Skript zur Vorlesung

Lineare Algebra I

gehalten von Prof. Dr. Markus Schweighofer
im Wintersemester 2017/2018
an der Universität Konstanz

entstanden aus einer
früheren elektronischen Mitschrift
von Michael Strecke

vorläufige Version vom
6. November 2017, 09:42 Uhr

Vorwort

Im Wintersemester 2009/2010 las ich zum ersten Mal die „Lineare Algebra I“ an der Universität Konstanz. Damals habe ich die Sätze, Definitionen etc. noch nicht durchnummeriert und es gab kein elektronisches Skript. Es gab noch die akademische Viertelstunde und man konnte oftmals bis zu 20 Minuten überziehen. Damals hatte der Hausdienst noch die Zeit und die Kapazität, die Tafeln in den noch vorhandenen Pausen zu wischen.

Als ich die Vorlesung im Wintersemester 2013/2014 zum zweiten Mal las, war die akademische Viertelstunde bereits abgeschafft (offiziell ist sie ausgesetzt, aber ich bezweifle, dass sie jemals wieder eingesetzt wird). Man musste mehr oder weniger pünktlich die Vorlesung beenden, ausserdem wird man als Dozent mit zunehmendem Alter immer langsamer. Ich musste einige Abschnitte der Vorlesung tippen und teilweise als Präsentation vorführen, um mit dem Stoff durchzukommen. Glücklicherweise erstellten die Hörer Thomas Schmidt und Michael Strecke darauf aufbauend weitgehend unabhängig voneinander jeweils sehr schöne, jedoch von mir nicht überprüfte, elektronische Mitschriften zu meiner Vorlesung. Diese Skripten zusammen mit den zugehörigen LaTeX-Quelltexten sind weiterhin zugänglich unter:

`http://www.math.uni-konstanz.de/~schweigh/lehre-.html`

Nun halte ich im Wintersemester 2017/2018 die Vorlesung mit voraussichtlich nur geringfügigen Änderungen zum dritten Mal. Da ich mittlerweile noch langsamer geworden bin und im zugeteilten Hörsaal die Tafelsteuerung teilweise defekt ist, habe ich mich entschieden, das Skript von Michael Strecke zu einem von mir autorisierten Skript umzubauen, um darauf verweisen zu können (die Wahl zwischen den beiden Skripten war nicht leicht). Dazu versuche ich, während des Fortschreitens der Vorlesung jeweils das Skript durchzusehen und eventuell gefundene Fehler zu verbessern. Im Text vermerke ich in roter Schrift, bis wohin das Skript nun von mir autorisiert sein soll. Trotzdem wird das Skript sicherlich nicht ohne Fehler sein. Ich bin für jegliche Hinweise zu Fehlern (auch Druckfehlern) und Anregungen dankbar und nehme diese gerne persönlich oder per Email an `markus.schweighofer@uni-konstanz` entgegen. Das Skript und den zugehörigen LaTeX-Quelltext, welches bis auf einzelne Stellen identisch mit dem Skript von Michael Strecke sein wird, mache ich verfügbar unter:

`http://www.math.uni-konstanz.de/~schweigh/lehre.html`

Konstanz, den 5. November 2017.

Fassung vom 6. November 2017

Inhaltsverzeichnis

Fassung vom 6. November 2017

§1 Mengen

[Georg Ferdinand Ludwig Philipp Cantor *1845, †1918]

§1.1 Mengen und Abbildungen

Pseudodefinition 1.1.1. Eine *Menge* ist eine gedachte ungeordnete Ansammlung von Objekten, die man die *Elemente* der Menge nennt. Jedes Element darf dabei nur einmal in der Ansammlung vorkommen. Eine Menge kann auch leer sein oder unendlich viele Elemente haben. Ihre Elemente können selber wieder Mengen sein.

Warnung 1.1.2. Aus logischen Gründen, die wir hier nicht erklären, sind bei der Bildung von Mengen gewisse Spielregeln einzuhalten. Zum Beispiel darf eine Menge nicht alle Mengen als Element haben, sehr wohl aber alle Mengen, die nur aus reellen Zahlen bestehen. Sollten Sie diese Spielregeln wirklich einmal verletzen, so sagen wir es Ihnen.

Notation 1.1.3. Sind $a_1, a_2, a_3, \ldots, a_n$ Objekte (z.B. Zahlen, Mengen, Wörter,...), so schreibt man

$$\{a_1, \ldots, a_n\}$$

für die Menge bestehend aus den Elementen $a_1, \ldots, a_n$. Die Reihenfolge von $a_1, \ldots, a_n$ spielt dabei keine Rolle. Auch dürfen mehrere a_i gleich sein. Obwohl eine mehrfache Aufzählung redundant ist (ein Element kann gemäß 1.1.1 ja nicht „mehrfach“ enthalten sein), vermeidet dies oft eine unnötige und lästige gesonderte Behandlung von Spezialfällen. Die Menge $\{a_1, \ldots, a_n\}$ kann also auch weniger als n Elemente haben.

$$\emptyset \underbrace{:=}_{\text{„wird definiert durch“}} \{\} \qquad \text{„}\textit{