q-i} \Big) \cdot \Big(\prod_{p=j;\ q>j} \frac{q-i}{q-j} \Big) \\
&\quad \cdot \Big(\prod_{p<i;\ q=i} \frac{j-p}{i-p} \Big) \cdot \Big(\prod_{p<i;\ q=j} \frac{i-p}{j-p} \Big) \cdot \Big(\prod_{i<p<j;\ q=j} \frac{i-p}{j-p} \Big) \\
&= \, -\, 1.
\end{aligned} \tag{II.63}$$

□

Lemma II.12. *Sind $n \in \mathbb{N}$, $n \geq 2$, und $\pi, \kappa \in \mathcal{S}_n$ zwei Permutationen, so gilt*

$$(-1)^{\pi \circ \kappa} \;=\; (-1)^\pi \,\cdot\, (-1)^\kappa. \tag{II.64}$$

Beweis.

$$(-1)^{\pi\circ\kappa} \;=\; \prod_{1 \leq i < j \leq n} \frac{\pi(\kappa(i)) - \pi(\kappa(j))}{i-j} \;=\; \prod_{1 \leq i < j \leq n} \frac{\big[\pi(\kappa(i)) - \pi(\kappa(j))\big] \cdot \big[\kappa(i) - \kappa(j)\big]}{\big[\kappa(i) - \kappa(j)\big] \cdot [i-j]} \tag{II.65}$$

$$\begin{aligned}
&= \Big(\prod_{1 \leq i < j \leq n} \frac{\pi(\kappa(i)) - \pi(\kappa(j))}{\kappa(i) - \kappa(j)} \Big) \cdot \Big(\prod_{1 \leq i < j \leq n} \frac{\kappa(i) - \kappa(j)}{i-j} \Big) \\
&= \Big(\prod_{1 \leq i < j \leq n} \frac{\pi(i) - \pi(j)}{i-j} \Big) \cdot \Big(\prod_{1 \leq i < j \leq n} \frac{\kappa(i) - \kappa(j)}{i-j} \Big) \\
&= (-1)^\pi \,\cdot\, (-1)^\kappa.
\end{aligned}$$

□

Aus Lemmata II.11 und II.12 folgt nun Glg. (II.35), die wir nochmal als Korollar formulieren.

Korollar II.13. *Sind $n \in \mathbb{N}$ und $\pi = \sigma_1 \circ \sigma_2 \circ \ldots \circ \sigma_m \in \mathcal{S}_n$ eine Permutation, die als Komposition von m Transpositionen $\sigma_1, \sigma_2, \ldots, \sigma_m \in \mathcal{S}_n$ dargestellt werden kann, so ist*

$$(-1)^\pi \;=\; (-1)^m. \tag{II.66}$$

Beweis. Nach Lemmata II.11 und II.12 ist

$$(-1)^\pi \;=\; (-1)^{\sigma_1} \cdot (-1)^{\sigma_2} \cdots (-1)^{\sigma_m} \;=\; (-1)^m. \tag{II.67}$$

□

II.4.3. Der Polynomring $R[x]$ über einem kommutativen Ring R

Sei R ein kommutativer Ring. Wir betrachten Folgen $\underline{a} = (a_n)_{n=0}^\infty \in R^{\mathbb{N}_0}$, wobei nur endlich viele Folgeglieder von 0 verschieden sind. D.h. es gibt ein $m = m(\underline{a}) \in \mathbb{N}$, so dass

$$\underline{a} \;=\; (a_0, a_1, a_2, \ldots, a_m, 0, 0, \ldots). \tag{II.68}$$

(Dabei hängt $m(\underline{a})$ im Allgemeinen von der betrachteten Folge $\underline{a}$ ab und ist nicht für alle Folgen gleich.) Wir sammeln diese Folgen in

$$R[x] \;:=\; \Big\{\underline{a} = (a_n)_{n=0}^\infty \in R^{\mathbb{N}_0} \;\Big|\; \exists\, m \in \mathbb{N}\; \forall\, n > m : a_n = 0\Big\}. \tag{II.69}$$

$R[x]$ wird zu einem kommutativen Ring bezüglich der Verknüpfungen

$$\underline{a} + \underline{b} \;:=\; (a_0 + b_0,\; a_1 + b_1,\; a_2 + b_2, \ldots) \tag{II.70}$$

und

$$\underline{a} \cdot \underline{b} \;=\; \underline{c}, \quad \text{wobei} \quad c_n := a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0. \tag{II.71}$$

Dies prüft man durch Nachrechnen der Ringaxiome leicht nach.

Wir führen nun x als formale Variable ein und identifizieren $\underline{a}$ mit dem Polynom

$$\underline{a}(x) \;:=\; a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m. \tag{II.72}$$

Dann sieht man sofort, dass die Addition und Multiplikation in $R[x]$ gerade der Addition und Multiplikation der zugehörigen Polynome entspricht:

$$\begin{aligned}(\underline{a} + \underline{b})(x) &= (a_0 + b_0) + (a_1 + b_1)x + \ldots + (a_N + b_N)x^N \\ &= \underline{a}(x) + \underline{b}(x),\end{aligned} \tag{II.73}$$

$$\begin{aligned}(\underline{a} \cdot \underline{b})(x) &= a_0 b_0 + (a_0 b_1 + a_1 b_0)x + \ldots + (a_0 b_N + a_1 b_{N-1} + \ldots + a_N b_0)x^N \\ &= \underline{a}(x) \cdot \underline{b}(x),\end{aligned} \tag{II.74}$$

wobei $N := m(\underline{a}) + m(\underline{b})$. Daher bezeichnet man $R[x]$ als <u>Ring der Polynome (in x) über R</u>.

II.4.4. Restklassenringe $\mathbb{Z}_p$ modulo Primzahlen p sind Körper

Definition II.14. Seien $a, b \in \mathbb{Z}$.

(i) Eine Zahl $g \in \mathbb{Z}$ heißt **Teiler von a**, falls es ein $h \in \mathbb{Z}$ so gibt, dass $a = gh$ gilt.

(ii) Für $|a| + |b| > 0$ ist der **größte gemeinsame Teiler ggT(a, b)** $\in \mathbb{Z}$ von a und b die größte ganze Zahl, die ein Teiler sowohl von a als auch von b ist.

Bemerkungen und Beispiele.

- Da $0 = g \cdot 0$ für alle $g \in \mathbb{Z}$ gilt, sind alle ganzen Zahlen Teiler von 0. Deshalb müssen wir $a = b = 0$ bei der Definition des größten gemeinsamen Teiler ausschließen.
- Sind $a, b \in \mathbb{Z}$, so ist 1 stets ein Teiler sowohl von a als auch von b. Daher ist $\mathrm{ggT}(a, b) \geq 1$, d.h. der größte gemeinsame Teiler von $a \in \mathbb{Z}$ und $b \in \mathbb{Z}$, $|a|+|b| > 0$, ist stets eine natürliche Zahl.

Lemma II.15. *Sind $a, b \in \mathbb{Z}$ mit $|a| + |b| > 0$, so gibt es $k, \ell \in \mathbb{Z}$ so, dass*

$$\mathrm{ggT}(a, b) \ = \ ka + \ell b. \tag{II.75}$$

Beweis. Sei

$$M \ := \ \big\{ n \in \mathbb{N} \ \big| \ \exists \, k, \ell \in \mathbb{Z} : \quad n = ka + \ell b \big\}. \tag{II.76}$$

Dann ist $|a| + |b| \in M$, also $M \neq \emptyset$. Seien $m \in \mathbb{N}$ die kleinste Zahl in M, $m := \min M$, und $k, \ell \in \mathbb{Z}$ so, dass $m = ka + \ell b$. Da $d := \mathrm{ggT}(a, b)$ sowohl a als auch b teilt, teilt d auch m, also ist $d \leq m$.

Seien nun $a = qm + r$, mit $q \in \mathbb{Z}$ und $0 \leq r < m$. Wäre $r \geq 1$, so wäre

$$r \ = \ a - qm \ = \ (1 - qk)a - q\ell b \ \in \ M\,. \tag{II.77}$$

Da m die kleinste natürliche Zahl in M ist, kann dies nicht richtig sein, und es folgt $r = 0$, d.h. $a = qm$. Genauso erhält man $b = pm$. Also teilt m sowohl a als auch b. Damit ist $m \leq \mathrm{ggT}(a, b) = d$.

Insgesamt folgt, dass

$$\mathrm{ggT}(a, b) \ = \ m \ \in \ M. \tag{II.78}$$

$\square$

Satz II.16. *$\mathbb{Z}_p$ ist ein Körper genau dann, wenn $p \in \mathbb{N}$ eine Primzahl ist.*

Beweis.
Ist $p \in \mathbb{N}$ keine Primzahl, so gibt es $1 < a \leq b < p$ mit $p = ab$. Dann sind $[a]_p,\ [b]_p \neq [0]_p$, aber

$$[a]_p \cdot [b]_p = [ab]_p = [0]_p. \tag{II.79}$$

Also ist $\mathbb{Z}_p$ kein Körper.

Sind p eine Primzahl und $1 < a < p$, so ist $\mathrm{ggT}(a, p) = 1$. Nach Lemma II.15 gibt es $k, \ell \in \mathbb{Z}$ so, dass $1 = ka + \ell b$. Also ist

$$[1]_p \; = \; [k]_p \cdot [a]_p + [\ell]_p \cdot [p]_p \; = \; [k]_p \cdot [a]_p, \tag{II.80}$$

d.h. $[k]_p$ ist das Inverse zu $[a]_p$ bezüglich Multiplikation in $\mathbb{Z}_p$. $\square$

III. Reelle und komplexe Zahlen

Die detaillierte Einführung der reellen und der komplexen Zahlen ist traditionell Gegenstand der Einführungsvorlesung der Analysis. Wir geben hier nur ein paar Eckpunkte wieder. Dabei setzen wir Begriffe wie *Ordnungsrelationen*, *totale Ordnung*, *Beschränktheit nach oben* (von Teilmengen einer total geordneten Menge) und *geordneter Körper* als bekannt voraus.

III.1. Reelle Zahlen

Definition III.1. Seien $S \neq \emptyset$ eine total geordnete Menge und $T \subseteq S$ eine nach oben beschränkte Teilmenge. Ein Element $b \in S$ heißt **Supremum von $\boldsymbol{T}$** $:\Leftrightarrow$

$$(i) \qquad \forall t \in T: \qquad t \leq b, \tag{III.1}$$

$$(ii) \quad \forall a \in S,\ a < b\ \exists\, t \in T: \qquad a < t. \tag{III.2}$$

Bemerkungen und Beispiele.

- Besitzt eine Teilmenge $T \subseteq S$ einer total geordneten Menge S ein Supremum $b \in S$, so ist dieses eindeutig und wir schreiben

$$b \ =: \ \sup\big\{T\big\}, \tag{III.3}$$

  für *das* Supremum von T.
- Sind $S := \mathbb{Q}$ und $T := \{a \in \mathbb{Q} \mid a^2 < 2\}$, so ist S total geordnet und T nach oben beschränkt - z.B. 2 ist eine obere Schranke. Es existiert jedoch kein Supremum von T in S: Man sieht nämlich leicht ein, dass $b^2 = 2$ für das Supremum $b = \sup\{T\} \in S = \mathbb{Q}$ gelten müsste, diese Gleichung jedoch keine rationale Lösung besitzt.

Definition III.2. Eine total geordnete Menge $S \neq \emptyset$ erfüllt das **Supremumsaxiom** oder **L.U.B.-Axiom**[1]

$:\Leftrightarrow$ Jede nach oben beschränkte Teilmenge $T \subseteq S$ besitzt ein Supremum, $\sup\{T\} \in S$.

[1] "L.U.B." steht für "**l**east **u**pper **b**ound". Solche einfallsreichen Namensgebungen werden uns noch häufiger begegnen.

Satz III.3. *Es gibt einen eindeutigen geordneten Körper, in dem $a+c<b+c$ für alle $a,b,c\in\mathbb{F}$ mit $a<b$ gilt, in dem $ab>0$ für alle $a>0$ und $b>0$ gilt, der $\mathbb{Q}$ enthält und der das Supremumsaxiom erfüllt. Wir nennen diesen Körper die* ***reellen Zahlen*** *und bezeichnen ihn mit $\mathbb{R}$.*

III.2. Komplexe Zahlen

Mit Hilfe der Erweiterung der rationalen Zahlen $\mathbb{Q}$ durch die reellen Zahlen $\mathbb{R}\supset\mathbb{Q}$ können wir also die Lösungen der Gleichung $x^2=2$ in $\mathbb{R}$ finden. Wenn wir hingegen die Lösungen der Gleichung $x^2=-2$ suchen, werden wir erneut vor ein Problem gestellt: Da Quadrate reeller Zahlen stets positiv oder null sind, kann $x^2=-2$ keine reelle Lösung haben. Die Frage nach einer abermaligen Erweiterung des Zahlenbereichs, um auch Lösungen von $x^2=-2$ finden, führt uns von den reellen Zahlen $\mathbb{R}$ auf die komplexen Zahlen $\mathbb{C}$.

Definition III.4. Auf der Menge $\mathbb{R}^2:=\mathbb{R}\times\mathbb{R}$ seien Addition und Multiplikation durch

$$(a,b)+(c,d) := (a+c,\ b+d), \tag{III.4}$$

$$(a,b)\cdot(c,d) := (a\cdot c-b\cdot d,\ b\cdot c+a\cdot d), \tag{III.5}$$

definiert.

Satz III.5. *Die Menge $\mathbb{R}^2:=\mathbb{R}\times\mathbb{R}$ bildet bezüglich der Addition* (III.4) *und der Multiplikation* (III.5) *einen Körper, den Körper $\mathbb{C}$ der* ***komplexen Zahlen****. Dabei sind* $(0,0)$ *das neutrale Element bezüglich Addition und* $(1,0)$ *das neutrale Element bezüglich Multiplikation. Die inversen Elemente in $\mathbb{C}$ sind wie folgt gegeben,*

$$-(a,b) = (-a,-b), \tag{III.6}$$

$$\frac{1}{(a,b)} = \left(\frac{a}{a^2+b^2},\ \frac{-b}{a^2+b^2}\right), \tag{III.7}$$

wobei zu beachten ist, dass

$$\Big[(a,b)\neq(0,0)\Big] \ \Leftrightarrow\ \Big[(a\neq 0\vee b\neq 0)\Big] \ \Leftrightarrow\ \Big[a^2+b^2>0\Big]. \tag{III.8}$$

Beweis. Der Beweis, dass $\mathbb{C}$ ein Körper ist, erfolgt durch Nachprüfen der Eigenschaften $(R_1)-(R_3)$ und $(K_1)-(K_3)$. □

Zu jeder komplexen Zahl $z=(a,b)\in\mathbb{C}$ definieren wir

ihren **Realteil**	$\mathrm{Re}\{z\} := a \in \mathbb{R}$,	(III.9)
ihren **Imaginärteil**	$\mathrm{Im}\{z\} := b \in \mathbb{R}$,	(III.10)
ihren **Betrag**	$\lvert z\rvert := \sqrt{a^2+b^2} \in \mathbb{R}_0^+$,	(III.11)
und die **konjugiert komplexe Zahl**	$\bar{z} := (a,-b) \in \mathbb{C}$.	(III.12)

Bemerkungen und Beispiele.

- Sind $w = (2,1)$ und $z = (-3,4)$, so sind

$$\mathrm{Re}\{w\} \;=\; 2,\; \mathrm{Im}\{w\} \;=\; 1\,, \quad \text{und} \quad \mathrm{Re}\{z\} \;=\; -3,\; \mathrm{Im}\{z\} \;=\; 4\,, \tag{III.13}$$

$$|w| \;=\; \sqrt{2^2+1^2} \;=\; \sqrt{5}\,, \quad \text{und} \quad |z| \;=\; \sqrt{(-3)^2+4^2} \;=\; \sqrt{25} \;=\; 5\,, \tag{III.14}$$

$$\overline{w} \;=\; \begin{pmatrix} 2 \\ -1 \end{pmatrix} \quad \text{und} \quad \overline{z} \;=\; \begin{pmatrix} -3 \\ -4 \end{pmatrix}. \tag{III.15}$$

- Weiterhin sind mit $w = (2,1)$ und $z = (-3,4)$ auch

$$w + z \;=\; (-1,5)\,, \qquad w - z \;=\; (5,-3)\,, \tag{III.16}$$

$$w \cdot z \;=\; (2\cdot(-3) - 1\cdot 4\;,\; 2\cdot 4 + 1\cdot(-3)) \;=\; (-10,5)\,, \tag{III.17}$$

$$\frac{1}{z} \;=\; \left(\frac{-3}{3^2+4^2}\,,\,\frac{-4}{3^2+4^2}\right) \;=\; \left(\frac{-3}{25},\frac{-4}{25}\right) \quad \text{und} \tag{III.18}$$

$$\frac{w}{z} \;=\; w\cdot\frac{1}{z} \;=\; (2,1)\cdot\left(\frac{-3}{25},\frac{-4}{25}\right) \tag{III.19}$$

$$= \left(\frac{2\cdot(-3) - 1\cdot(-4)}{25},\frac{1\cdot(-3) + 2\cdot(-4)}{25}\right) \;=\; \left(-\frac{2}{25}\,,\,-\frac{11}{25}\right).$$

- Die Definition (III.5)

$$(a,b)\cdot(c,d) \;=\; (ac - bd\,,\, bc + ad) \tag{III.20}$$

  des Produkts komplexer Zahlen wirkt auf den ersten Blick merkwürdig und unmotiviert. Tatsächlich ergibt sie sich jedoch zwingend aus den Körperaxiomen $(R_1) - (R_3)$ und $(K_1) - (K_3)$ und der zusätzlichen Forderung, dass

$$1\cdot z \;=\; z \quad \text{und} \quad |w\cdot z| \;=\; |w|\cdot|z| \tag{III.21}$$

  gelten möge: Das Produkt (III.20) ist das einzige auf $\mathbb{R}^2$, mit dem $\mathbb{R}^2$ die Körperaxiome und (III.21) erfüllt.

III.2.1. Imaginäre Einheit i und $\mathbb{R} \subset \mathbb{C}$ als Teilkörper

Es ist bequem, die sogenannte **imaginäre Einheit** $i := (0,1)$ einzuführen. Identifizieren wir weiterhin $\mathbb{1} := (1,0)$, so kann man jede komplexe Zahl $(a,b) \in \mathbb{C}$ als $(a,b) = a\cdot\mathbb{1} + b\cdot i$ schreiben. Damit ist

$$z \;=\; \mathrm{Re}\{z\}\cdot\mathbb{1} \;+\; \mathrm{Im}\{z\}\cdot i. \tag{III.22}$$

Weiterhin sind mit dieser Schreibweise (III.4) und (III.5) äquivalent zu

$$(a\cdot\mathbb{1} + b\cdot i) \;+\; (c\cdot\mathbb{1} + d\cdot i) \;=\; (a+c)\cdot\mathbb{1} \;+\; (b+d)\cdot i, \tag{III.23}$$

$$(a\cdot\mathbb{1} + b\cdot i) \;\cdot\; (c\cdot\mathbb{1} + d\cdot i) \;=\; (ac-bd)\cdot\mathbb{1} \;+\; (bc+ad)\cdot i. \tag{III.24}$$

Insbesondere ist

$$i^2 \;=\; (0,1) \cdot (0,1) \;=\; (-1,0) \;=\; -\mathbb{1}, \tag{III.25}$$

d.h. die imaginäre Einheit ist eine Quadratwurzel aus $-\mathbb{1}$. Das ist auch die einzige zusätzliche Rechenregel, die man beim Rechnen mit komplexen Zahlen im Vergleich zu den reellen beachten muss.

Eine sehr nützliche Beobachtung ist, dass die komplexen Zahlen mit verschwindendem Imaginärteil mit den reellen Zahlen identifiziert werden können. Dies rechtfertigt auch den Namen "Realteil" für die erste Komponente einer komplexen Zahl. Definieren wir

$$\mathrm{Re}\,\mathbb{C} \;:=\; \{(a,0) \;=\; a \cdot \mathbb{1} \in \mathbb{C} \mid a \in \mathbb{R}\}, \tag{III.26}$$

so sieht man leicht, dass $\mathrm{Re}\,\mathbb{C} \subseteq \mathbb{C}$ ein Teilkörper ist, d.h. $\mathrm{Re}\,\mathbb{C} \subseteq \mathbb{C}$ ist eine Teilmenge, die selbst ein Körper ist. Außerdem ist $\mathrm{Re}\,\mathbb{C} \subseteq \mathbb{C}$ *isomorph* (als Körper) zu $\mathbb{R}$. Etwas genauer formuliert, ist

$$J \;:\; \mathbb{R} \to \mathrm{Re}\,\mathbb{C}, \quad a \mapsto a \cdot \mathbb{1} \tag{III.27}$$

eine Bijektion, die die Körpereigenschaften erhält, d.h. es gilt $J(a+b) = J(a) + J(b)$ und $J(a \cdot b) = J(a) \cdot J(b)$. Mit (III.27) können wir $\mathbb{R}$ und $\mathrm{Re}\,\mathbb{C} = J[\mathbb{R}]$ miteinander identifizieren und die reellen Zahlen als Teilmenge der komplexen Zahlen auffassen. Konkret geschieht diese Identifikation einfach durch das Weglassen von "$\mathbb{1}$", also indem wir $a + ib := a \cdot \mathbb{1} + b \cdot i$ schreiben und Einsen $\mathbb{1}$ als Faktoren auslassen. Mit dieser Identifikation wird dann auch

$$|z|^2 \;=\; \bar{z} \cdot z, \tag{III.28}$$

für jede komplexe Zahl $z \in \mathbb{C}$, sowie

$$(a+ib) \,+\, (c+id) \;= (a+c) \,+\, i(b+d), \tag{III.29}$$

$$(a+ib) \,\cdot\, (c+id) \;= (ac-bd) \,+\, i(bc+ad). \tag{III.30}$$

III.2.2. Polardarstellung, Potenzen und Wurzeln komplexer Zahlen

Von großer Bedeutung ist die Polardarstellung für komplexe Zahlen. Für eine komplexe Zahl $z = a + ib \neq 0$, die nicht null ist, definieren wir einen Winkel $\varphi \in \mathbb{R}$ so, dass

$$\cos(\varphi) \;=\; \frac{a}{\sqrt{a^2+b^2}} \;=\; \frac{\mathrm{Re}\{z\}}{|z|} \qquad \text{und} \qquad \sin(\varphi) \;=\; \frac{b}{\sqrt{a^2+b^2}} \;=\; \frac{\mathrm{Im}\{z\}}{|z|}. \tag{III.31}$$

gelten. Mit diesem Winkel wird dann

$$z \;=\; |z| \big(\cos(\varphi) \,+\, i\, \sin(\varphi)\big). \tag{III.32}$$

Wegen der Periodizität $\cos(\varphi + 2\pi) = \cos(\varphi)$ der Kosinusfunktion und der Periodizität $\sin(\varphi + 2\pi) = \sin(\varphi)$ der Sinusfunktion ändert sich die Zahl z in (III.32) nicht, wenn man 2π zu φ hinzuaddiert. Daher kann der Winkel φ stets zwischen 0 und 2π gewählt werden,

$$0 \;\leq\; \varphi \;<\; 2\pi, \tag{III.33}$$

und man bezeichnet φ in diesem Zusammenhang als **Argument** der Zahl z oder als eine **Phase**. Die oben genannte Periodizität ist auch in der (Gaußschen) Zahlenebene $\mathbb{C}$ klar erkennbar: Eine Addition von π zu φ bedeutet geometrisch ein Schwenken des zu z gehörigen Vektors genau um einen geschlossenen Kreis $360° = 2\pi$ um den Ursprung.

Ist nun $\tilde{z} = |\tilde{z}|\{\cos(\psi) + i\sin(\psi)\}$ eine weitere komplexe Zahl mit zugehöriger Phase $\psi \in \mathbb{R}$, so erhalten wir aus den *Additionstheoremen für Sinus und Kosinus*,

$$\cos(\alpha+\beta) \;=\; \cos(\alpha)\cos(\beta) \;-\; \sin(\alpha)\sin(\beta), \tag{III.34}$$

$$\sin(\alpha+\beta) \;=\; \sin(\alpha)\cos(\beta) \;+\; \cos(\alpha)\sin(\beta), \tag{III.35}$$

die für beliebige Winkel $\alpha, \beta \in \mathbb{R}$ gelten, dass

$$\begin{aligned} z\cdot\tilde{z} \;&= |z|\cdot|\tilde{z}|\cdot\big\{\cos(\varphi) + i\,\sin(\varphi)\big\}\cdot\big\{\cos(\psi) + i\,\sin(\psi)\big\} \\ &= |z|\,|\tilde{z}|\,\big\{\big[\cos(\varphi)\cos(\psi) \;-\; \sin(\varphi)\sin(\psi)\big] \;+\; i\,\big[\sin(\varphi)\cos(\psi) \;+\; \cos(\varphi)\sin(\psi)\big]\big\} \\ &= |z|\,|\tilde{z}|\,\big\{\cos(\varphi+\psi) \;+\; i\,\sin(\varphi+\psi)\big\}. \end{aligned} \tag{III.36}$$

Die Multiplikation zweier komplexer Zahlen erfolgt also durch Multiplikation ihrer Beträge und Addition der Phasen, und es gelten sogar

$$e^{i\varphi} \;=\; \cos(\varphi) \;+\; i\,\sin(\varphi) \quad \text{und} \quad z \;=\; |z|\cdot e^{i\varphi}, \tag{III.37}$$

sodass (III.36) nur das Potenzgesetz $e^x e^y = e^{x+y}$ widerspiegelt.

$$z\cdot\tilde{z} \;=\; \big(|z|\cdot e^{i\varphi}\big)\cdot\big(|\tilde{z}|\cdot e^{i\psi}\big) \;=\; |z|\cdot|\tilde{z}|\cdot\big(e^{i\varphi}\cdot e^{i\psi}\big) \;=\; |z|\cdot|\tilde{z}|\cdot e^{i(\varphi+\psi)}. \tag{III.38}$$

Glg. (III.37) lässt sich durch Entwicklungen der Exponential-, der Kosinus- und der Sinusfunktion in *Potenzreihen* mathematisch streng begründen – dies ist jedoch Gegenstand der Analysis und wird in dieser Vorlesung nicht behandelt.

Mit $r := |z|$ wird (III.37) zur **Polardarstellung**

$$z \;=\; r\cdot e^{i\varphi}. \tag{III.39}$$

der komplexen Zahl z. (Diese ist auch für $z = 0$ gültig, da in diesem Fall $r = 0$ ist und die Phase φ beliebig gewählt werden kann.)

Bemerkungen und Beispiele.

- In der Polardarstellung haben dann Addition und Multiplikation zweier komplexer Zahlen jeweils folgende grafische Interpretation.
- Die Addition zweier komplexer Zahlen $z_1 = a + ib = (a, b)$ und $z_2 = c + id = (c, d)$ erfolgt komponentenweise, wie bei Vektoren in der Ebene $\mathbb{R}^2$,

$$z_1 + z_2 \;=\; (a, b) + (c, d) \;=\; (a + c,\; b + d) \;=\; (a + c) + i(b + d). \tag{III.40}$$

- Die Multiplikation zweier in Polardarstellung gegebener komplexer Zahlen $z_1 = r_1 \cdot e^{i\varphi_1}$ und $z_2 = r_2 \cdot e^{i\varphi_2}$ erfolgt durch Multiplikation ihrer Beträge und Addition ihrer Phasen,

$$z_1 \cdot z_2 \;=\; (r_1\, e^{i\varphi_1}) \cdot (r_2\, e^{i\varphi_2}) \;=\; r_1\, r_2\, e^{i(\varphi_1+\varphi_2)}. \tag{III.41}$$

- Die Division zweier in Polardarstellung gegebener komplexer Zahlen $z_1 = r_1 \cdot e^{i\varphi_1}$ und $z_2 = r_2 \cdot e^{i\varphi_2} \neq 0$ erfolgt durch Division ihrer Beträge und Subtraktion ihrer Phasen,

$$\frac{z_1}{z_2} \;=\; \frac{r_1\, e^{i\varphi_1}}{r_2\, e^{i\varphi_2}} \;=\; \frac{r_1}{r_2}\, e^{i(\varphi_1-\varphi_2)}. \tag{III.42}$$

- Wir belegen die obigen Aussagen mit einem Zahlenbeispiel. Sind $w = 2 + i$ und $z = -3 - 4i$, so sind $w = |w| \cdot e^{i\alpha}$ und $z = |z| \cdot e^{i\beta}$ mit

$$|w| \;=\; \sqrt{5}, \quad \cos(\alpha) = \frac{2}{\sqrt{5}},\ \sin(\alpha) = \frac{1}{\sqrt{5}} \;\Rightarrow\; \alpha \approx 0,46 \;\Rightarrow w \;\approx\; \sqrt{5} \cdot e^{i\cdot(0,46)}\ , \tag{III.43}$$

$$|z| \;=\; 5, \quad \cos(\beta) = \frac{-3}{5},\ \sin(\beta) = \frac{-4}{5} \;\Rightarrow\; \beta \approx 4,07 \;\Rightarrow z \;\approx\; 5 \cdot e^{i\cdot(2,21)}. \tag{III.44}$$

Andererseits können wir direkt ausrechnen, dass $w \cdot z = |w \cdot z| \cdot e^{i\gamma}$ mit

$$w \cdot z \;=\; -2 - 11i, \tag{III.45}$$

$$|w \cdot z| \;=\; \sqrt{(-2)^2 + (-11)^2} \;=\; \sqrt{125} \;=\; 5 \cdot \sqrt{5}, \tag{III.46}$$

$$\cos(\gamma) = \frac{-2}{\sqrt{125}}, \quad \sin(\gamma) = \frac{-11}{\sqrt{125}} \quad\Rightarrow\quad \gamma \approx 4,53 \tag{III.47}$$

$$\Rightarrow\quad w \cdot z \;\approx\; \sqrt{125} \cdot e^{i\cdot(4,53)} \;=\; \sqrt{5} \cdot e^{i\cdot(0,46)} \cdot 5 \cdot e^{i\cdot(4,07)}. \tag{III.48}$$

Die wahre Stärke der Polardarstellung $z = r \cdot e^{i\varphi}$ einer komplexen Zahl z liegt in der Vereinfachung der Berechnung von Potenzen z^n und Wurzeln $\sqrt[n]{z}$ von z. Für eine nichtnegative ganze Zahl $n \in \mathbb{N}_0$ ist nämlich nach (III.41)

$$z^n \;=\; \big(r \cdot e^{i\varphi}\big)^n \;=\; r^n \cdot \big(e^{i\varphi}\big)^n \;=\; r^n \cdot e^{in\cdot\varphi}, \tag{III.49}$$

und für $z \neq 0$ gilt dies auch für negative n, also für alle ganzen Zahlen $n \in \mathbb{Z}$.

Bemerkungen und Beispiele.
Sind etwa $z = 3 + 4i$ und $n = 4$, so kann man ausmultiplizieren und erhält

$$z^4 \;=\; (3+4i)^4 \;=\; \big((3+4i)^2\big)^2 \;=\; \big(3^2 + 2\cdot 3\cdot 4i + (4i)^2\big)^2 \;=\; \big(9 + 24i - 16\big)^2 \tag{III.50}$$

$$= (-7+24i)^2 \;=\; (-7)^2 + 2\cdot(-7)\cdot 24i + (24i)^2 \;=\; 49 - 336i - 576 \;=\; -527 - 336i$$

durch mühsame Rechnung. Schreibt man aber z in Polardarstellung, $z \approx 5 \cdot e^{i \cdot 0{,}927}$, so erhält man mit (III.41) ganz einfach

$$z^4 \;\approx\; \big(5 \cdot e^{i \cdot 0{,}927}\big)^4 \;=\; 5^4 \cdot \big(e^{i \cdot 0{,}927}\big)^4 \;=\; 5^4 \cdot e^{i \cdot 4 \cdot 0{,}927} \;=\; 625 \cdot e^{i \cdot 3{,}71}. \tag{III.51}$$

Um die Übereinstimmung der Ergebnisse zu überprüfen, beobachtet man, dass $\cos(3,71) \approx -0,843$ und $\sin(3,71) \approx -0,538$ sind und daher

$$\begin{aligned} 625 \cdot e^{i \cdot 3{,}71} &= 625 \cdot \cos(3,71) + 625 \cdot \sin(3,71) i \;\approx\; 625 \cdot (-0,843) + i\, 625 \cdot (-0,538) \\ &= -\, 526,875 - (336,25) i \;\approx\; -527 - 336 i \end{aligned} \tag{III.52}$$

gilt, wobei die Fehler nur durch das Runden von Sinus und Kosinus auf zwei signifikante Stellen hinter dem Komma entstehen – und nicht durch falsche Rechnung.

Auch das Wurzelziehen ist bei komplexen Zahlen leicht und funktioniert vor allem immer. Ist $z = r \cdot e^{i\varphi}$ eine komplexe Zahl mit $r > 0$ und $0 \leq \varphi < 2\pi$, so sind die Quadratwurzeln von z alle komplexen Zahlen w, für die $w^2 = z$ gilt. In Polardarstellung $w = s \cdot e^{i\alpha}$ muss also

$$s^2 \cdot e^{i2\alpha} \;=\; r \cdot e^{i\varphi} \tag{III.53}$$

gelten, wobei $s \geq 0$ und $0 \leq \alpha < 2\pi$. Somit ist $s = \sqrt{r} > 0$, und es muss $e^{i2\alpha} = e^{i\varphi}$ gelten. Letztere Gleichung hat zwei Lösungen im Intervall $[0, 2\pi)$, nämlich

$$\alpha_1 \;=\; \frac{\varphi}{2} \quad \text{und} \quad \alpha_2 \;=\; \frac{\varphi}{2} + \pi, \tag{III.54}$$

da $e^{2\pi i} = 1$. Außerdem ist $e^{\pi i} = -1$, und wir erhalten $z = w_1^2 = w_2^2$ mit

$$w_1 \;=\; \sqrt{r} \cdot e^{i\varphi/2} \quad \text{und} \quad w_2 \;=\; \sqrt{r} \cdot e^{i(\pi + \varphi/2)} \;=\; \sqrt{r} \cdot e^{i\pi} \cdot e^{i\varphi/2} \;=\; -\sqrt{r} \cdot e^{i\varphi/2} \;=\; -w_1, \tag{III.55}$$

wie gewohnt.

Auch das Ziehen der allgemeinen n. Wurzel ist nicht schwer: Sind $z = r \cdot e^{i\varphi} \neq 0$ eine komplexe und $n \geq 2$ eine natürliche Zahl, wobei wir abermals $r > 0$ und $0 \leq \varphi < 2\pi$ annehmen können, so sind die n. Wurzeln von z alle komplexen Zahlen w, für die $w^n = z$ gilt. In Polardarstellung $w = s \cdot e^{i\alpha}$ mit $s \geq 0$ und $0 \leq \alpha < 2\pi$ muss also wieder $s^n \cdot e^{in\alpha} = r \cdot e^{i\varphi}$ gelten, was auf $s = \sqrt[n]{r} = r^{1/n} > 0$ und

$$e^{in\alpha} \;=\; e^{i\varphi} \quad \Leftrightarrow \quad e^{i(n\alpha - \varphi)} \;=\; 1 \tag{III.56}$$

führt. Daher muss $n\alpha - \varphi$ ein ganzzahliges Vielfaches von 2π sein, und die möglichen Lösungen haben die Gestalt

$$\alpha_k \;=\; \frac{\varphi}{n} + \frac{2\pi(k-1)}{n}, \tag{III.57}$$

wobei $k \in \mathbb{Z}$ eine ganze Zahl ist. Da wir außerdem $0 \leq \alpha_k < 2\pi$ fordern, gibt es für α genau n Lösungen, nämlich

$$\alpha_1 \;=\; \frac{\varphi}{n}, \;\; \alpha_2 \;=\; \frac{\varphi}{n} + \frac{2\pi}{n}, \;\; \ldots, \;\; \alpha_n \;=\; \frac{\varphi}{n} + \frac{2\pi(n-1)}{n}. \tag{III.58}$$

Daher besitzt z genau die n. Wurzeln

$$w_1 \;=\; \sqrt[n]{r} \cdot e^{\frac{i\varphi}{n}}, \quad w_2 \;=\; \sqrt[n]{r} \cdot e^{\frac{i\varphi}{n} + \frac{2\pi i}{n}}, \quad \ldots \quad , \quad w_n \;=\; \sqrt[n]{r} \cdot e^{\frac{i\varphi}{n} + \frac{2\pi i(n-1)}{n}}. \tag{III.59}$$

Die obigen Überlegungen zeigen, dass eine komplexe Zahl $z \neq 0$ stets n (voneinander verschiedene) n. Wurzeln besitzt – im Gegensatz zu den reellen Zahlen, in denen etwa -1 keine reelle Quadratwurzel besitzt (sondern nur die komplexen Wurzeln $\pm i$, wobei wir dann aber schon wieder -1 als komplexe Zahl mit Imaginärteil null betrachten). Diese bemerkenswerte Eigenschaft der komplexen Zahlen führt letztendlich dazu, dass jedes komplexe Polynom in Linearfaktoren zerfällt. Genauer gilt der folgende Satz.

Satz III.6 (Fundamentalsatz der Algebra)**.** *Für $n \in \mathbb{N}$ seien $c_0, c_1, \ldots, c_{n-1} \in \mathbb{C}$ komplexe Zahlen und p das Polynom*

$$p(z) \;:=\; z^n + c_{n-1} z^{n-1} + \ldots + c_1 z + c_0. \tag{III.60}$$

Dann gibt es komplexe Zahlen $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ so, dass

$$p(z) \;=\; (z - \lambda_1) \cdot (z - \lambda_2) \cdots (z - \lambda_n) \tag{III.61}$$

für alle $z \in \mathbb{C}$ gilt, d.h. $\lambda_1, \lambda_2, \ldots, \lambda_n$ sind die (nicht notwendig voneinander verschiedenen) Nullstellen des Polynoms p.

Bemerkungen und Beispiele.

- Das *reelle* Polynom $p(x) = x^2 + 1$ besitzt keine *reellen* Nullstellen und lässt sich nicht in der Form $p(x) = (x - x_1)(x - x_2)$ mit $x_1, x_2 \in \mathbb{R}$ zerlegen. Der Fundamentalsatz der Algebra ist für reelle Zahlen also falsch.
- Das *komplexe* Polynom $p(z) = z^2 + 1$ (mit $1 = 1 + 0 \cdot i$ als komplexe Zahl) besitzt die *komplexen* Nullstellen $\pm i$ und zerfällt zu $p(z) = (z - i)(z + i)$ in Linearfaktoren, wie es der Fundamentalsatz der Algebra behauptet.
- Der Fundamentalsatz der Algebra, Satz III.6, sichert zwar für jedes Polynom $p(z) = z^n + c_{n-1} z^{n-1} + \ldots + c_1 z + c_0$ vom Grad n die Existenz von n Nullstellen $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$, beinhaltet aber keine Lösungsformel oder ein anderes Verfahren zu ihrer Bestimmung.
- Für Grad $n = 1$ ist $p(z) = z + c_0$, und offensichtlich gilt die Lösungsformel $\lambda_1 = -c_0$.
- Für Grad $n = 2$ ist $p(z) = z^2 + c_1 z + c_0$, und λ_1 und λ_2 können mit Hilfe der p-q-Formel bestimmt werden, $\lambda_1 = -\frac{c_1}{2} + \sqrt{\frac{c_1^2}{4} - c_0}$ und $\lambda_2 = -\frac{c_1}{2} - \sqrt{\frac{c_1^2}{4} - c_0}$.
- Für Grad $n = 3, 4$ gibt es gibt es allgemeine Lösungsformeln für die Nullstellen. Sie sind jedoch zu kompliziert um wirklich nützlich zu sein.
- $n \geq 5$: Vor etwa zweihundert Jahren haben die zwei (leider jung verstorbenen) Mathematiker Niels Henrik Abel (1802-1829) und Evariste Galois (1811-1832) die bemerkenswerte Tatsache bewiesen, dass es Lösungsformeln für die Bestimmung der Nullstellen von Polynomen vom Grad $n \geq 5$ *prinzipiell nicht geben kann!* Dies werden wir hier nicht behandeln, sondern ist Gegenstand einer Vorlesung über *Galoistheorie*.

IV. Vektorräume

Wir beginnen diesen Abschnitt mit der allgemeinen Definition eines Vektorraums. Wir notieren im Folgenden mit $\mathbb{K}$ den Körper $\mathbb{R}$ der reellen Zahlen oder den Körper $\mathbb{C}$ der komplexen Zahlen, d.h. es ist (durchgehend) $\mathbb{K} = \mathbb{R}$ oder (durchgehend) $\mathbb{K} = \mathbb{C}$.

Die Theorie der Vektorräume lässt sich weitgehend auch genauso für einen allgemeinen Körper $\mathbb{F}$ entwickeln. Beispielsweise würden wir mit $\mathbb{F} = \mathbb{Q}$ auch Vektorräume über den rationalen Zahlen oder mit $\mathbb{F} = \mathbb{Z}_p$ Vektorräume über den Restklassenring $\mathbb{Z}_p$ erhalten, falls $p \in \mathbb{N}$ eine Primzahl und somit $\mathbb{Z}_p$ ein Körper sind.

IV.1. Definition eines Vektorraums

Definition IV.1. Eine Menge X heißt **Vektorraum über** $\mathbb{K}$ oder $\mathbb{K}$**-Vektorraum** $:\Leftrightarrow$
Auf X sind Addition $+ : X \times X \to X$ und Multiplikation mit einem Skalar $(\cdot) : \mathbb{K} \times X \to X$ definiert, die die folgenden Eigenschaften besitzen:

$$(i) \quad \forall \vec{a}, \vec{b}, \vec{c} \in X : \quad \vec{a} + \vec{b} = \vec{b} + \vec{a}, \qquad (\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c}), \tag{IV.1}$$

$$(ii) \quad \exists \vec{0} \in X \ \forall \vec{a} \in X : \quad \vec{a} = \vec{a} + \vec{0} = \vec{0} + \vec{a}, \tag{IV.2}$$

$$(iii) \quad \forall \vec{a} \in X \ \exists (-\vec{a}) \in X : \quad \vec{a} + (-\vec{a}) = (-\vec{a}) + \vec{a} = \vec{0}, \tag{IV.3}$$

$$(iv) \quad \forall \vec{a}, \vec{b} \in X, \ \alpha, \beta \in \mathbb{K} : \quad \begin{cases} \alpha \cdot (\vec{a} + \vec{b}) = \alpha \cdot \vec{a} + \alpha \cdot \vec{b}, \\ (\alpha + \beta) \cdot \vec{a} = \alpha \cdot \vec{a} + \beta \cdot \vec{a}, \end{cases} \tag{IV.4}$$

$$(v) \quad \forall \vec{a} \in X, \ \alpha, \beta \in \mathbb{K} : \quad 1 \cdot \vec{a} = \vec{a}, \qquad \alpha(\beta \vec{a}) = (\alpha\beta)\vec{a}. \tag{IV.5}$$

Die Elemente von X heißen **Vektoren**, $\vec{0} \in X$ ist der **Nullvektor**. Vektorräume über $\mathbb{R}$ bezeichnet man auch als **reelle Vektorräume**, Vektorräume über $\mathbb{C}$ als **komplexe Vektorräume**.

Bemerkungen und Beispiele.

- Gleichungen (IV.1), (IV.2) und (IV.3) besagen zusammen, dass X bezüglich der Addition $+$ eine abelsche Gruppe bildet.

- *Multiplikation mit einem Skalar* und *Skalarprodukt* ist nicht dasselbe.
- $\mathbb{K}$ ist ein Vektorraum über $\mathbb{K}$. (Dieses Beispiel ist allerdings etwas künstlich.)
- Sei $N \in \mathbb{N}$. Schreiben wir

$$A^N \ := \ \underbrace{A \times A \times \ldots \times A}_{\text{N Faktoren}}, \tag{IV.6}$$

für die Menge aller N-Tupel in A [s. (I.51)], so wird

$$\mathbb{K}^N \ = \ \left\{ \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \ \middle| \ x_1, x_2, \ldots, x_N \ \in \ \mathbb{K} \right\} \tag{IV.7}$$

bezüglich komponentenweiser Addition und Multiplikation mit einem Skalar,

$$\gamma \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} \ := \ \begin{pmatrix} \gamma x_1 + y_1 \\ \vdots \\ \gamma x_N + y_N \end{pmatrix} \tag{IV.8}$$

zu einem Vektorraum über $\mathbb{K}$. Dabei sind

$$\vec{0} \ = \ \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{und} \quad - \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \ = \ \begin{pmatrix} -x_1 \\ \vdots \\ -x_N \end{pmatrix}. \tag{IV.9}$$

- Für $N = 1, 2, 3$ gewinnt man durch Zeichnungen eine gewisse Vorstellung von den Vektoren in $\mathbb{R}^N$.
- Schränkt man bei den komplexen Zahlen die Multiplikation auf die reellen Elemente ein, so ist $\mathbb{C} = \mathbb{R}^2$ ein Vektorraum über $\mathbb{R}$ (mit Dimension 2, s.u.).
- Ein weiteres Beispiel eines Vektorraums ist die Menge der reellen Funktionen auf $[\alpha, \beta]$, $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$,

$$\mathcal{F} \ := \ \{ f : [\alpha, \beta] \to \mathbb{R} \}. \tag{IV.10}$$

Mit $f, g \in \mathcal{F}$ und $\gamma \in \mathbb{R}$ wird $\mathcal{F}$ durch

$$\forall\, x \in [\alpha, \beta] : \quad (f + \gamma \cdot g)(x) \ := \ f(x) + \gamma \cdot g(x) \tag{IV.11}$$

zum reellen Vektorraum. Diese Festlegung wird auch als *punktweise Addition und Multiplikation mit einem Skalar* bezeichnet. Dabei sind

$$\forall\, x \in [\alpha, \beta] : \quad 0_{\mathcal{F}}(x) \ := \ 0_{\mathbb{R}}, \quad (-f)(x) \ := \ -f(x). \tag{IV.12}$$

Für diesen Vektorraum $\mathcal{F}$ liefert eine Zeichnung eine gewisse Anschauung.
- Das obige Beispiel lässt sich noch wie folgt verallgemeinern. Sind $A \neq \emptyset$ eine Menge und X ein Vektorraum über $\mathbb{K}$, so wird durch punktweise Addition und punktweise Multiplikation mit einem Skalar wie in (IV.11) auch der Funktionenraum

$$\mathcal{F}' \ := \ \{ f : A \to X \} \tag{IV.13}$$

zum Vektorraum über $\mathbb{K}$.

- Insbesondere ist $\mathcal{F}'$ für $A := \mathbb{Z}_1^N$ und $X := \mathbb{K}^1 = \mathbb{K}$ gleich (genauer: *isomorph zu*) $\mathbb{K}^N$,

$$\mathbb{K}^N = \{f : \mathbb{Z}_1^N \to \mathbb{K}\}. \tag{IV.14}$$

IV.2. Unterräume

Definition IV.2. Sei X ein $\mathbb{K}$-Vektorraum. Eine Teilmenge $Y \subseteq X$, die selbst ein $\mathbb{K}$-Vektorraum ist, heißt **Unterraum** oder **Teilraum**.

Lemma IV.3 (Unterraumkriterium)**.** *Seien X ein $\mathbb{K}$-Vektorraum und $Y \subseteq X$ eine Teilmenge. Dann gilt folgende Äquivalenz:*

$$\Big\{Y \textit{ ist ein Unterraum von } X\Big\} \tag{IV.15}$$

$$\Leftrightarrow \quad \vec{0} \in Y \quad \textit{und} \quad \Big\{\forall\, \alpha \in \mathbb{K}\ \forall \vec{x}, \vec{y} \in Y : \ (\vec{x} + \alpha\vec{y}) \in Y\Big\}.$$

Beweis.
$\underline{\Rightarrow}$: ist trivial.

$\underline{\Leftarrow}$: Setzen wir $\alpha := 1$, so folgt mit $\vec{x}, \vec{y} \in Y$ auch $\vec{x} + \vec{y} \in Y$. Wegen $\vec{0} \in Y$ ist weiterhin mit $\alpha \in \mathbb{K}$ und $\vec{x} \in Y$ auch $\alpha\vec{x} = \vec{0} + \alpha\vec{x} \in Y$. Schließlich ist mit $\vec{x} \in Y$ auch $-\vec{x} = (-1)\vec{x} \in Y$. Die anderen Vektorraumeigenschaften übertragen sich von X auf Y. □

Bemerkungen und Beispiele.

- Ist X ein $\mathbb{K}$-Vektorraum, so sind $\{\vec{0}\} \subseteq X$ und X selbst Unterräume von X, die man als **triviale Unterräume** bezeichnet.
- Sind X ein $\mathbb{K}$-Vektorraum und $\vec{x} \in X$, so definiert

$$\mathbb{K} \cdot \vec{x} := \big\{\alpha\, \vec{x} \;\big|\; \alpha \in \mathbb{K}\big\} \tag{IV.16}$$

  einen Unterraum von X.
- Sind $\mathbb{K}$ ein Körper und $\vec{a} := (a_1, \ldots, a_N)^T \in \mathbb{K}^N$, so definiert

$$Y_{\vec{a}} := \big\{\vec{x} = (x_1, x_2, \ldots, x_N)^T \in \mathbb{K}^N \;\big|\; a_1x_1 + a_2x_2 + \ldots + a_Nx_N = 0\big\} \subseteq \mathbb{K}^N \tag{IV.17}$$

  einen Unterraum. Dabei notieren wir für

$$(x_1, x_2, \ldots, x_N)^T := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \tag{IV.18}$$

  den zum Spaltenvektor jeweils gehörigen Zeilenvektor. (Der Superskript “T” bedeutet *Transposition*, was später noch eingeführt wird.)

- Sei $\mathcal{F} = \{f : [\alpha, \beta] \to \mathbb{R}\}$ der in (IV.10) definierte $\mathbb{R}$-Vektorraum der reellen Funktionen auf $[\alpha, \beta]$, dann bilden die Mengen der auf $[\alpha, \beta]$ stetigen Funktionen oder der auf $[\alpha, \beta]$ stetigen und auf (α, β) differenzierbaren Funktionen in $\mathcal{F}$ jeweils einen Unterraum.

Lemma IV.4. *Sei X ein $\mathbb{K}$-Vektorraum.*

(i) Sind $I \neq \emptyset$ eine Indexmenge und $\{Y_i\}_{i \in I}$ eine Familie von Unterräumen von X, so ist auch ihr Durchschnitt $\bigcap_{i \in I} Y_i$ wieder ein Unterraum von X.

(ii) Sind $N \in \mathbb{N}$ und $\{Y_n\}_{n=1}^N$ eine endliche Familie von Unterräumen von X, so ist auch ihre ***Summe***

$$Y_1 + Y_2 + \ldots + Y_N := \Big\{ \vec{y}_1 + \vec{y}_2 + \ldots + \vec{y}_N \;\Big|\; \vec{y}_1 \in Y_1,\ \vec{y}_2 \in Y_2, \ldots, \vec{y}_N \in Y_N \Big\} \tag{IV.19}$$

wieder ein Unterraum von X.

Beweis. Das Lemma ist eine einfache Folgerung aus dem Unterraumkriterium. □

Bemerkungen und Beispiele.

- Sind $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_M \in \mathbb{K}^N$ mit $\vec{a}_i := (a_{i,1}, a_{i,2}, \ldots, a_{i,N})^T$, so definiert

$$Y_{\vec{a}_1, \ldots, \vec{a}_M} := Y_{\vec{a}_1} \cap \ldots \cap Y_{\vec{a}_M} \tag{IV.20}$$

  nach Lemma IV.4 (i) einen Unterraum in $\mathbb{K}^N$, der durch die Lösungen $\vec{x} = (x_1, \ldots, x_N)^T$ des Gleichungssystems

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \ldots + a_{1,N}x_N &= 0, \\ a_{2,1}x_1 + a_{2,2}x_2 + \ldots + a_{2,N}x_N &= 0, \\ &\vdots \\ a_{M,1}x_1 + a_{M,2}x_2 + \ldots + a_{M,N}x_N &= 0 \end{aligned} \tag{IV.21}$$

  gegeben ist.

- Sind X ein $\mathbb{K}$-Vektorraum und $\vec{x}_1, \ldots, \vec{x}_M \in X$, so definiert

$$\mathbb{K} \cdot \vec{x}_1 + \ldots + \mathbb{K} \cdot \vec{x}_M = \big\{ \alpha_1 \vec{x}_1 + \ldots + \alpha_M \vec{x}_M \;\big|\; \alpha_1, \ldots, \alpha_M \in \mathbb{K} \big\} \tag{IV.22}$$

  nach Lemma IV.4 (ii) einen Unterraum in X.

Definition IV.5. Sei X ein $\mathbb{K}$-Vektorraum.

(i) Ist $S \subseteq X$ eine Teilmenge in X, so heißt der durch

$$\operatorname{span}(S) := \bigcap \Big\{ Y \;\Big|\; S \subseteq Y \subseteq X,\ Y \text{ ist ein Unterraum} \Big\} \tag{IV.23}$$

definierte Unterraum **lineare Hülle von *S*** oder der **von *S* erzeugte Unterraum**.

(ii) Sind $Y \subseteq X$ ein Unterraum, $S \subseteq X$ eine Teilmenge in X und $Y = \mathrm{span}(S)$, so heißt S **Erzeugendensystem von $\boldsymbol{Y}$**.

Bemerkungen und Beispiele.

- $\mathrm{span}(S) \subseteq X$ ist der kleinste Unterraum in X, der S enthält. D.h. ist $Y \subseteq X$ ein Unterraum mit $Y \supseteq S$, so ist auch $Y \supseteq \mathrm{span}(S)$.
- Sind X ein $\mathbb{K}$-Vektorraum und $\vec{x}_1, \ldots, \vec{x}_M \in X$, so ist

$$\mathrm{span}\Big(\{\vec{x}_1, \ldots, \vec{x}_M\}\Big) \; = \; \mathbb{K} \cdot \vec{x}_1 + \ldots + \mathbb{K} \cdot \vec{x}_M. \tag{IV.24}$$

IV.3. Lineare Unabhängigkeit und Basen

Definition IV.6. Seien X ein $\mathbb{K}$-Vektorraum und $\{\vec{x}_1, \ldots, \vec{x}_M\} \subseteq X$ eine endliche Teilmenge.

(i) Sind $\alpha_1, \ldots, \alpha_M \in \mathbb{K}$, so heißt die Summe $\alpha_1 \vec{x}_1 + \ldots + \alpha_M \vec{x}_M \in X$ **Linearkombination** (der Vektoren $\vec{x}_1, \ldots, \vec{x}_M$). Speziell bezeichnet man $0 \cdot \vec{x}_1 + \ldots + 0 \cdot \vec{x}_M = \vec{0}$ als **triviale Linearkombination**.

(ii) Die endliche Teilmenge $\{\vec{x}_1, \ldots, \vec{x}_M\} \subseteq X$ heißt **linear unabhängig** $:\Leftrightarrow$

Der Nullvektor lässt sich nur als triviale Linearkombination der $\vec{x}_1, \ldots, \vec{x}_M$ darstellen, d.h.

$$\forall \alpha_1, \ldots, \alpha_M \in \mathbb{K}: \quad \Big\{\alpha_1 \vec{x}_1 + \ldots + \alpha_M \vec{x}_M \; = \; \vec{0}\Big\} \; \Rightarrow \; \Big\{\alpha_1 = \ldots = \alpha_M = 0\Big\}. \tag{IV.25}$$

Die leere Menge $\emptyset$ definieren wir als linear unabhängig.

(iii) Eine (nicht notwendig endliche) Teilmenge $S \subseteq X$ heißt **linear unabhängig** $:\Leftrightarrow$ Jede endliche Teilmenge von S ist linear unabhängig.

(iv) Ist $\{\vec{x}_1, \ldots, \vec{x}_M\}$ nicht linear unabhängig, so sagt man, die Menge $\{\vec{x}_1, \ldots, \vec{x}_M\}$ oder die Vektoren $\vec{x}_1, \ldots, \vec{x}_M$ seien **linear abhängig**.

Bemerkungen und Beispiele.

- Die lineare Abhängigkeit zweier Vektoren $\vec{x}_1, \vec{x}_2 \in X$ ist gleichwertig mit ihrer Parallelität, d.h. es gibt ein $\alpha \in \mathbb{K}$, so dass $\vec{x}_1 = \alpha \vec{x}_2$ oder dass $\vec{x}_2 = \alpha \vec{x}_1$ gilt.
- Die lineare Abhängigkeit dreier Vektoren $\vec{x}_1, \vec{x}_2, \vec{x}_3 \in X$ ist gleichwertig mit ihrer Planarität, d.h. es gibt Zahlen $\alpha, \beta \in \mathbb{K}$, so dass $\vec{x}_1 = \alpha \vec{x}_2 + \beta \vec{x}_3$, dass $\vec{x}_2 = \alpha \vec{x}_3 + \beta \vec{x}_1$ oder dass $\vec{x}_3 = \alpha \vec{x}_1 + \beta \vec{x}_2$ gilt.
- Ist $\{\vec{x}_1, \ldots, \vec{x}_M\}$ linear unabhängig, so gilt $\vec{x}_i \neq \vec{0}$, für alle $i = 1, \ldots, M$. Wäre nämlich z.B. $\vec{x}_i = \vec{0}$, dann gälte $\alpha_1 \vec{x}_1 + \ldots + \alpha_M \vec{x}_M = \vec{0}$ mit $\alpha_i = 1$ und $\alpha_1 = \ldots = \alpha_{i-1} = \alpha_{i+1} = \ldots = \alpha_M = 0$.
- Sei $X = \mathbb{R}^3$. Es soll festgestellt werden, ob die Vektoren

$$\vec{x}_1 \; = \; \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; \vec{x}_2 \; = \; \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \; \vec{x}_3 \; = \; \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \tag{IV.26}$$

linear unabhängig sind. Ist $\alpha_1\vec{x}_1+\alpha_2\vec{x}_2+\alpha_3\vec{x}_2 = \vec{0}$ eine Darstellung des Nullvektors als Linearkombination dieser Vektoren, so muss

$$\alpha_1 \begin{pmatrix}1\\0\\0\end{pmatrix} + \alpha_2 \begin{pmatrix}1\\2\\1\end{pmatrix} + \alpha_3 \begin{pmatrix}0\\1\\1\end{pmatrix} = \begin{pmatrix}\alpha_1+\alpha_2\\2\alpha_2+\alpha_3\\\alpha_2+\alpha_3\end{pmatrix} = \vec{0} = \begin{pmatrix}0\\0\\0\end{pmatrix} \tag{IV.27}$$

gelten, d.h. es müssen

$$\left.\begin{array}{rc} (i): & \alpha_1+\alpha_2=0 \\ (ii): & 2\alpha_2+\alpha_3=0 \\ (iii): & \alpha_2+\alpha_3=0 \end{array}\right\} \Rightarrow \left\{\begin{array}{rl} (ii)-(iii): & \alpha_2 = 0 \\ (i): & \alpha_1 = -\alpha_2 \\ (iii): & \alpha_3 = -\alpha_2 \end{array}\right. \tag{IV.28}$$

gelten. Also sind $\alpha_1 = \alpha_2 = \alpha_3 = 0$, und damit ist $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ linear unabhängig.

- Mit $\{\vec{x}_1, \ldots, \vec{x}_M\}$ sind auch alle ihre Teilmengen linear unabhängig. Sind z.B. $1 \leq k < M$ und $\alpha_1\vec{x}_1 + \ldots + \alpha_k\vec{x}_k = \vec{0}$, so ist mit $\alpha_{k+1} := \ldots := \alpha_M := 0$ auch $\alpha_1\vec{x}_1 + \ldots + \alpha_M\vec{x}_M = \vec{0}$, und daraus folgt dann $\alpha_1 = \ldots = \alpha_k = 0$.
- Sei $\vec{x} \in \mathrm{span}\{\vec{x}_1, \ldots, \vec{x}_M\} \subseteq X$. Dann lässt sich $\vec{x}$ als Linearkombination

$$\vec{x} = \alpha_1\,\vec{x}_1 + \ldots + \alpha_M\,\vec{x}_M \tag{IV.29}$$

  darstellen. Ist $\{\vec{x}_1, \ldots, \vec{x}_M\}$ linear unabhängig, so sind die Koeffizienten α_i eindeutig. Ist nämlich $\vec{x} = \beta_1\vec{x}_1 + \ldots + \beta_M\vec{x}_M$, so folgt aus

$$(\alpha_1 - \beta_1)\,\vec{x}_1 + \ldots + (\alpha_M - \beta_M)\,\vec{x}_M = \vec{0}, \tag{IV.30}$$

  dass $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_M = \beta_M$.
- Ist $\{\vec{x}_1, \ldots, \vec{x}_M\}$ linear abhängig, so gibt es in (IV.29) mehrere verschiedene Linearkombinationen, die $\vec{x}$ darstellen.
- Es gilt folgende Äquivalenz:

$$\begin{aligned} &\{\vec{x}_1, \ldots, \vec{x}_M\} \text{ ist linear abhängig} \\ \Leftrightarrow\quad & \exists\, i \in \mathbb{Z}_1^M,\ \beta_1, \ldots, \beta_{i-1}, \beta_{i+1}, \ldots, \beta_M : \\ & \quad \vec{x}_i = \beta_1\vec{x}_1 + \ldots + \beta_{i-1}\vec{x}_{i-1} + \beta_{i+1}\vec{x}_{i+1} + \ldots + \beta_M\vec{x}_M. \end{aligned} \tag{IV.31}$$

- Aus dieser Äquivalenz ergibt sich unmittelbar die folgende Äquivalenz:

$$\begin{aligned} &\{\vec{x}_1, \ldots, \vec{x}_M\} \text{ ist linear abhängig } \Leftrightarrow \\ &\exists i \in \mathbb{Z}_1^M : \ \mathrm{span}\Big(\{\vec{x}_1, \ldots, \vec{x}_{i-1}, \vec{x}_{i+1}, \ldots, \vec{x}_M\}\Big) = \mathrm{span}\Big(\{\vec{x}_1, \ldots, \vec{x}_M\}\Big). \end{aligned} \tag{IV.32}$$

Definition IV.7. Sei X ein $\mathbb{K}$-Vektorraum. Eine Teilmenge $S \subseteq X$ heißt **Basis von *X***

$$:\Leftrightarrow S \text{ ist linear unabhängig und } X = \mathrm{span}(S).$$

Satz IV.8. *Seien X ein $\mathbb{K}$-Vektorraum und $Y = \mathrm{span}(S) \subseteq X$ ein endlich erzeugter Unterraum, $|S| < \infty$. Dann gibt es eine Teilmenge $\widetilde{S} \subseteq S$, die eine Basis von Y bildet. Jede andere Basis von Y hat dieselbe Elementzahl.*

Beweis. Nach dem Ergänzungssatz IV.12 kann $M = \emptyset$ mit Vektoren aus S zu einer Basis $\widetilde{S}$ von Y ergänzt werden.

Ist $\widetilde{T}$ eine andere Basis von Y, so folgt wegen $Y = \mathrm{span}(\widetilde{T})$ nach Satz IV.13, dass $|\widetilde{T}| \geq |\widetilde{S}|$. Genauso folgt natürlich umgekehrt $|\widetilde{S}| \geq |\widetilde{T}|$, also

$$|\widetilde{T}| \ = \ |\widetilde{S}|. \tag{IV.33}$$

□

IV.4. Dimension

Definition IV.9. Seien X ein $\mathbb{K}$-Vektorraum und $Y \subseteq X$ ein Unterraum.

(i) Sind Y endlich erzeugt und S eine Basis von Y, so setzen wir

$$\dim(Y) \ := \ |S| \ < \ \infty. \tag{IV.34}$$

(ii) Ist Y nicht endlich erzeugt, so setzen wir

$$\dim(Y) \ := \ \infty. \tag{IV.35}$$

Die Zahl $\dim(Y)$ bezeichnet man als **Dimension von Y (über $\mathbb{K}$)**.

Bemerkungen und Beispiele.

- Es ist $\mathrm{span}(\emptyset)$ der Durchschnitt aller Unterräume, die $\emptyset$ enthalten.
  Also ist $\mathrm{span}(\emptyset) = \{\vec{0}\}$, und somit

$$\dim\big(\{\vec{0}\}\big) \ = \ 0. \tag{IV.36}$$

- Die **kanonischen Basisvektoren** $\vec{e}_1, \ldots, \vec{e}_N \in \mathbb{K}^N$,

$$\vec{e}_1 \ := \ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 \ := \ \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ \ldots, \ \vec{e}_N \ := \ \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \tag{IV.37}$$

  bilden eine Basis in $\mathbb{K}^N$, die so genannte **Standardbasis** $\mathcal{E}_N := \{\vec{e}_1, \ldots, \vec{e}_N\}$. Also ist

$$\dim\big(\mathbb{K}^N\big) \ = \ N. \tag{IV.38}$$

- Der Vektorraum $\mathcal{F}$ der reellen Funktionen auf $[a, b]$ ist unendlichdimensional,

$$\dim \mathcal{F} \ = \ \infty. \tag{IV.39}$$

  So sind z.B. mit $f_1(x) := 1$, $f_2(x) := x^1$, $f_3(x) := x^2$, $f_4(x) := x^3$, ... die Mengen $\big\{f_1, \ldots, f_K\big\}$ linear unabhängig für jedes (noch so große) $K \in \mathbb{N}$.

- Seien $Y', Y \subseteq X$ zwei endlich erzeugte Unterräume eines $\mathbb{K}$-Vektorraums X und $Y' \subseteq Y$. Aus dem Ergänzungssatz folgt sofort, dass
$$\dim(Y') \;\leq\; \dim(Y). \tag{IV.40}$$
- Gelten hingegen $Y' \subseteq Y$ und $\dim(Y') = \dim(Y) < \infty$, so folgt umgekehrt
$$Y' \;=\; Y. \tag{IV.41}$$
- Sind $Y = \mathrm{span}(S)$ endlich erzeugt und $|S| > \dim(Y)$, so ist S linear abhängig.
- Ist umgekehrt $S \subseteq Y$ eine linear unabhängige Teilmenge und $|S| = \dim(Y)$, so ist S eine Basis von Y.
- Ist $\dim(X) = \infty$, so gibt es in X linear unabhängige Teilmengen mit beliebig großer Elementzahl.
- Sind $\dim(X) = \infty$ und $S \subseteq X$ eine Basis, so ist
$$X \;=\; \{\alpha_1 \vec{x}_1 + \cdots + \alpha_N \vec{x}_N \mid N \in \mathbb{N}, \alpha_1, \ldots, \alpha_N \in \mathbb{K},\ \vec{x}_1, \ldots, \vec{x}_N \in S\}. \tag{IV.42}$$
D.h. X enthält alle (endlichen) Linearkombinationen von Vektoren aus S. Unendliche Linearkombinationen sind (zunächst) nicht definiert.

Satz IV.10. *Jeder Vektorraum besitzt eine Basis.*

Beweis. Anwendung des Zornschen Lemmas bzw. des Auswahlaxioms. □

Bemerkungen und Beispiele.

- Satz IV.10 gilt auch für Vektorräume, die nicht endlich erzeugt sind.

Satz IV.11 (Dimensionsformel)**.** *Seien $Y, Y' \subseteq X$ zwei Unterräume eines $\mathbb{K}$-Vektorraums X. Dann gilt die **Dimensionsformel***
$$\dim(Y) + \dim(Y') \;=\; \dim(Y \cap Y') + \dim(Y + Y'). \tag{IV.43}$$
(Wobei auf beiden Seiten ∞ steht, falls $\dim(Y) = \infty$ oder $\dim(Y') = \infty$.)

IV.5. Ergänzungen

IV.5.1. Ergänzungssatz und Austauschsatz von Steinitz

Der Beweis von Satz IV.8 benutzt die beiden folgenden wichtigen Aussagen:

Lemma IV.12 (Ergänzungssatz). *Seien* $\mathbb{K}$ *ein Körper,* X *ein* $\mathbb{K}$*-Vektorraum und* $Y \subseteq X$ *ein Unterraum. Seien weiterhin* $M = \{\vec{x}_1, \ldots, \vec{x}_k\} \subseteq Y$ *eine endliche, linear unabhängige (möglicherweise leere) Teilmenge und* $S = \{\vec{y}_1, \ldots, \vec{y}_n\} \subseteq Y$ *ein endliches Erzeugendensystem,* $Y = \mathrm{span}(S)$*. Dann gibt es eine Teilmenge* $\widetilde{S} \subseteq S$*, so dass* $M \cup \widetilde{S}$ *eine Basis von* Y *bildet.*

Beweis. Betrachte das System von Teilmengen

$$\mathcal{T} \ := \ \big\{T \subseteq M \cup S \;\big|\; T \supseteq M, \quad Y = \mathrm{span}(T)\big\}. \tag{IV.44}$$

Wegen $M \cup S \in \mathcal{T}$ ist $\mathcal{T} \neq \emptyset$. Weiterhin gilt offensichtlich $k \leq |T| \leq k + n$ für jedes $T \in \mathcal{T}$. Sei nun $T \in \mathcal{T}$ eine Menge mit minimaler Elementzahl. Dann erzeugt T den Unterraum Y. Zeigen wir nun noch, dass T linear unabhängig ist, so bildet T eine Basis von Y. Seien also $\vec{y}_1, \ldots, \vec{y}_\ell \in S$ so, dass $T = \{\vec{x}_1, \ldots, \vec{x}_k, \vec{y}_1, \ldots, \vec{y}_\ell\}$. Seien weiterhin $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_\ell \in \mathbb{K}$, so dass

$$\alpha_1 \vec{x}_1 + \ldots + \alpha_k \vec{x}_k \ + \ \beta_1 \vec{y}_1 + \ldots + \beta_\ell \vec{y}_\ell \ = \ \vec{0}. \tag{IV.45}$$

Wäre etwa $\beta_i \neq 0$, so ließe sich $\vec{y}_i$ als Linearkombination der Vektoren in $\widetilde{T} := T \setminus \{\vec{y}_i\}$ darstellen. Dann wäre aber $\mathrm{span}(\widetilde{T}) = \mathrm{span}(T) = Y$ und somit $\widetilde{T} \in \mathcal{T}$. Andererseits wäre dann $|\widetilde{T}| < |T|$, und weil die Zahl der Elemente von T minimal ist, folgte auch $|\widetilde{T}| \geq |T|$. Widerspruch. Es folgt also, dass

$$\beta_1 \ = \ \beta_2 \ = \ \ldots \ = \ \beta_\ell \ = \ 0, \tag{IV.46}$$

was dann

$$\alpha_1 \vec{x}_1 + \ldots + \alpha_k \vec{x}_k \ = \ \vec{0} \tag{IV.47}$$

impliziert. Aus der linearen Unabhängigkeit von M erhalten wir dann auch noch

$$\alpha_1 \ = \ \alpha_2 \ = \ \ldots \ = \ \alpha_k \ = \ 0. \tag{IV.48}$$

Somit ist T linear unabhängig. □

Satz IV.13 (Austauschsatz von Steinitz). *Seien* $\mathbb{K}$ *ein Körper,* X *ein* $\mathbb{K}$*-Vektorraum und* $Y \subseteq X$ *ein Unterraum mit Basis* $\{\vec{x}_1, \ldots, \vec{x}_k\} \subseteq Y$*. Sei weiterhin* $S \subseteq Y$ *ein Erzeugendensystem von* $Y = \mathrm{span}(S)$*. Dann gibt es einen Vektor* $\vec{y} \in S$*, so dass auch* $\{\vec{y}, \vec{x}_2, \ldots, \vec{x}_k\}$ *eine Basis von* Y *bildet.*

Beweis. Setzen wir $M := \{\vec{x}_2, \ldots, \vec{x}_k\}$, so ist $M \subseteq Y$ linear unabhängig und kann nach Lemma IV.12 mit $\widetilde{S} \subseteq S$ zu einer Basis ergänzt werden. Dabei ist $\widetilde{S} \neq \emptyset$, da

$\vec{x}_1 \notin \operatorname{span}(M)$. Nehmen wir an, $\widetilde{S}$ enthielte mindestens zwei Vektoren $\vec{y}, \vec{z} \neq \vec{0}$, $\vec{y} \neq \vec{z}$. Weil $\{\vec{x}_1, \ldots, \vec{x}_k\}$ eine Basis in Y ist, gäbe es eindeutige α_i und β_i so, dass

$$\vec{y} \;=\alpha_1 \vec{x}_1 \;+\; \ldots \;+\; \alpha_k \vec{x}_k, \tag{IV.49}$$

$$\vec{z} \;=\beta_1 \vec{x}_1 \;+\; \ldots \;+\; \beta_k \vec{x}_k. \tag{IV.50}$$

Da $\vec{y}, \vec{z} \notin \operatorname{span}(M)$, wären $\alpha_1, \beta_1 \neq 0$. Damit wäre

$$-\beta_1 \vec{y} \;+\; \alpha_1 \vec{z} \;+\; (\beta_1\alpha_2 - \alpha_1\beta_2)\vec{x}_2 \;+\; \ldots \;+\; (\beta_1\alpha_k - \alpha_1\beta_k)\vec{x}_k \;=\; \vec{0}\,, \tag{IV.51}$$

und aus der linearen Unabhängigkeit von $\{\vec{y}, \vec{z}, \vec{x}_2, \ldots, \vec{x}_k\}$ folgte, dass alle Koeffizienten verschwinden und insbesondere auch $\beta_1 = \alpha_1 = 0$ gälte, was in Widerspruch zu $\alpha_1, \beta_1 \neq 0$ stünde. Also enthält $\widetilde{S}$ genau ein Element. □

IV.5.2. Beweis der Dimensionsformel

Beweis. In dieser Ergänzung führen wir den Beweis von Satz IV.11, d.h. wir beweisen die Dimensionsformel (IV.43). Dazu können wir o.B.d.A. $\dim(Y), \dim(Y') < \infty$ annehmen. Dann sind auch

$$\dim(Y \cap Y') \;\leq\; \dim(Y) \;<\; \infty, \tag{IV.52}$$

$$\dim(Y + Y') \;\leq\; \dim\Big(\operatorname{span}\big(\{Y \cup Y'\}\big)\Big) \;\leq\; \dim(Y) + \dim(Y') \;<\; \infty. \tag{IV.53}$$

Sei nun $\{\vec{x}_1, \ldots, \vec{x}_r\}$ eine Basis von $Y \cap Y'$. Nach dem Ergänzungssatz IV.12 gibt es Vektoren $\vec{y}_1, \ldots, \vec{y}_m \in Y$ und $\vec{z}_1, \ldots, \vec{z}_n \in Y'$, so dass $\{\vec{x}_1, \ldots, \vec{x}_r, \vec{y}_1, \ldots, \vec{y}_m\}$ eine Basis von Y und $\{\vec{x}_1, \ldots, \vec{x}_r, \vec{z}_1, \ldots, \vec{z}_n\}$ eine Basis von Y' bilden. Offenbar ist dann auch

$$Y + Y' \;=\; \operatorname{span}\big\{\vec{x}_1, \ldots, \vec{x}_r,\, \vec{y}_1, \ldots, \vec{y}_m,\, \vec{z}_1, \ldots, \vec{z}_n\big\}. \tag{IV.54}$$

Wir zeigen, dass diese Menge linear unabhängig ist. Seien $\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_m \in \mathbb{K}$, sowie $\gamma_1, \ldots, \gamma_n \in \mathbb{K}$, so dass

$$\alpha_1\vec{x}_1 + \ldots + \alpha_r\vec{x}_r + \beta_1\vec{y}_1 + \ldots + \beta_m\vec{y}_m + \gamma_1\vec{z}_1 + \ldots + \gamma_n\vec{z}_n \;=\; 0, \tag{IV.55}$$

bzw.

$$\vec{w} \;:=\; -\big(\beta_1\vec{y}_1 + \ldots + \beta_m\vec{y}_m\big) \;=\; \alpha_1\vec{x}_1 + \ldots + \alpha_r\vec{x}_r + \gamma_1\vec{z}_1 + \ldots + \gamma_n\vec{z}_n. \tag{IV.56}$$

Aus der ersten Gleichung in (IV.56) folgt, dass $\vec{w} \in Y$, und aus der zweiten, dass $\vec{w} \in Y'$. Also ist $\vec{w} = -(\beta_1\vec{y}_1 + \ldots + \beta_m\vec{y}_m) \in Y \cap Y'$ und kann als Linearkombination der $\vec{x}_i$ dargestellt werden, d.h. es gibt $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_r \in \mathbb{K}$, so dass

$$\beta_1\vec{y}_1 + \ldots + \beta_m\vec{y}_m + \tilde{\alpha}_1\vec{x}_1 + \ldots + \tilde{\alpha}_r\vec{x}_r \;=\; 0. \tag{IV.57}$$

Da $\{\vec{y}_1, \ldots, \vec{y}_m,\; \vec{x}_1, \ldots, \vec{x}_r\}$ eine Basis in Y ist, folgt daraus jedoch, dass $\tilde{\alpha}_1 = \cdots = \tilde{\alpha}_r = 0$ und insbesondere

$$\beta_1 \;=\; \beta_2 \;=\; \cdots \;=\; \beta_m \;=\; 0 \tag{IV.58}$$

sowie

$$\vec{0} \;=\; \vec{w} \;=\; \alpha_1 \vec{x}_1 + \ldots + \alpha_r \vec{x}_r + \gamma_1 \vec{z}_1 + \ldots + \gamma_n \vec{z}_n. \tag{IV.59}$$

Da aber $\{\vec{x}_1, \ldots, \vec{x}_r, \vec{z}_1, \ldots, \vec{z}_n\}$ eine Basis von Y' bilden, folgt daraus, dass auch

$$\alpha_1 \;=\; \alpha_2 \;=\; \cdots \;=\; \alpha_r \;=\; \gamma_1 \;=\; \gamma_2 \;=\; \cdots \;=\; \gamma_n \;=\; 0. \tag{IV.60}$$

Also ist $\{\vec{x}_1, \ldots, \vec{x}_r,\, \vec{y}_1, \ldots, \vec{y}_m,\, \vec{z}_1, \ldots, \vec{z}_n\}$ eine Basis von $Y + Y'$, und (IV.43) ergibt sich nun durch Abzählen aus

$$\dim(Y \cap Y') \;=\; r, \qquad \dim(Y + Y') \;=\; r + m + n, \tag{IV.61}$$

$$\dim(Y) \;=\; r + m, \qquad \dim(Y') \;=\; r + n. \tag{IV.62}$$

$\square$

V. Lineare Abbildungen

V.1. Definitionen

Definition V.1. Seien X und Y zwei $\mathbb{K}$-Vektorräume. Eine Abbildung $\Phi : X \to Y$ heißt **linear** oder **Homomorphismus**

$$:\Leftrightarrow \ \forall\, \vec{x}, \vec{y} \in X,\ \alpha \in \mathbb{K} : \quad \Phi(\alpha\vec{x} + \vec{y}) \ = \ \alpha\Phi(\vec{x}) + \Phi(\vec{y}). \tag{V.1}$$

Die Menge der linearen Abbildungen $X \to Y$ bezeichnen wir mit $\mathcal{L}(X;Y)$. Ist $Y = X$, so schreiben wir $\mathcal{L}(X) := \mathcal{L}(X;X)$

Bemerkungen und Beispiele.

- Durch vollständige Induktion erhält man aus (V.1) leicht, dass
$$\Phi\Big(\sum_{i=1}^{k} \alpha_i \vec{x}_i\Big) \ = \ \sum_{i=1}^{k} \alpha_i \Phi(\vec{x}_i), \tag{V.2}$$
für alle $\vec{x}_1, \ldots, \vec{x}_k \in X$ und $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$.
- $\Phi(\vec{0}) \ = \ \Phi(0 \cdot \vec{0}) \ = \ 0 \cdot \Phi(\vec{0}) \ = \ \vec{0}$.
- Sind $\vec{x}_1, \ldots, \vec{x}_k \in X$ linear abhängig, so sind auch $\Phi(\vec{x}_1), \ldots, \Phi(\vec{x}_k) \in Y$ linear abhängig, denn
$$\sum_{i=1}^{k} \alpha_i \vec{x}_i \ = \ \vec{0} \qquad \Rightarrow \qquad \sum_{i=1}^{k} \alpha_i \Phi(\vec{x}_i) \ = \ \Phi\Big(\sum_{i=1}^{k} \alpha_i \vec{x}_i\Big) \ = \ \vec{0}. \tag{V.3}$$
- Ist $U \subseteq X$ ein Unterraum, so ist auch $\Phi(U) \subseteq Y$ ein Unterraum.
- Ist $U = \operatorname{span}\{\vec{x}_1, \ldots, \vec{x}_k\}$, so ist $\Phi(U) = \operatorname{span}\{\Phi(\vec{x}_1), \ldots, \Phi(\vec{x}_k)\}$, denn für $\vec{y} \in \Phi(U)$ gibt es ein $\vec{x} \in U$, so dass $\vec{y} = \Phi(\vec{x})$. Weiterhin ist $\vec{x} = \alpha_1\vec{x}_1 + \ldots + \alpha_k\vec{x}_k$, für geeignete $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$. Damit ist dann aber
$$\vec{y} \ = \ \Phi\Big(\sum_{i=1}^{k} \alpha_i \vec{x}_i\Big) \ = \ \sum_{i=1}^{k} \alpha_i \Phi(\vec{x}_i) \ \in \ \operatorname{span}\big\{\Phi(\vec{x}_1), \ldots, \Phi(\vec{x}_k)\big\}. \tag{V.4}$$

- Letzteres ergibt auch sofort $\dim[\Phi(U)] \leq \dim[U]$.
- Sind X und Y zwei $\mathbb{K}$-Vektorräume, so bildet die Menge $\mathcal{L}(X;Y)$ der linearen Abbildungen selbst einen $\mathbb{K}$-Vektorraum bezüglich punktweiser Operationen, d.h. für $\alpha \in \mathbb{K}$ und $\Phi, \Psi \in \mathcal{L}(X;Y)$ ist $(\alpha\Phi + \Psi) \in \mathcal{L}(X;Y)$ definiert durch
$$\forall\, \vec{x} \in X: \quad (\alpha\Phi + \Psi)(\vec{x}) \ := \ \alpha\Phi(\vec{x}) + \Psi(\vec{x}). \tag{V.5}$$
- Sind X, Y, Z drei $\mathbb{K}$-Vektorräume und $\Phi \in \mathcal{L}(X;Y)$ sowie $\Psi \in \mathcal{L}(Y;Z)$, so ist auch deren Komposition linear, $(\Psi \circ \Phi) \in \mathcal{L}(X;Z)$.
- Insbesondere ist für $\Phi, \Psi \in \mathcal{L}(X)$ auch $(\Psi \circ \Phi) \in \mathcal{L}(X)$.

Satz V.2. *Seien X und Y zwei $\mathbb{K}$-Vektorräume. Sind $\{\vec{x}_1, \ldots, \vec{x}_N\} \subseteq X$ eine Basis und $\{\vec{y}_1, \ldots, \vec{y}_N\} \subseteq Y$ eine Teilmenge, dann gibt es genau eine lineare Abbildung $\Phi \in \mathcal{L}(X;Y)$ so, dass*

$$\Phi(\vec{x}_1) \ = \ \vec{y}_1, \quad \Phi(\vec{x}_2) \ = \ \vec{y}_2, \quad \ldots, \quad \Phi(\vec{x}_N) \ = \ \vec{y}_N. \tag{V.6}$$

Beweis. Existenz: Für $\vec{x} = \alpha_1\vec{x}_1 + \ldots + \alpha_N\vec{x}_N$ definieren wir

$$\Phi(\vec{x}) := \alpha_1\vec{y}_1 + \ldots + \alpha_N\vec{y}_N. \tag{V.7}$$

Wegen der linearen Unabhängigkeit von $\{\vec{x}_1, \ldots, \vec{x}_N\} \subseteq X$ sind die Koeffizienten $\alpha_1, \ldots, \alpha_N \in \mathbb{K}$ in (V.7) eindeutig, und Φ ist wohldefiniert. Weiterhin erfüllt Φ offensichtlich die Gleichungen (V.6).

Seien nun $\vec{x}, \vec{x}' \in X$ und $\beta \in \mathbb{K}$. Dann gibt es $\alpha_1, \ldots, \alpha_N,\ \alpha_1', \ldots, \alpha_N' \in \mathbb{K}$, so dass

$$\vec{x} \ = \ \alpha_1\vec{x}_1 + \cdots + \alpha_N\vec{x}_N, \quad \vec{x}' \ = \ \alpha_1'\vec{x}_1 + \ldots + \alpha_N'\vec{x}_N. \tag{V.8}$$

Daher gilt auch

$$\beta\vec{x} + \vec{x}' \ = \ (\beta\alpha_1 + \alpha_1')\vec{x}_1 + \ldots + (\beta\alpha_N + \alpha_N')\vec{x}_N. \tag{V.9}$$

Also ist

$$\Phi(\beta\vec{x} + \vec{x}') \ = \ \sum_{i=1}^N (\beta\alpha_i + \alpha_i')\vec{y}_i \ = \ \beta\Big(\sum_{i=1}^N \alpha_i\vec{y}_i\Big) + \Big(\sum_{i=1}^N \alpha_i'\vec{y}_i\Big) \ = \ \beta\Phi(\vec{x}) + \Phi(\vec{x}'), \tag{V.10}$$

und Φ ist linear.

Eindeutigkeit: Sei $\widetilde{\Phi} \in \mathcal{L}(X;Y)$ mit $\widetilde{\Phi}(\vec{x}_1) = \vec{y}_1, \ldots, \widetilde{\Phi}(\vec{x}_N) = \vec{y}_N$. Ist $\vec{x} \in X$, so gibt es eindeutige $\alpha_1, \ldots, \alpha_N \in \mathbb{K}$ so, dass $\vec{x} = \alpha_1\vec{x}_1 + \ldots + \alpha_N\vec{x}_N$. Wegen der Linearität von sowohl Φ als auch $\widetilde{\Phi}$ ist dann

$$\Phi(\vec{x}) \ = \ \sum_{i=1}^N \alpha_i\vec{y}_i \ = \ \widetilde{\Phi}(\vec{x}). \tag{V.11}$$

Dies gilt für alle $\vec{x} \in X$, also ist $\Phi = \widetilde{\Phi}$. □

Bemerkungen und Beispiele.

- Satz V.2 besagt, dass jede lineare Abbildung Φ durch die Bilder $\Phi(\vec{x}_i)$ der Basisvektoren $\vec{x}_i$ eindeutig bestimmt ist.
- $\mathbb{K} = \mathbb{R}$, $X = \mathbb{R}^3$ mit Standardbasis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\} \subseteq X$,

$$\vec{e}_1 \; = \; \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 \; = \; \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{e}_3 \; = \; \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}; \tag{V.12}$$

  $Y \; = \; \mathbb{R}^2$ und

$$\vec{b}_1 := \binom{2}{1}, \quad \vec{b}_2 := \binom{1}{1}, \quad \vec{b}_3 := \binom{1}{0}. \tag{V.13}$$

  Führen wir $x, y, z \in \mathbb{R}$ als Koordinaten (statt $\alpha_1, \alpha_2, \alpha_3$) ein, so wird die lineare Abbildung $\Phi \in \mathcal{L}(X; Y)$, für die $\Phi(\vec{e}_i) = \vec{b}$ gilt, zu

$$\Phi\left[\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right] \; = \; \Phi\big(x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3\big) \; = \; x\vec{b}_1 + y\vec{b}_2 + z\vec{b}_3 \; = \; \binom{2x + y + z}{x + y}. \tag{V.14}$$

V.2. Kern und Bild

Definition V.3. Seien X und Y zwei $\mathbb{K}$-Vektorräume und $\Phi \in \mathcal{L}(X; Y)$ eine lineare Abbildung. Dann heißen

$$\mathrm{Ker}[\Phi] \; := \; \big\{\vec{x} \in X \;\big|\; \Phi(\vec{x}) = \vec{0}\big\} \tag{V.15}$$

der **Kern von Φ** und

$$\mathrm{Ran}[\Phi] \; := \; \Phi(X) \; = \; \big\{\vec{y} \in Y \;\big|\; \exists\, \vec{x} \in X : \vec{y} = \Phi(\vec{x})\big\} \tag{V.16}$$

das **Bild von Φ**[1].

Als einfache Anwendung des Unterraumkriteriums ergibt sich sofort

Lemma V.4. *Seien X und Y zwei $\mathbb{K}$-Vektorräume und $\Phi \in \mathcal{L}(X; Y)$ eine lineare Abbildung. Dann sind $\mathrm{Ker}[\Phi] \subseteq X$ und $\mathrm{Ran}[\Phi] \subseteq Y$ Unterräume.*

Satz V.5. *Seien X und Y zwei $\mathbb{K}$-Vektorräume und $\Phi \in \mathcal{L}(X; Y)$ eine lineare Abbildung. Dann gilt*

$$\dim\big(\mathrm{Ker}[\Phi]\big) + \dim\big(\mathrm{Ran}[\Phi]\big) \; = \; \dim(X). \tag{V.17}$$

[1] Ker steht für "kernel", Ran steht für "range".

Beweis. Wir zeigen (V.17) nur für $\dim(X) = N < \infty$. Ist $\{\vec{x}_1, \ldots, \vec{x}_k\} \subseteq \mathrm{Ker}[\Phi]$ eine Basis, so können wir diese durch $\{\vec{x}_{k+1}, \ldots, \vec{x}_N\} \subseteq X$ zu einer Basis von X ergänzen. Wegen

$$\Phi\Big(\sum_{i=1}^N \alpha_i \vec{x}_i\Big) \; = \; \sum_{i=1}^N \alpha_i \Phi(\vec{x}_i) \; = \; \sum_{i=k+1}^N \alpha_i \Phi(\vec{x}_i) \tag{V.18}$$

ist $\mathrm{Ran}[\Phi] = \mathrm{span}\{\Phi(\vec{x}_{k+1}), \ldots, \Phi(\vec{x}_N)\}$. Sind nun $\alpha_{k+1}, \ldots, \alpha_N \in \mathbb{K}$ Koeffizienten, sodass $\alpha_{k+1}\Phi(\vec{x}_{k+1}) + \ldots + \alpha_N \Phi(\vec{x}_N) = \vec{0}$, so folgt $\vec{0} = \Phi(\alpha_{k+1}\vec{x}_{k+1} + \ldots + \alpha_N \vec{x}_N)$. Also ist

$$\alpha_{k+1}\vec{x}_{k+1} + \ldots + \alpha_N \vec{x}_N \in \mathrm{Ker}[\Phi]. \tag{V.19}$$

Somit gibt es $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$, so dass

$$\alpha_{k+1}\vec{x}_{k+1} + \ldots + \alpha_N \vec{x}_N \; = \; \alpha_1 \vec{x}_1 + \ldots + \alpha_k \vec{x}_k. \tag{V.20}$$

Wegen der linearen Unabhängigkeit von $\{\vec{x}_1, \ldots, \vec{x}_N\}$ ergibt sich daraus $\alpha_1 = \alpha_2 = \cdots = \alpha_N = 0$. Insbesondere verschwinden $\alpha_{k+1} = \cdots = \alpha_N = 0$. Damit ist jedoch die Menge $\{\Phi(\vec{x}_{k+1}), \ldots, \Phi(\vec{x}_N)\} \subseteq \mathrm{Ran}[\Phi]$ eine Basis, und (V.17) folgt trivial. $\square$

Bemerkungen und Beispiele.

- Seien wieder $X = \mathbb{R}^3$, $Y = \mathbb{R}^2$ und $\Phi\big[(x, y, z)^T\big] := (2x + y + z, x + y)^T$, wie in (V.14).
  - Für $\vec{b} = (\beta_1, \beta_2)^T \in \mathbb{R}^2$ ist dann $\vec{b} = \Phi[\vec{x}] \in \mathrm{Ran}[\Phi]$ mit $\vec{x} = (0, \beta_2, \beta_1 - \beta_2)^T$. Da $\vec{b} \in \mathbb{R}^2$ beliebig gewählt werden kann, folgt daraus
  $$\mathrm{Ran}[\Phi] \; = \; \mathbb{R}^2. \tag{V.21}$$
  - Ist $\vec{a} = (\alpha_1, \alpha_2, \alpha_3)^T \in \mathrm{Ker}[\Phi]$, so gelten $2\alpha_1 + \alpha_2 + \alpha_3 = 0$ und $\alpha_1 + \alpha_2 = 0$. Wählen wir $\gamma =: \alpha_1 \in \mathbb{R}$ beliebig, so folgt dass $\alpha_2 = -\alpha_1 = -\gamma$ und dass $\alpha_3 = -2\alpha_1 - \alpha_2 = -\gamma$, d.h. dass $\vec{a} = (\gamma, -\gamma, -\gamma)^T$. Da $\gamma \in \mathbb{R}$ frei gewählt werden kann, ergibt sich daraus, dass
  $$\mathrm{Ker}[\Phi] \; = \; \mathbb{R} \cdot \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}. \tag{V.22}$$
  - Wie in Satz V.5 behauptet, sind $\dim \mathrm{Ker}[\Phi] + \dim \mathrm{Ran}[\Phi] = 1 + 2 = 3 = \dim(\mathbb{R}^3)$.
- Seien $\mathbb{K} = \mathbb{R}$, $X = \mathbb{R}^3$, $Y = \mathbb{R}^3$,
$$\vec{x}_1 \; := \; \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{x}_2 \; := \; \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \quad \vec{x}_3 \; := \; \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}, \tag{V.23}$$
$$\vec{y}_1 \; := \; \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{y}_2 \; := \; \begin{pmatrix} 4 \\ 0 \\ 4 \end{pmatrix}, \quad \vec{y}_3 \; := \; \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}, \tag{V.24}$$

und $\Phi \in \mathcal{L}(\mathbb{R}^3; \mathbb{R}^3)$ mit

$$\Phi(\vec{x}_1) \;:=\; \vec{y}_1, \quad \Phi(\vec{x}_2) \;:=\; \vec{y}_2, \quad \Phi(\vec{x}_3) \;:=\; \vec{y}_3. \tag{V.25}$$

Wir wollen

$$\Phi\left[\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix}\right], \quad \mathrm{Ker}[\Phi], \quad \text{und} \quad \mathrm{Ran}[\Phi] \tag{V.26}$$

für beliebige $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ berechnen. Dazu bestimmen wir zunächst $\eta_1, \eta_2, \eta_3 \in \mathbb{R}$ so, dass $\vec{e}_1 = \eta_1\vec{x}_1 + \eta_2\vec{x}_2 + \eta_3\vec{x}_3$, d.h. wir lösen

$$\begin{pmatrix}1\\0\\0\end{pmatrix} \;=\; \begin{pmatrix} \eta_1 & + & 0 & + & 4\eta_3\\ \eta_1 & + & 2\eta_2 & + & 0\\ 0 & + & \eta_2 & + & \eta_3\end{pmatrix}. \tag{V.27}$$

Dies führt auf $\eta_1 = -2\eta_2 = 2\eta_3$ und $1 = \eta_1 + 4\eta_3 = 3\eta_1$, also $\eta_1 = \frac{1}{3}$, $\eta_2 = -\frac{1}{6}$ und $\eta_3 = \frac{1}{6}$. Somit ist

$$\vec{e}_1 \;=\; \frac{1}{3}\vec{x}_1 - \frac{1}{6}\vec{x}_2 + \frac{1}{6}\vec{x}_3. \tag{V.28}$$

Genauso erhalten wir

$$\vec{e}_2 \;=\; \frac{2}{3}\vec{x}_1 + \frac{1}{6}\vec{x}_2 - \frac{1}{6}\vec{x}_3, \tag{V.29}$$

$$\vec{e}_3 \;=\; -\frac{4}{3}\vec{x}_1 + \frac{2}{3}\vec{x}_2 + \frac{1}{3}\vec{x}_3. \tag{V.30}$$

Somit sind $X = \mathrm{span}(\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}) \subseteq \mathrm{span}(\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}) \subseteq X$ und insbesondere $\mathrm{span}(\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}) = X$. Aus $|\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}| = 3$ folgt dann, dass $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\} \subseteq X$ eine Basis ist.

Gleichungen (V.28)-(V.30) setzen wir ein, nutzen die Linearität von Φ aus und erhalten

$$\begin{aligned}
\Phi\left[\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix}\right] &= \Phi\big(\alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \alpha_3\vec{e}_3\big) \qquad \text{(V.31)}\\
&= \alpha_1\Phi\big(\vec{e}_1\big) + \alpha_2\Phi\big(\vec{e}_2\big) + \alpha_3\Phi\big(\vec{e}_3\big)\\
&= \alpha_1\Phi\big(\tfrac{1}{3}\vec{x}_1 - \tfrac{1}{6}\vec{x}_2 + \tfrac{1}{6}\vec{x}_3\big) + \alpha_2\Phi\big(\tfrac{2}{3}\vec{x}_1 + \tfrac{1}{6}\vec{x}_2 - \tfrac{1}{6}\vec{x}_3\big)\\
&\qquad + \alpha_3\Phi\big(-\tfrac{4}{3}\vec{x}_1 + \tfrac{2}{3}\vec{x}_2 + \tfrac{1}{3}\vec{x}_3\big)\\
&= \big(\tfrac{1}{3}\alpha_1 + \tfrac{2}{3}\alpha_2 - \tfrac{4}{3}\alpha_3\big)\,\Phi\big(\vec{x}_1\big) + \big(-\tfrac{1}{6}\alpha_1 + \tfrac{1}{6}\alpha_2 + \tfrac{2}{3}\alpha_3\big)\,\Phi\big(\vec{x}_2\big)\\
&\qquad + \big(\tfrac{1}{6}\alpha_1 - \tfrac{1}{6}\alpha_2 + \tfrac{1}{3}\alpha_3\big)\,\Phi\big(\vec{x}_3\big)\,,
\end{aligned}$$

also

$$\Phi\left[\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix}\right] = \big(\tfrac{1}{3}\alpha_1 + \tfrac{2}{3}\alpha_2 - \tfrac{4}{3}\alpha_3\big)\begin{pmatrix}2\\1\\0\end{pmatrix} + \big(-\tfrac{1}{6}\alpha_1 + \tfrac{1}{6}\alpha_2 + \tfrac{2}{3}\alpha_3\big)\begin{pmatrix}4\\0\\4\end{pmatrix} \quad (V.32)$$

$$+ \big(\tfrac{1}{6}\alpha_1 - \tfrac{1}{6}\alpha_2 + \tfrac{1}{3}\alpha_3\big)\begin{pmatrix}2\\-1\\4\end{pmatrix}$$

$$= \begin{pmatrix}\tfrac{1}{3}\alpha_1 + \tfrac{5}{3}\alpha_2 + \tfrac{2}{3}\alpha_3\\ \tfrac{1}{6}\alpha_1 + \tfrac{5}{6}\alpha_2 - \tfrac{5}{3}\alpha_3\\ 4\alpha_3\end{pmatrix}.$$

Weiterhin sehen wir, dass

$$\vec{x} = \begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} \in \mathrm{Ker}[\Phi] \quad\Leftrightarrow\quad \begin{pmatrix}\tfrac{1}{3}\alpha_1 + \tfrac{5}{3}\alpha_2 + \tfrac{2}{3}\alpha_3\\ \tfrac{1}{6}\alpha_1 + \tfrac{5}{6}\alpha_2 - \tfrac{5}{3}\alpha_3\\ 4\alpha_3\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}, \qquad (V.33)$$

woraus $\alpha_3 = 0$ und $\alpha_2 = -\frac{1}{5}\alpha_1$ folgen. Da keine weiteren Einschränkungen vorliegen, gilt somit

$$\vec{x} = \begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} \in \mathrm{Ker}[\Phi] \quad\Leftrightarrow\quad \vec{x} = \begin{pmatrix}\alpha_1\\ -\frac{1}{5}\alpha_1\\ 0\end{pmatrix},\ \alpha_1 \in \mathbb{R} \text{ beliebig wählbar, d.h.} \qquad (V.34)$$

$$\mathrm{Ker}[\Phi] = \mathbb{R}\cdot\begin{pmatrix}5\\-1\\0\end{pmatrix}, \quad \dim\big(\mathrm{Ker}[\Phi]\big) = 1. \qquad (V.35)$$

Nach Satz V.5 ist damit $\dim\big(\mathrm{Ran}[\Phi]\big) = 2$ und daher *muss* $\{\vec{y}_1, \vec{y}_2, \vec{y}_3\} \subseteq \mathbb{R}^3$ linear abhängig sein. In der Tat sieht man leicht, dass $\vec{y}_3 = \vec{y}_2 - \vec{y}_1$ und somit ist

$$\mathrm{Ran}[\Phi] = \mathrm{span}\big(\{\vec{y}_1, \vec{y}_2, \vec{y}_3\}\big) = \mathrm{span}\big(\{\vec{y}_1, \vec{y}_2\}\big) \qquad (V.36)$$

$$= \big\{\beta_1\vec{y}_1 + \beta_2\vec{y}_2 \big|\ \beta_1, \beta_2 \in \mathbb{R}\big\} = \left\{\begin{pmatrix}2\beta_1 + 4\beta_2\\ \beta_1\\ 4\beta_2\end{pmatrix}\ \middle|\ \beta_1, \beta_2 \in \mathbb{R}\right\}.$$

Weiterhin *muss* $\{\vec{y}_1, \vec{y}_2\} \subseteq \mathrm{Ran}[\Phi]$ wegen $\dim\big(\mathrm{Ran}[\Phi]\big) = 2$ auch eine Basis und insbesondere linear unabhängig sein.

- Satz V.5 ist auch für $\dim(X) = \infty$ richtig, d.h. für $\dim(X) = \infty$ gilt $\dim(\mathrm{Ker}[\Phi]) = \infty$ oder $\dim(\mathrm{Ran}[\Phi]) = \infty$ (oder beides).
- Wir beobachten, dass für $\Phi \in \mathcal{L}(X;Y)$

$$\Phi(\vec{x}) = \Phi(\vec{x}\,') \Leftrightarrow \qquad \Phi(\vec{x} - \vec{x}\,') = \vec{0}, \qquad (V.37)$$

$$\vec{x} = \vec{x}\,' \Leftrightarrow \qquad \vec{x} - \vec{x}\,' = \vec{0}. \qquad (V.38)$$

Daher gilt

$$\Big(\Phi \in \mathcal{L}(X;Y) \text{ ist injektiv}\Big) \quad \Leftrightarrow \quad \Big(\mathrm{Ker}[\Phi] = \{\vec{0}\}\Big). \tag{V.39}$$

- Offensichtlich gilt außerdem

$$\Big(\Phi \in \mathcal{L}(X;Y) \text{ ist surjektiv}\Big) \quad \Leftrightarrow \quad \Big(\mathrm{Ran}[\Phi] = Y\Big). \tag{V.40}$$

Satz V.6. *Seien X und Y zwei $\mathbb{K}$-Vektorräume mit $\dim(X) = \dim(Y) < \infty$ und $\Phi \in \mathcal{L}(X;Y)$. Dann gilt*

$$\Big(\Phi \textit{ ist injektiv}\Big) \quad \Leftrightarrow \quad \Big(\Phi \textit{ ist surjektiv}\Big) \quad \Leftrightarrow \quad \Big(\Phi \textit{ ist bijektiv}\Big). \tag{V.41}$$

Beweis. Es reicht, die erste Äquivalenz zu zeigen. Seien dazu

$$N_1 := \dim(\mathrm{Ker}[\Phi]), \quad N_2 := \dim(\mathrm{Ran}[\Phi]), N := \dim(X) \;=\; \dim(Y). \tag{V.42}$$

Dann gelten

$$\Big(\Phi \text{ ist injektiv}\Big) \quad \Leftrightarrow \quad \Big(\mathrm{Ker}[\Phi] = \{\vec{0}\}\Big) \Leftrightarrow \Big(N_1 = 0\Big), \tag{V.43}$$

$$\Big(\Phi \text{ ist surjektiv}\Big) \quad \Leftrightarrow \quad \Big(\mathrm{Ran}[\Phi] = Y\Big) \Leftrightarrow \Big(N_2 = N\Big). \tag{V.44}$$

Nach Satz V.5 ist $N_1 + N_2 = N$, was sofort die Gleichwertigkeit der Injektivität und der Surjektivität von Φ nach sich zieht. $\square$

Bemerkungen und Beispiele.

- Für $\dim(X) = \infty$ gilt (V.41) im Allgemeinen nicht. Ist etwa $X = \mathbb{C}[x]$ der $\mathbb{C}$-Vektorraum der Polynome in x über $\mathbb{C}$,

$$\mathbb{C}[x] \;=\; \big\{\alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n \;\big|\; n \in \mathbb{N},\, \alpha_0, \ldots, \alpha_n \in \mathbb{C}\big\}, \tag{V.45}$$

mit Vektorraumstruktur definiert durch

$$\begin{aligned} &\gamma(\alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n) \;+\; (\beta_0 + \beta_1 x + \ldots + \beta_m x^m) \\ &\qquad := \; (\gamma\alpha_0 + \beta_0) + (\gamma\alpha_1 + \beta_1)x \;+ \ldots + (\gamma\alpha_{m+n} + \beta_{m+n})x^{m+n}, \end{aligned} \tag{V.46}$$

wobei $\gamma \in \mathbb{C}$ und $\alpha_{n+1} := \cdots := \alpha_{n+m} := \beta_{m+1} := \cdots := \beta_{m+n} := 0$, so ist

$$\dim(\mathbb{C}[x]) \;=\; \infty, \tag{V.47}$$

denn für jedes $n \in \mathbb{N}$ ist die Menge $\{1, x, x^2, \ldots, x^n\} \subseteq \mathbb{C}[x]$ linear unabhängig.

Seien weiterhin $\Phi, \Psi \in \mathcal{L}(\mathbb{C}[x]; \mathbb{C}[x])$, wobei

$$\Phi(\alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n) \;:=\; \alpha_0 x + \alpha_1 x^2 + \ldots + \alpha_n x^{n+1}, \tag{V.48}$$

$$\Psi(\alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n) \;:=\; \alpha_1 + \alpha_2 x + \ldots + \alpha_n x^{n-1}. \tag{V.49}$$

Dann gelten:

$$\mathrm{Ran}[\Phi] \not\ni 1 \quad \Rightarrow \quad \Phi \text{ ist injektiv, aber nicht surjektiv}, \tag{V.50}$$

$$\mathrm{Ker}[\Psi] \ni 1 \quad \Rightarrow \quad \Psi \text{ ist surjektiv, aber nicht injektiv}. \tag{V.51}$$

Definition V.7. Seien X und Y zwei $\mathbb{K}$-Vektorräume. Eine lineare Abbildung $\Phi \in \mathcal{L}(X;Y)$ heißt **Isomorphismus**, falls Φ bijektiv ist. Existiert ein Isomorphismus $\Phi \in \mathcal{L}(X;Y)$, so heißen X und Y zueinander **isomorph**.

Bemerkungen und Beispiele.

- Isomorphie zwischen $\mathbb{K}$-Vektorräumen definiert eine Äquivalenzrelation. (Den Nachweis führt man durch Nachprüfen der drei Eigenschaften Reflexivität, Symmetrie und Transitivität.)

Satz V.8. *Für $N \in \mathbb{N}$ ist jeder N-dimensionale Vektorraum über $\mathbb{K}$ isomorph zu $\mathbb{K}^N$.*

Beweis. Sind X ein N-dimensionaler $\mathbb{K}$-Vektorraum mit Basis $\{\vec{x}_1, \ldots, \vec{x}_N\} \subseteq X$, und ist $\{\vec{e}_1, \ldots, \vec{e}_N\} \subseteq \mathbb{K}^N$ die Standardbasis, so kann man leicht zeigen, dass $\Phi \in \mathcal{L}(X;\mathbb{K}^N)$, mit

$$\Phi(\alpha_1\vec{x}_1 + \ldots + \alpha_N\vec{x}_N) \;=\; \alpha_1\vec{e}_1 + \ldots + \alpha_N\vec{e}_N \;=\; \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{pmatrix}, \tag{V.52}$$

surjektiv ist und somit nach Satz V.6 einen Isomorphismus definiert. □

Korollar V.9. *Seien X und Y zwei $\mathbb{K}$-Vektorräume.*

(i) Sind X und Y isomorph, so gilt $\dim(X) = \dim(Y)$.

(ii) Gilt $\dim(X) = \dim(Y) < \infty$, so sind X und Y isomorph.

Beweis. <u>(i)</u>: Ist $\Phi \in \mathcal{L}(X;Y)$ ein Isomorphismus, so sind $\mathrm{Ker}[\Phi] = \{\vec{0}\}$ und $\mathrm{Ran}[\Phi] = Y$. Nach Satz V.5 gilt also

$$\dim(X) \;=\; \dim\big(\underbrace{\mathrm{Ker}[\Phi]}_{=0}\big) \;+\; \dim\big(\underbrace{(\mathrm{Ran}[\Phi]}_{=Y}\big) \;=\; \dim(Y). \tag{V.53}$$

<u>(ii)</u>: Seien $\dim(X) = \dim(Y) =: N < \infty$ und $\{\vec{x}_1, \ldots, \vec{x}_N\} \subseteq X$, $\{\vec{y}_1, \ldots, \vec{y}_N\} \subseteq Y$ Basen in X bzw. in Y. Dann definieren wir $\Phi \in \mathcal{L}(X;Y)$ durch

$$\Phi(\alpha_1\vec{x}_1 + \ldots + \alpha_N\vec{x}_N) \;:=\; \alpha_1\vec{y}_1 + \ldots + \alpha_N\vec{y}_N. \tag{V.54}$$

Offensichtlich ist Φ surjektiv und nach Satz V.6 wegen $N < \infty$ auch bijektiv. □

Bemerkungen und Beispiele.

- Satz V.8 garantiert also, dass wir einen abstrakten $\mathbb{K}$-Vektorraum X der Dimension N stets o.B.d.A. durch $\mathbb{K}^N$ ersetzen dürfen.
- Letzteres bedeutet aber nicht, dass sich auch andere Strukturen, außer der Vektorraumstruktur, die X besitzt, ohne Weiteres auf $\mathbb{K}^N$ übertragen lassen. So ist beispielsweise $\mathbb{C}$ als reeller Vektorraum isomorph zu $\mathbb{R}^2$, aber die Multiplikation komplexer Zahlen ist eine zusätzliche Struktur auf $\mathbb{C}$.

VI. Matrizen

VI.1. Definitionen

Definition VI.1. Seien $M, N \in \mathbb{N}$ und $\mathbb{K}$ der Körper $\mathbb{R}$ der reellen Zahlen oder $\mathbb{C}$ der komplexen Zahlen.

(i)
$$\mathbb{K}^{N\times M} := \left\{ (a_{i,j})_{\substack{i=1,\ldots,N \\ j=1,\ldots,M}} \;\middle|\; a_{i,j} \in \mathbb{K} \right\} \tag{VI.1}$$

definiert die Menge der $N \times M$-**Matrizen über** $\mathbb{K}$. Dabei ist $(a_{i,j})_{\substack{i=1,\ldots,N \\ j=1,\ldots,M}}$ eine abkürzende Schreibweise für eine Tabelle von Zahlen,

$$(a_{i,j})_{\substack{i=1,\ldots,N \\ j=1,\ldots,M}} := \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,M} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,M} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} & \ldots & a_{N,M} \end{pmatrix}. \tag{VI.2}$$

(ii) Seien X und Y zwei $\mathbb{K}$-Vektorräume mit $\dim(X) = M$, $\dim(Y) = N$ und $\mathcal{X} := \{\vec{x}_1, \ldots, \vec{x}_M\} \subseteq X$, $\mathcal{Y} := \{\vec{y}_1, \ldots, \vec{y}_N\} \subseteq Y$ Basen in X bzw. in Y. Ist nun $\Phi \in \mathcal{L}(X;Y)$ mit

$$\Phi(\vec{x}_j) = a_{1,j}\vec{y}_1 + a_{2,j}\vec{y}_2 + \ldots + a_{N,j}\vec{y}_N, \tag{VI.3}$$

für $j \in \mathbb{Z}_1^M$, so heißt

$$_{\mathcal{Y}}\mathcal{M}_{\mathcal{X}}[\Phi] := \begin{pmatrix} a_{1,1} & \ldots & a_{1,M} \\ \vdots & & \vdots \\ a_{N,1} & \ldots & a_{N,M} \end{pmatrix} \in \mathbb{K}^{N\times M} \tag{VI.4}$$

Matrixdarstellung von Φ bezüglich der Basen $\boldsymbol{\mathcal{X}} = \{\vec{x}_j\}_{j=1}^M \subseteq X$ und $\boldsymbol{\mathcal{Y}} = \{\vec{y}_i\}_{i=1}^N \subseteq Y$.

(iii) Für $\Phi \in \mathcal{L}(X) = \mathcal{L}(X;X)$ schreiben wir $\mathcal{M}_{\mathcal{X}}[\Phi] := {_{\mathcal{X}}\mathcal{M}_{\mathcal{X}}}[\Phi]$.

(iv) Sind $X = \mathbb{K}^M$ und $Y = \mathbb{K}^N$ mit Standardbasen $\mathcal{E}_M = \{\vec{e}_1, \ldots, \vec{e}_M\} \subseteq X$ und $\mathcal{E}_N = \{\vec{e}_1, \ldots, \vec{e}_N\} \subseteq Y$, so bezeichnet

$$\mathcal{M}[\Phi] := {_{\mathcal{E}_N}\mathcal{M}_{\mathcal{E}_M}}[\Phi] \tag{VI.5}$$

die **Matrixdarstellung von $\Phi \in \mathcal{L}(\mathbb{K}^M; \mathbb{K}^N)$ bezüglich der Standardbasen**.

Bemerkungen und Beispiele.

- $\mathbb{K} = \mathbb{R}$, $M = 3$, $N = 2$, $\mathcal{X} = \{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$, $\mathcal{Y} = \{\vec{y}_1, \vec{y}_2\}$, $\Phi \in \mathcal{L}(X;Y)$ mit

$$\Phi(\vec{x}_1) := 2\vec{y}_1, \qquad \Phi(\vec{x}_2) := \vec{y}_1 - \vec{y}_2, \qquad \Phi(\vec{x}_3) := -2\vec{y}_2, \tag{VI.6}$$

$$\Rightarrow \quad {}_{\mathcal{Y}}\mathcal{M}_{\mathcal{X}}[\Phi] \;=\; \begin{pmatrix} 2 & 1 & 0 \\ 0 & -1 & -2 \end{pmatrix}. \tag{VI.7}$$

- Seien $X = \mathbb{K}^M$, $Y = \mathbb{K}^N$ mit Standardbasen $\mathcal{E}_M = \{\vec{e}_1^{(X)}, \ldots, \vec{e}_M^{(X)}\} \subseteq \mathbb{K}^M$ und $\mathcal{E}_N = \{\vec{e}_1^{(Y)}, \ldots, \vec{e}_N^{(Y)}\} \subseteq \mathbb{K}^N$ und $\Phi \in \mathcal{L}(\mathbb{K}^M;\mathbb{K}^N)$ mit

$$\Phi(\vec{e}_i^{(X)}) \;:=\; a_{1,i}\vec{e}_1^{(Y)} + \ldots + a_{N,i}\vec{e}_N^{(Y)} \;=\; \begin{pmatrix} a_{1,i} \\ \vdots \\ a_{N,i} \end{pmatrix}, \tag{VI.8}$$

für $i = 1, \ldots, M$. Dann ist

$$\mathcal{M}[\Phi] \;\equiv\; {}_{\mathcal{E}_N}\mathcal{M}_{\mathcal{E}_M}[\Phi] \;=\; \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,M} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} & & a_{N,M} \end{pmatrix}$$

$$= \Big(\Phi(\vec{e}_1^{(X)}), \Phi(\vec{e}_2^{(X)}), \ldots, \Phi(\vec{e}_M^{(X)})\Big), \tag{VI.9}$$

d.h. die Matrixdarstellung von Φ bzgl. der Standardbasis erhält man durch Aneinanderreihung der Bilder $\Phi(\vec{e}_1^{(X)}), \ldots, \Phi(\vec{e}_M^{(X)})$ (als Spaltenvektoren) der Basisvektoren $\vec{e}_1^{(X)}, \ldots, \vec{e}_M^{(X)}$.

Lemma VI.2. *Seien $M, N \in \mathbb{N}$ und X, Y zwei $\mathbb{K}$-Vektorräume mit Dimensionen $M = \dim(X)$, $N = \dim(Y)$ und Basen $\mathcal{X} \subseteq X$ und $\mathcal{Y} \subseteq Y$.*

(i) $\mathbb{K}^{N\times M}$ *ist ein $\mathbb{K}$-Vektorraum der Dimension* $\dim[\mathbb{K}^{N\times M}] = M \cdot N$ *bezüglich*

$$\gamma \begin{pmatrix} a_{1,1} & \ldots & a_{1,M} \\ \vdots & & \vdots \\ a_{N,1} & \ldots & a_{N,M} \end{pmatrix} + \begin{pmatrix} b_{1,1} & \ldots & b_{1,M} \\ \vdots & & \vdots \\ b_{N,1} & \ldots & b_{N,M} \end{pmatrix} :=$$

$$\begin{pmatrix} (\gamma a_{1,1} + b_{1,1}) & \ldots & (\gamma a_{1,M} + b_{1,M}) \\ \vdots & & \vdots \\ (\gamma a_{N,1} + b_{N,1}) & \ldots & (\gamma a_{N,M} + b_{N,M}) \end{pmatrix}. \tag{VI.10}$$

(ii) ${}_{\mathcal{Y}}\mathcal{M}_{\mathcal{X}} : \mathcal{L}(X;Y) \to \mathbb{K}^{N\times M}$ *ist ein Isomorphismus.*

Beweis. *(i)* ist trivial.
(ii): Seien $\mathcal{X} = \{\vec{x}_1, \ldots, \vec{x}_M\}$ und $\mathcal{Y} = \{\vec{y}_1, \ldots, \vec{y}_N\}$. Zu festen $m \in \mathbb{Z}_1^M$ und $n \in \mathbb{Z}_1^N$ definieren wir $\Theta_{n,m} \in \mathcal{L}(X;Y)$ durch

$$\Theta_{n,m}(\vec{x}_j) \;:=\; \delta_{j,m}\,\vec{y}_n. \tag{VI.11}$$

Ist nun eine lineare Abbildung $\Phi \in \mathcal{L}(X;Y)$ durch die Bilder $\Phi(\vec{x}_j) = a_{1,j}\vec{y}_1 + \ldots + a_{N,j}\vec{y}_N$ der Basisvektoren in $\mathcal{X}$ gegeben, so ist

$$\sum_{m=1}^{M}\sum_{n=1}^{N} a_{n,m}\,\Theta_{n,m}\big(\vec{x}_j\big) \;=\; \sum_{m=1}^{M}\sum_{n=1}^{N} a_{n,m}\,\delta_{j,m}\,\vec{y}_n \;=\; \sum_{n=1}^{N} a_{n,j}\,\vec{y}_n \;=\; \Phi(\vec{x}_j), \tag{VI.12}$$

für alle $j \in \mathbb{Z}_1^M$. Also ist

$$\Phi \;=\; \sum_{m=1}^{M}\sum_{n=1}^{N} a_{n,m}\,\Theta_{n,m}, \tag{VI.13}$$

d.h. jede lineare Abbildung in $\mathcal{L}(X;Y)$ kann als Linearkombination der $\Theta_{n,m}$ dargestellt werden, $\mathcal{L}(X;Y) = \mathrm{span}\big(\{\Theta_{n,m} \mid m \in \mathbb{Z}_1^M, n \in \mathbb{Z}_1^N\}\big)$.

Um die lineare Unabhängigkeit der $\Theta_{n,m}$ zu zeigen, nehmen wir die an, die Nullabbildung sei als Linearkombination der $\Theta_{n,m}$ dargestellt, $\sum_{m=1}^{M}\sum_{n=1}^{N}\gamma_{n,m}\,\Theta_{n,m} = 0_{\mathcal{L}(X;Y)}$, d.h.

$$\forall \vec{x} \in X: \quad \sum_{m=1}^{M}\sum_{n=1}^{N} \gamma_{n,m}\,\Theta_{n,m}(\vec{x}) \;=\; \vec{0} \tag{VI.14}$$

gilt, was äquivalent zu

$$\forall j \in \mathbb{Z}_1^M: \quad \vec{0} \;=\; \sum_{m=1}^{M}\sum_{n=1}^{N} \gamma_{n,m}\,\Theta_{n,m}(\vec{x}_j) \;=\; \sum_{m=1}^{M}\sum_{n=1}^{N} \gamma_{n,m}\,\delta_{j,m}\,\vec{y}_n \;=\; \sum_{n=1}^{N} \gamma_{n,j}\,\vec{y}_n \tag{VI.15}$$

ist. Weil $\{\vec{y}_1, \ldots, \vec{y}_N\} \subseteq Y$ eine Basis ist, folgt daraus $\gamma_{1,j} = \gamma_{2,j} = \cdots = \gamma_{n,j} = 0$, für alle $j \in \mathbb{Z}_1^M$. Also verschwinden alle $\gamma_{n,m}$ und $\{\Theta_{n,m} \mid m \in \mathbb{Z}_1^M, n \in \mathbb{Z}_1^N\} \subseteq \mathcal{L}(X;Y)$ ist eine Basis.

Damit sind $\dim[\mathbb{K}^{N\times M}] = M\,N = \dim[\mathcal{L}(X;Y)] < \infty$, und $\mathbb{K}^{N\times M}$ und $\mathcal{L}(X;Y)$ sind gemäß Korollar V.9*(ii)* zueinander isomorph. Die zugehörigen Matrixdarstellungen

$$E_{n,m} \;:=\; {}_{\mathcal{Y}}\mathcal{M}_{\mathcal{X}}[\Theta_{n,m}] \;=\; \begin{array}{c} \begin{pmatrix} 0 & & \vdots & & 0 \\ & & 0 & & \\ \cdots & 0 & 1 & 0 & \cdots \\ & & 0 & & \\ 0 & & \vdots & & 0 \end{pmatrix} \leftarrow n.\ \text{Zeile} \\ \uparrow \\ m.\ \text{Spalte} \end{array} \tag{VI.16}$$

bezeichnet man als **Matrixeinheiten**. Die Menge der Matrixeinheiten $\{E_{n,m} \mid m \in \mathbb{Z}_1^M, n \in \mathbb{Z}_1^N\} \subseteq \mathbb{K}^{N\times M}$ bildet eine Basis. □

Bemerkungen und Beispiele.

- Seien $\mathbb{K} = \mathbb{R}$, $M = 3$, $N = 2$, so ist z.B.

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

$$\begin{pmatrix} (3\cdot 1+2) & (3\cdot 2-1) & (3\cdot 3-1) \\ (3\cdot 1+3) & (3\cdot 0+2) & (3\cdot 2+1) \end{pmatrix} \;=\; \begin{pmatrix} 5 & 5 & 8 \\ 6 & 2 & 7 \end{pmatrix}. \tag{VI.17}$$

VI.2. Matrixmultiplikation

Definition VI.3. Seien $L, M, N \in \mathbb{N}$. Das **Matrixprodukt** ist eine Verknüpfung

$$(\cdot): \ \mathbb{K}^{N\times M} \times \mathbb{K}^{M\times L} \quad \to \quad \mathbb{K}^{N\times L}, \tag{VI.18}$$

$$\left(\begin{pmatrix} a_{1,1} & \dots & a_{1,M} \\ \vdots & & \vdots \\ a_{N,1} & \dots & a_{N,M} \end{pmatrix}, \begin{pmatrix} b_{1,1} & \dots & b_{1,L} \\ \vdots & & \vdots \\ b_{M,1} & \dots & b_{M,L} \end{pmatrix} \right) \longmapsto \begin{pmatrix} c_{1,1} & \dots & c_{1,L} \\ \vdots & & \vdots \\ c_{N,1} & \dots & c_{N,L} \end{pmatrix}, \tag{VI.19}$$

wobei

$$c_{n,\ell} := \sum_{m=1}^{M} a_{n,m}\, b_{m,\ell}. \tag{VI.20}$$

Bemerkungen und Beispiele.

- Zunächst ein Zahlenbeispiel:

$$\begin{aligned} \begin{pmatrix} 2 & 3 & 0 \\ 7 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{pmatrix} &= \begin{pmatrix} 2\cdot 1 + 3\cdot 2 + 0\cdot 0 & 2\cdot 0 + 3\cdot 1 + 0\cdot 1 \\ 7\cdot 1 + 1\cdot 2 + 1\cdot 0 & 7\cdot 0 + 1\cdot 1 + 1\cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} 8 & 3 \\ 9 & 2 \end{pmatrix}. \end{aligned} \tag{VI.21}$$

- Zur Berechnung des Matrixproduktes ist folgendes Schema hilfreich:

$$\begin{array}{c|c} & \begin{pmatrix} b_{1,1} & \cdots & b_{1,\ell} & \cdots & b_{1,L} \\ \vdots & & \downarrow & & \vdots \\ b_{M,1} & \cdots & b_{M,\ell} & \cdots & b_{M,L} \end{pmatrix} \\ \hline \begin{pmatrix} a_{1,1} & \cdots & a_{1,M} \\ \vdots & & \vdots \\ a_{n,1} & \longrightarrow & a_{n,M} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,M} \end{pmatrix} & \begin{pmatrix} c_{1,1} & \cdots & | & \cdots & c_{1,L} \\ \vdots & & \downarrow & & \vdots \\ — & \longrightarrow & c_{n,\ell} & & \\ \vdots & & & & \vdots \\ c_{N,1} & \cdots & & \cdots & c_{N,L} \end{pmatrix} \end{array}. \tag{VI.22}$$

- Auch Vektoren lassen sich als Matrizen auffassen. Es sind $\mathbb{K}^{1\times M}$, $\mathbb{K}^{M\times 1}$ und $\mathbb{K}^M$ isomorph, denn sie haben alle Dimension M. Identifiziert man $\mathbb{K}^M$ mit $\mathbb{K}^{M\times 1}$, d.h. liest man die Vektoren

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_M \end{pmatrix} \in \mathbb{K}^M \quad \longleftrightarrow \quad \mathcal{M}[\vec{x}] = \begin{pmatrix} x_1 \\ \vdots \\ x_M \end{pmatrix} \in \mathbb{K}^{M\times 1} \tag{VI.23}$$

als $(M \times 1)$-Matrix, so lässt sich die Anwendung von $\Phi \in \mathcal{L}(\mathbb{K}^M; \mathbb{K}^N)$ auch als Matrixmultiplikation schreiben. Ist nämlich

$$\mathcal{M}[\Phi] \;=\; \begin{pmatrix} a_{1,1} & \dots & a_{1,M} \\ \vdots & & \vdots \\ a_{N,1} & \dots & a_{N,M} \end{pmatrix}, \tag{VI.24}$$

so ist $\mathcal{M}[\Phi(\vec{x})] \in \mathbb{K}^{N\times 1}$ mit

$$\begin{aligned} \mathcal{M}[\Phi(\vec{x})] \;&=\; \mathcal{M}[\Phi] \cdot \mathcal{M}[\vec{x}] \qquad\qquad \text{(VI.25)} \\ &=\; \begin{pmatrix} a_{1,1} & \dots & a_{1,M} \\ \vdots & & \vdots \\ a_{N,1} & \dots & a_{N,M} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_M \end{pmatrix} \;=\; \begin{pmatrix} a_{1,1}x_1 & + \cdots + & a_{1,M}x_M \\ \vdots & & \vdots \\ a_{N,1}x_1 & + \cdots + & a_{N,M}x_M \end{pmatrix}. \end{aligned}$$

Wegen dieser Interpretation der Anwendung von Φ auf $\vec{x}$ als Matrixmultiplikation lässt man häufig die Argumentsklammern weg und schreibt

$$\Phi\vec{x} \;:=\; \Phi(\vec{x}). \tag{VI.26}$$

Der folgende Satz zeigt nun, dass die Matrixmultiplikation genau der Komposition von linearen Abbildungen entspricht.

Satz VI.4. *Seien X, Y, Z drei $\mathbb{K}$-Vektorräume mit Basen $\mathcal{X} := \{\vec{x}_\ell\}_{\ell=1}^L \subseteq X$, $\mathcal{Y} := \{\vec{y}_m\}_{m=1}^M \subseteq Y$, $\mathcal{Z} := \{\vec{z}_n\}_{n=1}^N \subseteq Z$, wobei $L, M, N \in \mathbb{N}$. Sind weiterhin $\Phi \in \mathcal{L}(X; Y)$ und $\Psi \in \mathcal{L}(Y; Z)$, so gilt*

$${}_{\mathcal{Z}}\mathcal{M}_{\mathcal{X}}[\Psi \circ \Phi] \;=\; {}_{\mathcal{Z}}\mathcal{M}_{\mathcal{Y}}[\Psi] \cdot {}_{\mathcal{Y}}\mathcal{M}_{\mathcal{X}}[\Phi]. \tag{VI.27}$$

Beweis. Sind

$$\forall m \in \mathbb{Z}_1^M : \qquad \Psi(\vec{y}_m) \;= a_{1,m}\,\vec{z}_1 + \ldots + a_{N,m}\,\vec{z}_N, \tag{VI.28}$$

$$\forall \ell \in \mathbb{Z}_1^L : \qquad \Phi(\vec{x}_\ell) \;= b_{1,\ell}\,\vec{y}_1 + \ldots + b_{M,\ell}\,\vec{y}_M, \tag{VI.29}$$

so ist, für jedes $\ell \in \mathbb{Z}_1^L$,

$$\begin{aligned} (\Psi \circ \Phi)(\vec{x}_\ell) \;&= \Psi\big(b_{1,\ell}\vec{y}_1 + \ldots + b_{M,\ell}\vec{y}_M\big) \\ &= b_{1,\ell}\Psi(\vec{y}_1) + \ldots + b_{M,\ell}\Psi(\vec{y}_M) \qquad\qquad \text{(VI.30)} \\ &= b_{1,\ell} \cdot \big(a_{1,1}\vec{z}_1 + \ldots + a_{N,1}\vec{z}_N\big) + \ldots + b_{M,\ell} \cdot \big(a_{1,M}\vec{z}_1 + \ldots + a_{N,M}\vec{z}_N\big) \\ &= \big(a_{1,1}b_{1,\ell} + a_{1,2}b_{2,\ell} + \ldots + a_{1,M}b_{M,\ell}\big)\,\vec{z}_1 + \\ &\quad \ldots + \big(a_{N,1}b_{1,\ell} + a_{N,2}b_{2,\ell} + \ldots + a_{N,M}b_{M,\ell}\big)\,\vec{z}_N. \end{aligned}$$

Die Behauptung ergibt sich nun sofort durch den Vergleich von (VI.30), (VI.20) und (VI.3). $\square$

VI.3. Der Matrixring $\mathbb{K}^{N\times N}$

Definition VI.5. Sei X ein $\mathbb{K}$-Vektorraum.

(i) Man bezeichnet die linearen Abbildungen $X \to X$ als **Endomorphismen** und schreibt

$$\mathcal{L}(X) \;:=\; \mathcal{L}(X;X). \tag{VI.31}$$

(ii) Die bijektiven Elemente in $\mathcal{L}(X)$ bilden die Familie der **Automorphismen**, [1]

$$GL(X) \;:=\; \big\{\Phi \in \mathcal{L}(X) \;\big|\; \Phi \text{ ist bijektiv}\big\}. \tag{VI.32}$$

Lemma VI.6. *Sei X ein $\mathbb{K}$-Vektorraum.*

(i) Die Endomorphismen $\mathcal{L}(X)$ bilden einen Ring bezüglich punktweiser Addition und Komposition als Multiplikation, d.h. für $\Phi, \Psi, \Theta \in \mathcal{L}(X)$ ist $(\Phi \circ \Psi + \Theta) \in \mathcal{L}(X)$ definiert durch

$$\big(\Phi \circ \Psi + \Theta\big)(\vec{x}) \;:=\; \Phi\big(\Psi(\vec{x})\big) + \Theta(\vec{x}). \tag{VI.33}$$

Die Identität $\mathbb{1}_X \in \mathcal{L}(X)$, $\mathbb{1}_X(\vec{x}) = \vec{x}$, operiert als Einselement.

(ii) Die Automorphismen $GL(X) \subseteq \mathcal{L}(X)$ bilden eine Gruppe bezüglich Multiplikation (d.h. Komposition).

Beweis.
(i) erhält man durch Nachprüfen der Ringaxiome.

(ii) Mit $\Phi, \Psi \in \mathcal{L}(X)$ bijektiv ist auch $\Phi \circ \Psi$ bijektiv, also ist $\circ : GL(X) \times GL(X) \to GL(X)$. Die Komposition von Abbildungen ist auch assoziativ, und es gilt (G1). Die Identitätsabbildung $\mathbb{1}_X$ ist offensichtlich auch ein Automorphismus, also gilt (G2). Es verbleibt die Existenz des Inversen, (G3), nachzuprüfen. Seien dazu $\Phi \in GL(X)$ und $\Phi^{-1} : X \to X$ die inverse Abbildung, d.h.

$$\forall\, \vec{x} \in X: \quad \big(\Phi \circ \Phi^{-1}\big)(\vec{x}) \;=\; \big(\Phi^{-1} \circ \Phi\big)(\vec{x}) \;=\; \vec{x}. \tag{VI.34}$$

Sind nun $\vec{x}, \vec{y} \in X$ und $\alpha \in \mathbb{K}$, so gibt es $\vec{w}, \vec{z} \in X$ mit $\vec{x} = \Phi(\vec{w})$, $\vec{y} = \Phi(\vec{z})$. Somit ist

$$\begin{aligned} \Phi^{-1}(\alpha\vec{x} + \vec{y}) \;&= \Phi^{-1}\big(\alpha\Phi(\vec{w}) + \Phi(\vec{z})\big) \;=\; \Phi^{-1}\big(\Phi(\alpha\vec{w} + \vec{z})\big) \;=\; \alpha\vec{w} + \vec{z} \\ &= \alpha\Phi^{-1}(\vec{x}) + \Phi^{-1}(\vec{y})\,, \end{aligned} \tag{VI.35}$$

und Φ^{-1} ist linear. Offensichtlich ist Φ^{-1} auch bijektiv, also $\Phi^{-1} \in GL(X)$. □

Satz VI.7. *Seien $N \in \mathbb{N}$ und X ein $\mathbb{K}$-Vektorraum mit Basis $\mathcal{X} := \{\vec{x}_1, \ldots, \vec{x}_N\} \subseteq X$.*

(i) $\mathbb{K}^{N\times N}$ ist ein Ring (bezüglich komponentenweiser Addition und Matrixmultiplikation) mit Einselement

$$\mathbb{1} \;=\; \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}, \tag{VI.36}$$

*der **Einheitsmatrix**.*

[1] GL steht für "general linear group"

(ii) $\mathcal{M}_{\mathcal{X}} : \mathcal{L}(X) \to \mathbb{K}^{N \times N}$ *ist ein Ringisomorphismus, d.h.* $\mathcal{M}_{\mathcal{X}}$ *ist bijektiv und erhält die Ringeigenschaften,*

$$\forall \, \Phi, \Psi, \Theta \in \mathcal{L}(X) : \quad \mathcal{M}_{\mathcal{X}}[\Phi \circ \Psi + \Theta] \; = \; \mathcal{M}_{\mathcal{X}}[\Phi] \cdot \mathcal{M}_{\mathcal{X}}[\Psi] + \mathcal{M}_{\mathcal{X}}[\Theta]. \qquad \text{(VI.37)}$$

Beweis. $\underline{(i)}$ erhält man wieder durch Nachprüfen der Ringaxiome, und $\underline{(ii)}$ ergibt sich trivial aus Satz VI.4 und Lemma VI.6. □

Bemerkungen und Beispiele.

- $\mathbb{K}^{N \times N}$ ist nicht kommutativ für $N \geq 2$, etwa

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \qquad \text{(VI.38)}$$

- Für $\Phi \in GL(X)$ erhält man aus (VI.37) sofort, dass

$$\mathcal{M}_{\mathcal{X}}[\Phi^{-1}] \; = \; \big(\mathcal{M}_{\mathcal{X}}[\Phi]\big)^{-1}. \qquad \text{(VI.39)}$$

Die Berechnung der inversen Matrix auf der rechten Seite in (VI.39) ist allerdings nicht ganz einfach. (Siehe Satz VII.8.)

- Wir wählen $N = 2$, $\mathbb{K} = \mathbb{R}$, $X = \mathbb{R}^2$ und Basen

$$\mathcal{X} \; = \; \left\{ \vec{x}_1 := \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \vec{x}_2 := \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}, \quad \mathcal{E} \; = \; \left\{ \vec{e}_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \vec{e}_2 := \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \; \subseteq \; \mathbb{R}^2, \qquad \text{(VI.40)}$$

und eine lineare Abbildung $\Phi \in \mathcal{L}(\mathbb{R}^2)$ mit

$$\Phi(\vec{x}_1) \; = \; \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \qquad \Phi(\vec{x}_2) \; = \; \begin{pmatrix} 3 \\ 3 \end{pmatrix}. \qquad \text{(VI.41)}$$

Dann sind $\Phi(\vec{x}_1) = 2\vec{x}_1 + 0\vec{x}_2$ und $\Phi(\vec{x}_2) = 0\vec{x}_1 + 3\vec{x}_2$, also ist

$$\mathcal{M}_{\mathcal{X}}[\Phi] \; = \; \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}. \qquad \text{(VI.42)}$$

Andererseits sind

$$\vec{e}_1 \; = \; \begin{pmatrix} 1 \\ 0 \end{pmatrix} \; = \; \tfrac{1}{2}\vec{x}_1, \qquad \vec{e}_1 \; = \; \begin{pmatrix} 0 \\ 1 \end{pmatrix} \; = \; \vec{x}_2 - \tfrac{1}{2}\vec{x}_1, \qquad \text{(VI.43)}$$

und daher sind

$$\Phi(\vec{e}_1) \; = \tfrac{1}{2}\Phi(\vec{x}_1) \; = \; \tfrac{1}{2}\begin{pmatrix} 4 \\ 0 \end{pmatrix} \; = \; \begin{pmatrix} 2 \\ 0 \end{pmatrix} \; = \; 2\vec{e}_1, \qquad \text{(VI.44)}$$

$$\Phi(\vec{e}_2) \; = \Phi(\vec{x}_2) - \tfrac{1}{2}\Phi(\vec{x}_1) \; = \; \begin{pmatrix} 3 \\ 3 \end{pmatrix} - \tfrac{1}{2}\begin{pmatrix} 4 \\ 0 \end{pmatrix} \; = \; \begin{pmatrix} 1 \\ 3 \end{pmatrix} \; = \; 1\vec{e}_1 + 3\vec{e}_2, \qquad \text{(VI.45)}$$

also ist

$$\mathcal{M}_{\mathcal{E}}[\Phi] \; = \; \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} \; \neq \; \mathcal{M}_{\mathcal{X}}[\Phi]. \qquad \text{(VI.46)}$$

VI.4. Transformation von Matrizen bei Basiswechsel

Seien X und Y zwei $\mathbb{K}$-Vektorräume mit $\dim(X) = M$, $\dim(Y) = N$ und Basen $\mathcal{X} := \{\vec{x}_1, \ldots, \vec{x}_M\} \subseteq X$, $\mathcal{Y} := \{\vec{y}_1, \ldots, \vec{y}_N\} \subseteq Y$. Die Matrixdarstellung $\mathcal{M}_{\mathcal{Y},\mathcal{X}}[\Phi]$ einer linearen Abbildung $\Phi \in \mathcal{L}(X;Y)$ bezüglich dieser Basen ist basisabhängig, so wie die Koordinaten eines Vektors basisabhängig sind. Sind $\mathcal{W} := \{\vec{w}_1, \ldots, \vec{w}_M\} \subseteq X$ und $\mathcal{Z} := \{\vec{z}_1, \ldots, \vec{z}_N\} \subseteq Y$ auch Basen von X bzw. Y, so stellt sich die Frage, wie man ${}_{\mathcal{Y}}\mathcal{M}_{\mathcal{X}}[\Phi]$ in ${}_{\mathcal{Z}}\mathcal{M}_{\mathcal{W}}[\Phi]$ umrechnet.

Dazu bemerken wir, dass es nach Satz V.2 eindeutige lineare Abbildungen $\Xi \in \mathcal{L}(X)$, $\Theta \in \mathcal{L}(Y)$ gibt, mit

$$\Xi(\vec{x}_1) = \vec{w}_1, \quad \Xi(\vec{x}_2) = \vec{w}_2, \quad \ldots, \quad \Xi(\vec{x}_M) = \vec{w}_M, \tag{VI.47}$$

$$\Theta(\vec{y}_1) = \vec{z}_1, \quad \Theta(\vec{y}_2) = \vec{z}_2, \quad \ldots, \quad \Theta(\vec{y}_N) = \vec{z}_N. \tag{VI.48}$$

Ξ und Θ sind beide surjektiv, also nach Satz V.6 bijektiv und somit Automorphismen, $\Xi \in GL(X)$, $\Theta \in GL(Y)$. Die inversen Abbildungen sind damit ebenfalls Automorphismen $\Xi^{-1} \in GL(X)$, $\Theta^{-1} \in GL(Y)$ und besitzen die Eigenschaften

$$\Xi^{-1}(\vec{w}_1) = \vec{x}_1, \quad \Xi^{-1}(\vec{w}_2) = \vec{x}_2, \quad \ldots, \quad \Xi^{-1}(\vec{w}_M) = \vec{x}_M, \tag{VI.49}$$

$$\Theta^{-1}(\vec{z}_1) = \vec{y}_1, \quad \Theta^{-1}(\vec{z}_2) = \vec{y}_2, \quad \ldots, \quad \Theta^{-1}(\vec{z}_N) = \vec{y}_N. \tag{VI.50}$$

Die linearen Abbildungen Ξ, Θ, Ξ^{-1}, Θ^{-1} bezeichnet man als **Basistransformationen**.

Satz VI.8. *Seien $M, N \in \mathbb{N}$, X und Y zwei $\mathbb{K}$-Vektorräume mit Basen $\mathcal{X} := \{\vec{x}_m\}_{m=1}^M \subseteq X$, $\mathcal{Y} := \{\vec{y}_n\}_{n=1}^N \subseteq Y$, $\mathcal{W} := \{\vec{w}_m\}_{m=1}^M \subseteq X$, $\mathcal{Z} := \{\vec{z}_n\}_{n=1}^N \subseteq Y$ und zugehörigen Basistransformationen $\Xi \in \mathcal{L}(X)$, $\Theta \in \mathcal{L}(Y)$ mit $\Xi(\vec{x}_m) = \vec{w}_m$ und $\Theta(\vec{y}_n) = \vec{z}_n$. Dann gilt für jede lineare Abbildung $\Phi \in \mathcal{L}(X;Y)$, dass*

$${}_{\mathcal{Z}}\mathcal{M}_{\mathcal{W}}[\Phi] = {}_{\mathcal{Y}}\mathcal{M}_{\mathcal{X}}[\Theta^{-1} \circ \Phi \circ \Xi]. \tag{VI.51}$$

Beweis. Seien $a_{n,m} \in \mathbb{K}$ so gegeben, dass

$$\forall m \in \mathbb{Z}_1^M : \quad \Phi(\vec{w}_m) = a_{1,m}\vec{z}_1 + \ldots + a_{N,m}\vec{z}_N, \tag{VI.52}$$

d.h.

$${}_{\mathcal{Z}}\mathcal{M}_{\mathcal{W}}[\Phi] = \begin{pmatrix} a_{1,1} & \cdots & a_{1,M} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,M} \end{pmatrix}. \tag{VI.53}$$

Dann ist, für alle $m \in \mathbb{Z}_1^M$,

$$\begin{aligned} \Phi\big(\Xi(\vec{x}_m)\big) = \Phi(\vec{w}_m) &= a_{1,m}\Theta(\vec{y}_1) + \ldots + a_{N,m}\Theta(\vec{y}_N) \\ &= \Theta\big(a_{1,m}\vec{y}_1 + \ldots + a_{N,m}\vec{y}_N\big). \end{aligned} \tag{VI.54}$$

Durch Anwendung von Θ^{-1} erhalten wir dann

$$\forall\, m \in \mathbb{Z}_1^M : \quad \big(\Theta^{-1} \circ \Phi \circ \Xi\big)(\vec{x}_m) = a_{1,m}\vec{y}_1 + \ldots + a_{N,m}\vec{y}_N, \tag{VI.55}$$

d.h.

$$ {}_{\mathcal{Y}}\mathcal{M}_{\mathcal{X}}\big[\Theta^{-1} \circ \Phi \circ \Xi\big] \;=\; \begin{pmatrix} a_{1,1} & \cdots & a_{1,M} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,M} \end{pmatrix}. \tag{VI.56} $$

□

Korollar VI.9. *Gelten die Voraussetzungen wie in Satz VI.8, so ist*

$$ {}_{\mathcal{Z}}\mathcal{M}_{\mathcal{W}}[\Phi] \;=\; \big(\mathcal{M}_{\mathcal{Y}}[\Theta]\big)^{-1} \cdot {}_{\mathcal{Y}}\mathcal{M}_{\mathcal{X}}[\Phi] \cdot \mathcal{M}_{\mathcal{X}}[\Xi]. \tag{VI.57} $$

Beweis. Ergibt sich sofort aus Satz VI.4 und (VI.39). □

Korollar VI.10. *Seien $N \in \mathbb{N}$, X ein $\mathbb{K}$-Vektorraum mit Basen $\mathcal{X} := \{\vec{x}_n\}_{n=1}^N$, $\mathcal{W} := \{\vec{w}_n\}_{n=1}^N \subseteq X$ und zugehöriger Basistransformation $\Xi \in GL(X)$, mit $\Xi(\vec{x}_n) = \vec{w}_n$. Für alle $\Phi \in \mathcal{L}(X)$ gilt*

$$ \mathcal{M}_{\mathcal{W}}[\Phi] \;=\; \big(\mathcal{M}_{\mathcal{X}}[\Xi]\big)^{-1} \cdot \mathcal{M}_{\mathcal{X}}[\Phi] \cdot \mathcal{M}_{\mathcal{X}}[\Xi]. \tag{VI.58} $$

Bemerkungen und Beispiele.

- Zur Berechnung der inversen Matrix braucht man die Determinante. Diese ist Gegenstand des nächsten Kapitels. Erst wenn wir Matrizen invertieren können, sind (VI.57) und (VI.58) zu etwas nütze.
- Wir illustrieren Korollar VI.10 durch ein Beispiel und kommen dazu nochmal auf (VI.40)–(VI.46) zurück, wählen also $N = 2$, $\mathbb{K} = \mathbb{R}$, $X = \mathbb{R}^2$ und Basen

$$ \mathcal{E} \;=\; \left\{ \vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \quad \mathcal{X} \;=\; \left\{ \vec{x}_1 := \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \vec{x}_2 := \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \;\subseteq\; \mathbb{R}^2. \tag{VI.59} $$

Der Basis $\mathcal{E}$ entspricht nun also $\mathcal{X}$ in Korollar VI.10 und $\mathcal{X}$ in (VI.59) entspricht $\mathcal{W}$ in Korollar VI.10.

Wir berechnen nun die Matrixdarstellungen $\mathcal{M}_{\mathcal{E}}[\Xi]$, $\big(\mathcal{M}_{\mathcal{E}}[\Xi]\big)^{-1} = \mathcal{M}_{\mathcal{E}}[\Xi^{-1}] \in \mathbb{K}^{2\times 2}$ zur Basistransformation $\Xi \in GL(X)$, $\Xi(\vec{e}_1) = \vec{x}_1$, $\Xi(\vec{e}_2) = \vec{x}_2$, und ihrer Inversen $\Xi^{-1} \in GL(X)$, $\Xi^{-1}(\vec{x}_1) = \vec{e}_1$, $\Xi^{-1}(\vec{x}_2) = \vec{e}_2$. Dazu beobachten wir, dass

$$ \Xi(\vec{e}_1) \;=\; \vec{x}_1 \;=\; \begin{pmatrix} 2 \\ 0 \end{pmatrix} \;=\; 2\vec{e}_1, \tag{VI.60} $$

$$ \Xi(\vec{e}_2) \;=\; \vec{x}_2 \;=\; \begin{pmatrix} 1 \\ 1 \end{pmatrix} \;=\; \vec{e}_1 + \vec{e}_2, \tag{VI.61} $$

und

$$ \vec{e}_1 \;=\; \Xi^{-1}(\vec{x}_1) \;=\; 2\Xi^{-1}(\vec{e}_1) \;\Rightarrow\; \Xi^{-1}(\vec{e}_1) \;=\; \tfrac{1}{2}\vec{e}_1, \tag{VI.62} $$

$$ \vec{e}_2 \;=\; \Xi^{-1}(\vec{x}_2) \;=\; \Xi^{-1}(\vec{e}_1 + \vec{e}_2) \;=\; \tfrac{1}{2}\vec{e}_1 + \Xi^{-1}(\vec{e}_2) \;\Rightarrow\; \Xi^{-1}(\vec{e}_2) \;=\; \vec{e}_2 - \tfrac{1}{2}\vec{e}_1. \tag{VI.63} $$

Somit sind

$$\mathcal{M}_{\mathcal{E}}[\Xi] \;=\; \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad \mathcal{M}_{\mathcal{E}}[\Xi^{-1}] \;=\; \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 \end{pmatrix}. \tag{VI.64}$$

Nun überprüfen wir $\big(\mathcal{M}_{\mathcal{E}}[\Xi]\big)^{-1} = \mathcal{M}_{\mathcal{E}}[\Xi^{-1}]$ durch Matrixmultiplikation. In der Tat sind

$$\mathcal{M}_{\mathcal{E}}[\Xi] \cdot \mathcal{M}_{\mathcal{E}}[\Xi^{-1}] \;=\; \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 \end{pmatrix} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \;=\; \mathbb{1}, \tag{VI.65}$$

$$\mathcal{M}_{\mathcal{E}}[\Xi^{-1}] \cdot \mathcal{M}_{\mathcal{E}}[\Xi] \;=\; \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \;=\; \mathbb{1}. \tag{VI.66}$$

Wir kommen nun auf die in (VI.41) definierte lineare Abbildung $\Phi \in \mathcal{L}(\mathbb{R}^2)$ zurück, die gemäß (VI.42) und (VI.46) die Matrixdarstellungen

$$\mathcal{M}_{\mathcal{E}}[\Phi] \;=\; \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} \quad \text{und} \quad \mathcal{M}_{\mathcal{X}}[\Phi] \;=\; \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \tag{VI.67}$$

besitzt. Durch Matrixmultiplikation erhalten wir nun sofort, dass

$$\begin{aligned} \big(\mathcal{M}_{\mathcal{E}}[\Xi]\big)^{-1} \cdot \mathcal{M}_{\mathcal{E}}[\Phi] \cdot \mathcal{M}_{\mathcal{E}}[\Xi] \;&=\; \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \\ &=\; \begin{pmatrix} 1 & -1 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \;=\; \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \;=\; \mathcal{M}_{\mathcal{X}}[\Phi], \end{aligned} \tag{VI.68}$$

wie in Korollar VI.10 behauptet.

VII. Determinanten

VII.1. Der Leibnizsche Entwicklungssatz

Definition VII.1. Für $N \in \mathbb{N}$ sei $\det : \mathbb{K}^{N \times N} \to \mathbb{K}$ definiert durch

$$\det[A] \; := \; \sum_{\pi \in \mathcal{S}_N} (-1)^\pi \, a_{\pi(1),1} \cdot a_{\pi(2),2} \cdots a_{\pi(N),N}, \tag{VII.1}$$

für

$$A \; = \; \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix} \; \in \; \mathbb{K}^{N \times N}. \tag{VII.2}$$

$\det[A]$ heißt **Determinante von $\boldsymbol{A}$**. Die Formel (VII.1) heißt **Leibnizscher Entwicklungssatz**.

Bemerkungen und Beispiele.

- In (VII.1) ist $(-1)^\pi$ das Signum der Permutation π; ist π das Produkt von t Transpositionen, so ist $(-1)^\pi = (-1)^t$. (Siehe Kapitel II.)
- Sei $N = 2$. Dann ist $\mathcal{S}_2 = \{\mathbb{1}, \sigma\}$, mit

$$\mathbb{1} \; = \; \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \qquad \sigma \; = \; \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \quad \Rightarrow \quad (-1)^{\mathbb{1}} = +1, \quad (-1)^\sigma = -1. \tag{VII.3}$$

  Also ist

$$\begin{aligned} \det\left[\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}\right] \; &= \; \sum_{\pi \in \mathcal{S}_2} (-1)^\pi a_{\pi(1),1} a_{\pi(2),2} \\ &= (+1) \cdot a_{\mathbb{1}(1),1} \cdot a_{\mathbb{1}(2),2} \; + \; (-1) \cdot a_{\sigma(1),1} \cdot a_{\sigma(2),2} \\ &= a_{1,1} \cdot a_{2,2} - a_{2,1} \cdot a_{1,2}. \end{aligned} \tag{VII.4}$$

  <u>Merkregel</u>:

$$\det\left[\begin{pmatrix} a & & b \\ & & \\ c & & d \end{pmatrix}\right] \; = \; ad - bc. \tag{VII.5}$$

- Sei $N = 3$. Dann ist $\mathcal{S} = \{\pi_1, \pi_2, \ldots, \pi_6\}$ mit

$$\pi_1 := \mathbb{1} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \pi_2 := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \pi_3 := \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \tag{VII.6}$$

$$\pi_4 := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \pi_5 := \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \pi_6 := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \tag{VII.7}$$

und

$$(-1)^{\pi_1} = (-1)^{\pi_2} = (-1)^{\pi_3} = +1, \quad (-1)^{\pi_4} = (-1)^{\pi_5} = (-1)^{\pi_6} = -1. \tag{VII.8}$$

Somit wird

$$\det\left[\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}\right] \tag{VII.9}$$

$$\begin{aligned} = \ &+ \underbrace{a_{1,1} \cdot a_{2,2} \cdot a_{3,3}}_{\pi_1} + \underbrace{a_{2,1} \cdot a_{3,2} \cdot a_{1,3}}_{\pi_2} + \underbrace{a_{3,1} \cdot a_{1,2} \cdot a_{2,3}}_{\pi_3} \\ &- \underbrace{a_{2,1} \cdot a_{1,2} \cdot a_{3,3}}_{\pi_4} - \underbrace{a_{3,1} \cdot a_{2,2} \cdot a_{1,3}}_{\pi_5} - \underbrace{a_{1,1} \cdot a_{3,2} \cdot a_{2,3}}_{\pi_6}. \end{aligned}$$

<u>Merkregel (Sarrussche Regel)</u>: $\searrow = +1$, $\swarrow = -1$.

$$\left|\begin{array}{ccccc} a_{1,1} & & a_{1,2} & & a_{1,3} \\ & \searrow & & \searrow\!\!\!\!\swarrow & \\ a_{2,1} & & a_{2,2} & & a_{2,3} \\ & \swarrow & & \searrow\!\!\!\!\swarrow & \\ a_{3,1} & & a_{3,2} & & a_{3,3} \end{array}\right|\begin{array}{ccccc} & a_{1,1} & & a_{1,2} \\ \searrow\!\!\!\!\swarrow & & \swarrow & \\ & a_{2,1} & & a_{2,2} \\ \searrow\!\!\!\!\swarrow & & \searrow & \\ & a_{3,1} & & a_{3,2} \end{array}$$

- So sind beispielsweise

$$\det\left[\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}\right] = 2 \cdot 0 - 1 \cdot 1 = -1 \tag{VII.10}$$

und

$$\begin{aligned} \det&\left[\begin{pmatrix} 3 & 3 & 2 \\ 1 & 0 & 1 \\ 1 & 2 & 2 \end{pmatrix}\right] \\ &= 3 \cdot 0 \cdot 2 + 3 \cdot 1 \cdot 1 + 2 \cdot 1 \cdot 2 - 3 \cdot 1 \cdot 2 - 2 \cdot 0 \cdot 1 - 3 \cdot 1 \cdot 2 \\ &= -5. \end{aligned} \tag{VII.11}$$

- Es ist

$$\mathbb{1} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = \begin{pmatrix} \delta_{1,1} & \cdots & \delta_{1,N} \\ \vdots & & \vdots \\ \delta_{N,1} & \cdots & \delta_{N,N} \end{pmatrix}, \tag{VII.12}$$

wobei das **Kroneckersymbol $\boldsymbol{\delta_{i,j}}$** definiert ist als

$$\forall i, j \in \mathbb{Z}_1^N : \quad \delta_{i,j} := \begin{cases} 1 & \text{falls} \quad i = j, \\ 0 & \text{falls} \quad i \neq j. \end{cases} \tag{VII.13}$$

Beachte nun, dass für $\pi \in \mathcal{S}_N$

$$\big\{\delta_{\pi(1),1} \cdot \delta_{\pi(2),2} \cdots \delta_{\pi(N),N} \neq 0\big\} \;\Leftrightarrow\; \big\{\pi(1) = 1, \ldots, \pi(N) = N\big\} \;\Leftrightarrow\; \big\{\pi = \mathbb{1}\big\}. \tag{VII.14}$$

Also ist

$$\det[\mathbb{1}] \;=\; \sum_{\pi \in \mathcal{S}_N} (-1)^\pi \delta_{\pi(1),1} \cdots \delta_{\pi(N),N} \;=\; \delta_{1,1} \cdots \delta_{N,N} \;=\; 1. \tag{VII.15}$$

Wichtig ist auch, dass man die Determinante auch aus drei sie determinierenden Eigenschaften gewinnt, wie der folgende Satz behauptet.

Satz VII.2. *Für $N \in \mathbb{N}$ besitzt die Abbildung $\det : \mathbb{K}^{N \times N} \to \mathbb{K}$ folgende drei Eigenschaften (i)–(iii) und ist durch sie eindeutig bestimmt. Seien dazu $\gamma \in \mathbb{K}$ und*

$$\forall i \in \mathbb{Z}_1^N : \quad \vec{x}_i := \begin{pmatrix} a_{1,i} \\ \vdots \\ a_{N,i} \end{pmatrix}, \qquad \vec{y}_i := \begin{pmatrix} b_{1,i} \\ \vdots \\ b_{N,i} \end{pmatrix} \in \mathbb{K}^N. \tag{VII.16}$$

(i) Die Determinante ist linear in jeder Spalte, d.h. für jedes $i \in \mathbb{Z}_1^N$ gilt

$$\begin{aligned} &\det\big[\big(\vec{x}_1, \ldots, \vec{x}_{i-1}, \gamma\vec{x}_i + \vec{y}_i, \vec{x}_{i+1}, \ldots, \vec{x}_N\big)\big] \\ &= \gamma \cdot \det\big[\big(\vec{x}_1, \ldots, \vec{x}_{i-1}, \vec{x}_i, \vec{x}_{i+1}, \ldots, \vec{x}_N\big)\big] + \det\big[\big(\vec{x}_1, \ldots, \vec{x}_{i-1}, \vec{y}_i, \vec{x}_{i+1}, \ldots, \vec{x}_N\big)\big]. \end{aligned} \tag{VII.17}$$

(ii) Sind zwei Spalten gleich, so verschwindet die Determinante, d.h. gibt es ein Paar $i, j \in \mathbb{Z}_1^N$, $i \neq j$, so dass $\vec{z} := \vec{x}_i = \vec{x}_j$, so gilt

$$\det\big[\big(\vec{x}_1, \ldots, \vec{z}, \ldots \vec{z}, \ldots, \vec{x}_n\big)\big] \;=\; 0. \tag{VII.18}$$

(iii)

$$\det[\mathbb{1}] \;= 1. \tag{VII.19}$$

Beweis. S. Ergänzung im Abschnitt VII.5.1. □

Eine weitere wichtige Eigenschaft von Determinanten ist ihre Multiplikativität.

Satz VII.3. *Seien $N \in \mathbb{N}$ und $A, B \in \mathbb{K}^{N \times N}$. Dann gilt*

$$\det[A \cdot B] \;=\; \det[A] \cdot \det[B]. \tag{VII.20}$$

Beweis. Seien $A =: (a_{i,j})_{i,j=1}^N$, $B =: (b_{i,j})_{i,j=1}^N$ und $A \cdot B =: (\gamma_{i,j})_{i,j=1}^N$, also

$$\gamma_{i,j} \;=\; \sum_{k=1}^N a_{i,k}\, b_{k,j}. \tag{VII.21}$$

Sei $\kappa \in \mathcal{S}_N$ fest gewählt. Dann ist, für alle $\pi \in \mathcal{S}_N$,

$$a_{\pi(1),1} \cdots a_{\pi(N),N} \;=\; a_{(\pi\circ\kappa)(1),\kappa(1)} \cdots a_{(\pi\circ\kappa)(N),\kappa(N)}. \tag{VII.22}$$

Also ist, mit $\eta := \pi \circ \kappa$,

$$\begin{aligned}
\det[A] \cdot \det[B] \;&=\; \sum_{\pi,\kappa\in\mathcal{S}_N} \underbrace{(-1)^\pi \cdot (-1)^\kappa}_{=\,(-1)^{\pi\circ\kappa}} a_{\pi(1),1} \cdots a_{\pi(N),N} \cdot b_{\kappa(1),1} \cdots b_{\kappa(N),N} \\
&=\; \sum_{\kappa\in\mathcal{S}_N} \Big\{ \sum_{\pi\in\mathcal{S}_N} (-1)^{\pi\circ\kappa}\, a_{\pi\circ\kappa(1),\kappa(1)} \cdots a_{\pi\circ\kappa(N),\kappa(N)}\, b_{\kappa(1),1} \cdots b_{\kappa(N),N} \Big\} \\
&=\; \sum_{\kappa\in\mathcal{S}_N} \Big\{ \sum_{\eta\in\mathcal{S}_N} (-1)^\eta a_{\eta(1),\kappa(1)} \cdots a_{\eta(N),\kappa(N)} \Big\} b_{\kappa(1),1} \cdots b_{\kappa(N),N} \\
&=\; \sum_{\kappa\in\mathcal{S}_N} \det\big[\big(\vec{x}_{\kappa(1)}, \ldots, \vec{x}_{\kappa(N)}\big)\big] \cdot b_{\kappa(1),1} \cdots b_{\kappa(N),N}, \qquad \text{(VII.23)}
\end{aligned}$$

wobei

$$\forall\, \ell \in \mathbb{Z}_1^N : \quad \vec{x}_\ell := \begin{pmatrix} a_{1,\ell} \\ \vdots \\ a_{N,\ell} \end{pmatrix}. \tag{VII.24}$$

In (VII.23) ist $\kappa \in \mathcal{S}_N$, also ist $\kappa : \mathbb{Z}_1^N \to \mathbb{Z}_1^N$ eine bijektive Abbildung. Ist $\kappa : \mathbb{Z}_1^N \to \mathbb{Z}_1^N$ nicht bijektiv, so gibt es $i, j \in \mathbb{Z}_1^N$ mit $\kappa(i) = \kappa(j)$, und dann ist

$$\det\big[\big(\vec{x}_{\kappa(1)}, \ldots, \vec{x}_{\kappa(i)}, \ldots, \vec{x}_{\kappa(j)}, \ldots, \vec{x}_{\kappa(N)}\big)\big] \;=\; 0. \tag{VII.25}$$

Also ist

$$\begin{aligned}
&\sum_{\kappa\in\mathcal{S}_N} \det\big[\big(\vec{x}_{\kappa(1)}, \ldots, \vec{x}_{\kappa(N)}\big)\big] \cdot b_{\kappa(1),1} \cdots b_{\kappa(N),N} \\
&\qquad =\; \sum_{\substack{\kappa:\mathbb{Z}_1^N\to\mathbb{Z}_1^N,\\ \kappa \text{ bijektiv}}} \det\big[\big(\vec{x}_{\kappa(1)}, \ldots, \vec{x}_{\kappa(N)}\big)\big] \cdot b_{\kappa(1),1} \cdots b_{\kappa(N),N} \\
&\qquad =\; \sum_{\kappa:\mathbb{Z}_1^N\to\mathbb{Z}_1^N} \det\big[\big(\vec{x}_{\kappa(1)}, \ldots, \vec{x}_{\kappa(N)}\big)\big] \cdot b_{\kappa(1),1} \cdots b_{\kappa(N),N} \\
&\qquad =\; \sum_{\kappa(1),\ldots,\kappa(N)=1}^N \det\big[\big(\vec{x}_{\kappa(1)}, \ldots, \vec{x}_{\kappa(N)}\big)\big] \cdot b_{\kappa(1),1} \cdots b_{\kappa(N),N}. \qquad \text{(VII.26)}
\end{aligned}$$

Damit wird

$$\begin{aligned}
\det[A]\cdot\det[B] &= \sum_{\kappa(1),\ldots,\kappa(N)=1}^{N} \sum_{\pi\in\mathcal{S}_N} (-1)^\pi a_{\pi(1),\kappa(1)}\cdots a_{\pi(N),\kappa(N)}\, b_{\kappa(1),1}\cdots b_{\kappa(N),N} \\
&= \sum_{\pi\in\mathcal{S}_N} (-1)^\pi \Big(\sum_{\kappa(1)=1}^{N} a_{\pi(1),\kappa(1)} b_{\kappa(1),1}\Big)\cdots\Big(\sum_{\kappa(N)=1}^{N} a_{\pi(N),\kappa(N)} b_{\kappa(N),N}\Big) \\
&= \sum_{\pi\in\mathcal{S}_N} (-1)^\pi \gamma_{\pi(1),1}\cdots\gamma_{\pi(N),N} \;=\; \det[A\cdot B]. \qquad \text{(VII.27)}
\end{aligned}$$

□

VII.2. Das Inverse einer Matrix

Definition VII.4. Sei $A = (a_{i,j})_{i\in\mathbb{Z}_1^N, j\in\mathbb{Z}_1^M} \in \mathbb{K}^{N\times M}$ eine $N\times M$-Matrix, dann heißt $A^T := (b_{i,j})_{i\in\mathbb{Z}_1^M, j\in\mathbb{Z}_1^N} \in \mathbb{K}^{M\times N}$ mit $b_{i,j} := a_{j,i}$, die **zu *A* transponierte Matrix**,

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,M} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,M} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,M} \end{pmatrix} \;\Rightarrow\; A^T = \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{N,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{N,2} \\ \vdots & \vdots & & \vdots \\ a_{1,M} & a_{2,M} & \cdots & a_{N,M} \end{pmatrix}. \qquad \text{(VII.28)}$$

Bemerkungen und Beispiele.

- Man erhält A^T aus A durch Spiegelung an der Diagonalen.
- Zum Beispiel ist
$$A \;=\; \begin{pmatrix} 8 & 3 \\ 7 & 9 \\ 1 & 2 \end{pmatrix} \Rightarrow \qquad A^T \;=\; \begin{pmatrix} 8 & 7 & 1 \\ 3 & 9 & 2 \end{pmatrix}. \qquad \text{(VII.29)}$$
- Sind $A\in\mathbb{K}^{N\times M}$ und $B\in\mathbb{K}^{M\times L}$, so ist
$$(A\cdot B)^T \;=\; B^T\cdot A^T. \qquad \text{(VII.30)}$$
(Beachte vertauschte Reihenfolge!)
- Ist $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \in \mathbb{K}^N$ ein Spaltenvektor, den wir als $N\times 1$-Matrix der Form $\mathcal{M}(\vec{x}) = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \in \mathbb{K}^{N\times 1}$ auffassen, so entspricht die zu $\mathcal{M}(\vec{x})$ transponierte Matrix
$$\mathcal{M}(\vec{x})^T \;=\; (x_1,\ldots,x_N) \;\in\; \mathbb{K}^{1\times N} \qquad \text{(VII.31)}$$
gerade der Darstellung von $\vec{x}$ als Zeilenvektor.

Lemma VII.5. *Ist $A \in \mathbb{K}^{N \times N}$, so gilt*

$$\det[A] \;=\; \det[A^T]. \tag{VII.32}$$

Beweis. S. Ergänzung im Abschnitt VII.5.2. □

Definition VII.6. Seien $N \in \mathbb{N}$ und $A = (a_{i,j})_{i,j=1}^N \in \mathbb{K}^{N \times N}$. Definiere die Matrix $A_{\mathrm{minor}} = (b_{i,j})_{i,j=1}^N \in \mathbb{K}^{N \times N}$ der **Minoren** $b_{i,j}$ durch

$$b_{i,j} \;:=\; \sum_{\substack{\pi \in \mathcal{S}_N,\\ \pi(j)=i}} (-1)^\pi \, a_{\pi(1),1} \cdots a_{\pi(j-1),j-1} \, a_{\pi(j+1),j+1} \cdots a_{\pi(N),N}. \tag{VII.33}$$

Bemerkungen und Beispiele.

- Man erhält $b_{i,j}$ als $b_{i,j} = (-1)^{i+j} \cdot \det\left[A^{(i,j)}\right]$ wobei $A^{(i,j)} \in \mathbb{K}^{(N-1)\times(N-1)}$ gegeben ist durch

$$A^{(i,j)} \;=\; \begin{pmatrix} \boldsymbol{a_{1,1}} & \cdots & \boldsymbol{a_{1,j-1}} & a_{1,j} & \boldsymbol{a_{1,j+1}} & \cdots & \boldsymbol{a_{1,N}} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \boldsymbol{a_{i-1,1}} & \cdots & \boldsymbol{a_{i-1,j-1}} & a_{i-1,j} & \boldsymbol{a_{i-1,j+1}} & \cdots & \boldsymbol{a_{i-1,N}} \\ a_{i,1} & \cdots & a_{i,j-1} & a_{i,j} & a_{i,j+1} & \cdots & a_{i,N} \\ \boldsymbol{a_{i+1,1}} & \cdots & \boldsymbol{a_{i+1,j-1}} & a_{i+1,j} & \boldsymbol{a_{i+1,j+1}} & \cdots & \boldsymbol{a_{i+1,N}} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \boldsymbol{a_{N,1}} & \cdots & \boldsymbol{a_{N,j-1}} & a_{N,j} & \boldsymbol{a_{N,j+1}} & \cdots & \boldsymbol{a_{N,N}} \end{pmatrix}, \tag{VII.34}$$

  also aus A durch Streichen der i. Zeile und der j. Spalte hervorgeht.

- Für

$$A \;=\; \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}, \quad A_{\mathrm{minor}} \;=\; \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{pmatrix} \tag{VII.35}$$

  sind beispielsweise

$$b_{1,2} \;= (-1)^{1+2} \cdot \det\left[\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ \boldsymbol{a_{2,1}} & a_{2,2} & \boldsymbol{a_{2,3}} \\ \boldsymbol{a_{3,1}} & a_{3,2} & \boldsymbol{a_{3,3}} \end{pmatrix}\right] \;=\; -\det\left[\begin{pmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{pmatrix}\right], \tag{VII.36}$$

$$b_{3,3} \;= (-1)^{3+3} \cdot \det\left[\begin{pmatrix} \boldsymbol{a_{1,1}} & \boldsymbol{a_{1,2}} & a_{1,3} \\ \boldsymbol{a_{2,1}} & \boldsymbol{a_{2,2}} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}\right] \;=\; \det\left[\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}\right]. \tag{VII.37}$$

- Ein konkretes Zahlenbeispiel:

$$A \;=\; \begin{pmatrix} 3 & 3 & 2 \\ 1 & 0 & 1 \\ 1 & 2 & 2 \end{pmatrix}. \tag{VII.38}$$

Dann sind

$$b_{1,1} \;=\; (-1)^2 \cdot \det\left[\begin{pmatrix} 0 & 1 \\ 2 & 2 \end{pmatrix}\right] \;=\; (+1)\cdot(0-2) \;=\; -2, \tag{VII.39}$$

$$b_{1,2} \;=\; (-1)^3 \cdot \det\left[\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\right] \;=\; (-1)\cdot(2-1) \;=\; -1, \tag{VII.40}$$

$$b_{1,3} \;=\; (-1)^4 \cdot \det\left[\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}\right] \;=\; (+1)\cdot(2-0) \;=\; 2, \tag{VII.41}$$

$$b_{2,1} \;=\; (-1)^3 \cdot \det\left[\begin{pmatrix} 3 & 2 \\ 2 & 2 \end{pmatrix}\right] \;=\; (-1)\cdot(6-4) \;=\; -2, \tag{VII.42}$$

$$b_{2,2} \;=\; (-1)^4 \cdot \det\left[\begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}\right] \;=\; (+1)\cdot(6-2) \;=\; 4, \tag{VII.43}$$

$$b_{2,3} \;=\; (-1)^5 \cdot \det\left[\begin{pmatrix} 3 & 3 \\ 1 & 2 \end{pmatrix}\right] \;=\; (-1)\cdot(6-3) \;=\; -3, \tag{VII.44}$$

$$b_{3,1} \;=\; (-1)^4 \cdot \det\left[\begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix}\right] \;=\; (+1)\cdot(3-0) \;=\; 3, \tag{VII.45}$$

$$b_{3,2} \;=\; (-1)^5 \cdot \det\left[\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}\right] \;=\; (-1)\cdot(3-2) \;=\; -1, \tag{VII.46}$$

$$b_{3,3} \;=\; (-1)^6 \cdot \det\left[\begin{pmatrix} 3 & 3 \\ 1 & 0 \end{pmatrix}\right] \;=\; (+1)\cdot(0-3) \;=\; -3, \tag{VII.47}$$

und somit ist

$$A_{\mathrm{minor}} \;=\; \begin{pmatrix} -2 & -1 & 2 \\ -2 & 4 & -3 \\ 3 & -1 & -3 \end{pmatrix}. \tag{VII.48}$$

- Für $N = 2$ kann A_{minor} sofort direkt angegeben werden, da die Determinante einer 1×1-Matrix die Zahl selbst ist:

$$A \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad\Rightarrow\quad A_{\mathrm{minor}} \;=\; \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}. \tag{VII.49}$$

Satz VII.7. *Seien $N \in \mathbb{N}$ und $A = (a_{i,j})_{i,j=1}^N \in \mathbb{K}^{N\times N}$. Dann ist*

$$A \cdot A_{\mathrm{minor}}^T \;=\; A_{\mathrm{minor}}^T \cdot A \;=\; \det[A] \cdot \mathbb{1}. \tag{VII.50}$$

Der Beweis des Satzes VII.7 ähnelt dem Beweis des Satzes VII.3, s. die Ergänzung VII.5.3. Wir ziehen hier nur die wichtigen Folgerungen.

Satz VII.8. *Für $N \in \mathbb{N}$ und $A \in \mathbb{K}^{N\times N}$ gelten folgende Aussagen:*

(i) A ist genau dann invertibel, wenn $\det[A] \neq 0$ ist;

(ii) Ist A invertibel und bezeichnet $A^{-1} \in \mathbb{K}^{N\times N}$ mit $A \cdot A^{-1} = A^{-1} \cdot A = \mathbb{1}$ die zu A inverse Matrix, so gilt

$$A^{-1} \;=\; \frac{1}{\det[A]} \cdot A_{\mathrm{minor}}^T. \tag{VII.51}$$

Beweis. Sei $\det[A] \neq 0$. Dann setzen wir $\widetilde{A} := (\det[A])^{-1} \cdot A^T_{\text{minor}}$ und erhalten

$$A \cdot \widetilde{A} \;=\; \widetilde{A} \cdot A \;=\; \mathbb{1} \tag{VII.52}$$

aus Satz VII.7. Somit ist A invertibel und $A^{-1} = \widetilde{A}$.

Sei umgekehrt $\det[A] = 0$. Wäre A invertibel mit Inverser $A^{-1} \in \mathbb{K}^{N \times N}$, so wäre nach Satz VII.3

$$1 \;=\; \det[\mathbb{1}] \;=\; \det[A \cdot A^{-1}] \;=\; \underbrace{\det[A]}_{=0} \cdot \det[A^{-1}] \;=\; 0, \tag{VII.53}$$

was einen Widerspruch ergibt. □

Bemerkungen und Beispiele.

- Seien $N = 2$ und $a, b, c, d \in \mathbb{K}$, so dass $ad - bc \neq 0$. Mit $A := \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ ist also $\det[A] \neq 0$, und A ist invertibel mit

$$A^{-1} \;=\; \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{VII.54}$$

- Die Bestimmung der Inversen einer 3×3-Matrix ist schon aufwändiger. Dazu müssen wir ihre Determinante und 9 Minoren (Determinanten von 2×2-Matrizen) berechnen. Beispielsweise sind nach (VII.11) und (VII.38)–(VII.48)

$$A \;=\; \begin{pmatrix} 3 & 3 & 2 \\ 1 & 0 & 1 \\ 1 & 2 & 2 \end{pmatrix} \quad\Rightarrow\quad \det[A] \;=\; -5, \quad A_{\text{minor}} \;=\; \begin{pmatrix} -2 & -1 & 2 \\ -2 & 4 & -3 \\ 3 & -1 & -3 \end{pmatrix} \tag{VII.55}$$

$$\Rightarrow \quad A^{-1} \;=\; -\frac{1}{5} \begin{pmatrix} -2 & -2 & 3 \\ -1 & 4 & -1 \\ 2 & -3 & -3 \end{pmatrix} \;=\; \begin{pmatrix} \frac{2}{5} & \frac{2}{5} & -\frac{3}{5} \\ \frac{1}{5} & -\frac{4}{5} & \frac{1}{5} \\ -\frac{2}{5} & \frac{3}{5} & \frac{3}{5} \end{pmatrix}.$$

- Die Menge der invertiblen $N \times N$-Matrizen über $\mathbb{K}$ bezeichnen wir mit

$$GL(N, \mathbb{K}) \;:=\; \big\{ A \in \mathbb{K}^{N \times N} \;\big|\; \det[A] \neq 0 \big\}. \tag{VII.56}$$

VII.3. Lineare Gleichungssysteme

Eine wichtige Anwendung von Matrizen und ihrer Inversen sind **lineare Gleichungssysteme** (**LGS**). Sind $A = (a_{i,j})_{i,j=1}^N \in \mathbb{K}^{N \times N}$ eine $N \times N$-Matrix über $\mathbb{K}$ und $\vec{y} = (y_i)_{i=1}^N \in \mathbb{K}^N \cong \mathbb{K}^{N \times 1}$ ein Spaltenvektor, so betrachten wir das LGS

$$a_{1,1}x_1 + a_{1,2}x_2 + \ldots + a_{1,N}x_N \;= y_1, \tag{VII.57}$$

$$a_{2,1}x_1 + a_{2,2}x_2 + \ldots + a_{2,N}x_N \;= y_2, \tag{VII.58}$$

$$\vdots \tag{VII.59}$$

$$a_{N,1}x_1 + a_{N,2}x_2 + \ldots + a_{N,N}x_N \;= y_N. \tag{VII.60}$$

Wir möchten Bedingungen finden, wann das LGS (VII.57)-(VII.60) eine Lösung $\vec{x} = (x_i)_{i=1}^N \in \mathbb{K}^N$ besitzt und ggf. Aussagen über die Lösungsmenge machen. Offensichtlich ist das LGS (VII.57)-(VII.60) äquivalent zur Vektorgleichung

$$A\vec{x} \;=\; \vec{y}. \tag{VII.61}$$

Satz VII.9.

(i) ***Homogene LGS*** $(\vec{y} = \vec{0})$*:*

$$\big\{ \det[A] = 0 \big\} \quad \Leftrightarrow \quad \big\{ \exists\, \vec{x} \in \mathbb{K}^N \setminus \{\vec{0}\} : \; A\vec{x} = \vec{0} \big\}. \tag{VII.62}$$

(ii) ***Inhomogene LGS****:*

$$\big\{ \det[A] \neq 0, \; A\vec{x} = \vec{y} \big\} \quad \Rightarrow \quad \big\{ \vec{x} = A^{-1}\vec{y} \big\}. \tag{VII.63}$$

Beweis.
(ii) folgt trivial aus Satz VII.8. Um *(i)* zu beweisen, nehmen wir zunächst $\det[A] \neq 0$ an. Dann ist A invertibel und $\vec{x} = A^{-1}\vec{0} = \vec{0}$ die eindeutige Lösung von $A\vec{x} = \vec{0}$. Also gibt es keine nichttriviale Lösung von $A\vec{x} = \vec{0}$.

Ist umgekehrt $\det[A] \;=\; 0$, so ist A nicht invertibel. Nun ist $\mathcal{M} : \mathcal{L}(\mathbb{K}^N) \to \mathbb{K}^{N \times N}$ gemäß Satz VI.7, *(ii)* ein Ringisomorphismus, und es gibt eine eindeutige lineare Abbildung $\Phi \in \mathcal{L}(\mathbb{K}^N)$, so dass $A = \mathcal{M}[\Phi]$. Weil A nicht invertibel ist, ist Φ nicht bijektiv und somit auch nicht injektiv. Es gibt also $\vec{x}, \vec{x}\,' \in \mathbb{K}^N$, $\vec{x} \neq \vec{x}\,'$, so dass $A\vec{x} = A\vec{x}\,'$. Damit ist $(\vec{x} - \vec{x}\,') \neq \vec{0}$ eine Lösung: $A(\vec{x} - \vec{x}\,') = \vec{0}$. □

VII.4. Determinante und Spur einer linearen Abbildung

In den vorigen Abschnitten dieses Kapitels haben wir die Determinante $\det[A]$ einer $N \times N$-Matrix A kennengelernt. Wir haben gesehen, dass A genau dann invertibel ist, wenn $\det[A] \neq 0$ gilt. Wir wollen den Zusammenhang zwischen einer linearen Abbildung $\Phi \in \mathcal{L}(X)$ und der Determinante ihrer Matrixdarstellung bezgl. einer Basis bestimmen. Doch zuvor führen wir den Begriff der Spur einer Matrix ein.

Definition VII.10. Für $N \in \mathbb{N}$ sei $\mathrm{Tr} : \mathbb{K}^{N\times N} \to \mathbb{K}$ definiert durch

$$\mathrm{Tr}\{A\} := \sum_{i=1}^{N} a_{i,i} \quad \text{für} \quad A \;=\; \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix} \;\in\; \mathbb{K}^{N\times N}. \tag{VII.64}$$

$\mathrm{Tr}\{A\}$ heißt **Spur von *A***.

Lemma VII.11 (Zyklizität der Spur)**.** *Für $N \in \mathbb{N}$ seien $G_1, G_2, \ldots, G_k \in \mathbb{K}^{N\times N}$. Dann ist*

$$\mathrm{Tr}\{G_1\, G_2 \cdots G_k\} \;=\; \mathrm{Tr}\{G_k\, G_1\, G_2 \cdots G_{k-1}\} \tag{VII.65}$$

Beweis. Sei zunächst $k = 2$, also etwa

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix}, \quad B = \begin{pmatrix} b_{1,1} & \cdots & b_{1,N} \\ \vdots & & \vdots \\ b_{N,1} & \cdots & b_{N,N} \end{pmatrix}. \tag{VII.66}$$

Dann ist

$$\mathrm{Tr}\{A \cdot B\} = \sum_{i=1}^{N} \left(\sum_{j=1}^{N} a_{i,j} b_{j,i} \right) = \sum_{j=1}^{N} \left(\sum_{i=1}^{N} b_{j,i} a_{i,j} \right) = \mathrm{Tr}\{B \cdot A\}. \tag{VII.67}$$

Setzt man $A := G_1 G_2 \cdots G_{k-1}$ und $B := G_k$, so folgt (VII.65) für $k \geq 2$ direkt aus (VII.67). □

Satz VII.12. *Seien $N \in \mathbb{N}$ und X ein N-dimensionaler Vektorraum über $\mathbb{K}$ mit Basen $\mathcal{X} := \{\vec{x}_1, \ldots, \vec{x}_N\} \subseteq X$ und $\mathcal{W} := \{\vec{w}_1, \ldots, \vec{w}_N\} \subseteq X$. Ist $\Phi \in \mathcal{L}(X)$ eine lineare Abbildung, so sind*

$$\det\big[\mathcal{M}_{\mathcal{X}}[\Phi]\big] = \det\big[\mathcal{M}_{\mathcal{W}}[\Phi]\big], \tag{VII.68}$$

$$\mathrm{Tr}\big\{\mathcal{M}_{\mathcal{X}}[\Phi]\big\} = \mathrm{Tr}\big\{\mathcal{M}_{\mathcal{W}}[\Phi]\big\}. \tag{VII.69}$$

Beweis. Ist $\Xi \in \mathcal{L}(X)$ die Basistransformation zwischen $\mathcal{X}$ und $\mathcal{W}$, also $\Xi(\vec{x}_i) = \vec{w}_i$, für alle $i = 1, \ldots, N$, so gilt nach (VI.58)

$$\widetilde{A} = H^{-1} A\, H, \tag{VII.70}$$

wobei $A := \mathcal{M}_{\mathcal{X}}[\Phi]$, $\widetilde{A} := \mathcal{M}_{\mathcal{W}}[\Phi]$ und $H := \mathcal{M}_{\mathcal{X}}[\Xi]$. Nach Satz VII.3 ist dann

$$\det[\widetilde{A}] = \det[H^{-1}] \cdot \det[A] \cdot \det[H] = \det[H^{-1} H] \cdot \det[A] = \det[A]. \tag{VII.71}$$

Ähnlich erhalten wir mit Lemma VII.11

$$\mathrm{Tr}\{\widetilde{A}\} = \mathrm{Tr}\{H^{-1} A\, H\} = \mathrm{Tr}\{H\, H^{-1} A\} = \mathrm{Tr}\{A\}. \tag{VII.72}$$

□

Wir sehen also, dass Determinante und Spur einer Matrix nicht durch Basistransformationen beeinflusst werden.

Definition VII.13. Für $N \in \mathbb{N}$ sei X ein N-dimensionaler Vektorraum über $\mathbb{K}$. Definiere die Abbildungen $\det, \mathrm{Tr} : \mathcal{L}(X) \to \mathbb{K}$ durch

$$\det[\Phi] := \det\big[\mathcal{M}_{\mathcal{X}}[\Phi]\big], \tag{VII.73}$$

$$\mathrm{Tr}\{\Phi\} := \mathrm{Tr}\big\{\mathcal{M}_{\mathcal{X}}[\Phi]\big\}, \tag{VII.74}$$

wobei $\mathcal{X} := \{\vec{x}_1, \ldots, \vec{x}_N\} \subseteq X$ (irgend)eine Basis von X ist.

Bemerkungen und Beispiele.

- Satz VII.12 sichert die Wohldefiniertheit (Basisunabhängigkeit) der Abbildungen $\det : \mathcal{L}(X) \to \mathbb{K}$ und $\mathrm{Tr} : \mathcal{L}(X) \to \mathbb{K}$.
- Man sagt auch, Determinante und Spur seien <u>invariant</u> (unter Basistransformationen).
- Es ist

$$\mathrm{Tr}\left\{ \begin{pmatrix} 3 & 3 & 2 \\ 1 & 0 & 1 \\ 1 & 2 & 2 \end{pmatrix} \right\} = 3 + 0 + 2 = 5. \tag{VII.75}$$

VII.5. Ergänzungen

VII.5.1. Die drei determinierenden Eigenschaften der Determinante

Beweis. Wir zeigen nur, dass det die Eigenschaften (i)–(iii) aus Satz VII.2 besitzt, aber nicht, dass sie die einzige Abbildung mit diesen Eigenschaften ist.

$\underline{(i)}$: ergibt sich sofort aus der Definition,

$$\begin{aligned}
&\sum_{\pi \in \mathcal{S}_N} (-1)^\pi a_{\pi(1),1} \cdots \big(\gamma a_{\pi(i),i} + b_{\pi(i),i}\big) \cdots a_{\pi(N),N} \\
&= \gamma \Big(\sum_{\pi \in \mathcal{S}_N} (-1)^\pi a_{\pi(1),1} \cdots a_{\pi(i),i} \cdots a_{\pi(N),N} \Big) \\
&\qquad + \Big(\sum_{\pi \in \mathcal{S}_N} (-1)^\pi a_{\pi(1),1} \cdots b_{\pi(i),i} \cdots a_{\pi(N),N} \Big).
\end{aligned} \tag{VII.76}$$

$\underline{(iii)}$: haben wir schon in (VII.15) gezeigt.

$\underline{(ii)}$: Sei $\vec{x}_i = \vec{x}_j$, etwa für (i,j) mit $1 \le i < j \le N$. Dann ist, mit $A := (\vec{x}_1, \ldots, \vec{x}_N) \in \mathbb{K}^{N \times N}$,

$$\begin{aligned}
\det[A] &= \sum_{\pi \in \mathcal{S}_N} (-1)^\pi \prod_{\ell=1}^{N} a_{\pi(\ell),\ell} \\
&= \sum_{\pi \in \mathcal{S}_N} (-1)^\pi a_{\pi(i),i} a_{\pi(j),j} \prod_{\ell(\neq i,j)} a_{\pi(\ell),\ell} \\
&= \sum_{\pi \in \mathcal{S}_N} (-1)^\pi \gamma_{\pi(i)} \gamma_{\pi(j)} \prod_{\ell(\neq i,j)} a_{\pi(\ell),\ell},
\end{aligned} \tag{VII.77}$$

wobei $\gamma_k := a_{k,i} = a_{k,j}$ und "$\prod_{\ell(\neq i,j)}$" das Produkt über alle $\ell \in \mathbb{Z}_1^N \setminus \{i,j\}$ notiert. Sei nun $\sigma \in \mathcal{S}_N$ die Transposition (i,j), d.h.

$$\sigma := \begin{pmatrix} 1 & 2 & \ldots & i & \ldots & j & \ldots & N \\ 1 & 2 & \ldots & j & \ldots & i & \ldots & N \end{pmatrix}. \tag{VII.78}$$

Dann sind $(-1)^{\pi\circ\sigma} = -(-1)^\pi$ und

$$\gamma_{(\pi\circ\sigma)(i)}\, \gamma_{(\pi\circ\sigma)(j)} \prod_{\ell(\neq i,j)} a_{(\pi\circ\sigma)(\ell),\ell} = \gamma_{\pi(j)} \cdot \gamma_{\pi(i)} \cdot \prod_{\ell(\neq i,j)} a_{\pi(\ell),\ell} = \gamma_{\pi(i)}\, \gamma_{\pi(j)} \prod_{\ell(\neq i,j)} a_{\pi(\ell),\ell}. \tag{VII.79}$$

Daher ist

$$\begin{aligned}
\det[A] &= -\sum_{\pi \in \mathcal{S}_N} (-1)^{\pi\circ\sigma} \gamma_{(\pi\circ\sigma)(i)}\, \gamma_{(\pi\circ\sigma)(j)} \prod_{\ell(\neq i,j)} a_{(\pi\circ\sigma)(\ell),\ell} \\
&= -\sum_{\kappa \in \mathcal{S}_N} (-1)^{\kappa} \gamma_{\kappa(i)}\, \gamma_{\kappa(j)} \prod_{\ell(\neq i,j)} a_{\kappa(\ell),\ell} \\
&= -\det[A],
\end{aligned} \tag{VII.80}$$

wobei wir $\kappa := \pi \circ \sigma$ als neue Summationsvariable einführen und benutzen, dass $\mathcal{S}_N$ eine Gruppe ist und insbesondere $\mathcal{S}_N = \mathcal{S}_N \circ \sigma$ gilt. Also ist $\det[A] = -\det[A] = 0$. □

VII.5.2. Erhaltung der Determinante unter Transposition

Beweis. (Beweis von Lemma VII.5): Sei $A = (a_{i,j})_{i,j=1}^N$. Dann erhalten wir durch Umsortieren des Produkts, dass

$$\begin{aligned} a_{\pi(1),1} \cdot a_{\pi(2),2} \cdots a_{\pi(N),N} &= a_{\pi(1),\pi^{-1}(\pi(1))} \cdot a_{\pi(2),\pi^{-1}(\pi(2))} \cdots a_{\pi(N),\pi^{-1}(\pi(N))} \\ &= a_{1,\pi^{-1}(1)} \cdot a_{2,\pi^{-1}(2)} \cdots a_{N,\pi^{-1}(N)}. \end{aligned} \tag{VII.81}$$

Weil $(-1)^\pi = (-1)^{\pi^{-1}}$, und weil $\mathcal{S}_N$ eine Gruppe ist, erhalten wir, mit $\kappa := \pi^{-1}$ als neue Summationsvariable, dass

$$\det[A] \;=\; \sum_{\pi \in \mathcal{S}_N} (-1)^{\pi^{-1}} a_{1,\pi^{-1}(1)} \cdots a_{N,\pi^{-1}(N)} \;=\; \sum_{\kappa \in \mathcal{S}_N} a_{1,\kappa(1)} \cdots a_{N,\kappa(N)} \;=\; \det[A^T]. \tag{VII.82}$$

□

VII.5.3. Beweis von Satz VII.7

Bezeichnen wir die Produktmatrizen mit

$$(\gamma_{i,j})_{i,j=1}^N := A_{\mathrm{minor}}^T \cdot A \quad \text{und} \quad (\tilde{\gamma}_{i,j})_{i,j=1}^N := A \cdot A_{\mathrm{minor}}^T, \tag{VII.83}$$

d.h.

$$\gamma_{i,j} \;=\; \sum_{k=1}^N b_{k,i}\, a_{k,j} \quad \text{und} \quad \tilde{\gamma}_{i,j} \;=\; \sum_{k=1}^N a_{i,k}\, b_{j,k}, \tag{VII.84}$$

so ist also zu zeigen, dass $\gamma_{i,j} = \tilde{\gamma}_{i,j} = \det[A] \cdot \delta_{i,j}$. Wir zeigen nur $\gamma_{i,j} = \det[A] \cdot \delta_{i,j}$. Setzen wir (VII.33) ein, so erhalten wir für festes $i, j \in \mathbb{Z}_1^N$

$$\begin{aligned} \gamma_{i,j} \;&=\; \sum_{k=1}^N a_{k,j} b_{k,i} \;=\; \sum_{k=1}^N a_{k,j} \sum_{\substack{\pi \in \mathcal{S}_N, \\ \pi(i)=k}} (-1)^\pi a_{\pi(1),1} \cdots a_{\pi(i-1),i-1} \cdot a_{\pi(i+1),i+1} \cdots a_{\pi(N),N} \\ &=\; \sum_{k=1}^N \sum_{\pi \in \mathcal{S}_N} \delta_{\pi(i),k} (-1)^\pi a_{\pi(1),1} \cdots a_{\pi(i-1),i-1} \cdot \underbrace{a_{k,j}}_{=a_{\pi(i),j}} \cdot a_{\pi(i+1),i+1} \cdots a_{\pi(N),N} \\ &=\; \sum_{\pi \in \mathcal{S}_N} \Big\{ (-1)^\pi a_{\pi(1),1} \cdots a_{\pi(i-1),i-1} \cdot a_{\pi(i),j} \cdot a_{\pi(i+1),i+1} \cdots a_{\pi(N),N} \underbrace{\Big(\sum_{k=1}^N \delta_{\pi(i),k} \Big)}_{=1} \Big\} \\ &=\; \sum_{\pi \in \mathcal{S}_N} (-1)^\pi a_{\pi(1),1} \cdots a_{\pi(i-1),i-1} \cdot a_{\pi(i),j} \cdot a_{\pi(i+1),i+1} \cdots a_{\pi(N),N} \\ &=\; \det\left[(\vec{x}_1, \ldots, \vec{x}_{i-1}, \vec{x}_j, \vec{x}_{i+1}, \ldots, \vec{x}_j, \ldots, \vec{x}_N)\right], \end{aligned} \tag{VII.85}$$

wobei

$$\forall\, \ell \in \mathbb{Z}_1^N : \quad \vec{x}_\ell \;:=\; \begin{pmatrix} a_{1,\ell} \\ \vdots \\ a_{N,\ell} \end{pmatrix}. \tag{VII.86}$$

Für $\underline{i \neq j}$ sind dann zwei Spalten in (VII.85) gleich, und aus Satz VII.2 (ii), folgt $\gamma_{i,j} = 0$. Für $\underline{i = j}$ ist aber $\gamma_{i,j} = \det[A]$, also

$$\gamma_{i,j} \;=\; \det[A] \cdot \delta_{i,j}. \tag{VII.87}$$

VIII. Der Gauß-Algorithmus

Zur Berechnung von Determinanten und von inversen Matrizen ist der Gauß-Algorithmus viel praktischer als der Leibnizsche Entwicklungssatz und die Matrix der Minoren. Er basiert auf der Multiplikativität der Determinante und dem sogenannten *Kästchensatz.*

Satz VIII.1 (Kästchensatz)**.** *Seien* $N = n_1 + n_2 + \ldots + n_L \in \mathbb{N}$ *mit* $n_1, n_2, \ldots, n_L \in \mathbb{N}$*. Seien weiterhin* $g_{k,\ell} \in \mathbb{K}^{n_k \times n_\ell}$ *und*

$$A \;=\; \left(\begin{array}{c|c|c|c} g_{1,1} & g_{1,2} & & g_{1,L} \\ \hline g_{2,1} & g_{2,2} & & \\ \hline \vdots & & \ddots & \\ \hline g_{L,1} & \cdots & & g_{L,L} \end{array}\right) \quad \in \mathbb{K}^{N \times N}. \tag{VIII.1}$$

Sind $g_{k,\ell} = 0$*, falls* $k > \ell$*, also*

$$A \;=\; \left(\begin{array}{c|c|c|c} g_{1,1} & g_{1,2} & \cdots & g_{1,L} \\ \hline 0 & g_{2,2} & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & g_{L-1,L} \\ \hline 0 & \cdots & 0 & g_{L,L} \end{array}\right), \tag{VIII.2}$$

oder sind $g_{k,\ell} = 0$*, falls* $k < \ell$*, also*

$$A \;=\; \left(\begin{array}{c|c|c|c} g_{1,1} & 0 & \cdots & 0 \\ \hline g_{2,1} & g_{2,2} & \ddots & \vdots \\ \hline \vdots & & \ddots & 0 \\ \hline g_{L,1} & \cdots & g_{L,L-1} & g_{L,L} \end{array}\right), \tag{VIII.3}$$

so gilt in beiden Fällen

$$\det[A] \;=\; \det[g_{1,1}] \cdot \det[g_{2,2}] \cdots \det[g_{L,L}]. \tag{VIII.4}$$

Beweis. Es genügt, die Aussage für $L = 2$ zu zeigen. Für $L > 2$ folgt sie dann leicht per vollständiger Induktion.

Seien also $A = (a_{i,j})_{i,j=1}^N$, $g_{1,1} = (a_{i,j})_{i,j=1}^n$, $g_{2,2} = (a_{i,j})_{i,j=n+1}^N$ und $a_{i,j} = 0$, falls $i \geq n+1$ und $j \leq n$. D.h.

$$A = \left(\begin{array}{ccc|ccc} a_{1,1} & \cdots & a_{1,n} & a_{1,n+1} & \cdots & a_{1,N} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} & a_{n,n+1} & \cdots & a_{n,N} \\ \hline 0 & \cdots & 0 & a_{n+1,n+1} & \cdots & a_{n+1,N} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{N,n+1} & \cdots & a_{N,N} \end{array} \right). \tag{VIII.5}$$

Seien nun $\pi \in \mathcal{S}_N$ eine Permutation und $j \leq n$. Ist $\pi(j) \geq n+1$, so ist $a_{\pi(j),j} = 0$ und somit auch

$$(-1)^\pi a_{\pi(1),1} \cdots a_{\pi(N),N} = 0. \tag{VIII.6}$$

Also tragen nur die Permutationen $\pi \in \mathcal{S}_N$ zu $\det[A]$ bei, die die Mengen $\{1, \ldots, n\}$ und $\{n+1, \ldots, N\}$ jeweils auf sich abbilden. Diese sind aber gerade die Produkte aus Permutationen von $\{1, \ldots, n\}$ und $\{n+1, \ldots, N\}$,

$$\pi = \kappa \circ \eta, \tag{VIII.7}$$

wobei $\kappa \in \mathcal{S}_n$ und $\eta \in \mathcal{S}_{N-n}$ (auf $\{n+1, \ldots, N\}$ wirkend). Also ist

$$\begin{aligned} \det[A] &= \sum_{\pi \in \mathcal{S}_N} (-1)^\pi a_{\pi(1),1} \cdots a_{\pi(n),n} \cdot a_{\pi(n+1),n+1} \cdots a_{\pi(N),N} \\ &= \sum_{\kappa \in \mathcal{S}_n, \eta \in \mathcal{S}_{N-n}} (-1)^{\kappa \circ \eta} a_{\kappa(1),1} \cdots a_{\kappa(n),n} \cdot a_{\eta(n+1),n+1} \cdots a_{\eta(N),N} \\ &= \det[g_{1,1}] \cdot \det[g_{2,2}]. \end{aligned} \tag{VIII.8}$$

□

Ein Spezialfall des Kästchensatzes ist $L = N$ und $n_1 = n_2 = \ldots = n_N = 1$, d.h. wenn A eine *rechte obere Dreiecksmatrix* oder eine *linke untere Dreiecksmatrix* ist. In diesem Fall ist die Determinante durch das Produkt der Matrixelemente auf der Diagonalen gegeben.

Korollar VIII.2. *Sei $N \in \mathbb{N}$. Ist*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,N} \\ 0 & a_{2,2} & a_{2,3} & \cdots & a_{2,N} \\ 0 & 0 & a_{3,3} & \cdots & a_{3,N} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & a_{N-1,N-1} & a_{N-1,N} \\ 0 & \cdots & 0 & 0 & a_{N,N} \end{pmatrix} \in \mathbb{K}^{N \times N}, \tag{VIII.9}$$

*eine **rechte obere Dreiecksmatrix** oder*

$$A = \begin{pmatrix} a_{1,1} & 0 & 0 & \cdots & \cdots & 0 \\ a_{2,1} & a_{2,2} & 0 & \cdots & \cdots & \vdots \\ a_{3,1} & a_{3,2} & a_{3,3} & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ a_{N-1,1} & a_{N-1,2} & a_{N-1,3} & \cdots & a_{N-1,N-1} & 0 \\ a_{N,1} & a_{N,2} & a_{N,3} & \cdots & a_{N,N-1} & a_{N,N} \end{pmatrix} \in \mathbb{K}^{N \times N}, \qquad \text{(VIII.10)}$$

*eine **linke untere Dreiecksmatrix**, so gilt in beiden Fällen*

$$\det[A] = a_{1,1} \cdot a_{2,2} \cdots a_{N,N}. \qquad \text{(VIII.11)}$$

VIII.1. Elementare Zeilen- und Spaltenoperationen

Der später vorgestellte Gauß-Algorithmus besteht aus sukzessiven Anwendungen von *elementaren Zeilenoperationen* auf eine zu untersuchende Matrix $A \in \mathbb{K}^{N \times N}$, die ihre Determinante bis auf ihr Vorzeichen unverändert lassen.

Definition VIII.3. Sei $N \in \mathbb{N}$ und $A \in \mathbb{K}^{N \times N}$ eine $N \times N$-Matrix.

(Z) Als **elementare Zeilenoperationen** bezeichnet man

- das Vertauschen der k. Zeile mit der ℓ. Zeile in A und
- das Addieren des η-fachen der k. Zeile zur ℓ. Zeile in A,

für alle $k, \ell \in \mathbb{Z}_1^N$ mit $k \neq \ell$ und $\eta \in \mathbb{K}$.

(S) Als **elementare Spaltenoperationen** bezeichnet man

- das Vertauschen der k. Spalte mit der ℓ. Spalte in A und
- das Addieren des η-fachen der k. Spalte zur ℓ. Spalte in A,

für alle $k, \ell \in \mathbb{Z}_1^N$ mit $k \neq \ell$ und $\eta \in \mathbb{K}$.

Bemerkungen und Beispiele.

- Sind $A = (a_{i,j})_{i,j=1}^N \in \mathbb{K}^{N \times N}$ eine $N \times N$-Matrix und $k, \ell \in \mathbb{Z}_1^N$ mit $k \neq \ell$ und $\eta \in \mathbb{K}$, so bewirken die *elementaren Zeilenoperationen*, dass

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{k,1} & \cdots & a_{k,N} \\ \vdots & & \vdots \\ a_{\ell,1} & \cdots & a_{\ell,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix} \qquad \text{(VIII.12)}$$

übergeht in

$$A_Z^{(k,\ell)} := \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{\ell,1} & \cdots & a_{\ell,N} \\ \vdots & & \vdots \\ a_{k,1} & \cdots & a_{k,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix}, \quad A_Z^{(\ell+\eta k)} := \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{k,1} & \cdots & a_{k,N} \\ \vdots & & \vdots \\ a_{\ell,1} + \eta\, a_{k,1} & \cdots & a_{\ell,N} + \eta\, a_{k,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix}. \tag{VIII.13}$$

- Sind $A = (a_{i,j})_{i,j=1}^N \in \mathbb{K}^{N \times N}$ eine $N \times N$-Matrix und $k, \ell \in \mathbb{Z}_1^N$ mit $k \neq \ell$ und $\eta \in \mathbb{K}$, so bewirken die *elementaren Spaltenoperationen*, dass

$$A \;=\; \begin{pmatrix} a_{1,1} & \cdots & a_{1,k} & \cdots & a_{1,\ell} & \cdots & a_{1,N} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,k} & \cdots & a_{N,\ell} & \cdots & a_{N,N} \end{pmatrix} \tag{VIII.14}$$

übergeht in

$$A_S^{(k,\ell)} := \begin{pmatrix} a_{1,1} & \cdots & a_{1,\ell} & \cdots & a_{1,k} & \cdots & a_{1,N} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,\ell} & \cdots & a_{N,k} & \cdots & a_{N,N} \end{pmatrix},$$

$$A_S^{(\ell+\eta k)} := \begin{pmatrix} a_{1,1} & \cdots & a_{1,k} & \cdots & a_{1,\ell} + \eta\, a_{1,k} & \cdots & a_{1,N} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,k} & \cdots & a_{N,\ell} + \eta\, a_{N,k} & \cdots & a_{N,N} \end{pmatrix}. \tag{VIII.15}$$

Elementare Zeilen- und Spaltenoperationen bewirken höchstens einen Vorzeichenwechsel der Determinante der Matrix A, wie das folgende Lemma zeigt.

Lemma VIII.4. *Seien $N \in \mathbb{N}$, $A \in \mathbb{K}^{N\times N}$ eine $N \times N$-Matrix und $k, \ell \in \mathbb{Z}_1^N$ mit $k \neq \ell$ sowie $\eta \in \mathbb{K}$.*

(i) Gehen $A_Z^{(k,\ell)} \in \mathbb{K}^{N\times N}$ aus A durch Vertauschen der k. und der ℓ. Zeilen und $A_S^{(k,\ell)} \in \mathbb{K}^{N\times N}$ aus A durch Vertauschen der k. und der ℓ. Spalten hervor, so gilt

$$\det\big[A_Z^{(k,\ell)}\big] \;=\; \det\big[A_S^{(k,\ell)}\big] \;=\; -\det[A]\,. \tag{VIII.16}$$

(ii) Gehen $A_Z^{(\ell+\eta k)} \in \mathbb{K}^{N\times N}$ aus A durch Addieren des η-fachen der k. Zeile zur ℓ. Zeile und $A_S^{(\ell+\eta k)} \in \mathbb{K}^{N\times N}$ aus A durch Addieren des η-fachen der k. Spalte zur ℓ. Spalte $A_S^{(k,\ell)} \in \mathbb{K}^{N\times N}$ hervor, so gilt

$$\det\big[A_Z^{(\ell+\eta k)}\big] \;=\; \det\big[A_S^{(\ell+\eta k)}\big] \;=\; \det[A]\,. \tag{VIII.17}$$

Beweis. Wir beweisen tatsächlich noch etwas allgemeinere Identitäten als (VIII.16) und (VIII.17). Dazu führen wir ein paar Rechnungen durch.

Seien $\gamma_2, \gamma_3, \ldots, \gamma_N \in \mathbb{K}$ und $A, G, H \in \mathbb{K}^{N\times N}$ gegeben durch

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix}, \tag{VIII.18}$$

$$G = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ \gamma_2 & 1 & \ddots & & \vdots \\ \gamma_3 & 0 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \gamma_N & 0 & \cdots & 0 & 1 \end{pmatrix}, \tag{VIII.19}$$

$$H = \begin{pmatrix} 1 & \eta_2 & \eta_3 & \cdots & \eta_N \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix}. \tag{VIII.20}$$

Offensichtlich sind G eine linke untere bzw. H eine rechte obere Dreieckmatrix, und aus Korollar VIII.2 folgt sofort, dass

$$\det[G] = \det[H] = 1. \tag{VIII.21}$$

Durch Matrixmultiplikation erhalten wir außerdem, dass $G \cdot G^{-1} = G^{-1} \cdot G = \mathbb{1}$ und $H \cdot H^{-1} = H^{-1} \cdot H = \mathbb{1}$ mit

$$G^{-1} = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ -\gamma_2 & 1 & \ddots & & \vdots \\ -\gamma_3 & 0 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -\gamma_N & 0 & \cdots & 0 & 1 \end{pmatrix}, \tag{VIII.22}$$

$$H^{-1} = \begin{pmatrix} 1 & -\eta_2 & -\eta_3 & \cdots & -\eta_N \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix}, \tag{VIII.23}$$

und weiterhin

$$G \cdot A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} + \gamma_2 a_{1,1} & a_{2,2} + \gamma_2 a_{1,2} & \cdots & a_{2,N} + \gamma_2 a_{1,N} \\ \vdots & \vdots & & \vdots \\ a_{N,1} + \gamma_N a_{1,1} & a_{N,2} + \gamma_N a_{1,2} & \cdots & a_{N,N} + \gamma_N a_{1,N} \end{pmatrix}, \tag{VIII.24}$$

$$A \cdot H = \begin{pmatrix} a_{1,1} & a_{1,2} + \eta_2 a_{1,1} & \cdots & a_{1,N} + \eta_N a_{1,1} \\ a_{2,1} & a_{2,2} + \eta_2 a_{2,1} & \cdots & a_{2,N} + \eta_N a_{2,1} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} + \eta_2 a_{N,1} & \cdots & a_{N,N} + \eta_N a_{N,1} \end{pmatrix}. \tag{VIII.25}$$

Insbesondere sind

$$\det\left[\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} + \gamma_2 a_{1,1} & a_{2,2} + \gamma_2 a_{1,2} & \cdots & a_{2,N} + \gamma_2 a_{1,N} \\ \vdots & \vdots & & \vdots \\ a_{N,1} + \gamma_N a_{1,1} & a_{N,2} + \gamma_N a_{1,2} & \cdots & a_{N,N} + \gamma_N a_{1,N} \end{pmatrix}\right] = \det[G \cdot A] \tag{VIII.26}$$

$$= \det[G] \cdot \det[A] = \det[A] = \det\left[\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{pmatrix}\right],$$

und

$$\det\left[\begin{pmatrix} a_{1,1} & a_{1,2} + \eta_2 a_{1,1} & \cdots & a_{1,N} + \eta_N a_{1,1} \\ a_{2,1} & a_{2,2} + \eta_2 a_{2,1} & \cdots & a_{2,N} + \eta_N a_{2,1} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} + \eta_2 a_{N,1} & \cdots & a_{N,N} + \eta_N a_{N,1} \end{pmatrix}\right] = \det[A \cdot H] \tag{VIII.27}$$

$$= \det[A] \cdot \det[H] = \det[A] = \det\left[\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{pmatrix}\right],$$

Wir definieren auch für $1 \le i < j \le N$ die Permutationsmatrix $S^{(i,j)} \in \mathbb{K}^{N \times N}$ durch

$$\big(S^{(i,j)}\big)_{k,\ell} = \begin{cases} 1, & \text{falls } k = \ell \notin \{i,j\}, \\ 1, & \text{falls } k = j \text{ und } \ell = i, \\ 1, & \text{falls } k = i \text{ und } \ell = j, \\ 0, & \text{sonst.} \end{cases} \tag{VIII.28}$$

und setzen $S^{(i,i)} := \mathbb{1}$. Dann bewirkt die Multiplikation mit $S^{(i,j)}$ von links die Vertauschung der i. mit der j. Zeile von A,

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{i,1} & \cdots & a_{i,N} \\ \vdots & & \vdots \\ a_{j,1} & \cdots & a_{j,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix}, \qquad S^{(i,j)} \cdot A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{j,1} & \cdots & a_{j,N} \\ \vdots & & \vdots \\ a_{i,1} & \cdots & a_{i,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix}, \tag{VIII.29}$$

und die Multiplikation mit $S^{(i,j)}$ von rechts vertauscht die i. mit der j. Spalte von A,

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,i} & \cdots & a_{1,j} & \cdots & a_{1,N} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,i} & \cdots & a_{N,j} & \cdots & a_{N,N} \end{pmatrix}, \tag{VIII.30}$$

$$A \cdot S^{(i,j)} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,j} & \cdots & a_{1,i} & \cdots & a_{1,N} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,j} & \cdots & a_{N,i} & \cdots & a_{N,N} \end{pmatrix}. \tag{VIII.31}$$

Da $S^{(i,j)}$ für $i < j$ wie eine Transposition und für $i = j$ wie die Identität wirkt, ist

$$\forall\, i \leq j : \qquad \det\big[S^{(i,j)}\big] = (-1)^{1-\delta_{i,j}}, \tag{VIII.32}$$

wobei $\delta_{i,j}$ das Kroneckersymbol notiert.

Sind nun $k, \ell \in \mathbb{Z}_1^N$ mit $k \neq \ell$ und $\eta \in \mathbb{K}$, so folgt (VIII.16) für $A_Z^{(k,\ell)}$ sofort aus (VIII.32) und $A_Z^{(k,\ell)} = S^{(k,\ell)} \cdot A$,

$$\det\big[A_Z^{(k,\ell)}\big] = \det\big[S^{(k,\ell)} \cdot A\big] = \det\big[S^{(k,\ell)}\big] \cdot \det\big[A\big] = -\det\big[A\big] \tag{VIII.33}$$

und natürlich analog für $A_S^{(k,\ell)}$.

Den Beweis von (VIII.17) führen wir nur für den Fall, dass $\ell > 1$ ist. Dann definieren wir G wie in (VIII.18), wobei wir $\gamma_i :=$ für $i \in \mathbb{Z}_2^N \setminus \{\ell\}$ und $\gamma_\ell := \eta$ wählen, also

$$G = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & & \vdots \\ 0 & & \ddots & \ddots & \ddots & & & \vdots \\ \eta & 0 & & \ddots & \ddots & \ddots & & \vdots \\ 0 & & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}. \tag{VIII.34}$$

Ist nun $k = 1$, so ist $\ell > 2$, und wir erhalten aus (VIII.24), dass

$$G \cdot A \;=\; \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ \vdots & \vdots & & \vdots \\ a_{\ell-1,1} & a_{\ell-1,2} & \cdots & a_{\ell-1,N} \\ a_{\ell,1} + \eta a_{1,1} & a_{\ell,2} + \eta a_{1,2} & \cdots & a_{\ell,N} + \eta a_{1,N} \\ a_{\ell+1,1} & a_{\ell+1,2} & \cdots & a_{\ell+1,N} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{pmatrix} \;=\; A_Z^{(\ell+\eta 1)}, \tag{VIII.35}$$

was in diesem Fall $(k, \ell) = (1, \ell)$ gerade die Behauptung ergibt,

$$\det\big[A_Z^{(\ell+\eta 1)}\big] \;=\; \det\big[G \cdot A\big] \;=\; \det\big[G\big] \cdot \det\big[A\big] \;=\; \det\big[A\big]. \tag{VIII.36}$$

Ist $k \in \mathbb{Z}_2^N \setminus \{\ell\}$, so tauschen wir erst die k. Zeile in die erste Zeile, wenden dann (VIII.35) an und tauschen anschließend die erste Zeile mit der k. Zeile. Dann ist also $A_Z^{(\ell+\eta k)} = S^{(1,k)} \cdot G \cdot S^{(1,k)} \cdot A$ und somit

$$\det\big[A_Z^{(\ell+\eta k)}\big] \;=\; \big(\det\big[S^{(1,k)}\big]\big)^2 \cdot \det\big[G\big] \cdot \det\big[A\big] \;=\; \det\big[A\big]. \tag{VIII.37}$$

□

VIII.2. Der Gauß-Algorithmus für Determinanten

Wir kommen nun zur Beschreibung des Gauß-Algorithmus' zur Berechnung von Determinanten. Seien $N \in \mathbb{N}$, und

$$A \;=\; \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix} \;\in\; \mathbb{K}^{N \times N} \tag{VIII.38}$$

eine $N \times N$-Matrix, deren Determinante wir berechnen wollen.

1. Wir betrachten zunächst die erste Spalte $(a_{1,1}, \ldots, a_{N,1})$ von A und unterscheiden drei Fälle:

   1.1 Verschwindet die erste Spalte identisch, $(a_{1,1}, \ldots, a_{N,1}) = (0, \ldots, 0)$, so gilt nach dem Kästchensatz

$$\begin{aligned} \det[A] \;&=\; \det\left[\left(\begin{array}{c|ccc} 0 & a_{1,2} & \cdots & a_{1,N} \\ \hline 0 & a_{2,2} & \cdots & a_{2,N} \\ \vdots & \vdots & & \vdots \\ 0 & a_{N,2} & \cdots & a_{N,N} \end{array}\right)\right] \\ &=\; 0 \cdot \det\left[\begin{pmatrix} a_{2,2} & \cdots & a_{2,N} \\ \vdots & & \vdots \\ a_{N,2} & \cdots & a_{N,N} \end{pmatrix}\right] \;=\; 0, \end{aligned} \tag{VIII.39}$$

   und wir sind fertig.

1.2 Ist $a_{1,1} \neq 0$, so setzen wir $j := 1$ und $B := A = S^{(1,1)} \cdot A$.

1.3 Sind schließlich $(a_{1,1}, \ldots, a_{N,1}) \neq (0, \ldots, 0)$, aber $a_{1,1} = 0$, so gibt es ein $j \in \mathbb{Z}_2^N$, sodass $a_{j,1} \neq 0$ (anderenfalls wären wir im Fall 1.1). Nun vertauschen wir die j. Zeile mit der 1. Zeile und nennen die resultierende Matrix wieder $B := S^{(1,j)} \cdot A$. Da $\det[S^{(1,j)}] = -1$ ist, wechselt die Determinante von A dabei nur das Vorzeichen, d.h. $\det[B] = -\det[A]$.

Beachte, dass dieser Schritt gerade die Anwendung der elementaren Zeilenoperation *Vertauschung der* 1. *Zeile mit der* j. *Zeile* ist.

Beide Fälle 1.2 und 1.3 zusammenfassend stellen wir fest, dass danach $b_{1,1} \neq 0$ gilt, wobei

$$B \;=\; S^{(1,j)} \cdot A \;=\; \begin{pmatrix} b_{1,1} & \cdots & b_{1,N} \\ \vdots & & \vdots \\ b_{N,1} & \cdots & b_{N,N} \end{pmatrix}, \tag{VIII.40}$$

und dass weiterhin

$$\det[A] \;=\; (-1)^{1-\delta_{1,j}} \; \det[B] \tag{VIII.41}$$

gilt.

2. Wir setzen nun $\eta_k := -b_{k,1}/b_{1,1}$, für $k = 2, 3, \ldots, N$, multiplizieren B von links mit

$$G \;=\; \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ \eta_2 & 1 & \ddots & & 0 \\ \eta_3 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \eta_N & 0 & \cdots & 0 & 1 \end{pmatrix}, \tag{VIII.42}$$

und erhalten nach (VIII.24)

$$\begin{aligned} G \cdot B \;&=\; \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,N} \\ b_{2,1} + \eta_2 b_{1,1} & b_{2,2} + \eta_2 b_{1,2} & \cdots & b_{2,N} + \eta_2 b_{1,N} \\ \vdots & \vdots & & \vdots \\ b_{N,1} + \eta_N b_{1,1} & b_{N,2} + \eta_N b_{1,2} & \cdots & b_{N,N} + \eta_N b_{1,N} \end{pmatrix} \\ &=\; \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,N} \\ 0 & b_{2,2} + \eta_2 b_{1,2} & \cdots & b_{2,N} + \eta_2 b_{1,N} \\ \vdots & \vdots & & \vdots \\ 0 & b_{N,2} + \eta_N b_{1,2} & \cdots & b_{N,N} + \eta_N b_{1,N} \end{pmatrix}. \end{aligned} \tag{VIII.43}$$

Wir bemerken, dass auch dieser Schritt als $(N-1)$-mal ausgeführte elementaren Zeilenoperationen aufgefasst werden kann, nämlich *Addition des* η_j*-fachen der* 1. *Zeile zur* j. *Zeile* für $j = 2, 3, \ldots, N$ ist.

3. Nach dem Kästchensatz sind $\det[G] = 1$ und

$$\det[A] \;=\; (-1)^{1-\delta_{1,j}} \; \det[B] \;=\; (-1)^{1-\delta_{1,j}} \; \det[G \cdot B] \;=\; (-1)^{1-\delta_{1,j}} \; b_{1,1} \; \det[\widetilde{A}], \tag{VIII.44}$$

wobei $\widetilde{A} \in \mathbb{K}^{(N-1)\times(N-1)}$ eine $(N-1)\times(N-1)$-Matrix ist, die durch

$$\widetilde{A} \;=\; \begin{pmatrix} b_{2,2}+\eta_2 b_{1,2} & \cdots & b_{2,N}+\eta_2 b_{1,N} \\ \vdots & & \vdots \\ b_{N,2}+\eta_N b_{1,2} & \cdots & b_{N,N}+\eta_N b_{1,N} \end{pmatrix} \tag{VIII.45}$$

gegeben ist.

4. Damit haben wir die Berechnung der Determinante der $N\times N$-Matrix A auf die Berechnung der Determinante der $(N-1)\times(N-1)$-Matrix $\widetilde{A}$ zurückgeführt. Wenden wir dieses Verfahren N-mal hintereinander an, erhalten wir schließlich die Determinante von A.
5. Zur Illustration ein Beispiel:

$$\begin{aligned} \det\left[\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}\right] \;&=\; \det\left[\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix}\right] \;=\; \det\left[\begin{pmatrix} 1 & 2 & 2 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix}\right] \\ &= 0. \end{aligned} \tag{VIII.46}$$

VIII.3. Der Gauß-Algorithmus zur Berechnung des Inversen einer Matrix

Wir skizzieren nun noch den Gauß-Algorithmus zur Berechnung der Inversen einer Matrix. Seien $N \in \mathbb{N}$, und

$$A \;=\; \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix} \;\in\; \mathbb{K}^{N\times N} \tag{VIII.47}$$

eine invertible $N\times N$-Matrix, für die also $\det[A] \neq 0$ gelten soll (was wir vorher z. B. mit dem Gauß-Algorithmus zur Berechnung der Determinante überprüft haben).

- Wir setzen $A^{(0)} := A$ und $a_{i,j}^{(0)} := a_{i,j}$. Wir wenden zunächst dieselben Schritte wie zur Berechnung der Determinanten mit dem Gauß-Algorithmus an und erhalten

$$A^{(1)} \;=\; G_1 \cdot S^{(1,j_1)} \cdot A^{(0)}\,, \tag{VIII.48}$$

wobei $j_1 \geq 1$,

$$A^{(1)} \;=\; \begin{pmatrix} a_{1,1}^{(1)} & a_{1,2}^{(1)} & \cdots & a_{1,N}^{(1)} \\ 0 & a_{2,2}^{(1)} & \cdots & a_{2,N}^{(1)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{N,2}^{(1)} & \cdots & a_{N,N}^{(1)} \end{pmatrix} \tag{VIII.49}$$

und

$$G_1 \;=\; \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ \eta_2^{(1)} & 1 & \ddots & & 0 \\ \eta_3^{(1)} & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \eta_N^{(1)} & 0 & \cdots & 0 & 1 \end{pmatrix}. \tag{VIII.50}$$

- Anschließend wiederholen wir das Verfahren für die rechte untere $(N-1)\times(N-1)$-Teilmatrix in $A^{(1)}$. Wir erhalten

$$A^{(2)} \;=\; G_2 \cdot S^{(1,j_2)} \cdot A^{(1)} \;=\; G_2 \cdot S^{(1,j_2)} \cdot G_1 \cdot S^{(1,j_1)} \cdot A^{(0)}, \tag{VIII.51}$$

wobei $j_2 \geq 2$,

$$A^{(2)} \;=\; \begin{pmatrix} a_{1,1}^{(2)} & a_{1,2}^{(2)} & a_{1,3}^{(2)} & \cdots & a_{1,N}^{(2)} \\ 0 & a_{2,2}^{(2)} & a_{2,3}^{(2)} & \cdots & a_{2,N}^{(2)} \\ 0 & 0 & a_{3,3}^{(2)} & \cdots & a_{3,N}^{(2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{N,3}^{(2)} & \cdots & a_{N,N}^{(2)} \end{pmatrix} \tag{VIII.52}$$

und

$$G_2 \;=\; \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & & \\ 0 & \eta_3^{(2)} & 1 & \ddots & & \vdots \\ 0 & \eta_4^{(2)} & 0 & \ddots & \ddots & \vdots \\ 0 & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \eta_N^{(2)} & 0 & \cdots & 0 & 1 \end{pmatrix}. \tag{VIII.53}$$

- Wenden wir dieses Verfahren $(N-1)$-mal an, so erhalten wir eine rechte obere Dreiecksmatrix

$$A^{(N-1)} \;=\; G_{N-1} \cdot S^{(N-1,j_{N-1})} \cdots G_1 \cdot S^{(1,j_1)} \cdot A^{(0)}, \tag{VIII.54}$$

wobei $j_k \geq k$, für alle $k \in \mathbb{Z}_1^{N-1}$,

$$A^{(N-1)} \;=\; \begin{pmatrix} a_{1,1}^{(N-1)} & a_{1,2}^{(N-1)} & \cdots & a_{1,N}^{(N-1)} \\ 0 & a_{2,2}^{(N-1)} & \cdots & a_{2,N}^{(N-1)} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{N,N}^{(N-1)} \end{pmatrix} \tag{VIII.55}$$

und

$$G_k \;=\; \begin{pmatrix} 1 & 0 & & \cdots & & \cdots & & 0 \\ 0 & \ddots & \ddots & & & & & \vdots \\ \vdots & \ddots & 1 & 0 & & & & \vdots \\ 0 & & 0 & 1 & 0 & & & \vdots \\ \vdots & & \vdots & \eta_{k+1}^{(k)} & 1 & 0 & & \vdots \\ \vdots & & \vdots & \eta_{k+2}^{(k)} & 0 & \ddots & \ddots & \vdots \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \eta_N^{(k)} & 0 & \cdots & 0 & 1 \end{pmatrix}. \tag{VIII.56}$$

- Um die Notation übersichtlich zu halten, schreiben wir $\widehat{D} = (d_{i,j})_{i,j=1}^N := A^{(N-1)}$, also

$$\widehat{D} \;:=\; A^{(N-1)} \;=\; \begin{pmatrix} d_{1,1} & d_{1,2} & \cdots & d_{1,N} \\ 0 & d_{2,2} & \cdots & d_{2,N} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & d_{N,N} \end{pmatrix}, \tag{VIII.57}$$

und beobachten, dass

$$\forall\, j \in \mathbb{Z}_1^N : \qquad d_{j,j} \;\neq\; 0\,, \tag{VIII.58}$$

da sich $\det[\widehat{D}]$ und $\det[A]$ höchstens im Vorzeichen unterscheiden und wir eingangs $\det[A] \neq 0$ vorausgesetzt hatten (da A anderenfalls ohnehin nicht invertibel wäre).

- Nun wenden wir das analoge Verfahren zum eben durchgeführten an, diesmal aber zur Elimination der Spalten *oberhalb* der Diagonalen von $\widehat{D}$. Wegen der rechten oberen Dreiecksstruktur und (VIII.58) sind Zeilenvertauschungen nicht mehr notwendig.

- Auch die Definition der Matrizen zur Zeilenaddition ist einfach: Im ersten Schritt setzen wir $\hat{\eta}_i^{(N)} := -d_{i,N}/d_{N,N}$ und

$$\widehat{G}_N \;=\; \begin{pmatrix} 1 & 0 & \cdots & 0 & \hat{\eta}_1^{(N)} \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & & \ddots & 0 \\ 0 & & \ddots & \ddots & \hat{\eta}_{N-1}^{(N)} \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix}. \tag{VIII.59}$$

Multiplizieren wir eine Matrix A von links mit $\widehat{G}_N$, so wird das $\hat{\eta}_i^{(N)}$-fache der N. Zeile zur i. Zeile von A hinzuaddiert.

Hier multiplizieren wir $\widehat{D}$ von links mit $\widehat{G}_N$ und erreichen dadurch, dass die N. Spalte oberhalb der Diagonalen eliminiert und $\widehat{D}$ sonst aber unverändert gelassen wird,

$$\widehat{G}_N \cdot \widehat{D} \;=\; \begin{pmatrix} d_{1,1} & d_{1,2} & \cdots & d_{1,N-1} & 0 \\ 0 & d_{2,2} & \cdots & d_{2,N-1} & \vdots \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & d_{N-1,N-1} & 0 \\ 0 & \cdots & 0 & 0 & d_{N,N} \end{pmatrix}. \tag{VIII.60}$$

- Allgemein gehen wir $k \in \mathbb{Z}_2^N$ genauso vor: Wir definieren $\hat{\eta}_i^{(k)} := -d_{i,k}/d_{k,k}$ und

$$\widehat{G}_k \;=\; \begin{pmatrix} 1 & 0 & \cdots & 0 & \hat{\eta}_1^{(k)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & 1 & 0 & \hat{\eta}_{k-2}^{(k)} & \vdots & & \vdots \\ \vdots & & \ddots & 1 & \hat{\eta}_{k-1}^{(k)} & \vdots & & \vdots \\ \vdots & & & \ddots & 1 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & & \ddots & \ddots & 0 \\ 0 & & \cdots & & \cdots & \cdots & 0 & 1 \end{pmatrix} \tag{VIII.61}$$

und beobachten, dass wieder $\det[\widehat{G}_k] = 1$ und

$$\widehat{G}_k^{-1} \;=\; \begin{pmatrix} 1 & 0 & \cdots & 0 & -\hat{\eta}_1^{(k)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & 1 & 0 & -\hat{\eta}_{k-2}^{(k)} & \vdots & & \vdots \\ \vdots & & \ddots & 1 & -\hat{\eta}_{k-1}^{(k)} & \vdots & & \vdots \\ \vdots & & & \ddots & 1 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & & \ddots & \ddots & 0 \\ 0 & & \cdots & & \cdots & \cdots & 0 & 1 \end{pmatrix} \tag{VIII.62}$$

- Multiplizieren wir sukzessiv die Matrizen $\widehat{G}_N, \widehat{G}_{N-1}, \ldots, \widehat{G}_2$ von links an $\widehat{D}$, so erhalten wir die Diagonalmatrix

$$D \;:=\; \widehat{G}_2 \cdot \widehat{G}_3 \cdots \widehat{G}_N \cdot \widehat{D} \;=\; \begin{pmatrix} d_{1,1} & 0 & \cdots & 0 \\ 0 & d_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_{N,N} \end{pmatrix}. \tag{VIII.63}$$

Diese Matrix lässt sich jedoch leicht invertieren, nämlich

$$D^{-1} \;=\; \begin{pmatrix} d_{1,1}^{-1} & 0 & \cdots & 0 \\ 0 & d_{2,2}^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_{N,N}^{-1} \end{pmatrix}, \tag{VIII.64}$$

wobei (VIII.58) sicherstellt, dass die Kehrwerte $d_{j,j}^{-1} \in \mathbb{K}$ existieren. Insgesamt ist also

$$
\begin{aligned}
\mathbb{1} &= D^{-1} D \;=\; D^{-1} \widehat{G}_2 \cdots \widehat{G}_N \widehat{D} \;=\; D^{-1} \widehat{G}_2 \cdots \widehat{G}_N A^{(N-1)} \\
&= D^{-1} \widehat{G}_2 \cdots \widehat{G}_N G_{N-1} S^{(N-1,j_{N-1})} \cdots G_1 S^{(1,j_1)} A^{(0)} \\
&= D^{-1} \widehat{G}_2 \cdots \widehat{G}_N G_{N-1} S^{(N-1,j_{N-1})} \cdots G_1 S^{(1,j_1)} A , \qquad \text{(VIII.65)}
\end{aligned}
$$

und das Matrixprodukt links von A ist die gesuchte inverse Matrix,

$$
A^{-1} \;=\; D^{-1} \widehat{G}_2 \cdots \widehat{G}_N G_{N-1} S^{(N-1,j_{N-1})} G_{N-2} S^{(N-1,j_{N-1})} \cdots G_1 S^{(1,j_1)} . \qquad \text{(VIII.66)}
$$

- Zur konkreten Durchführung der oben beschriebenen Schritte gibt es ein praktisches Schema, das aus der Matrix A und rechts daneben der Einheitsmatrix besteht,

$$
\big| \, A \, \big| \, \mathbb{1} \, \big| \;=\; \left| \begin{array}{cccc|cccc} a_{1,1} & a_{1,2} & \cdots & a_{1,N} & 1 & 0 & \cdots & 0 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \ddots & 0 \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} & 0 & \cdots & 0 & 1 \end{array} \right| . \qquad \text{(VIII.67)}
$$

Nun wird die erste Multiplikation von links mit $S^{(1,j_1)}$ *beider Seiten* durchgeführt, d.h. man berechnet

$$
\big| \, S^{(1,j_1)} \cdot A \, \big| \, S^{(1,j_1)} \cdot \mathbb{1} \, \big| \;=\; \big| \, S^{(1,j_1)} A \, \big| \, S^{(1,j_1)} \, \big| . \qquad \text{(VIII.68)}
$$

Dies ist leicht durchzuführen, denn wir wissen, dass $S^{(1,j_1)}$ nur die Vertauschung der 1. mit der j_1. Zeile bewirkt. Wir führen also diese Vertauschung sowohl für A auf der linken als auch für die Einheitsmatrix $\mathbb{1}$ auf der rechten Seite durch.

Danach multiplizieren wir beide Matrizen mit G_1,

$$
\big| \, G_1 S^{(1,j_1)} A \, \big| \, G_1 S^{(1,j_1)} \, \big| . \qquad \text{(VIII.69)}
$$

Dadurch werden die $\eta_j^{(1)}$-fachen der ersten auf die j. Zeile addiert, wir brauchen also auch hier keine Matrixmultiplikationen vorzunehmen.

So fahren wir fort und erhalten mit $S_k := S^{(1,j_1)}$ schließlich

$$
\begin{aligned}
&\big| \, D^{-1} \widehat{G}_2 \cdots \widehat{G}_N G_{N-1} S_{N-1} \cdots G_1 S_1 A \, \big| \, D^{-1} \widehat{G}_2 \cdots \widehat{G}_N G_{N-1} S_{N-1} \cdots G_1 S_1 \, \big| \\
&\quad = \big| \, \mathbb{1} \, \big| \, A^{-1} \, \big| . \qquad \text{(VIII.70)}
\end{aligned}
$$

Durch die sukzessive Anwendung geschickt gewählter Zeilenoperationen haben wir also die linke Seite des Schemas von A in $\mathbb{1}$ und die rechte Seite von $\mathbb{1}$ in A^{-1} umgewandelt, und wir können die inverse Matrix auf der rechten Seite direkt ablesen. Man kann sich rechts auch die einzelnen Schritte notieren - dies bewahrt einem vor fehlerhaften Berechnungen.

- Ein konkretes Beispiel: $\mathbb{K} = \mathbb{R}$, $N = 3$ und

$$A = \begin{pmatrix} 0 & 1 & 4 \\ 1 & 1 & 1 \\ 2 & 3 & 5 \end{pmatrix} \in \mathbb{R}^{3\times 3}. \tag{VIII.71}$$

Das Schema sieht zu Beginn so aus,

$$\left| \, A \, \middle| \, \mathbb{1} \, \right| = \left| \begin{array}{ccc|ccc} 0 & 1 & 4 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 2 & 3 & 5 & 0 & 0 & 1 \end{array} \right| \quad \begin{array}{l} (i)_0 \\ (ii)_0 \\ (iii)_0 \end{array} . \tag{VIII.72}$$

Wir schreiben ab jetzt nur noch die Schemata auf. Mit $j_1 = 2$ wird

$$\longrightarrow \quad \left| \begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 4 & 1 & 0 & 0 \\ 2 & 3 & 5 & 0 & 0 & 1 \end{array} \right| \quad \begin{array}{l} (i)_1 = (ii)_0 \\ (ii)_1 = (i)_0 \\ (iii)_1 = (iii)_0 \end{array} \tag{VIII.73}$$

$$\longrightarrow \quad \left| \begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & -2 & 1 \end{array} \right| \quad \begin{array}{l} (i)_2 = (i)_1 \\ (ii)_2 = (ii)_1 \\ (iii)_2 = (iii)_1 - 2 \cdot (i)_1 \end{array} \tag{VIII.74}$$

$$\longrightarrow \quad \left| \begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 4 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & -2 & 1 \end{array} \right| \quad \begin{array}{l} (i)_3 = (i)_2 \\ (ii)_3 = (ii)_2 \\ (iii)_3 = (iii)_2 - (ii)_2 \end{array} \tag{VIII.75}$$

$$\longrightarrow \quad \left| \begin{array}{ccc|ccc} 1 & 1 & 0 & -1 & -1 & 1 \\ 0 & 1 & 0 & -3 & -8 & 4 \\ 0 & 0 & -1 & -1 & -2 & 1 \end{array} \right| \quad \begin{array}{l} (i)_4 = (i)_3 + (iii)_3 \\ (ii)_4 = (ii)_3 + 4 \cdot (iii)_3 \\ (iii)_4 = (iii)_3 \end{array} \tag{VIII.76}$$

$$\longrightarrow \quad \left| \begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 7 & -3 \\ 0 & 1 & 0 & -3 & -8 & 4 \\ 0 & 0 & -1 & -1 & -2 & 1 \end{array} \right| \quad \begin{array}{l} (i)_5 = (i)_4 - (ii)_4 \\ (ii)_5 = (ii)_4 \\ (iii)_5 = (iii)_4 \end{array} \tag{VIII.77}$$

$$\longrightarrow \quad \left| \begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 7 & -3 \\ 0 & 1 & 0 & -3 & -8 & 4 \\ 0 & 0 & 1 & 1 & 2 & -1 \end{array} \right| \quad \begin{array}{l} (i)_6 = (i)_5 \\ (ii)_6 = (ii)_5 \\ (iii)_6 = (iii)_5/(-1) \end{array} , \tag{VIII.78}$$

also ist

$$A^{-1} = \begin{pmatrix} 2 & 7 & -3 \\ -3 & -8 & 4 \\ 1 & 2 & -1 \end{pmatrix} . \tag{VIII.79}$$

IX. Skalarprodukte

In diesem Kapitel wollen wir wie schon zuvor voraussetzen, dass der zugrunde liegende Körper $\mathbb{K}$ der Körper $\mathbb{R}$ der reellen Zahlen oder der Körper $\mathbb{C}$ der komplexen Zahlen ist. In den vorigen Kapiteln war diese Voraussetzung jedoch entbehrlich, und die dort bewiesenen Resultate würde ihre Gültigkeit behalten, wenn $\mathbb{K}$ irgendein Körper wäre. Im Gegensatz dazu ist in diesem und den folgenden Kapiteln die Annahme, dass $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ gilt, wesentlich.

IX.1. Quadratische Formen und Skalarprodukte

Definition IX.1. Sei X ein $\mathbb{K}$-Vektorraum.

(i) Eine Abbildung $Q : X \times X \to \mathbb{K}$ heißt **quadratische Form (auf X)**

$$:\Leftrightarrow \quad \forall\, \alpha, \beta \in \mathbb{K},\ \vec{x}, \vec{y}, \vec{w}, \vec{z} \in X :$$

$$\begin{aligned} &Q(\alpha\vec{x} + \vec{y}, \beta\vec{w} + \vec{z}) \qquad \text{(IX.1)} \\ &= \begin{cases} \alpha\beta Q(\vec{x}, \vec{w}) + \alpha Q(\vec{x}, \vec{z}) + \beta Q(\vec{y}, \vec{w}) + Q(\vec{y}, \vec{z}), & \text{falls } \mathbb{K} = \mathbb{R}, \\ \bar{\alpha}\beta Q(\vec{x}, \vec{w}) + \bar{\alpha} Q(\vec{x}, \vec{z}) + \beta Q(\vec{y}, \vec{w}) + Q(\vec{y}, \vec{z}), & \text{falls } \mathbb{K} = \mathbb{C}. \end{cases} \end{aligned}$$

Für $\mathbb{K} = \mathbb{R}$ bezeichnet man die Eigenschaft (IX.1) als **Bilinearität**, für $\mathbb{K} = \mathbb{C}$ heißt (IX.1) **Sesquilinearität**.

(ii) Eine quadratische Form $Q : X \times X \to \mathbb{K}$ heißt **symmetrisch**

$$:\Leftrightarrow\ \forall\, \vec{x}, \vec{y} \in X : \quad Q(\vec{x}, \vec{y}) \ = \begin{cases} Q(\vec{y}, \vec{x}), & \text{falls } \mathbb{K} = \mathbb{R}, \\ \overline{Q(\vec{y}, \vec{x})}, & \text{falls } \mathbb{K} = \mathbb{C}. \end{cases} \qquad \text{(IX.2)}$$

(iii) Eine symmetrische quadratische Form $Q : X \times X \to \mathbb{K}$, für die für alle $\vec{x} \in X \setminus \{\vec{0}\}$

$Q(\vec{x}, \vec{x}) > 0$ gilt, heißt **positiv definit**, (IX.3)

$Q(\vec{x}, \vec{x}) \geq 0$ gilt, heißt **positiv semidefinit**, (IX.4)

$Q(\vec{x}, \vec{x}) < 0$ gilt, heißt **negativ definit**, (IX.5)

$Q(\vec{x}, \vec{x}) \leq 0$ gilt, heißt **negativ semidefinit**. (IX.6)

Gibt es Vektoren $\vec{x}, \vec{y} \in X$ so, dass $Q(\vec{x}, \vec{x}) > 0$ und $Q(\vec{y}, \vec{y}) < 0$ gelten, so heißt $Q : X \times X \to \mathbb{K}$ **indefinit**.

(iv) Eine Abbildung $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{K}$ heißt **Skalarprodukt (auf $\boldsymbol{X}$)**

$$:\Leftrightarrow \ \langle\cdot|\cdot\rangle : X \times X \to \mathbb{K} \text{ ist eine positiv definite symmetrische quadratische Form.} \tag{IX.7}$$

Bemerkungen und Beispiele.

- Quadratische Formen auf $\mathbb{R}$-Vektorräumen heißen auch Bilinearformen.
- Quadratische Formen auf $\mathbb{C}$-Vektorräumen heißen auch Sesquilinearformen. Die Eigenschaft $Q(\alpha\vec{x}, \vec{y}) = \overline{\alpha} Q(\vec{x}, \vec{y})$ nennt man auch Antilinearität.
- Die hier eingeführte Terminologie ist in der Literatur nicht ganz einheitlich. Häufig werden bilineare bzw. sesquilineare Abbildungen als Skalarprodukte bezeichnet, und Definitheit ist eine zusätzliche Eigenschaft.
- Aus (*ii*) folgt, dass für eine symmetrische quadratische Form $Q(\vec{x}, \vec{x}) \in \mathbb{R}$ stets reell ist (auch für $\mathbb{K} = \mathbb{C}$). Es gilt also stets $Q(\vec{x}, \vec{x}) > 0$, $Q(\vec{x}, \vec{x}) < 0$ oder $Q(\vec{x}, \vec{x}) = 0$.
- Für $d \in \mathbb{N}$ und $(\alpha_1, \ldots, \alpha_d)^T, (\beta_1, \ldots, \beta_d)^T \in \mathbb{R}^d$ ist

$$\left\langle \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} \middle| \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_d \end{pmatrix} \right\rangle := \alpha_1\beta_1 + \alpha_2\beta_2 + \ldots + \alpha_d\beta_d \tag{IX.8}$$

  das euklidische Skalarprodukt auf $\mathbb{R}^d$.
- Für $d \in \mathbb{N}$ und $(\alpha_1, \ldots, \alpha_d)^T, (\beta_1, \ldots, \beta_d)^T \in \mathbb{C}^d$ ist

$$\left\langle \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} \middle| \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_d \end{pmatrix} \right\rangle := \bar{\alpha}_1\beta_1 + \bar{\alpha}_2\beta_2 + \ldots + \bar{\alpha}_d\beta_d \tag{IX.9}$$

  das unitäre Skalarprodukt auf $\mathbb{C}^d$.
- Identifiziert man $\mathbb{R}^d$ mit $(\mathrm{Re}\,\mathbb{C})^d$ durch

$$\mathbb{R}^d \ni (\alpha_1, \ldots, \alpha_d)^T \mapsto (\alpha_1 + i0, \ldots, \alpha_d + i0)^T \in \mathbb{C}^d, \tag{IX.10}$$

  so fallen euklidisches und unitäres Skalarprodukt zusammen. Wir können also stets $\mathbb{R}^d$ als reellen Teilraum von $\mathbb{C}^d$ betrachten, auf dem das unitäre Skalarprodukt (IX.9) die Form (IX.8) des euklidischen annimmt.
- Auf $\mathbb{R}^4$ ist durch

$$Q\left[\begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}, \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} \right] := \alpha_0\beta_0 - \alpha_1\beta_1 - \alpha_2\beta_2 - \alpha_3\beta_3 \tag{IX.11}$$

  eine indefinite quadratische Form definiert, wobei $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)^T \in \mathbb{R}^4$ und $(\beta_0, \beta_1, \beta_2, \beta_3)^T \in \mathbb{R}^4$. Glg. (IX.11) wird auch als Minkowski-Skalarprodukt bezeichnet (obwohl es indefinit und somit kein Skalarprodukt ist); $\mathbb{R}^4$ nennt man in diesem Zusammenhang auch den Minkowskiraum und schreibt $\mathbb{M}^4$ oder $\mathbb{R}^{1,3}$ statt $\mathbb{R}^4$.

- Sei $X = \mathbb{C}[x]$ der $\mathbb{C}$-Vektorraum der Polynome in x mit komplexen Koeffizienten. Für $p, q \in \mathbb{C}[x]$ definiert

$$\langle p|q\rangle := \int_{-\infty}^{\infty} \overline{p(x)}\, q(x)\, \frac{e^{-x^2}}{\sqrt{2\pi}}\, dx \tag{IX.12}$$

ein Skalarprodukt auf $\mathbb{C}[x]$.

Satz IX.2 (Cauchy-Schwarzsche Ungleichung)**.** *Sei X ein $\mathbb{K}$-Vektorraum mit einer positiv semidefiniten symmetrischen quadratische Form $Q : X \times X \to \mathbb{K}$. Dann gilt*

$$\forall \vec{x}, \vec{y} \in X : \quad |Q(\vec{x}, \vec{y})| \;\le\; \sqrt{Q(\vec{x}, \vec{x})} \cdot \sqrt{Q(\vec{y}, \vec{y})}\,. \tag{IX.13}$$

Beweis. Wir führen den Beweis nur für $\mathbb{K} = \mathbb{C}$. Für $Q(\vec{x}, \vec{y}) = 0$ gilt (IX.13) trivialerweise, und wir können im Folgenden $Q(\vec{x}, \vec{y}) \neq 0$ annehmen. Für jedes $\lambda \in \mathbb{K}$ ist dann

$$\begin{aligned} 0 \;&\le Q(\lambda\vec{x} + \vec{y}\,,\, \lambda\vec{x} + \vec{y}) \;=\; \bar{\lambda}\,\lambda\, Q(\vec{x}, \vec{x}) \;+\; Q(\vec{y}, \vec{y}) \;+\; \bar{\lambda}\, Q(\vec{x}, \vec{y}) \;+\; \lambda\, Q(\vec{y}, \vec{x}) \\ &= |\lambda|^2\, Q(\vec{x}, \vec{x}) \;+\; Q(\vec{y}, \vec{y}) \;+\; 2\,\mathrm{Re}\{\bar{\lambda}\, Q(\vec{x}, \vec{y})\}\,. \end{aligned} \tag{IX.14}$$

Wir setzen

$$\lambda := -r \cdot \frac{Q(\vec{x}, \vec{y})}{|Q(\vec{x}, \vec{y})|}\,, \tag{IX.15}$$

wobei $r \ge 0$ später gewählt wird. Dann sind $|\lambda| = r$ und

$$2\mathrm{Re}\{\bar{\lambda}\, Q(\vec{x}, \vec{y})\} \;=\; 2\mathrm{Re}\bigg\{ -\frac{r\,|Q(\vec{x}, \vec{y})|^2}{|Q(\vec{x}, \vec{y})|} \bigg\} \;=\; -2r\,|Q(\vec{x}, \vec{y})|\,. \tag{IX.16}$$

Damit folgt aus (IX.14)

$$\forall\, r \in \mathbb{R}_0^+ : \qquad r\,|Q(\vec{x}, \vec{y})| \;\le\; \frac{r^2}{2}\, Q(\vec{x}, \vec{x}) \;+\; \frac{1}{2}\, Q(\vec{y}, \vec{y})\,. \tag{IX.17}$$

Wäre nun $Q(\vec{x}, \vec{x}) = 0$, so erhielten wir durch die Wahl $r := Q(\vec{y}, \vec{y})\, |Q(\vec{x}, \vec{y})|^{-1}$ die Aussage $|Q(\vec{y}, \vec{y})| \le |Q(\vec{y}, \vec{y})|/2$, was $|Q(\vec{y}, \vec{y})| = 0$ impliziert. Damit wären also $Q(\vec{x}, \vec{x}) = |Q(\vec{y}, \vec{y})| = 0$, und aus (IX.17) würde $r\,|Q(\vec{x}, \vec{y})| \le 0$ für jedes $r \ge 0$ folgen. Somit wäre auch $|Q(\vec{x}, \vec{y})| = 0$, was in Widerspruch zur anfangs gemachten Annahme $Q(\vec{x}, \vec{y}) \neq 0$ stünde. Es muss also $Q(\vec{x}, \vec{x}) > 0$ gelten. Durch die somit definierte Wahl $r := Q(\vec{y}, \vec{y})^{1/2}\, Q(\vec{x}, \vec{x})^{-1/2} \in \mathbb{R}_0^+$ erhalten wir die behauptete Ungleichung (IX.13). $\square$

Korollar IX.3 (Cauchy-Schwarzsche Ungleichung)**.** *Sei X ein $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{K}$. Dann gilt*

$$\forall \vec{x}, \vec{y} \in X : \quad |\langle\vec{x}|\vec{y}\rangle| \;\le\; \sqrt{\langle\vec{x}|\vec{x}\rangle} \cdot \sqrt{\langle\vec{y}|\vec{y}\rangle}. \tag{IX.18}$$

Bemerkungen und Beispiele.

- Satz IX.2 unterscheidet sich von Korollar IX.3 nur durch die etwas schwächere Annahme der positiven *Semi*definitheit.
- Dieser Unterschied ist jedoch von großer Bedeutung für verschiedene Beweise in der Mathematik, da die positive Semidefinitheit in vielen Situation einfach zu gewinnen ist, die positive Definitheit aber nicht.
- Ein prominentes Beispiel ist die GNS-Konstruktion (Gelfand, Naimark, Segal) der Theorie der C^*-Algebren.

IX.2. Skalarprodukte und Normen

Definition IX.4. Eine Abbildung $\|\cdot\| : X \to \mathbb{R}_0^+$ auf einem $\mathbb{K}$-Vektorraum X heißt **Norm**, falls folgende drei Eigenschaften erfüllt sind,

$$\begin{aligned} &(i) \quad \forall\, \vec{x} \in X: && \Big(\|\vec{x}\| = 0 \Big) \quad \Leftrightarrow \quad \Big(\vec{x} = \vec{0} \Big), && \text{(IX.19)} \\ &(ii) \quad \forall\, \alpha \in \mathbb{K}, \vec{x} \in X: && \|\alpha \vec{x}\| \;=\; |\alpha| \cdot \|\vec{x}\|, && \text{(IX.20)} \\ &(iii) \quad \forall\, \vec{x}, \vec{y} \in X: && \|\vec{x} + \vec{y}\| \;\leq\; \|\vec{x}\| + \|\vec{y}\|, && \text{(IX.21)} \end{aligned}$$

und in diesem Fall heißt $\big(X, \|\cdot\|\big)$ **normierter Vektorraum**.

Satz IX.5. *Sei X ein $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{K}$. Dann wird X ein normierter Vektorraum mit*

$$\forall\, x \in X: \quad \|\vec{x}\| := \sqrt{\langle \vec{x} | \vec{x} \rangle}. \tag{IX.22}$$

Beweis. Eigenschaft (IX.19) folgt aus der positiven Definitheit $\langle \vec{x}|\vec{x}\rangle > 0$, für $\vec{x} \neq \vec{0}$. (Natürlich gilt $\langle \vec{0}|\vec{0}\rangle = \langle \vec{0}|0 \cdot \vec{0}\rangle = 0\langle \vec{0}|\vec{0}\rangle = 0$.) Die Homogenität (IX.20) folgt direkt aus der Bilinearität bzw. Sesquilinearität, etwa

$$\sqrt{\langle \alpha\vec{x} | \alpha\vec{x} \rangle} \;=\; \sqrt{\bar{\alpha} \cdot \alpha \langle \vec{x}|\vec{x}\rangle} \;=\; |\alpha| \cdot \sqrt{\langle \vec{x}|\vec{x}\rangle}, \tag{IX.23}$$

für $\mathbb{K} = \mathbb{C}$. Die Dreiecksungleichung (IX.21) resultiert aus der Cauchy-Schwarzschen Ungleichung (IX.13), denn

$$\begin{aligned} \langle \vec{x}+\vec{y} | \vec{x}+\vec{y} \rangle &= \langle \vec{x}|\vec{x}\rangle + \langle \vec{y}|\vec{y}\rangle + \langle \vec{x}|\vec{y}\rangle + \langle \vec{y}|\vec{x}\rangle \\ &\leq \langle \vec{x}|\vec{x}\rangle + \langle \vec{y}|\vec{y}\rangle + 2|\langle \vec{x}|\vec{y}\rangle| \\ &\leq \langle \vec{x}|\vec{x}\rangle + \langle \vec{y}|\vec{y}\rangle + 2\sqrt{\langle \vec{x}|\vec{x}\rangle} \cdot \sqrt{\langle \vec{y}|\vec{y}\rangle} \\ &= \Big(\sqrt{\langle \vec{x}|\vec{x}\rangle} + \sqrt{\langle \vec{y}|\vec{y}\rangle} \Big)^2, \end{aligned} \tag{IX.24}$$

und man erhält (IX.21) durch Wurzelziehen. $\square$

Bemerkungen und Beispiele.

- Für $X = \mathbb{R}^d$ mit $d \in \mathbb{N}$ induziert das euklidische Skalarprodukt die euklidische Norm,

$$\forall\, \vec{x} \;=\; \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} \in \mathbb{R}^d : \quad \|\vec{x}\|_{\mathrm{eukl}} \;:=\; \sqrt{\langle \vec{x} | \vec{x} \rangle} \;=\; \sqrt{\alpha_1^2 + \ldots + \alpha_d^2}, \tag{IX.25}$$

  und wir erhalten mit $\| \cdot \|_{\mathrm{eukl}}$ den üblichen Abstandsbegriff,

$$\mathrm{dist}(\vec{x}, \vec{y}) \;=\; \|\vec{x} - \vec{y}\|_{\mathrm{eukl}}. \tag{IX.26}$$

- Analog findet man einen Abstandsbegriff auf $\mathbb{C}^d$ durch $\|\vec{x} - \vec{y}\|_{\mathrm{unit}}$, wobei

$$\vec{x} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_d \end{pmatrix}, \quad \|\vec{x} - \vec{y}\|_{\mathrm{unit}} \;=\; \sqrt{|\alpha_1 - \beta_1|^2 + \ldots + |\alpha_d - \beta_d|^2}. \tag{IX.27}$$

- Es gibt aber auch Normen, die nicht durch ein Skalarprodukt induziert werden. Für $d \in \mathbb{N}$ und $X = \mathbb{C}^d$ sind, mit $\vec{x} = (\alpha_1, \ldots, \alpha_d)^T \in \mathbb{C}^d$,

$$\|\vec{x}\|_1 \;=\; |\alpha_1| + \ldots + |\alpha_d|, \quad \|\vec{x}\|_\infty \;=\; \max\big\{|\alpha_1|, \ldots, |\alpha_d|\big\} \tag{IX.28}$$

  und allgemeiner

$$\forall\, 1 \le p < \infty : \quad \|\vec{x}\|_p \;:=\; \Big(|\alpha_1|^p + |\alpha_2|^p + \ldots + |\alpha_d|^p\Big)^{1/p} \tag{IX.29}$$

  alles Normen auf X. Außer für $p = 2$ ($\| \cdot \|_2 = \| \cdot \|_{\mathrm{unit}}$) ist keine dieser Normen durch ein Skalarprodukt erzeugt.

IX.3. Orthogonalität und Orthonormalbasen

Definition IX.6. Sei X ein $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle \cdot | \cdot \rangle : X \times X \to \mathbb{K}$.

(i) Zwei Vektoren $\vec{x}, \vec{y} \in X$ heißen **orthogonal** oder **senkrecht zueinander,** $\boldsymbol{\vec{x} \perp \vec{y}}$

$$:\Leftrightarrow \qquad \langle \vec{x} | \vec{y} \rangle \;=\; 0. \tag{IX.30}$$

(ii) Ist $A \subseteq X$ eine Teilmenge, so heißt

$$A^\perp \;=\; \Big\{\vec{x} \in X \;\Big|\; \forall\, \vec{a} \in A : \; \langle \vec{a} | \vec{x} \rangle = 0\Big\} \tag{IX.31}$$

das **orthogonale Komplement zu** $\boldsymbol{A}$**.**

(iii) Eine Teilmenge $A \subseteq X$ heißt **orthonormal**

$$:\Leftrightarrow \quad \forall\, \vec{x}, \vec{y} \in A,\; \vec{x} \neq \vec{y} : \quad \langle \vec{x} | \vec{x} \rangle \;=\; \langle \vec{y} | \vec{y} \rangle \;=\; 1, \;\; \langle \vec{x} | \vec{y} \rangle \;=\; 0. \tag{IX.32}$$

(iv) Eine Teilmenge $A \subseteq X$ heißt **Orthonormalbasis (ONB)**

$$:\Leftrightarrow \quad A \text{ ist orthonormal und eine Basis in } X. \tag{IX.33}$$

Bemerkungen und Beispiele.

- Für jede Teilmenge $A \subseteq X$ ist $A^\perp$ ein Unterraum in X, denn offensichtlich ist $\vec{0} \in A^\perp$, und mit $\vec{x}, \vec{y} \in A^\perp$ und $\alpha \in \mathbb{K}$ ist

$$\forall \, \vec{a} \in A: \quad \langle \vec{a} | \alpha \vec{x} + \vec{y} \rangle \; = \; \alpha \langle \vec{a} | \vec{x} \rangle + \langle \vec{a} | \vec{y} \rangle \; = \; 0. \tag{IX.34}$$

  (Anwendung des Unterraumkriteriums.)

- Sind $A \subseteq X$ und $\vec{x} \in A \cap A^\perp$, so gilt $\|\vec{x}\|^2 = \langle \vec{x} | \vec{x} \rangle = 0$ und somit $\vec{x} = \vec{0}$. Also ist stets

$$A \cap A^\perp \; \subseteq \; \{\vec{0}\}. \tag{IX.35}$$

- Für $A \subseteq B \subseteq X$ ist $A^\perp \supseteq B^\perp$.
- Für $A \subseteq X$ ist $A^\perp = (\mathrm{span}(A))^\perp$, denn einerseits ist $A^\perp \supseteq (\mathrm{span}(A))^\perp$ wegen $A \subseteq (\mathrm{span}(A))$. Sind andererseits $\vec{x} \in A^\perp$ und $\vec{a} = \alpha_1 \vec{a}_1 + \ldots + \alpha_N \vec{a}_N \in \mathrm{span}(A)$ mit $\alpha_j \in \mathbb{K}$ und $\vec{a}_j \in A$, so gilt

$$\langle \vec{a} | \vec{x} \rangle \; = \; \sum_{j=1}^N \bar{\alpha}_j \, \underbrace{\langle \vec{a}_j | \vec{x} \rangle}_{=0} \; = \; 0, \tag{IX.36}$$

  also ist auch $\vec{x} \in (\mathrm{span}(A))^\perp$.

- Seien $X = \mathbb{R}^3$ mit dem euklidschen Skalarprodukt und $A := \{\vec{a}\} \subseteq B := \{\vec{a}, \vec{b}\}$ sowie $\vec{x} \in X$, mit

$$\vec{a} \; = \; \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \vec{b} \; = \; \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \vec{x} \; = \; \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{IX.37}$$

  Dann sind $\langle \vec{a} | \vec{x} \rangle = x_1 + x_2$ und $\langle \vec{b} | \vec{x} \rangle = x_2$. Somit gelten

$$\vec{x} \in A^\perp \; \Leftrightarrow \; x_2 = -x_1, \qquad \vec{x} \in B^\perp \; \Leftrightarrow \; x_2 = -x_1, \; x_2 = 0, \tag{IX.38}$$

  also

$$\begin{aligned} A^\perp \; &= \; \left\{ \begin{pmatrix} \alpha \\ -\alpha \\ \beta \end{pmatrix} \middle| \; \alpha, \beta \in \mathbb{R} \right\} \; = \; \mathrm{span}\left[\left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \right] \\ &\supseteq \; \left\{ \begin{pmatrix} 0 \\ 0 \\ \beta \end{pmatrix} \middle| \; \beta \in \mathbb{R} \right\} \; = \; \mathrm{span}\left[\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \right] \; = \; B^\perp. \end{aligned} \tag{IX.39}$$

- Eine orthonormale Teilmenge $A \subseteq X$ ist stets linear unabhängig. Ist nämlich eine endliche Teilmenge von A gegeben durch $\{\vec{a}_1, \ldots, \vec{a}_N\} \subseteq A$, so gilt nach (IX.32)

$$\forall\, i, j \in \mathbb{Z}_1^N : \quad \langle \vec{a}_i | \vec{a}_j \rangle \;=\; \delta_{i,j}. \tag{IX.40}$$

Sind nun $\alpha_1, \ldots, \alpha_N \in \mathbb{K}$ so, dass $\alpha_1 \vec{a}_1 + \ldots + \alpha_N \vec{a}_N = \vec{0}$, dann folgt für jedes $i \in \mathbb{Z}_1^N$, dass

$$0 \;=\; \langle \vec{a}_i | \vec{0} \rangle \;=\; \langle \vec{a}_i | \alpha_1 \vec{a}_1 + \ldots + \alpha_N \vec{a}_N \rangle \;=\; \sum_{j=1}^N \alpha_j \underbrace{\langle \vec{a}_i | \vec{a}_j \rangle}_{\delta_{i,j}} \;=\; \alpha_i. \tag{IX.41}$$

Also sind $\alpha_1 = \cdots = \alpha_N = 0$, und $\{\vec{a}_1, \ldots, \vec{a}_N\}$ ist linear unabhängig.

- Die Standardbasis $\{\vec{e}_1, \ldots, \vec{e}_N\} \subseteq \mathbb{R}^N$, mit

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \; \ldots, \; \vec{e}_N = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \tag{IX.42}$$

ist eine ONB bezüglich des euklidischen/unitären Skalarproduktes in $\mathbb{K}^N$.

Satz IX.7. *Seien X ein Vektorraum der Dimension $N < \infty$ mit Skalarprodukt $\langle \cdot | \cdot \rangle : X \times X \to \mathbb{K}$ und Orthonormalbasis $\mathcal{A} = \{\vec{a}_1, \ldots, \vec{a}_N\} \subseteq X$.*

(i) Für alle $\vec{x} \in X$ ist

$$\vec{x} \;=\; \sum_{j=1}^N \langle \vec{a}_j | \vec{x} \rangle \, \vec{a}_j. \tag{IX.43}$$

(ii) Ist $\Phi \in \mathcal{L}(X)$ eine lineare Abbildung, so ist ihre Matrixdarstellung bzgl. $\mathcal{A}$ gegeben durch

$$M_{\mathcal{A}}[\Phi] \;=\; \begin{pmatrix} \langle \vec{a}_1 | \Phi \vec{a}_1 \rangle & \langle \vec{a}_1 | \Phi \vec{a}_2 \rangle & \cdots & \langle \vec{a}_1 | \Phi \vec{a}_N \rangle \\ \langle \vec{a}_2 | \Phi \vec{a}_1 \rangle & \langle \vec{a}_2 | \Phi \vec{a}_2 \rangle & \cdots & \langle \vec{a}_2 | \Phi \vec{a}_2 \rangle \\ \vdots & \vdots & & \vdots \\ \langle \vec{a}_N | \Phi \vec{a}_1 \rangle & \langle \vec{a}_N | \Phi \vec{a}_2 \rangle & \cdots & \langle \vec{a}_N | \Phi \vec{a}_N \rangle \end{pmatrix}. \tag{IX.44}$$

Beweis. *(i):* Da $\mathcal{A}$ eine Basis ist, besitzt der Vektor $\vec{x} \in X$ eine eindeutige Darstellung $\vec{x} = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_N \vec{a}_N \in X$ als Linearkombination der Basisvektoren in $\mathcal{A}$. Bilden wir das Skalarprodukt mit $\vec{a}_j$, so erhalten wir

$$\langle \vec{a}_j | \vec{x} \rangle \;=\; \sum_{i=1}^N \alpha_i \, \langle \vec{a}_j | \vec{a}_i \rangle \;=\; \sum_{i=1}^N \alpha_i \, \delta_{i,j} \;=\; \alpha_j, \tag{IX.45}$$

für $j \in \mathbb{Z}_1^N$, und wir erhalten *(i)*.

(ii): Ist $\mathcal{M}_{\mathcal{A}}[\Phi] = (a_{i,j})_{i,j=1}^N \in \mathbb{K}^{N\times N}$, so gilt

$$\forall\, j \in \mathbb{Z}_1^N : \quad \Phi\vec{a}_j \;=\; \sum_{i=1}^N a_{i,j}\vec{a}_i. \tag{IX.46}$$

Also ist, für alle $k, j \in \mathbb{Z}_1^N$,

$$\langle \vec{a}_k | \Phi\vec{a}_j\rangle \;=\; \sum_{i=1}^N a_{i,j}\, \underbrace{\langle \vec{a}_k|\vec{a}_i\rangle}_{=\,\delta_{k,i}} \;=\; a_{k,j}. \tag{IX.47}$$

□

Definition IX.8. Seien $N \in \mathbb{N}$ und $A = (a_{kj,j})_{k,j=1}^N \in \mathbb{K}^{N\times N}$. Die **zu *A* adjungierte Matrix** ist durch

$$A^* \;=\; \big(\overline{A}\big)^T \;=\; \begin{pmatrix} \overline{a_{1,1}} & \overline{a_{2,1}} & \cdots & \overline{a_{N,1}} \\ \overline{a_{1,2}} & \overline{a_{2,2}} & \cdots & \overline{a_{N,2}} \\ \overline{a_{1,N}} & \overline{a_{2,N}} & \cdots & \overline{a_{N,N}} \end{pmatrix} \tag{IX.48}$$

definiert.

Bemerkungen und Beispiele.

- Es sind

$$\begin{pmatrix} 1+i & 2i \\ 3 & 2-i \end{pmatrix}^* \;=\; \begin{pmatrix} 1-i & 3 \\ -2i & 2+i \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}^* \;=\; \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}. \tag{IX.49}$$

- Für $A \in \mathbb{R}^{N\times N}$ ist $A^* = A^T$.

Korollar IX.9. *Seien X ein $\mathbb{K}$-Vektorraum der Dimension $N < \infty$ mit Skalarprodukt $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{K}$.*

(i) Zu jeder linearen Abbildung $\Phi \in \mathcal{L}(X)$ existiert genau eine lineare Abbildung $\Phi^* \in \mathcal{L}(X)$, so dass

$$\forall\, \vec{x}, \vec{y} \in X : \quad \langle \vec{x}|\Phi\vec{y}\rangle \;=\; \langle \Phi^*\vec{x}|\vec{y}\rangle. \tag{IX.50}$$

*Dabei heißt Φ^* die **zu Φ adjungierte lineare Abbildung**.*

(ii) Ist $\mathcal{A} = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_N\} \subseteq X$ eine ONB, so gilt

$$\mathcal{M}_{\mathcal{A}}[\Phi^*] \;=\; \big(\mathcal{M}_{\mathcal{A}}[\Phi]\big)^*, \tag{IX.51}$$

für jede lineare Abbildung $\Phi \in \mathcal{L}(X)$.

Beweis. Wir führen den Beweis unter der Annahme, dass wir über eine ONB $\mathcal{A} = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_N\} \subseteq X$ verfügen, was wir erst später mit den Schmidtschen Orthonormierungverfahren rechtfertigen.

Seien $\Phi \in \mathcal{L}(X)$ eine lineare Abbildung mit Matrixdarstellung $M = (m_{k,j})_{k,j=1}^N = \mathcal{M}_\mathcal{A}[\Phi]$ bezüglich der ONB $\mathcal{A}$. Nach Satz IX.7 (ii) ist dann

$$\forall\, k, j \in \mathbb{Z}_1^N : \quad m_{k,j} \;=\; \langle \vec{a}_k \,|\, \Phi \vec{a}_j \rangle \,. \tag{IX.52}$$

Wir definieren nun eine lineare Abbildung $\Phi^* \in \mathcal{L}(X)$ durch ihre Matrixdarstellung $\mathcal{M}_\mathcal{A}[\Phi^*] := \big(\overline{m_{j,k}}\big)_{k,j=1}^N = \overline{M}^T$. Kraft dieser Definition erfüllt $\mathcal{M}_\mathcal{A}[\Phi^*]$ damit (IX.51). Außerdem gilt für alle $k, j \in \mathbb{Z}_1^N$, dass

$$\langle \Phi^* \vec{a}_k \,|\, \vec{a}_j \rangle \;=\; \overline{\langle \vec{a}_j \,|\, \Phi^* \vec{a}_k \rangle} \;=\; \overline{\big(\overline{M}^T\big)_{j,k}} \;=\; m_{k,j} \;=\; \langle \vec{a}_k \,|\, \Phi \vec{a}_j \rangle \,. \tag{IX.53}$$

Somit gilt (IX.50) für alle Paare $\vec{a}_k, \vec{a}_j \in \mathcal{A}$ von Basisvektoren und wegen der Sesquilinearität des Skalarprodukts damit auch für alle Paare $\vec{x}, \vec{y} \in X$. □

IX.4. Das Schmidtsche Orthonormierungsverfahren

Zur Vorbereitung des unten beschriebenen Schmidtschen Orthonormierungsverfahrens definieren wir noch orthogonale Projektionen und formulieren eine allgemeine Version des Satzes von Pythagoras. Seien dazu X ein $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle \cdot | \cdot \rangle$, $\mathcal{A} = \{\vec{a}_1, \ldots, \vec{a}_N\} \subseteq X$ eine orthonormale Teilmenge und $Y := \mathrm{span}(\mathcal{A})$. Wir definieren lineare Abbildungen $P, P^\perp \in \mathcal{L}(X)$ durch $P^\perp := \mathbb{1} - P$ und

$$\forall \vec{x} \in X : \quad P\vec{x} \;:=\; \sum_{n=1}^N \langle \vec{a}_n | \vec{x} \rangle \, \vec{a}_n. \tag{IX.54}$$

Dann ist

$$\forall m \in \mathbb{Z}_1^N : \quad P\vec{a}_m \;:\; \sum_{n=1}^N \underbrace{\langle \vec{a}_n | \vec{a}_m \rangle}_{=\delta_{m,n}} \, \vec{a}_n \;=\; \vec{a}_m. \tag{IX.55}$$

Also ist für alle $\vec{y} = \alpha_1 \vec{a}_1 + \ldots + \alpha_N \vec{a}_N \in Y$

$$P\vec{y} \;=\; \sum_{n=1}^N \alpha_n P \vec{a}_n \;=\; \sum_{n=1}^N \alpha_n \vec{a}_n \;=\; \vec{y}. \tag{IX.56}$$

Weiterhin ist für alle $\vec{x} \in X$

$$P^2 \vec{x} \;=\; P\big(P\vec{x}\big) \;=\; P\bigg(\sum_{n=1}^N \langle \vec{a}_n | \vec{x} \rangle \, \vec{a}_n \bigg) \;=\; \sum_{n=1}^N \langle \vec{a}_n | \vec{x} \rangle \, P\vec{a}_n \;=\; \sum_{n=1}^N \langle \vec{a}_n | \vec{x} \rangle \, \vec{a}_n \;=\; P\vec{x}, \tag{IX.57}$$

d.h. es gilt

$$P^2 \;=\; P \tag{IX.58}$$

und daher auch $(P^\perp)^2 = (\mathbb{1}-P)^2 = \mathbb{1}-2P+P^2 = \mathbb{1}-P = P^\perp$ sowie $PP^\perp = P(\mathbb{1}-P) = P - P^2 = 0$ und $P^\perp P = (\mathbb{1}-P)P = P - P^2 = 0$, also insgesamt

$$P^2 \; = \; P, \quad \big(P^\perp\big)^2 = P^\perp, \quad P\,P^\perp = 0, \quad P^\perp\,P = 0. \tag{IX.59}$$

Schließlich ist für alle $\vec{x}, \vec{z} \in X$

$$\langle \vec{x} | P\vec{z} \rangle \; = \; \left\langle \vec{x} \,\middle|\, \sum_{n=1}^{N} \langle \vec{a}_n | \vec{z} \rangle \, \vec{a}_n \right\rangle \; = \; \sum_{n=1}^{N} \langle \vec{a}_n | \vec{z} \rangle \, \langle \vec{x} | \vec{a}_n \rangle \; = \; \left\langle \sum_{n=1}^{N} \langle \vec{a}_n | \vec{x} \rangle \, \vec{a}_n \,\middle|\, \vec{z} \right\rangle \; = \; \langle P\vec{x} | \vec{z} \rangle, \tag{IX.60}$$

und wir erhalten für alle $\vec{x} \in X$

$$\begin{aligned} \|\vec{x}\|^2 \; &= \langle \vec{x} | \vec{x} \rangle \; = \; \langle P\vec{x} + P^\perp\vec{x} | P\vec{x} + P^\perp\vec{x} \rangle \\ &= \langle P\vec{x} | P\vec{x} \rangle + \langle P^\perp\vec{x} | P\vec{x} \rangle + \langle P\vec{x} | P^\perp\vec{x} \rangle + \langle P^\perp\vec{x} | P^\perp\vec{x} \rangle \\ &= \|P\vec{x}\|^2 + \langle PP^\perp\vec{x} | \vec{x} \rangle + \langle \vec{x} | PP^\perp\vec{x} \rangle + \|P^\perp\vec{x}\|^2 \\ &= \|P\vec{x}\|^2 + \|P^\perp\vec{x}\|^2 \,. \end{aligned} \tag{IX.61}$$

P heißt **orthogonale Projektion (auf $\boldsymbol{Y}$)** und (IX.61) ist eine Verallgemeinerung des Satzes von Pythagoras.

Satz IX.10 (Schmidtsches Orthonormierungsverfahren)**.** *Ist X ein $\mathbb{K}$-Vektorraum mit einem Skalarprodukt $\langle\cdot|\cdot\rangle : X \to \mathbb{K}$ und einer abzählbaren Basis $\{\vec{b}_j\}_{j=1}^J$, wobei $J \in \mathbb{N}$ oder $J = \infty$, so besitzt X auch eine abzählbare Orthonormalbasis $\{\vec{a}_j\}_{j=1}^J$.*

Beweis. Wir führen den Beweis nur für $\mathbb{K} = \mathbb{C}$ und $\dim(X) = \infty$. Der Fall $\dim_{\mathbb{K}}(X) < \infty$ geht analog. Für $N \in \mathbb{N}$ seien $\{\vec{a}_1, \ldots, \vec{a}_N\} \subseteq X$ eine orthonormale und $\{\vec{a}_1, \ldots, \vec{a}_N, \vec{b}_{N+1}\} \subseteq X$ eine linear unabhängige Teilmenge. Wie in (IX.54) definieren wir die orthogonale Projektion $P_N \in \mathcal{L}(X)$ durch

$$P_N\vec{x} \; := \; \sum_{n=1}^{N} \langle \vec{a}_n | \vec{x} \rangle \vec{a}_n, \tag{IX.62}$$

und verwenden nun (IX.56)–(IX.61). Weil $\{\vec{a}_1, \ldots, \vec{a}_N, \vec{b}_{N+1}\} \subseteq X$ linear unabhängig ist, ist

$$P_N^\perp \vec{b}_{N+1} \; = \; \vec{b}_{N+1} - \sum_{n=1}^{N} \langle \vec{a}_n | \vec{b}_{N+1} \rangle \vec{a}_n \; \neq \; \vec{0}, \tag{IX.63}$$

und daher können wir

$$\vec{a}_{N+1} \; := \; \frac{P_N^\perp \vec{b}_{N+1}}{\big\| P_N^\perp \vec{b}_{N+1} \big\|} \tag{IX.64}$$

bilden. Beachte, dass $\vec{a}_{N+1} \in \operatorname{span}\big(\{\vec{a}_1, \ldots, \vec{a}_N, \vec{b}_{N+1}\}\big)$. Wegen

$$\vec{b}_{N+1} \; = \; P_N\vec{b}_{N+1} + P_N^\perp\vec{b}_{N+1} \; = \; \sum_{n=1}^{N} \langle \vec{a}_n | \vec{b}_{N+1} \rangle \, \vec{a}_n + \big\| P_N^\perp \vec{b}_{N+1} \big\| \, \vec{a}_{N+1} \tag{IX.65}$$

gilt umgekehrt auch $\vec{b}_{N+1} \in \mathrm{span}\big(\{\vec{a}_1, \ldots, \vec{a}_N, \vec{a}_{N+1}\}\big)$. Mit $Y_{N+1} := \mathrm{span}\big(\{\vec{a}_1, \ldots, \vec{a}_{N+1}\}\big)$ ist also

$$Y_{N+1} \;=\; \mathrm{span}\big(\{\vec{a}_1, \ldots, \vec{a}_N, \vec{a}_{N+1}\}\big) \;=\; \mathrm{span}\big(\{\vec{a}_1, \ldots, \vec{a}_N, \vec{b}_{N+1}\}\big). \tag{IX.66}$$

Offensichtlich ist $\|\vec{a}_{N+1}\| = 1$. Weiterhin sind $\vec{a}_{N+1} = P_N^\perp \vec{a}_{N+1}$ und $\vec{a}_k = P_N \vec{a}_k$, für alle $k \in \mathbb{Z}_1^N$, und deshalb

$$\big\langle \vec{a}_k \,\big|\, \vec{a}_{N+1} \big\rangle \;=\; \big\langle P_N \vec{a}_k \,\big|\, P_N^\perp \vec{a}_{N+1} \big\rangle \;=\; \big\langle \vec{a}_k \,\big|\, P_N P_N^\perp \vec{a}_{N+1} \big\rangle \;=\; 0, \tag{IX.67}$$

da $P_N P_N^\perp = 0$. Also ist

$$\forall\, i, j \in \mathbb{Z}_1^{N+1} : \quad \langle \vec{a}_i | \vec{a}_j \rangle \;=\; \delta_{i,j}, \tag{IX.68}$$

und $\{\vec{a}_1, \ldots, \vec{a}_{N+1}\} \subseteq Y_{N+1}$ ist eine Orthonormalbasis von Y_{N+1}. Sei nun $\{\vec{b}_n\}_{n=1}^\infty$ eine Basis in X. Wir setzen $\vec{a}_1 := \vec{b}_1 / \|\vec{b}_1\|$ und für $N \in \mathbb{N}$ definieren wir rekursiv

$$\vec{a}_{N+1} \;=\; \frac{P_N^\perp \vec{b}_{N+1}}{\big\|P_N^\perp \vec{b}_{N+1}\big\|} \;=\; \frac{\vec{b}_{N+1} - \sum_{k=1}^N \langle \vec{a}_k | \vec{b}_{N+1} \rangle \vec{a}_k}{\Big\|\vec{b}_{N+1} - \sum_{k=1}^N \langle \vec{a}_k | \vec{b}_{N+1} \rangle \vec{a}_k\Big\|}, \tag{IX.69}$$

$$X_{N+1} \;=\; \mathrm{span}\big(\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_{N+1}\}\big). \tag{IX.70}$$

Wir behaupten, dass

$$X_{N+1} \;=\; Y_{N+1}. \tag{IX.71}$$

Für $\underline{N = 0}$ ist (IX.71) trivial, denn $X_1 = \mathbb{K} \cdot \vec{b}_1 = \mathbb{K} \cdot \|\vec{b}_1\| \cdot \vec{a}_1 = \mathbb{K} \cdot \vec{a}_1 = Y_1$.

Für $\underline{N \geq 1}$ folgt (IX.71) mit (IX.66) induktiv aus derselben Behauptung für N,

$$X_{N+1} \;=\; \mathrm{span}\big(X_N \cup \{\vec{b}_{N+1}\}\big) \;=\; \mathrm{span}\big(Y_N \cup \{\vec{b}_{N+1}\}\big) \;=\; Y_{N+1}. \tag{IX.72}$$

Daher gilt (IX.71) für alle $N \in \mathbb{N}$. Nach (IX.68) ist die durch (IX.69) definierte Folge $\{\vec{a}_n\}_{n=1}^\infty$ orthonormal und damit auch linear unabhängig. Sie erzeugt aber auch X, denn wenn $\vec{x} \in X$, so gibt es ein $N \in \mathbb{N}$ und Zahlen $\alpha_1, \alpha_2, \ldots, \alpha_N \in \mathbb{K}$, so dass $\vec{x} = \alpha_1 \vec{b}_1 + \ldots + \alpha_N \vec{b}_N$, d.h. $\vec{x} \in X_N$. (X enthält nur *endliche* Linearkombinationen der Basisvektoren $\{\vec{b}_n\}_{n=1}^\infty$, siehe (IV.35))). Wegen (IX.71) ist dann jedoch

$$\vec{x} \in Y_N \;=\; \mathrm{span}\big(\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_N\}\big). \tag{IX.73}$$

Also ist $\{\vec{a}_n\}_{n=1}^\infty \subseteq X$ eine Basis. $\square$

Korollar IX.11. *Sei X ein $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{K}$.*

(i) Sind $\dim_{\mathbb{K}}(X) = N \in \mathbb{N}$ und $\{\vec{a}_n\}_{n=1}^N \subseteq X$ eine Orthonormalbasis, so gilt

$$\forall\, \vec{x} \in X : \quad \vec{x} \;=\; \sum_{n=1}^N \langle \vec{a}_n | \vec{x} \rangle\, \vec{a}_n. \tag{IX.74}$$

(ii) Sind $\dim_{\mathbb{K}}(X) = \infty$ *und* $\{\vec{a}_n\}_{n=1}^{\infty} \subseteq X$ *eine abzählbare Orthonormalbasis, so gilt*

$$\forall\, \vec{x} \in X: \quad \vec{x} \;=\; \sum_{n=1}^{N} \langle \vec{a}_n | \vec{x} \rangle\, \vec{a}_n, \tag{IX.75}$$

wobei $N \in \mathbb{N}$ *genügend groß ist, sodass* $\vec{x} \in \operatorname{span}\big(\{\vec{a}_1, \ldots, \vec{a}_N\}\big)$.

Beweis. Ist $\vec{x} = \vec{\alpha}_1 \vec{a}_1 + \cdots + \vec{\alpha}_N \vec{a}_N$, so ist

$$\langle \vec{a}_n | \vec{x} \rangle \;=\; \sum_{k=1}^{N} \alpha_k \underbrace{\langle \vec{a}_n | \vec{a}_k \rangle}_{=\delta_{n,k}} \;=\; \alpha_n. \tag{IX.76}$$

□

Bemerkungen und Beispiele.

- Seien $\mathbb{K} = \mathbb{R}$, $X = \mathbb{R}^3$ mit dem euklidschen Skalarprodukt und

$$\vec{b}_1 \;=\; \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \vec{b}_2 \;=\; \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \qquad \vec{b}_3 \;=\; \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{IX.77}$$

  Wir wenden das Schmidtsche Orthonormierungsverfahren auf die Basis $\{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ von X an.

  Dann ist $\|\vec{b}_1\| = \sqrt{\langle \vec{b}_1 | \vec{b}_1 \rangle} = \sqrt{2}$ und somit

$$\vec{a}_1 \;:=\; \frac{\vec{b}_1}{\|\vec{b}_1\|} \;=\; \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}. \tag{IX.78}$$

  Also ist

$$\begin{aligned} P_1^{\perp} \vec{b}_2 \;&=\; \vec{b}_2 - P_1 \vec{b}_2 \;=\; \vec{b}_2 - \langle \vec{a}_1 | \vec{b}_2 \rangle \vec{a}_1 \;=\; \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} - \left\langle \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \middle| \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \right\rangle \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \\ &=\; \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} - \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} - \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} \;=\; \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 2 \end{pmatrix}, \end{aligned} \tag{IX.79}$$

  was $\|P_1^{\perp} \vec{b}_2\| = \sqrt{\frac{1}{4} + \frac{1}{4} + 4} = \frac{3}{\sqrt{2}}$ nach sich zieht. Also ist

$$\vec{a}_2 \;:=\; \frac{\vec{b}_2}{\|\vec{b}_2\|} \;=\; \frac{\sqrt{2}}{3} \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 2 \end{pmatrix} \;=\; \begin{pmatrix} -\frac{1}{3\sqrt{2}} \\ \frac{1}{3\sqrt{2}} \\ \frac{2\sqrt{2}}{3} \end{pmatrix}. \tag{IX.80}$$

Dann berechnen wir schließlich

$$\begin{aligned} P_2^\perp \vec{b}_3 &= \vec{b}_3 - P_2 \vec{b}_3 \;=\; \vec{b}_3 - \langle \vec{a}_1 | \vec{b}_3 \rangle \vec{a}_1 - \langle \vec{a}_2 | \vec{b}_3 \rangle \vec{a}_2 \\ &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \left\langle \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \middle| \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} - \left\langle \begin{pmatrix} -\frac{1}{3\sqrt{2}} \\ \frac{1}{3\sqrt{2}} \\ \frac{2\sqrt{2}}{3} \end{pmatrix} \middle| \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle \begin{pmatrix} -\frac{1}{3\sqrt{2}} \\ \frac{1}{3\sqrt{2}} \\ \frac{2\sqrt{2}}{3} \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \sqrt{2} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} - \frac{2\sqrt{2}}{3} \begin{pmatrix} -\frac{1}{3\sqrt{2}} \\ \frac{1}{3\sqrt{2}} \\ \frac{2\sqrt{2}}{3} \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} -\frac{2}{9} \\ \frac{2}{9} \\ \frac{8}{9} \end{pmatrix} \;=\; \begin{pmatrix} \frac{2}{9} \\ -\frac{2}{9} \\ \frac{1}{9} \end{pmatrix}, \end{aligned} \tag{IX.81}$$

also $\|P_2^\perp \vec{b}_3\| = \sqrt{\frac{4}{81} + \frac{4}{81} + \frac{1}{81}} = \frac{1}{3}$ und daher

$$\vec{a}_3 \;:=\; \frac{P_2^\perp \vec{b}_3}{\|P_2^\perp \vec{b}_3\|} \;=\; 3 \begin{pmatrix} \frac{2}{9} \\ -\frac{2}{9} \\ \frac{1}{9} \end{pmatrix} \;=\; \begin{pmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{pmatrix}. \tag{IX.82}$$

IX.5. Diskrete Fourier-Transformation und Diskrete Kosinustransformation

IX.5.1. Diskrete Fourier-Transformation

Seien $L \in \mathbb{N}$ und $\mathbb{Z}_L$ der Restklassenring modulo L mit vollständigem Repräsentantensystem $\{0, 1, 2, \ldots, L-1\}$ (siehe Abschnitt II.2.1). Wir betrachten den $\mathbb{C}$-Vektorraum $X := \{\mathbb{Z}_L \to \mathbb{C}\}$ der komplexwertigen Abbildungen auf $\mathbb{Z}_L$. Schreiben wir die Werte von $f \in X$ als Zeilenvektor,

$$f \;=\; \big(f(0),\, f(1),\, f(2),\, \ldots,\, f(L-1)\big) \;\in\; X \tag{IX.83}$$

auf, so ist klar, dass X isomorph zu $\mathbb{C}^L$ ist. Das unitäre Skalarprodukt auf $\mathbb{C}^L$ wird somit zum Skalarprodukt auf X durch

$$\forall f, g \in X: \quad \langle f | g \rangle \;:=\; \sum_{x=0}^{L-1} \overline{f(x)}\, g(x), \tag{IX.84}$$

und die Standardbasis $\{\delta_0, \delta_1, \ldots, \delta_{L-1}\} \subseteq X$ ist gegeben durch $\delta_z(x) := \delta_{z,x}$, wobei $\delta_{z,x}$ das Kroneckersymbol notiert.

Zu $\xi \in \mathbb{Z}_L$ definieren wir nun Vektoren $\varphi_\xi \in X$ durch

$$\forall x \in \mathbb{Z}_L: \quad \varphi_\xi(x) \;:=\; \frac{1}{\sqrt{L}} \exp\left[\tfrac{2\pi i}{L} \xi x\right]. \tag{IX.85}$$

Lemma IX.12. *Die Menge $\{\varphi_\xi \mid \xi \in \mathbb{Z}_L\} \subseteq X$ ist eine ONB.*

Beweis. Zur Berechnung der Skalarprodukte $\langle \varphi_\xi | \varphi_\eta \rangle$ benötigen wir die *geometrische Summe.* Ist $\alpha \in \mathbb{C} \setminus \{1\}$, so ist

$$(1-\alpha)\bigg(\sum_{x=0}^{L-1} \alpha^x\bigg) \;=\; \sum_{x=0}^{L-1} \alpha^x - \sum_{x=0}^{L-1} \alpha^{x+1} \;=\; 1 - \alpha^L, \tag{IX.86}$$

also

$$\forall\, \alpha \in \mathbb{C} \setminus \{1\} : \qquad \sum_{x=0}^{L-1} \alpha^x \;=\; \frac{1-\alpha^L}{1-\alpha}. \tag{IX.87}$$

Sind nun $\xi, \eta \in \mathbb{Z}_L$, so ist

$$\langle \varphi_\xi | \varphi_\eta \rangle \;=\; \frac{1}{L} \sum_{x=0}^{L-1} \exp\big[-\tfrac{2\pi i}{L} \xi x\big] \exp\big[\tfrac{2\pi i}{L} \eta x\big] \;=\; \frac{1}{L} \sum_{x=0}^{L-1} \exp\big[\tfrac{2\pi i}{L} (\eta - \xi) x\big]\,. \tag{IX.88}$$

Für $\xi = \eta$ ist $\exp[\frac{2\pi i}{L}(\eta - \xi)x] = 1$, für alle $x \in \mathbb{Z}_L$, und aus (IX.88) folgt sofort, dass $\langle \varphi_\xi | \varphi_\xi \rangle = \langle \varphi_\eta | \varphi_\eta \rangle = 1$. Ist umgekehrt $\xi \neq \eta$, so ist $\eta - \xi \in \mathbb{Z}_L \setminus \{0\} \cong \{1, 2, \ldots, L-1\}$. Dann sind

$$\exp[2\pi i(\eta - \xi)] \;=\; 0 \quad \text{und} \quad \exp\big[2\pi i\big(\tfrac{\eta-\xi}{L}\big)\big] \;\neq\; 0\,, \tag{IX.89}$$

und (IX.87) und (IX.88) implizieren, dass

$$\langle \varphi_\xi | \varphi_\eta \rangle \;=\; \frac{1}{L} \sum_{x=0}^{L-1} \Big(\exp\big[\tfrac{2\pi i}{L}(\eta - \xi)\big]\Big)^x \;=\; \frac{1}{L}\, \frac{1 - \exp[2\pi i(\eta - \xi)]}{1 - \exp\big[2\pi i\big(\frac{\eta-\xi}{L}\big)\big]} \;=\; 0\,. \tag{IX.90}$$

Somit ist die Menge $\{\varphi_\xi | \xi \in \mathbb{Z}_L\} \subseteq X$ orthonormal und insbesondere linear unabhängig. Weiterhin ist $|\{\varphi_\xi \mid \xi \in \mathbb{Z}_L\}| = L = \dim(X)$, und deshalb ist $\{\varphi_\xi \mid \xi \in \mathbb{Z}_L\}$ eine ONB von X. $\square$

Gemäß (IX.43) ist damit $f = \sum_{\xi \in \mathbb{Z}_L} \langle \varphi_\xi | f \rangle \varphi_\xi$, für jedes $f \in X$, d.h.

$$\forall x \in \mathbb{Z}_L : \quad f(x) \;=\; \sum_{\xi \in \mathbb{Z}_L} \hat{f}(\xi)\, \frac{e^{\frac{2\pi i}{L}\xi x}}{\sqrt{L}}, \tag{IX.91}$$

wobei

$$\forall \xi \in \mathbb{Z}_L : \quad \hat{f}(\xi) \;:=\; \langle \varphi_\xi | f \rangle \;=\; \sum_{x \in \mathbb{Z}_L} \frac{e^{-\frac{2\pi i}{L}\xi x}}{\sqrt{L}}\, f(x). \tag{IX.92}$$

Die Basistransformation von der durch die Standardbasis $\{\delta_z | z \in \mathbb{Z}_L\}$ gegebenen ONB auf die ONB $\{\varphi_\xi | \xi \in \mathbb{Z}_L\}$ von X bezeichnet man als **diskrete Fourier-Transformation (DFT)**. Zu gegebenem $f : \mathbb{Z}_L \to \mathbb{C}$ bezeichnet man die durch (IX.92) gegebene Funktion $\hat{f} : \mathbb{Z}_L \to \mathbb{C}$ als ihre **(diskrete) Fourier-Transformierte**. Die Glg. (IX.91) wird auch als **inverse diskrete Fourier-Transformation** bezeichnet, da man mit ihrer Hilfe f aus $\hat{f}$ zurückgewinnen kann.

IX.5.2. Diskrete Kosinustransformation

Analog zu (IX.85) definieren wir nun zu $\xi \in \Gamma_L := \{0, 1, 2, \ldots, L-1\}$ Vektoren $\psi_\xi \in Y := \{\Gamma_L \to \mathbb{C}\}$ durch

$$\forall x \in \Gamma_L : \quad \psi_\xi(x) \ := \ \begin{cases} \sqrt{\frac{2}{L}} \, \cos\left[\frac{\pi}{L}\xi(x+\frac{1}{2})\right], & \text{falls } \xi \neq 0, \\ \frac{1}{\sqrt{L}} & \text{falls } \xi = 0. \end{cases} \tag{IX.93}$$

Lemma IX.13. *Die Menge $\{\psi_\xi \mid \xi \in \Gamma_L\} \subseteq Y$ ist eine ONB.*

Beweis. Wir definieren zunächst $\tilde{\psi}_\xi \in Y$ durch

$$\forall x \in \Gamma_L : \quad \tilde{\psi}_\xi(x) \ := \ \cos\left[\tfrac{\pi}{L}\xi(x+\tfrac{1}{2})\right], \tag{IX.94}$$

sodass $\psi_0 = \sqrt{\frac{1}{L}}\tilde{\psi}_0$ und $\psi_\xi = \sqrt{\frac{2}{L}}\tilde{\psi}_\xi$, für $\xi \neq 0$.
Sind nun $\xi, \eta \in \Gamma_L$, so ist

$$\begin{aligned}
\langle \tilde{\psi}_\xi | \tilde{\psi}_\eta \rangle \ &= \ \sum_{x=0}^{L-1} \cos\left[\tfrac{\pi}{L}\xi(x+\tfrac{1}{2})\right] \, \cos\left[\tfrac{\pi}{L}\eta(x+\tfrac{1}{2})\right] \\
&= \frac{1}{4}\sum_{x=0}^{L-1} \Big\{ \Big(\exp\left[\tfrac{\pi i}{L}\xi(x+\tfrac{1}{2})\right] \, + \, \exp\left[-\tfrac{\pi i}{L}\xi(x+\tfrac{1}{2})\right] \Big) \\
&\qquad\qquad \cdot \Big(\exp\left[\tfrac{\pi i}{L}\eta(x+\tfrac{1}{2})\right] \, + \, \exp\left[-\tfrac{\pi i}{L}\eta(x+\tfrac{1}{2})\right] \Big) \Big\} \\
&= \frac{1}{4}\sum_{\sigma,\tau=\pm 1}\sum_{x=0}^{L-1} \exp\left[\tfrac{\pi i}{L}\sigma(\xi+\tau\eta)(x+\tfrac{1}{2})\right] \\
&= \frac{1}{2}\mathrm{Re}\Big\{ \sum_{\tau=\pm 1}\sum_{x=0}^{L-1} \exp\left[\tfrac{\pi i}{L}(\xi+\tau\eta)(x+\tfrac{1}{2})\right] \Big\} \\
&= \frac{1}{2}\mathrm{Re}\Big\{ \sum_{\tau=\pm 1} \exp\left[\tfrac{\pi i}{2L}(\xi+\tau\eta)\right] \sum_{x=0}^{L-1} \Big(\exp\left[\tfrac{\pi i}{L}(\xi+\tau\eta)\right] \Big)^x \Big\} .
\end{aligned} \tag{IX.95}$$

Wir unterscheiden jetzt drei Fälle.

(a) <u>$\eta \neq \xi$:</u> Da ξ und η nichtnegativ sind, muss $1 \leq \xi + \tau\eta \leq 2L-2$ für $\tau = 1$ gelten. Außerdem ist $\xi + \tau\eta = \xi - \eta \neq 0$ für $\tau = -1$. Also ist

$$\forall \tau \in \{-1, 1\} : \qquad \exp\left[\tfrac{\pi i}{L}(\xi+\tau\eta)\right] \ \neq \ 1 , \tag{IX.96}$$

und weiterhin ist

$$\exp\left[\pi i(\xi+\tau\eta)\right] \ = \ (-1)^{\xi+\tau\eta} . \tag{IX.97}$$

Zur Berechnung der Skalarprodukte $\langle \tilde{\psi}_\xi | \tilde{\psi}_\eta \rangle$ verwenden wir wieder die geometrische Summe (IX.87) und erhalten

$$\begin{aligned} \langle \tilde{\psi}_\xi | \tilde{\psi}_\eta \rangle \ &= \frac{1}{2} \mathrm{Re} \Big\{ \sum_{\tau=\pm 1} \exp\left[\tfrac{\pi i}{2L}(\xi + \tau\eta) \right] \frac{1 \ - \ \exp[\pi i(\xi+\tau\eta)]}{1 \ - \ \exp[\frac{\pi i}{L}(\xi+\tau\eta)]} \Big\} \\ &= \mathrm{Re} \Big\{ \frac{i}{2} \sum_{\tau=\pm 1} \frac{1 \ - \ (-1)^{\xi+\tau\eta}}{\sin[-\frac{\pi}{2L}(\xi+\tau\eta)]} \Big\} \ = \ 0 \, , \end{aligned} \tag{IX.98}$$

da die Zahl in geschweiften Klammern, deren Realteil genommen wird, rein imaginär ist. Daher gilt

$$\forall\, \xi, \eta \in \Gamma_L, \ \xi \neq \eta : \qquad \langle \psi_\xi | \psi_\eta \rangle \ = \ 0 \, . \tag{IX.99}$$

(b) Ist $\eta = \xi > 0$, so ist $\xi + \tau\eta = 0$, für $\tau = -1$. Eine Rechnung wie in (IX.96)-(IX.98) zeigt, dass dann der Summand für $\tau = 1$ abermals verschwindet, und wir erhalten

$$\langle \tilde{\psi}_\xi | \tilde{\psi}_\xi \rangle \ = \frac{1}{2} \mathrm{Re} \Big\{ \sum_{x=0}^{L-1} 1 \Big\} \ = \ \frac{L}{2} \, . \tag{IX.100}$$

Somit folgt

$$\forall\, \xi \in \Gamma_L \setminus \{0\} : \qquad \langle \psi_\xi | \psi_\xi \rangle \ = \ 1 \, . \tag{IX.101}$$

(c) Ist $\eta = \xi = 0$, so folgt direkt

$$\langle \psi_0 | \psi_0 \rangle \ = \ \frac{1}{L} \sum_{x=0}^{L-1} 1 \ = \ 1 \, . \tag{IX.102}$$

□

Gemäß (IX.43) ist damit $f = \sum_{\xi \in \Gamma_L} \langle \psi_\xi | f \rangle \psi_\xi$, für jedes $f \in Y$, d.h.

$$\forall x \in \Gamma_L : \quad f(x) \ = \ \frac{D[f](0)}{L} + \sum_{\xi=1}^{L-1} \frac{2\, D[f](\xi)}{L} \cos\left[\tfrac{\pi}{L} \xi (x + \tfrac{1}{2}) \right] , \tag{IX.103}$$

wobei

$$\forall \xi \in \Gamma_L : \quad D[f](\xi) \ := \sum_{x \in \Gamma_L} \cos\left[\tfrac{\pi}{L} \xi (x + \tfrac{1}{2}) \right] f(x) \, . \tag{IX.104}$$

Analog zur diskreten Fouriertransformation bezeichnet man die Basistransformation von der durch die Standardbasis $\{\delta_z | z \in \Gamma_L\}$ gegebenen ONB auf die ONB $\{\psi_\xi \mid \xi \in \Gamma_L\}$ von Y als **diskrete Kosinustransformation (DCT)**. Zu gegebenem $f : \Gamma_L \to \mathbb{C}$ bezeichnet man die durch (IX.92) gegebene Funktion $D[f] : \Gamma_L \to \mathbb{C}$ als ihre **diskrete Kosinustransformierte**. Die Glg. (IX.91) wird auch als **inverse diskrete Kosinustransformation** bezeichnet, da man mit ihrer Hilfe f aus $D[f]$ zurückgewinnen kann.

Die DCT spielt in der Signalverarbeitung und vor allem in der Bildverarbeitung eine wichtige Rolle. In digitalen Bildern (Fotos, Videos) sind die RGB- bzw. YUV-Farbwerte der Pixel durch reelle Zahlen repräsentiert. Zur Vereinfachung gehen wir von einem Schwarz-Weiß-Bild aus, in dem für jeden der ca. 1000×1000 Pixel ein reeller Grauwert angegeben ist. [Tatsächlich ist der Grauwert eine ganze Zahl zwischen 0 (schwarz) und 255 (weiß).] Das heißt also, dass jedes Bild eines Schwarz-Weiß-Videos durch 1.000.000 reelle Zahlen -die Grauwerte der Pixel- gegeben ist. Bei 50 Frames pro Sekunde kommen 50.000.000 reelle Zahlen zusammen, die pro Sekunde übertragen werden müssen. Sind noch weitere Signale zu übertragen, berücksichtigen wir die Farben der Pixel und bauen wir auch zur Vorbeugung von Übertragungsfehlern Redundanz ein, so wird die pro Sekunde zu übertragende Datenmenge zu groß für ein Streaming in Echtzeit.

Um dieses Problem zu lösen, macht man sich zunutze, dass sich in den meisten Bilden die Grauwerte benachbarter Pixel kaum unterscheiden. Um dies mathematisch zu formulieren, fassen wir ein Bild als reellwertige Funktion $f : \Lambda \times \Lambda \to \mathbb{R}$ mit $\Lambda = \{0, 1, 2, \ldots, 999\}$ auf. Für $x_1, x_2 \in \Lambda$ ist dann $f(x_1, x_2) \in \mathbb{R}$ der Grauwert am Pixel mit den Koordinaten $(x_1, x_2) \in \Lambda \times \Lambda$. Die langsame Variation der Grauwerte benachbarter Bildpunkte könnte man quantitativ durch

$$\begin{aligned} \forall (x_1, x_2), (y_1, y_2) \in & \Lambda \times \Lambda, \ \|(x_1, x_2) - (y_1, y_2)\|_\infty \leq L : \\ & |f(x_1, x_2) - f(y_1, y_2)| \ \leq \ \varepsilon \sum_{(z_1, z_2) \in \Lambda \times \Lambda} |f(z_1, z_2)| \end{aligned} \tag{IX.105}$$

ausdrücken, wobei $0 < \varepsilon \ll 1$ ist. In der Praxis wird $L = 8$ gewählt. Nun bildet man die DCT $D[f]$ auf jeder Kachel mit $L \times L$ Pixel. Die langsame Variation von f manifestiert sich nun darin, dass die DCT $D[f](\xi_1, \xi_2)$ nur für kleine $\xi_1, \xi_2 \in \Gamma_L$ nicht verschwindend klein ist und man deshalb beispielsweise nur $D[f](0, 0)$, $D[f](0, 1)$, $D[f](1, 0)$ und $D[f](1, 1)$ überträgt und die 60 Werte von $D[f](\xi_1, \xi_2)$ für $\max(\xi_1, \xi_2) \geq 2$ schlicht ignoriert. Ein solches Verfahren bezeichnet man in der Signalverarbeitung als *verlustbehaftete Kompression*.

Dies beschreibt das Vorgehen bei Anwendung des JPEG-Standards, der seit Mitte der 90er Jahren der maßgebliche Bildstandard für digitale Fotos und Videos ist. Das ungeübte menschliche Auge kann trotzdem praktisch keinen Unterschied erkennen, und die zu übertragende Datenmenge wurde um den Faktor $64/4 = 16$ reduziert. Mit etwas Übung kann man manchmal die einzelnen Kacheln auf Bildern als Artefakte erkennen.

IX.6. Ergänzungen

IX.6.1. Äquivalenz von Normen

Sei X ein $\mathbb{K}$-Vektorraum. Zwei Normen $\| \cdot \|, \| \cdot \|' : X \to \mathbb{R}_0^+$ heißen **äquivalent**, falls es Konstanten $0 < c < C < \infty$ gibt, sodass

$$\forall \vec{x} \in X : \quad c \, \|\vec{x}\| \ \leq \ \|\vec{x}\|' \ \leq \ C \, \|\vec{x}\| \tag{IX.106}$$

gilt. Die in Glg. (IX.29) definierten Normen sind alle paarweise äquivalent, da $\mathbb{K}^d$ ein Vektorraum endlicher Dimension $d = \dim[\mathbb{K}^d] = d \in \mathbb{N}$ ist.

Allgemein gilt, dass Normen auf X zueinander äquivalent sind, falls X ein $\mathbb{K}$-Vektorraum endlicher Dimension ist. Für $d \in \mathbb{N}$ und $X = \mathbb{C}^d$ bedeutet dies, dass es zu $1 \leq p, q \leq \infty$ Konstanten $0 < c_{d,p,q} < C_{d,p,q} < \infty$ so gibt, dass

$$\forall \vec{x} \in X : \quad c_{d,p,q} \, \|\vec{x}\|_p \ \leq \ \|\vec{x}\|_q \ \leq \ C_{d,p,q} \, \|\vec{x}\|_p \tag{IX.107}$$

gilt.

Die Äquivalenz zweier Normen ist wichtig für die Analysis, denn sie sichert, dass die mit ihnen formulierten Konvergenzbegriffe zusammenfallen. Eine wichtige Beobachtung, die wir hier ohne Beweis formulieren, ist die Tatsache, dass je zwei Normen auf einem $\mathbb{K}$-Vektorraum X stets äquivalent sind. Um Unterschiede im Konvergenzbegriff sehen zu können, muss der zugrunde Vektorraum also notwendig unendlichdimensional sein. Dies kann man als Ausgangspunkt des mathematischen Teilgebiets der Funktionalanalysis betrachten.

X. Eigenwerte und Diagonalisierbarkeit

In diesem Kapitel wollen wir $\mathbb{K} = \mathbb{C}$ annehmen. Eine analoge Theorie lässt sich auch für $\mathbb{K} = \mathbb{R}$ entwickeln, ist aber aufwändiger und letztendlich auch weniger wichtig.

X.1. Eigenwerte

Definition X.1. Seien X ein $\mathbb{C}$-Vektorraum und $\Phi \in \mathcal{L}(X)$ eine lineare Abbildung.

(i) Die **Resolventenmenge von $\boldsymbol{\Phi}$** ist definiert durch

$$\rho(\Phi) \; = \; \big\{ \lambda \in \mathbb{C} \, \big| \, (\Phi - \lambda \cdot \mathbf{1}) \text{ ist bijektiv} \big\}. \tag{X.1}$$

(ii) Das **Spektrum von $\boldsymbol{\Phi}$** ist definiert als

$$\sigma(\Phi) \; = \; \big\{ \lambda \in \mathbb{C} \, \big| \, \Phi - \lambda \cdot \mathbf{1} \text{ ist nicht bijektiv} \big\} \; = \; \mathbb{C} \setminus \rho(\Phi). \tag{X.2}$$

(iii) Eine Zahl $\lambda \in \mathbb{C}$ heißt **Eigenwert von $\boldsymbol{\Phi}$**

$$:\Leftrightarrow \; \exists \, \vec{x} \in X \setminus \{\vec{0}\} : \quad \Phi \vec{x} \; = \; \lambda \vec{x}. \tag{X.3}$$

Ist $\Phi\vec{x} = \lambda\vec{x}$ mit $\vec{x} \neq 0$, so heißt $\vec{x}$ **Eigenvektor von $\boldsymbol{\Phi}$ (zum Eigenwert $\boldsymbol{\lambda}$)**.

Satz X.2. *Sind X ein endlich-dimensionaler $\mathbb{C}$-Vektorraum und $\Phi \in \mathcal{L}(X)$ eine lineare Abbildung, so gilt*

$$\begin{aligned} \sigma(\Phi) \; &= \big\{ \lambda \in \mathbb{C} \, \big| \, \lambda \textit{ ist Eigenwert von } \Phi \big\} \\ &= \big\{ \lambda \in \mathbb{C} \, \big| \, \det[\Phi - \lambda \cdot \mathbf{1}] = 0 \big\} \; \neq \; \emptyset. \end{aligned} \tag{X.4}$$

Beweis. Die Gleichheit der drei Mengen folgt aus der Gleichwertigkeit der folgenden Aussagen,

$$\begin{aligned} & \Big(\operatorname{Ker}\{\Phi - \lambda \cdot \mathbf{1}\} = \{\vec{0}\} \Big) \quad \Leftrightarrow \quad \Big((\Phi - \lambda \cdot \mathbf{1}) \text{ ist injektiv} \Big) \\ \Leftrightarrow \quad & \Big((\Phi - \lambda \cdot \mathbf{1}) \text{ ist bijektiv} \Big) \quad \Leftrightarrow \Big(\det[\Phi - \lambda \cdot \mathbf{1}] \neq 0 \Big). \end{aligned} \tag{X.5}$$

Die letzten beiden Äquivalenzen ergeben sich aus Satz V.6 und Satz VII.8, (i), die beide nur für endlich-dimensionale Vektorräume gültig sind. Ist nun $\mathcal{M}_{\mathcal{X}}[\Phi] = (a_{i,j})_{i,j=1}^N \in \mathbb{C}^{N\times N}$ eine Matrixdarstellung von Φ, mit $N = \dim(X)$, so ist das **charakteristische Polynom von Φ**,

$$\det[\Phi - \lambda \cdot 1] \;=\; (-1)^N \big(\lambda^N + c_{N-1}\lambda^{N-1} + \ldots + c_1\lambda + c_0\big) \tag{X.6}$$

ein Polynom N. Grades in λ mit komplexen Koeffizienten (deren genauer Wert von den Matrixelementen $a_{i,j}$ abhängt). Nach dem Fundamentalsatz der Algebra, Satz III.6, zerfällt somit $\det[\Phi - \lambda \cdot 1]$ in Linearfaktoren, d.h. es gibt $L \in \mathbb{N}$, $\lambda_1, \ldots, \lambda_L \in \mathbb{C}$ und $n_1, \ldots, n_L \in \mathbb{N}$, mit $n_1 + \ldots + n_L = N$ so, dass

$$\det[\Phi - \lambda \cdot 1] \;=\; (-1)^N (\lambda - \lambda_1)^{n_1} \cdot (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_L)^{n_L}. \tag{X.7}$$

Damit ist

$$\sigma(\Phi) \;=\; \{\lambda_1, \lambda_2, \ldots, \lambda_L\} \neq \emptyset. \tag{X.8}$$

□

Bemerkungen und Beispiele.

- Seien $X = \mathbb{C}^2$ und $\Phi \in \mathcal{L}(X)$ eine lineare Abbildung mit Matrixdarstellung

$$A \;=\; \mathcal{M}(\Phi) \;=\; \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}. \tag{X.9}$$

  1. Wir berechnen die Eigenwerte mit Hilfe des charakteristischen Polynoms

$$\begin{aligned} \det\big[A - \lambda \mathbf{1}\big] \;&=\; \det\left[\begin{pmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{pmatrix}\right] \;=\; (4-\lambda)(3-\lambda) - 2 \\ &= \lambda^2 - 7\lambda + 10 \end{aligned} \tag{X.10}$$

     Die Eigenwerte $\sigma(\Phi)$ von Φ sind die Nullstellen des charakteristischen Polynoms, also $\sigma(\Phi) = \{\lambda_+, \lambda_-\}$ mit

$$\lambda_+ \;=\; 5\,, \quad \lambda_- \;=\; 2\,, \tag{X.11}$$

     wie man leicht mit Hilfe der p-q-Formel erhält (was allerdings nur in Dimension $\dim(X) = 2$ so geht; in höheren Dimensionen muss man die Eigenwerte i.A. numerisch berechnen und erhält sie somit nur näherungsweise).

  2. Wir berechnen die zu $\lambda_\pm$ gehörigen Eigenvektoren $\vec{x}_\pm := (\alpha_\pm, \beta_\pm)^T \in \mathbb{C}^2 \setminus \{\vec{0}\}$. Dazu lösen wir jeweils das zugehörige LGS:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \vec{0} \;=\; \big(A - \lambda_\pm \mathbf{1}\big)\vec{x}_\pm \;=\; \begin{pmatrix} 4-\lambda_\pm & 1 \\ 2 & 3-\lambda_\pm \end{pmatrix} \begin{pmatrix} \alpha_\pm \\ \beta_\pm \end{pmatrix}, \tag{X.12}$$

     also

$$(4-\lambda_\pm)\,\alpha_\pm \;+\; \beta_\pm \;=\; 0, \tag{X.13}$$
$$2\,\alpha_\pm \;+\; (3-\lambda_\pm)\,\beta_\pm \;=\; 0. \tag{X.14}$$

Für $\lambda_+ = 5$ liefert dies

$$\left.\begin{array}{rcl} -\alpha_+ + \beta_+ & = & 0 \\ 2\alpha_+ - 2\beta_+ & = & 0 \end{array}\right\} \;\Leftrightarrow\; \beta_+ \;=\; \alpha_+ \;\Leftrightarrow\; \vec{x}_+ \;=\; t \begin{pmatrix} 1 \\ 1 \end{pmatrix},\; t \in \mathbb{C} \setminus \{0\}, \tag{X.15}$$

und für $\lambda_- = 2$ erhalten wir

$$\left.\begin{array}{rcl} 2\alpha_- + \beta_- & = & 0, \\ 2\alpha_- + \beta_- & = & 0. \end{array}\right\} \;\Leftrightarrow\; \beta_- \;=\; -2\alpha_- \;\Leftrightarrow\; \vec{x}_- \;=\; t \begin{pmatrix} 1 \\ -2 \end{pmatrix},\; t \in \mathbb{C} \setminus \{0\}, \tag{X.16}$$

- Ist X unendlichdimensional, aber mit einer geeigneten topologischen Struktur (z.B. X ein Hilbertraum) versehen, so gilt auch allgemein $\sigma(\Phi) \neq \emptyset$. Die Punkte im Spektrum sind aber nicht notwendig Eigenwerte, sondern es gilt nur

$$\big\{\lambda \in \mathbb{C} \,\big|\, \lambda \text{ ist Eigenwert von } \Phi\big\} \subseteq \sigma(\Phi). \tag{X.17}$$

- Für $X = \ell^2(\mathbb{Z}) := \big\{\psi : \mathbb{Z} \to \mathbb{C} \,\big|\, \sum_{n=-\infty}^{\infty} |\psi_n|^2 < \infty\big\}$ und $(\Phi\psi)_n := \psi_{n+1} + \psi_{n-1}$ ist etwa $\sigma(\Phi) = [-2, 2]$, aber $\Phi \in \mathcal{L}(X)$ besitzt gar keinen Eigenwert.
- Eigenvektoren zu verschiedenen Eigenwerten sind stets linear unabhängig, wie das folgende Lemma X.3 zeigt.

Lemma X.3. *Seien X ein $\mathbb{C}$-Vektorraum, $\Phi \in \mathcal{L}(X)$ eine lineare Abbildung und $\lambda_1, \lambda_2, \ldots, \lambda_L \in \sigma(\Phi)$ paarweise verschiedene Eigenwerte von Φ mit zugehörigen Eigenvektoren $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_L \in X \setminus \{0\}$, also $\Phi\vec{x}_\ell = \lambda_\ell \vec{x}_\ell$, für $\ell \in \mathbb{Z}_1^L$. Dann ist $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_L\} \subseteq X$ linear unabhängig.*

Beweis. Siehe Abschnitt X.6.1. □

X.2. Diagonalisierbarkeit

Definition X.4. Sei $N \in \mathbb{N}$.

(i) Eine Matrix $A \in \mathbb{C}^{N \times N}$ heißt **diagonalisierbar**

$$:\Leftrightarrow \; \exists\, H \in GL(N, \mathbb{C}) \;\; \exists\, \lambda_1, \ldots, \lambda_N \in \mathbb{C} :$$

$$H^{-1} A\, H \;=\; \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{pmatrix}. \tag{X.18}$$

(ii) Sei X ein N-dimensionaler Vektorraum über $\mathbb{C}$. Eine lineare Abbildung $\Phi \in \mathcal{L}(X)$ heißt **diagonalisierbar**

$$:\Leftrightarrow \; \text{Es gibt eine Basis } \mathcal{X} \subseteq X, \text{ so dass } \mathcal{M}_{\mathcal{X}}[\Phi] \in \mathbb{C}^{N \times N} \text{ diagonalisierbar ist.} \tag{X.19}$$

Lemma X.5. *Seien $N \in \mathbb{N}$ und X ein $\mathbb{C}$-Vektorraum der Dimension N. Sei weiterhin $\Phi \in \mathcal{L}(X)$ mit Matrixdarstellung $A := \mathcal{M}_{\mathcal{W}}[\Phi] \in \mathbb{C}^{N \times N}$, wobei $\mathcal{W} \subseteq X$ eine Basis ist. Dann gilt folgende Äquivalenz:*

$$\big\{\Phi \text{ ist diagonalisierbar}\big\} \;\Leftrightarrow\; \big\{\mathcal{M}_{\mathcal{W}}[\Phi] \text{ ist diagonalisierbar}\big\}. \tag{X.20}$$

Beweis. Gleichung (X.20) ist offensichtlich, wenn $H = \mathcal{M}_{\mathcal{W}}[V]$ die Matrixdarstellung einer Basistransformation ist, die die Basis $\mathcal{W}$ auf die Basis $\mathcal{X}$ transformiert, für die (X.18) gilt. □

Satz X.6. *Seien X ein N-dimensionaler Vektorraum über $\mathbb{C}$, mit $N \in \mathbb{N}$, und $\Phi \in \mathcal{L}(X)$.*

$$\begin{aligned} &\big\{\Phi \text{ ist diagonalisierbar}\big\} && \text{(X.21)} \\ \Leftrightarrow\; &\big\{\text{Es gibt eine Basis } \mathcal{X} = \{\vec{x}_1, \ldots, \vec{x}_N\} \subseteq X \text{ aus Eigenvektoren von } \Phi\big\}. \end{aligned}$$

Beweis. Sind Φ diagonalisierbar und $\mathcal{X} = \{\vec{x}_1, \ldots, \vec{x}_N\} \subseteq X$ eine Basis, so dass

$$\mathcal{M}_{\mathcal{X}}[\Phi] \;=\; \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}, \tag{X.22}$$

so gilt $\Phi\vec{x}_1 = \lambda_1\vec{x}_1, \ldots, \Phi\vec{x}_N = \lambda_N\vec{x}_N$, und $\mathcal{X}$ ist die gesuchte Basis aus Eigenvektoren von Φ. Dies gilt offensichtlich auch umgekehrt. □

Korollar X.7. *Seien $N \in \mathbb{N}$ und X ein $\mathbb{C}$-Vektorraum der Dimension* $\dim(X) = N$ *sowie $\Phi \in \mathcal{L}(X)$. Besitzt Φ N paarweise verschiedene Eigenwerte, so ist Φ diagonalisierbar.*

Beweis. Sind $\lambda_1, \ldots, \lambda_N \subseteq \sigma(\Phi)$ die Eigenwerte von Φ mit zugehörigen Eigenvektoren $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_L \in X \setminus \{0\}$, also $\Phi\vec{x}_\ell = \lambda_\ell\vec{x}_\ell$, für $\ell \in \mathbb{Z}_1^L$, so ist $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_L\} \subseteq X$ nach Lemma X.3 linear unabhängig und wegen $\dim(X) = N$ auch eine Basis. □

Bemerkungen und Beispiele.

- Nicht alle linearen Abbildungen bzw. Matrizen sind diagonalisierbar. Sei etwa

$$A \;=\; \mathcal{M}(\Phi) \;=\; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{X.23}$$

  so ist

$$\det[A - \lambda\mathbf{1}] \;=\; \det\left[\begin{pmatrix} -\lambda & 1 \\ 0 & -\lambda \end{pmatrix}\right] \;=\; \lambda^2, \tag{X.24}$$

  und $\lambda = 0$ ist der einzige Eigenwert von Φ. Die Eigenvektoren von Φ sind dann genau die nichtverschwindenden Vektoren im Kern von Φ. Dieser ist jedoch eindimensional und somit verschieden von $\mathbb{C}^2$, denn

$$\begin{aligned} \mathrm{Ker}[\Phi] \;&=\; \big\{\vec{x} \in \mathbb{C}^2 \;\big|\; A\vec{x} = \vec{0}\big\} && \text{(X.25)} \\ &=\; \left\{\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in \mathbb{C}^2 \;\middle|\; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \;=\; \begin{pmatrix} \beta \\ 0 \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right\} \;=\; \mathbb{C} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

X.3. Diagonalisierbarkeit selbstadjungierter Matrizen und der Spektralsatz

Zu Beginn dieses Abschnitts erinnern wir an den in Korollar IX.9 gebildeten Begriff der Adjungierten $\Phi^* \in \mathcal{L}(X)$ einer linearen Abbildung $\Phi \in \mathcal{L}(X)$. Für einen $\mathbb{C}$-Vektorraum X mit Skalarprodukt $\langle\cdot|\cdot\rangle$ ist Φ^* definiert durch

$$\forall\, \vec{x}, \vec{y} \in X: \quad \langle \vec{x} | \Phi \vec{y} \rangle \;=\; \langle \Phi^* \vec{x} | \vec{y} \rangle. \tag{X.26}$$

Analog ist die adjungierte Matrix $A^* = (b_{i,j})_{i,j=1}^N \in \mathbb{C}^{N\times N}$ einer gegebenen Matrix $A = (a_{i,j})_{i,j=1}^N \in \mathbb{C}^{N\times N}$ definiert durch

$$b_{i,j} \;=\; \overline{a_{j,i}}. \tag{X.27}$$

Der Zusammenhang zwischen dem Adjungieren einer linearen Abbildung und der adjungierten Matrix besteht in der Identität $\mathcal{M}_{\mathcal{A}}[\Phi^*] = (\mathcal{M}_{\mathcal{A}}[\Phi])^*$, die für jede ONB $\mathcal{A}$ in X gilt.

Definition X.8.

(i) Sei X ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{C}$. Eine lineare Abbildung $\Phi \in \mathcal{L}(X)$ heißt **selbstadjungiert**

$$:\Leftrightarrow \qquad \Phi \;=\; \Phi^*. \tag{X.28}$$

(ii) Eine Matrix $A \in \mathbb{C}^{N\times N}$, $N \in \mathbb{N}$ heißt **selbstadjungiert**

$$:\Leftrightarrow \qquad A \;=\; A^*. \tag{X.29}$$

Bemerkungen und Beispiele.

- Der Begriff der Selbstadjungiertheit im Fall, dass Φ ein unbeschränkter Operator auf einem unendlichdimensionalen Hilbertraum X ist, ist erheblich subtiler. Aus dem *Satz vom abgeschlossenen Graphen* (siehe Vorlesung Funktionalanalysis) folgt nämlich, dass ein linearer Operator $\Phi \in \mathcal{L}(X)$, der (X.28) genügt, notwendig auch beschränkt ist. Scheinbar gibt es also gar keine unbeschränkten, selbstadjungierten linearen Operatoren.
- Die Auflösung dieses scheinbaren Widerspruchs gelang J. von Neumann: unbeschränkte Operatoren sind gar nicht auf ganz X definiert, sondern nur auf einem, in X dichtem Unterraum $\mathrm{dom}(\Phi) \subseteq X$. Selbstadjungiertheit definiert man dann durch $\mathrm{dom}(\Phi^*) = \mathrm{dom}(\Phi)$ und

$$\forall\, \vec{x}, \vec{y} \in \mathrm{dom}(\Phi): \quad \langle \vec{x} | \Phi \vec{y} \rangle \;=\; \langle \Phi^* \vec{x} | \vec{y} \rangle. \tag{X.30}$$

- Sind $\dim(X) =: N < \infty$ und $\mathcal{A} = \{\vec{a}_1, \ldots, \vec{a}_N\} \subseteq X$ eine ONB, so gilt

$$\big\{\Phi \;=\; \Phi^*\big\} \;\Leftrightarrow\; \big\{\mathcal{M}[\Phi] \;=\; \mathcal{M}[\Phi]^*\big\}. \tag{X.31}$$

- Beispiele selbstadjungierter Matrizen in $\mathbb{C}^{3\times 3}$ sind

$$\begin{pmatrix} 1 & 2 & 9 \\ 2 & 3 & 7i \\ 9 & -7i & 5 \end{pmatrix}, \quad \begin{pmatrix} 2 & 4 & 8+3i \\ 4 & 10 & 2i \\ 8-3i & -2i & 12 \end{pmatrix}. \tag{X.32}$$

Definition X.9. Seien X eine Menge und $\Phi : X \to X$ eine Abbildung. Eine Teilmenge $B \subseteq X$ heißt **unter Φ invariant**

$$:\Leftrightarrow \qquad \Phi(B) \ \subseteq \ B, \quad \text{d.h. } \forall\, x \in B : \quad \Phi(x) \in B\,. \tag{X.33}$$

Lemma X.10. *Seien X ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt $\langle \cdot | \cdot \rangle : X \times X \to \mathbb{C}$ und $\Phi = \Phi^* \in \mathcal{L}(X)$ eine selbstadjungierte lineare Abbildung. Ist $Y \subseteq X$ ein unter Φ invarianter Unterraum, so ist auch $Y^\perp \subseteq X$ unter Φ invariant,*

$$\Phi(Y) \subseteq Y \ \Rightarrow \ \Phi(Y^\perp) \subseteq Y^\perp. \tag{X.34}$$

Beweis. Seien $\vec{x} \in Y^\perp$ und $\vec{y} \in Y$. Mit $\vec{y} \in Y$ und der Invarianz von Y unter Φ ist dann auch $\Phi^* \vec{y} = \Phi \vec{y} \in Y$, und deshalb gilt

$$\langle \Phi \vec{x} | \vec{y} \rangle \ = \ \langle \vec{x} | \Phi^* \vec{y} \rangle \ = \ \langle \vec{x} | \Phi \vec{y} \rangle \ = \ 0\,. \tag{X.35}$$

Weil $\vec{y} \in Y$ beliebig ist, folgt $\Phi \vec{x} \in Y^\perp$. □

Lemma X.11. *Seien X ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt $\langle \cdot | \cdot \rangle : X \times X \to \mathbb{C}$ und $Y \subseteq X$ ein Unterraum. Dann ist*

$$X \ = \ Y + Y^\perp \ =: \ Y \oplus Y^\perp. \tag{X.36}$$

Beweis. Sei $\{\vec{a}_1, \ldots, \vec{a}_M\} \subseteq Y$ eine ONB in Y, die wir uns nötigenfalls durch das Schmidtsche Orthonormierungsverfahren beschafft haben. Wir definieren $P \in \mathcal{L}(X)$ durch

$$P\vec{x} \ = \ \sum_{i=1}^{M} \langle \vec{a}_i | \vec{x} \rangle \, \vec{a}_i, \tag{X.37}$$

und bemerken, dass $\mathrm{Ran}(P) = Y$. Offenbar ist $P = P^* = P^2$ die orthogonale Projektion auf Y. Schreiben wir nun, für $\vec{x} \in X$,

$$\vec{x} \ = \ P\vec{x} + P^\perp \vec{x}, \tag{X.38}$$

so ist $P\vec{x} \in Y$. Für jedes $\vec{y} \in Y$ können wir wegen $Y = \mathrm{Ran}(P)$ ein $\vec{x}\,'$ finden, so dass $\vec{y} = P\vec{x}\,'$. Damit ist

$$\langle \vec{y} \,|\, P^\perp \vec{x} \rangle \ = \ \langle P\vec{x}\,' | P^\perp \vec{x} \rangle \ = \ \langle \vec{x}\,' | P P^\perp \vec{x} \rangle \ = \ 0, \tag{X.39}$$

da $PP^\perp = 0$. Es folgt, dass $P^\perp \vec{x} \in Y^\perp$. □

Satz X.12 (Spektralsatz)**.** *Seien $N \in \mathbb{N}$ und X ein N-dimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt $\langle \cdot | \cdot \rangle : X \times X \to \mathbb{C}$ und $\Phi = \Phi^* \in \mathcal{L}(X)$ eine selbstadjungierte lineare Abbildung. Dann gibt es eine Orthonormalbasis $\mathcal{A} = \{\vec{a}_1, \ldots, \vec{a}_N\} \subseteq X$ von X aus Eigenvektoren von Φ.*

Beweis. Wir führen den Beweis durch Induktion in N. Für $N = 1$ ist $X = \mathbb{C} \cdot \vec{a}_1$ für einen geeigneten normierten Vektor $\vec{a}_1$ und wegen $\Phi\vec{a}_1 \in \mathbb{C} \cdot \vec{a}_1$ gibt es ein $\lambda_1 \in \mathbb{C}$ so, dass $\Phi\vec{a}_1 = \lambda_1\vec{a}_1$, d.h. $\vec{a}_1$ ist trivialerweise ein Eigenvektor von Φ.

Seien nun $N \geq 2$ und die Behauptung für $N-1$ richtig. Nach Satz X.2 ist $\sigma(\Phi) \neq \emptyset$, und Φ besitzt einen Eigenwert $\lambda \in \mathbb{C}$. Es gibt also ein $\vec{y} \in X \setminus \{\vec{0}\}$, so dass $\Phi\vec{y} = \lambda\vec{y}$. Wir setzen nun

$$Y \;=\; \mathbb{C} \cdot \vec{y} \tag{X.40}$$

und stellen fest, dass $\Phi(Y) = \lambda Y \subseteq Y$. Weil $\Phi = \Phi^*$ selbstadjungiert ist, impliziert Lemma X.10, dass Φ auch $Y^\perp$ invariant lässt. Nach Lemma X.11 ist $\dim(Y^\perp) = N-1$, und nach Induktionsannahme besitzt $Y^\perp$ eine ONB $\mathcal{A}_{N-1} = \{\vec{a}_1, \ldots, \vec{a}_{N-1}\} \subseteq Y^\perp$ aus Eigenvektoren von (der Restriktion auf $Y^\perp$ von) Φ. $\square$

Bemerkungen und Beispiele.

- Sind $\Phi = \Phi^* \in \mathcal{L}(X)$ und $\lambda \in \mathbb{C}$ ein Eigenwert von Φ, so ist $\lambda \in \mathbb{R}$ reell. Ist nämlich $\vec{a}_\lambda$ ein zugehöriger Eigenvektor, $\Phi\vec{a}_\lambda = \lambda\vec{a}_\lambda$, so folgt

$$\lambda\langle\vec{a}_\lambda|\vec{a}_\lambda\rangle \;=\; \langle\vec{a}_\lambda|\Phi\vec{a}_\lambda\rangle \;=\; \langle\Phi\vec{a}_\lambda|\vec{a}_\lambda\rangle \;=\; \bar{\lambda}\langle\vec{a}_\lambda|\vec{a}_\lambda\rangle. \tag{X.41}$$

- Eigenvektoren zu verschiedenen Eigenwerten von $\Phi = \Phi^* \in \mathcal{L}(X)$ sind automatisch orthogonal. Sind etwa $\Phi\vec{a}_\lambda = \lambda\vec{a}_\lambda$ und $\Phi\vec{a}_\mu = \mu\vec{a}_\mu$, mit $\vec{a}_\lambda, \vec{a}_\mu \neq \vec{0}$ und $\lambda \neq \mu$, so folgt

$$\lambda\langle\vec{a}_\mu|\vec{a}_\lambda\rangle \;=\; \langle\vec{a}_\mu|\lambda\vec{a}_\lambda\rangle \;=\; \langle\Phi\vec{a}_\mu|\vec{a}_\lambda\rangle \;=\; \mu\langle\vec{a}_\mu|\vec{a}_\lambda\rangle. \tag{X.42}$$

  Also ist $\langle\vec{a}_\mu|\vec{a}_\lambda\rangle = \frac{0}{\lambda-\mu} = 0$.

X.4. Isometrien, orthogonale und unitäre Abbildungen

In diesem Abschnitt untersuchen wir Basistransformationen, die die Länge der Basisvektoren und ihre Winkel zueinander erhalten.

Definition X.13. Sei X ein $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{K}$. Eine Abbildung $\phi : X \to X$ heißt **Isometrie**

$$:\Leftrightarrow \quad \forall\, \vec{x}, \vec{y} \in X: \quad \langle\phi(\vec{x})|\phi(\vec{y})\rangle \;=\; \langle\vec{x}|\vec{y}\rangle. \tag{X.43}$$

Definition X.14. Sei $N \in \mathbb{N}$.

(i) Eine Matrix $D \in \mathbb{R}^{N\times N}$ heißt **orthogonal**

$$:\Leftrightarrow \qquad D^{-1} \;=\; D^T. \tag{X.44}$$

$$O(N) \;:=\big\{D \in \mathbb{R}^{N\times N} \;\big|\; D \text{ ist orthogonal}\big\}, \tag{X.45}$$

$$SO(N) \;:=\big\{D \in O(N) \;\big|\; \det[D] = 1\big\}. \tag{X.46}$$

$O(N)$ bezeichnet man als **orthogonale Gruppe**, $SO(N)$ heißt **eigentliche orthogonale Gruppe**.

(ii) Eine Matrix $U \in \mathbb{C}^{N \times N}$ heißt **unitär**

$$:\Leftrightarrow \qquad U^{-1} \;=\; U^*. \tag{X.47}$$

$$U(N) \;:=\; \big\{ U \in \mathbb{C}^{N \times N} \;\big|\; U \text{ ist unitär} \big\}, \tag{X.48}$$

$$SU(N) \;:=\; \big\{ U \in U(N) \;\big|\; \det[U] = 1 \big\}. \tag{X.49}$$

$U(N)$ heißt **unitäre Gruppe**, $SU(N)$ heißt **eigentliche unitäre Gruppe**.

Bemerkungen und Beispiele.

- Eine Isometrie erhält die Längen von Vektoren,

$$\|\phi(\vec{x})\| \;=\; \sqrt{\langle \phi(\vec{x}) | \phi(\vec{x}) \rangle} \;=\; \sqrt{\langle \vec{x} | \vec{x} \rangle} \;=\; \|\vec{x}\|, \tag{X.50}$$

und Orthogonalität,

$$\Big(\langle \vec{x} | \vec{y} \rangle \;=\; 0 \Big) \quad \Rightarrow \quad \Big(\langle \phi(\vec{x}) | \phi(\vec{y}) \rangle \;=\; 0 \Big). \tag{X.51}$$

- Für $N := \dim_{\mathbb{K}}(X) < \infty$ ist also mit $\mathcal{A} = \{\vec{a}_1, \ldots, \vec{a}_N\} \subseteq X$ auch $\phi(\mathcal{A}) = \{\phi(\vec{a}_1), \ldots, \phi(\vec{a}_N)\} \subseteq X$ eine ONB.
- Für $N := \dim_{\mathbb{K}}(X) < \infty$ ist jede Isometrie automatisch bijektiv und linear. Sind nämlich $\alpha \in \mathbb{K}$, $\vec{x}, \vec{y} \in X$ und $\mathcal{A} = \{\vec{a}_1, \ldots, \vec{a}_N\} \subseteq X$ eine ONB, so ist, für alle $j \in \mathbb{Z}_1^N$,

$$\begin{aligned} &\langle \phi(\vec{a}_j) | \phi(\alpha\vec{x} + \vec{y}) - \alpha\phi(\vec{x}) - \phi(\vec{y}) \rangle \\ &\qquad = \langle \phi(\vec{a}_j) | \phi(\alpha\vec{x} + \vec{y}) \rangle - \alpha \langle \phi(\vec{a}_j) | \phi(\vec{x}) \rangle - \langle \phi(\vec{a}_j) | \phi(\vec{y}) \rangle \\ &\qquad = \langle \vec{a}_j | \alpha\vec{x} + \vec{y} \rangle - \alpha \langle \vec{a}_j | \vec{x} \rangle - \langle \vec{a}_j | \vec{y} \rangle \;=\; \langle \vec{a}_j | \vec{0} \rangle \;=\; 0\,. \end{aligned} \tag{X.52}$$

Da $\phi(\mathcal{A})$ eine ONB in X ist, folgt $\phi(\alpha\vec{x} + \vec{y}) = \alpha\phi(\vec{x}) + \phi(\vec{y})$, und ϕ ist linear. Da $\phi(\mathcal{A})$ eine ONB in X ist, ist ϕ surjektiv und daher auch bijektiv.

- Ist $\phi \in \mathcal{L}(X)$ als linear vorausgesetzt, so ist (X.43) sogar gleichwertig mit (X.50), denn

$$\langle \phi(\vec{x}) | \phi(\vec{y}) \rangle \;=\; \frac{1}{2}\Big(\|\phi(\vec{x} + \vec{y})\|^2 - \|\phi(\vec{x})\|^2 - \|\phi(\vec{y})\|^2 \Big), \text{ für } \mathbb{K} = \mathbb{R}, \tag{X.53}$$

$$\langle \phi(\vec{x}) | \phi(\vec{y}) \rangle \;=\; \frac{1}{2}\Big(\|\phi(\vec{x} + \vec{y})\|^2 + \|\phi(\vec{x} - i\vec{y})\|^2 \Big) - \|\phi(\vec{x})\|^2 - \|\phi(\vec{y})\|^2, \text{ für } \mathbb{K} = \mathbb{C}. \tag{X.54}$$

- Ist $U \in U(N)$ unitär, so gilt

$$1 \;=\; \det[U^* U] \;=\; \overline{\det[U^T]} \cdot \det[U] \;=\; |\det[U]|^2, \tag{X.55}$$

d.h. die Determinante einer unitären Matrix ist im Betrag gleich 1.

- Dasselbe gilt für die Determinante einer orthogonalen Matrix D. Weil D eine reelle Matrix ist, folgt außerdem

$$\det[D] \;\in\; \{-1, +1\}. \tag{X.56}$$

Satz X.15. *Seien X ein N-dimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{K}$ und Orthonormalbasis $\mathcal{A} = \{\vec{a}_1, \ldots, \vec{a}_N\} \subseteq X$.*

(i) Für $\mathbb{K} = \mathbb{R}$ sind die Matrixdarstellungen bezüglich $\mathcal{A}$ der Isometrien in $\mathcal{L}(X)$ genau die orthogonalen Matrizen,

$$O(N) \;=\; \big\{\mathcal{M}[\phi] \;\big|\; \phi : X \to X \text{ ist Isometrie}\big\}. \tag{X.57}$$

(ii) Für $\mathbb{K} = \mathbb{C}$ sind die Matrixdarstellungen bezüglich $\mathcal{A}$ der Isometrien in $\mathcal{L}(X)$ genau die unitären Matrizen,

$$U(N) \;=\; \big\{\mathcal{M}[\phi] \;\big|\; \phi : X \to X \text{ ist Isometrie}\big\}. \tag{X.58}$$

Beweis. Wir beweisen nur (ii). Ist $\phi : X \to X$ eine Isometrie, so ist ϕ automatisch linear und bijektiv. Nach Korollar IX.9 ist $U := \mathcal{M}[\phi] := (u_{i,j})_{i,j=1}^N \in \mathbb{C}^{N\times N}$ gegeben durch $u_{i,j} = \langle\vec{a}_i|\phi\vec{a}_j\rangle$. Mit $U^*U =: (\gamma_{i,j})_{i,j=1}^N$ und $UU^* =: (\tilde{\gamma}_{i,j})_{i,j=1}^N$ sind also

$$\begin{aligned}\gamma_{i,j} \;&=\; \sum_{k=1}^N \overline{u_{k,i}}\, u_{k,j} \;=\; \sum_{k=1}^N \overline{\langle\vec{a}_k|\phi\vec{a}_i\rangle}\, \langle\vec{a}_k|\phi\vec{a}_j\rangle \\ &=\; \sum_{k=1}^N \langle\phi\vec{a}_i|\vec{a}_k\rangle\langle\vec{a}_k|\phi\vec{a}_j\rangle \;=\; \left\langle \phi\vec{a}_i \;\middle|\; \sum_{k=1}^N \langle\vec{a}_k|\phi\vec{a}_j\rangle\vec{a}_k \right\rangle\end{aligned} \tag{X.59}$$

und analog

$$\tilde{\gamma}_{i,j} \;=\; \left\langle \vec{a}_i \;\middle|\; \sum_{k=1}^N \langle\phi\vec{a}_k|\vec{a}_j\rangle\phi\vec{a}_k \right\rangle. \tag{X.60}$$

Nach (IX.43) sind jedoch

$$\sum_{k=1}^N \langle\vec{a}_k|\phi\vec{a}_j\rangle\, \vec{a}_k \;=\; \phi\vec{a}_j \quad \text{und} \quad \sum_{k=1}^N \langle\phi\vec{a}_k|\vec{a}_j\rangle\, \phi\vec{a}_k \;=\; \vec{a}_j, \tag{X.61}$$

weil $\mathcal{A} = \{\vec{a}_1, \ldots, \vec{a}_N\}$ und auch $\phi\mathcal{A} = \{\phi\vec{a}_1, \ldots, \phi\vec{a}_N\}$ ONB in X sind. Setzen wir dies in (X.59) und (X.60) ein, so erhalten wir

$$\gamma_{i,j} \;=\; \langle\phi\vec{a}_i|\phi\vec{a}_j\rangle \;=\; \delta_{i,j} \quad \text{und} \quad \tilde{\gamma}_{i,j} \;=\; \langle e_i|e_j\rangle \;=\; \delta_{i,j}. \tag{X.62}$$

Das bedeutet aber nichts anderes als

$$U^*U \;=\; UU^* \;=\; \mathbb{1}\,, \tag{X.63}$$

d.h. $U \in U(N)$. Somit ist

$$\big\{\mathcal{M}[\phi] \;\big|\; \phi : X \to X \text{ ist Isometrie}\big\} \;\subseteq\; U(N). \tag{X.64}$$

Ist umgekehrt $U \in U(N)$, also $U \in \mathbb{C}^{N\times N}$ mit $U^{-1} = U^*$, und ist $\Phi \in \mathcal{L}(X)$ die eindeutige lineare Abbildung, so dass $\mathcal{M}[\Phi] = U$, dann ist $\Phi^{-1} = \Phi^*$, weil $\mathcal{M} : \mathcal{L}(X) \to \mathbb{C}^{N\times N}$ ein Ringisomorphismus ist. Für alle $\vec{x}, \vec{y} \in X$ gilt also

$$\big\langle \vec{x} \big| \vec{y} \big\rangle \;=\; \big\langle \vec{x} \big| \Phi^{-1}\Phi\vec{y} \big\rangle \;=\; \big\langle \vec{x} \big| \Phi^*\Phi\vec{y} \big\rangle \;=\; \big\langle \Phi\vec{x} \big| \Phi\vec{y} \big\rangle. \tag{X.65}$$

Also ist Φ eine Isometrie, und damit gilt

$$\mathrm{U(N)} \;\subseteq\; \big\{ \mathcal{M}[\phi] \;\big|\; \phi : X \to X \text{ ist Isometrie} \big\}. \tag{X.66}$$

□

Bemerkungen und Beispiele.

- Schreibt man $U = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_N\} \in \mathbb{K}^{N\times N}$ mit

$$\vec{u}_1 \;=\; \begin{pmatrix} u_{1,1} \\ \vdots \\ u_{N,1} \end{pmatrix}, \ldots,\; \vec{u}_N \;=\; \begin{pmatrix} u_{1,N} \\ \vdots \\ u_{N,N} \end{pmatrix} \;\in\; \mathbb{K}^N, \tag{X.67}$$

  so ist U genau dann unitär (orthogonal), wenn $\{\vec{u}_1, \ldots, \vec{u}_N\} \subseteq \mathbb{K}^N$ eine ONB darstellt,

$$\Big\{ U \in U(N) \,\big(\in O(N)\big) \Big\} \quad \Leftrightarrow \quad \Big\{ \forall\, i,j \in \mathbb{Z}_1^N : \; \langle \vec{u}_i | \vec{u}_j \rangle = \delta_{i,j} \Big\}, \tag{X.68}$$

  wobei hier $\langle\cdot|\cdot\rangle$ das unitäre (euklidische) Skalarprodukt notiert.

Korollar X.16. *Seien $N \in \mathbb{N}$ und $A = A^* \in \mathbb{C}^{N\times N}$ eine selbstadjungierte Matrix. Dann ist A diagonalisierbar, und es gibt eine unitäre Matrix $U \in U(N)$ und Eigenwerte $\lambda_1, \lambda_2, \ldots, \lambda_N \in \mathbb{R}$, so dass*

$$U^* A\, U \;=\; \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ 0 & & & \lambda_N \end{pmatrix}. \tag{X.69}$$

Beweis. Sind $\mathcal{E} = \{\vec{e}_1, \ldots, \vec{e}_N\} \subseteq \mathbb{C}^N$ die Standardbasis und $\Phi \in \mathcal{L}(\mathbb{C}^N)$ die eindeutige lineare Abbildung, sodass $A = \mathcal{M}_{\mathcal{E}}[\Phi]$ die Matrixdarstellung von Φ ist, so ist $\Phi = \Phi^*$ selbstadjungiert. Nach dem Spektralsatz gibt es eine ONB $\mathcal{W} = \{\vec{w}_1, \ldots, \vec{w}_N\} \subseteq \mathbb{C}^N$ aus Eigenvektoren von Φ mit zugehörigen Eigenwerten λ_j, sodass also $\Phi\vec{w}_j = \lambda_j\vec{w}_j$. Ist nun $\Theta \in \mathcal{L}(\mathbb{C}^N)$ die Transformation von der ONB $\mathcal{E}$ auf die ONB $\mathcal{W}$, d.h. gilt $\Theta\vec{e}_j = \vec{w}_j$, so ist $U := \mathcal{M}_{\mathcal{E}}[\Theta] \in U(N)$ unitär, und es gilt

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix} \;=\; \mathcal{M}_{\mathcal{W}}[\Phi] \;=\; \mathcal{M}_{\mathcal{E}}[\Theta]^{-1} \cdot \mathcal{M}_{\mathcal{E}}[\Phi] \cdot \mathcal{M}_{\mathcal{E}}[\Theta] \;=\; U^* A\, U\,. \tag{X.70}$$

□

X.5. Anwendung des Spektralsatzes zur Lösung von Systemen linearer Differenzialgleichungen

Zur Illustration des Nutzens des Werkzeugs, das der Spektralsatz für Anwendungen darstellt, wollen wir zeigen, wie man ein System gewöhnlicher Differenzialgleichungen mit ihm löst. Ein solches System ist für $N \in \mathbb{N}$ durch N lineare Differenzialgleichungen

$$\dot{x}_1(t) = a_{1,1}x_1(t) + a_{1,2}x_2(t) + \ldots + a_{1,N}x_N(t)\,, \tag{X.71}$$

$$\dot{x}_2(t) = a_{2,1}x_1(t) + a_{2,2}x_2(t) + \ldots + a_{2,N}x_N(t)\,, \tag{X.72}$$

$$\vdots$$

$$\dot{x}_N(t) = a_{N,1}x_1(t) + a_{N,2}x_2(t) + \ldots + a_{N,N}x_N(t)\,, \tag{X.73}$$

gegeben. Gesucht sind Lösungen $x_1, x_2, \ldots, x_N : \mathbb{R}_0^+ \to \mathbb{C}$, die $x_1(0) = \alpha_1, x_2(0) = \alpha_2, \ldots, x_N(0) = \alpha_N$ genügen, wobei $\alpha_1, \alpha_2, \ldots, \alpha_N \in \mathbb{C}$ eine Anfangsbedingung definieren. Mit

$$\vec{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_N(t) \end{pmatrix}, \quad A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\ \vdots & a_{2,2} & & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{pmatrix}, \quad \text{und} \quad \vec{x}_0 = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{pmatrix}, \tag{X.74}$$

kann man das Differenzialgleichungssystem (X.71)-(X.73) als eine einzige Differenzialgleichung

$$\forall\, t > 0 : \quad \dot{\vec{x}}(t) = A\vec{x}(t)\,, \quad \vec{x}(0) = \vec{x}_0\,, \tag{X.75}$$

für eine differenzierbare vektorwertige Funktion $\vec{x} : \mathbb{R}_0^+ \to \mathbb{C}^N$ und einen Anfangswert $\vec{x}_0 \in \mathbb{C}^N$ schreiben. Wir wollen dabei annehmen, dass die die Gleichung definierende komplexe $N \times N$-Matrix $A := (a_{i,j})_{i,j=1}^N \in \mathbb{C}^{N\times N}$ selbstadjungiert ist, $A = A^*$.

Bemerkungen und Beispiele.

- Vermutlich wäre die Bezeichung „Anfangsvektor„ für $\vec{x}_0$ treffender, ist aber nicht üblich.
- Die Beschränkung des obigen AWP auf selbstadjungierte Matrizen erfasst zwar viele, aber nicht alle Beispielanwendungen. Tatsächlich kann man das AWP (X.75) für allgemeine komplexe $N \times N$-Matrizen ähnlich lösen; wir kommen am Schluss darauf zurück.
- Auch die Festlegung, dass die Matrix A in (X.75) unabhängig von der Variablen t (die wir als Zeitvariable interpretieren) sein soll, ist eine der Kürze der für die Vorlesung zur Verfügung stehenden Zeit geschuldete Vereinfachung - aber kein prinzipielles mathematisches Hindernis.
- Im AWP (X.75) tritt die Zeitableitung einer vektorwertigen Funktion auf, die in der Schule nicht behandelt wird.

- Wir erinnern zunächst an die Definition der Differenzierbarkeit einer reellen Funktion, die für nichtnegative Zeiten definiert ist: Eine Funktion $f : \mathbb{R}_0^+ \to \mathbb{R}$ heißt *differenzierbar bei* $t > 0$, falls der Grenzwert
$$\dot{f}(t) \ := \ \lim_{h \to 0} \left\{ \frac{f(t+h) - f(t)}{h} \right\} \ \in \ \mathbb{R} \tag{X.76}$$
existiert. In diesem Fall nennen wir $\dot{f}(t)$ die *Ableitung von* f *(bei* t*).*
- Ist f differenzierbar bei t für alle $t > 0$, so heißt f *differenzierbar auf* $\mathbb{R}^+$ und die Ableitung definiert eine Abbildung $\dot{f} : \mathbb{R}^* \to \mathbb{R}$.
- Man beachte, dass die Ableitung bei $t = 0$ nicht gebildet werden kann, weil die Grenzwertbildung in (X.76) bedingt, dass man sich t von links *und* von rechts nähert.
- Die Definition der Differenzierbarkeit einer Funktion $\vec{x} : \mathbb{R}_0^+ \to \mathbb{C}^N$ und ihrer Ableitung ist die natürliche Verallgemeinerung von (X.76): $\vec{x} : \mathbb{R}_0^+ \to \mathbb{C}^N$ heißt *differenzierbar bei* $t > 0$, falls der Grenzwert
$$\dot{\vec{x}}(t) \ := \ \lim_{h \to 0} \left\{ \frac{\vec{x}(t+h) - \vec{x}(t)}{h} \right\} \ \in \ \mathbb{C}^N \tag{X.77}$$
existiert. In diesem Fall nennen wir $\dot{\vec{x}}(t)$ die *Ableitung von* $\vec{x}$ *(bei* t*).*
- Ist $\vec{x}$ differenzierbar bei t für alle $t > 0$, so heißt $\vec{x}$ *differenzierbar auf* $\mathbb{R}^+$ und die Ableitung definiert eine Abbildung $\dot{\vec{x}} : \mathbb{R}^* \to \mathbb{C}^N$.
- Schreiben wir $\vec{x}(t) = \big(x_1(t), \ldots, x_N(t)\big)^T$ mit $x_n(t) \in \mathbb{C}$, so ist
$$\begin{aligned} \frac{\vec{x}(t+h) - \vec{x}(t)}{h} \ &= \frac{1}{h} \left\{ \begin{pmatrix} x_1(t+h) \\ \vdots \\ x_N(t+h) \end{pmatrix} - \begin{pmatrix} x_1(t) \\ \vdots \\ x_N(t) \end{pmatrix} \right\} \\ &= \begin{pmatrix} \frac{\mathrm{Re}[x_1(t+h)] - \mathrm{Re}[x_1(t)]}{h} + i \frac{\mathrm{Im}[x_1(t+h)] - \mathrm{Im}[x_1(t)]}{h} \\ \vdots \\ \frac{\mathrm{Re}[x_N(t+h)] - \mathrm{Re}[x_N(t)]}{h} + i \frac{\mathrm{Im}[x_N(t+h)] - \mathrm{Im}[x_N(t)]}{h} \end{pmatrix}, \end{aligned} \tag{X.78}$$
und es folgt sofort, dass die Ableitung der vektorwertigen Funktion $t \mapsto \vec{x}(t)$ auf die übliche Ableitungen reeller Funktionen, nämlich der Real- und Imaginärteile der Komponenten von $\vec{x}(t)$, zurückgeführt werden kann,
$$\dot{\vec{x}}(t) \ = \ \begin{pmatrix} \mathrm{Re}[\dot{x}_1(t)] + i\mathrm{Im}[\dot{x}_1(t)] \\ \vdots \\ \mathrm{Re}[\dot{x}_N(t)] + i\mathrm{Im}[\dot{x}_N(t)] \end{pmatrix} \ = \ \begin{pmatrix} \dot{x}_1(t) \\ \vdots \\ \dot{x}_N(t) \end{pmatrix}. \tag{X.79}$$
Setzt man auf der rechten Seite die Gleichungen (X.71)-(X.73) ein, erhält man sofort (X.75).
- Sind $B \in \mathbb{C}^{N \times N}$ eine feste komplexe $N \times N$-Matrix und $\vec{y}(t) := B\vec{x}(t)$, so beobachten wir, dass
$$\frac{\vec{y}(t+h) - \vec{y}(t)}{h} \ = \ \frac{B\vec{x}(t+h) - B\vec{x}(t)}{h} \ = \ B\left(\frac{\vec{x}(t+h) - \vec{x}(t)}{h} \right). \tag{X.80}$$

Im Limes $h \to 0$ folgt, dass mit $\vec{x}$ auch $\vec{y}$ differenzierbar ist und

$$\dot{\vec{y}}(t) \;=\; B\,\dot{\vec{x}}(t)\,. \tag{X.81}$$

Zur Lösung der Differenzialgleichung (X.75) wenden wir den Spektralsatz und genauer Korollar X.16 auf die selbstadjungierte Matrix A an: Nach Korollar X.16 ist A diagonalisierbar, und es gibt eine unitäre Matrix $U \in U(N)$ und Eigenwerte $\lambda_1, \lambda_2, \ldots, \lambda_N \in \mathbb{R}$, so dass

$$U^*\,A\,U \;=\; \Lambda \;:=\; \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ 0 & & & \lambda_N \end{pmatrix} \tag{X.82}$$

gilt. Durch Multiplikation mit U von links und U^* von rechts wird diese Gleichung äquivalent überführt in

$$A \;=\; U\,\Lambda\,U^*\,. \tag{X.83}$$

Sei nun $\vec{x} : \mathbb{R}_0^* \to \mathbb{C}^N$ eine Lösung der Differenzialgleichung (X.75). Wir definieren

$$\vec{z}(t) \;:=\; U^*\vec{x}(t) \quad \text{und} \quad \vec{z}_0 \;:=\; U^*\vec{x}_0 \tag{X.84}$$

und beobachten, dass $\vec{z}(0) = \vec{z}_0$ und für $t > 0$

$$\dot{\vec{z}}(t) \;=\; U^*\dot{\vec{x}}(t) \;=\; U^*\,A\vec{x}(t) \;=\; U^*\,A\,U\,U^*\vec{x}(t) \;=\; \Lambda\vec{z}(t) \tag{X.85}$$

gilt. Glg. (X.85) und die Anfangsbedingung $\vec{z}(0) = \vec{z}_0 =: (\beta_1, \ldots, \beta_N)^T$ ist aber gleichwertig mit dem System folgender N Differenzialgleichungen,

$$\dot{z}_1(t) \;=\; \lambda_1\,z_1(t)\,, \qquad z_1(0) \;=\; \beta_1 \tag{X.86}$$

$$\dot{z}_2(t) \;=\; \lambda_2\,z_2(t)\,, \qquad z_2(0) \;=\; \beta_2 \tag{X.87}$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$\dot{z}_N(t) \;=\; \lambda_N\,z_N(t)\,, \qquad z_N(0) \;=\; \beta_N\,. \tag{X.88}$$

Diese Differenzialgleichungen besitzen aber die eindeutigen Lösungen

$$\forall\, n \in \mathbb{Z}_1^N\,,\; t \geq 0 : z_n(t) \;=\; e^{\lambda_n t}\,\beta_n\,, \tag{X.89}$$

bzw. in Vektorschreibweise

$$\forall\, t \geq 0 : \vec{z}(t) \;=\; \exp[t\Lambda]\,\vec{z}_0\,, \qquad \exp[t\Lambda] \;:=\; \begin{pmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & 0 \\ & & \ddots & \\ 0 & & & e^{\lambda_N t} \end{pmatrix}. \tag{X.90}$$

Damit erhalten wir für die ursprünglich gesuchte Vektorfunktion

$$\forall\, t \geq 0 : \vec{x}(t) \;=\; U\,\exp[t\Lambda]\,U^*\vec{x}_0\,. \tag{X.91}$$

Dies ist eine Lösung der Differenzialgleichung (X.75) und weil die Lösungen $z_n : \mathbb{R}_0^+ \to \mathbb{C}$ der Differenzialgleichungen $\dot{z}_n(t) = \lambda_n z_n(t)$ mit Anfangswerten $z_n(0) = \beta_0$ eindeutig sind, ist (X.91) auch die einzige Lösung von (X.75).

X.6. Ergänzungen

X.6.1. Beweis von Lemma X.3

Beweis. Für $n \in \mathbb{Z}_1^L$ definieren wir

$$A_n \ := \ \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\} \tag{X.92}$$

und beobachten, dass $A_1 = \{\vec{x}_1\}$ wegen $\vec{x}_1 \neq 0$ linear unabhängig ist. Wir zeigen nun, dass

$$\forall\, n \in \mathbb{Z}_1^{L-1}: \quad \Big(A_n \text{ ist linear unabhängig}\Big) \ \Rightarrow \ \Big(A_{n+1} \text{ ist linear unabhängig}\Big) \tag{X.93}$$

gilt. Dies liefert dann induktiv die lineare Unabhängigkeit von A_L und somit die Behauptung.

Um (X.93) zu zeigen, nehmen wir an, dass A_n linear unabhängig ist und A_{n+1} linear abhängig wäre und führen diese Annahme zum Widerspruch. Die lineare Abhängigkeit von A_{n+1} bedeutet, dass es $(\alpha_1, \ldots, \alpha_n, \alpha_{n+1}) \neq (0, \ldots, 0, 0)$ so gibt, dass

$$\alpha_1 \vec{x}_1 + \ldots + \alpha_n \vec{x}_n + \alpha_{n+1} \vec{x}_{n+1} \ = \ \vec{0}\,. \tag{X.94}$$

Wäre nun $\alpha_{n+1} = 0$, so folgte aus der linearen Unabhängigkeit von A_n, dass mit $\alpha_1 \vec{x}_1 + \ldots + \alpha_n \vec{x}_n = \vec{0}$ auch $\alpha_1 = \cdots = \alpha_n = 0$ wären, also $(\alpha_1, \ldots, \alpha_n, \alpha_{n+1}) = (0, \ldots, 0, 0)$ gälte, was in Widerspruch zu $(\alpha_1, \ldots, \alpha_n, \alpha_{n+1}) \neq (0, \ldots, 0, 0)$ stünde. Also ist $\alpha_{n+1} \neq 0$.

Wir setzen nun $\beta_\ell := -\alpha_\ell / \alpha_{n+1}$ und erhalten

$$\vec{x}_{n+1} \ = \ \beta_1 \vec{x}_1 + \ldots + \beta_n \vec{x}_n\,. \tag{X.95}$$

Wenden wir nun $\Phi - \lambda_{n+1}\mathbf{1}$ auf (X.95) an, so erhalten wir

$$\begin{aligned} \vec{0} \ &= \big(\Phi - \lambda_{n+1}\mathbf{1}\big)\vec{x}_{n+1} \ = \ \sum_{\ell=1}^{n} \beta_\ell \,\big(\Phi - \lambda_{n+1}\mathbf{1}\big)\vec{x}_\ell \ = \ \sum_{\ell=1}^{n} \beta_\ell \,\big(\Phi\vec{x}_\ell - \lambda_{n+1}\vec{x}_\ell\big) \\ &= \sum_{\ell=1}^{n} \beta_\ell (\lambda_\ell - \lambda_{n+1})\, \vec{x}_\ell. \end{aligned} \tag{X.96}$$

Da A_n linear unabhängig ist, impliziert dies, dass $\beta_\ell(\lambda_\ell - \lambda_{n+1}) = 0$ für alle $\ell \in \mathbb{Z}_1^n$ sind. Wegen der paarweisen Verschiedenheit der Eigenwerte sind außerdem $\lambda_\ell - \lambda_{n+1} \neq 0$, und deshalb muss sogar $\beta_1 = \cdots = \beta_n = 0$ gelten, d.h. es ist $\vec{x}_{n+1} = \vec{0}$. Widerspruch; also ist A_{n+1} linear unabhängig. □

X.6.2. Die Drehgruppe $\mathrm{SO}(N)$

Bemerkungen und Beispiele.

- Ist $D = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SO(2)$, so gilt nach (VII.56) und mit $\det[D] = 1$, dass

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \ = \ D^{-1} \ = \ D^T \ = \ \begin{pmatrix} a & c \\ b & d \end{pmatrix}, \tag{X.97}$$

also $a = d$ und $b = -c$. Außerdem ist dann $1 = \det[D] = a^2 + b^2$. Wählen wir $\varphi \in [0, 2\pi]$ so, dass

$$\cos(\varphi) = \frac{a}{\sqrt{a^2 + b^2}} = a \quad \text{und} \quad \sin(\varphi) = \frac{b}{\sqrt{a^2 + b^2}} = b\,, \tag{X.98}$$

dann ist also

$$D = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix}. \tag{X.99}$$

Betrachtet man D als Basistransformation, so *dreht* D das kartesische Koordinatensystem um φ

- Ein typisches Element $S \in O(2) \setminus SO(2)$ ist

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{X.100}$$

Als Basistransformation *spiegelt* S die $\vec{a}_2$-Achse und lässt die $\vec{a}_1$-Achse unberührt. Die Spiegelungseigenschaft wird mathematisch durch $S^2 = \mathbb{1}$ ausgedrückt.

- Jede Matrix $M \in O(2)$ ist entweder eine Drehung, $M \in SO(2)$, oder das Produkt einer Drehung $D \in SO(2)$ und einer Spiegelung $S \in O(2) \setminus SO(2)$, d.h. $M = D \cdot S$.
- Weiterhin bilden $O(N) \subseteq GL(N, \mathbb{R})$ eine Untergruppe der reellen, invertiblen $N \times N$-Matrizen und $SO(N) \subseteq O(N)$ eine Untergruppe von O(N) bezüglich Matrixmultiplikation, wie man leicht durch Nachprüfen der Gruppenaxiome einsieht.
- Man bezeichnet SO(N) auch als Drehgruppe.
- Tatsächlich lässt sich auch jedes $D \in SO(3)$ für geeignete $\alpha, \beta, \gamma \in [0, 2\pi]$ als $D = D_3(\alpha) \cdot D_1(\beta) \cdot D_3(\gamma)$ schreiben, wobei

$$D_1(\beta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta) & \sin(\beta) \\ 0 & -\sin(\eta) & \cos(\beta) \end{pmatrix}, \quad D_3(\alpha) = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{X.101}$$

Dabei sind α, β, γ die eulerschen Winkel. Diese Parametrisierung spielt in der Theorie des Kreisels eine wichtige Rolle. Jede Drehung in $\mathbb{R}^3$ ist also die Komposition einer Drehung um die $\vec{a}_3$-Achse, dann einer weiteren um die $\vec{a}_1$-Achse und anschließend nochmal eine um die $\vec{a}_3$-Achse.

- Natürlich bilden auch $SU(N) \subseteq U(N) \subseteq GL(N, \mathbb{C})$ Untergruppen der komplexen, invertiblen $(N \times N)$-Matrizen bezüglich Matrixmultiplikation.

X.6.3. Diagonalisierbarkeit normaler Operatoren

Definition X.17. Seien X ein endlich dimensionaler Vektorraum über $\mathbb{C}$ und $\Phi, \Psi \in \mathcal{L}(X)$ zwei lineare Abbildungen. Der **Kommutator von Φ und Ψ** ist die lineare Abbildung

$$[\Phi, \Psi] := \Phi\,\Psi - \Psi\,\Phi \in \mathcal{L}(X). \tag{X.102}$$

Satz X.18 (Kommutierende Observablen)**.** *Seien* $N \in \mathbb{N}, X$ *ein* $\mathbb{C}$*-Vektorraum der Dimension* $\dim(X) = N$ *mit Skalarprodukt* $\langle \cdot | \cdot \rangle : X \times X \to \mathbb{C}$ *und* $\Phi_1 = \Phi_1^*, \Phi_2 = \Phi_2^*, \ldots, \Phi_L = \Phi_L^* \in \mathcal{L}(X)$ *selbstadjungierte lineare Abbildungen mit Eigenwerten* $\sigma(\Phi_\ell) = \{\lambda_i^{(\ell)}\}_{i=1}^N$*, für* $\ell \in \mathbb{Z}_1^L$*. Kommutieren* $[\Phi_k, \Phi_\ell] = 0$*, für alle* $k, \ell \in \mathbb{Z}_1^L$*, so gibt es eine ONB* $\mathcal{A} = \{\vec{a}_1, \ldots, \vec{a}_N\} \subseteq X$ *gemeinsamer Eigenvektoren aller* $\Phi_1, \Phi_2, \ldots, \Phi_L$*, d.h.*

$$\forall\, \ell \in \mathbb{Z}_1^L \;\forall\, n \in \mathbb{Z}_1^N : \quad \Phi_\ell \vec{a}_n \;=\; \lambda_n^{(\ell)} \vec{a}_n. \tag{X.103}$$

Beweis. Wir führen den Beweis nur für $L = 2$ und genauer zwei selbstadjungierte linearen Abbildungen $\Phi = \Phi^*, \Psi = \Psi^* \in \mathcal{L}(X)$, die kommutieren, $\Phi\Psi = \Psi\Phi$.

Seien $\lambda_1, \lambda_2, \ldots, \lambda_K \in \sigma(\Phi)$ die paarweise verschiedenen Eigenwerte von Φ und $Y_k := \mathrm{Ker}[\Phi - \lambda_k \mathbf{1}]$ die von den zugehörigen Eigenvektoren aufgespannten Unterräume, die **Eigenräume von $\boldsymbol{\Phi}$**. Da Eigenvektoren Φ zu verschiedenen Eigenwerten zueinander orthogonal sind, folgt

$$X \;=\; Y_1 \oplus Y_2 \oplus \ldots \oplus Y_K. \tag{X.104}$$

Seien nun $k \in \mathbb{Z}_1^K$ und $\vec{x} \in Y_k$, also $\Phi\vec{x} = \lambda_k \vec{x}$. Wegen

$$\Phi(\Psi\vec{x}) \;=\; \Psi(\Phi\vec{x}) \;=\; \lambda_k\, \Psi\vec{x} \tag{X.105}$$

ist auch $\Psi\vec{x} \in Y_k$, d.h. die Eigenräume Y_k sind auch unter Ψ invariant,

$$\forall k \in \mathbb{Z}_1^K : \quad \Psi Y_k \;\subseteq\; Y_k. \tag{X.106}$$

Bezeichnen wir die Restriktion von Ψ auf Y_k mit $\Psi_k := \Psi \restriction_{Y_k} \in \mathcal{L}(Y_k)$, so überträgt sich die Selbstadjungiertheit von Ψ auf Ψ_k und nach dem Spektralsatz gibt es eine ONB in Y_k aus Eigenvektoren von Ψ_k, die auch alle Eigenvektoren von Φ zum Eigenwert λ_k sind. Die Vereinigung dieser ONB über alle $k \in \mathbb{Z}_1^K$ ist dann die gesuchte ONB von X aus gemeinsamen Eigenvektoren von Φ und Ψ. □

Definition X.19. Seien X ein endlich-dimensionaler Vektorraum über $\mathbb{C}$ mit Skalarprodukt $\langle \cdot | \cdot \rangle : X \times X \to \mathbb{C}$. Eine lineare Abbildung $\Phi \in \mathcal{L}(X)$ heißt **normal**

$$:\Leftrightarrow \qquad \Phi\Phi^* \;=\; \Phi^*\Phi. \tag{X.107}$$

Bemerkungen und Beispiele.

- Selbstadjungierte lineare Abbildungen sind normal, da $\Phi\Phi^* = \Phi^2 = \Phi^*\Phi$.
- Unitäre lineare Abbildungen sind normal, da $UU^* = \mathbf{1} = U^*U$.
- Zu jeder linearen Abbildung $\Phi \in \mathcal{L}(X)$ bilden wir ihren **Realteil**

  $$\mathrm{Re}\{\Phi\} \;=\; \big(\mathrm{Re}\{\Phi\}\big)^* \;:=\; \frac{1}{2}\left(\Phi + \Phi^*\right) \;\in\; \mathcal{L}(X), \tag{X.108}$$

  und ihren **Imaginärteil**

  $$\mathrm{Im}\{\Phi\} \;=\; \big(\mathrm{Im}\{\Phi\}\big)^* \;:=\; \frac{1}{2i}\left(\Phi - \Phi^*\right) \;\in\; \mathcal{L}(X). \tag{X.109}$$

  Dann gilt

  $$\Phi \;=\; \mathrm{Re}\{\Phi\} \;+\; i\,\mathrm{Im}\{\Phi\}. \tag{X.110}$$

- Es gilt folgende Äquivalenz:

$$\Big(\Phi \in \mathcal{L}(X) \text{ ist normal}\Big) \quad \Leftrightarrow \quad \Big([\mathrm{Re}\{\Phi\}, \mathrm{Im}\{\Phi\}] \;=\; 0\Big). \tag{X.111}$$

  Dies ergibt sich aus

$$\begin{aligned} [\mathrm{Re}\{\Phi\},\ \mathrm{Im}\{\Phi\}] \;&=\; \frac{1}{4i}\,\big\{(\Phi+\Phi^*)\,(\Phi-\Phi^*) - (\Phi-\Phi^*)\,(\Phi+\Phi^*)\big\} \\ &=\; \frac{1}{4i}\,\big\{\Phi^2 - (\Phi^*)^2 + \Phi^*\Phi - \Phi\Phi^* - \Phi^2 + (\Phi^*)^2 + \Phi^*\Phi - \Phi\Phi^*\big\} \\ &=\; \frac{1}{2i}\,\big\{\Phi^*\Phi - \Phi\Phi^*\big\}. \end{aligned} \tag{X.112}$$

Aus Satz X.18 und der Äquivalenz (X.111) ergibt sich noch das folgende Korollar

Korollar X.20. *Seien X ein endlich-dimensionaler Vektorraum über $\mathbb{C}$ mit Skalarprodukt $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{C}$ und $\Phi \in \mathcal{L}(X)$ eine normale lineare Abbildung. Dann gibt es eine ONB $\mathcal{A} = \{\vec{a}_1, \ldots, \vec{a}_N\} \subseteq X$ von X aus Eigenvektoren von Φ. Somit ist Φ diagonalisierbar.*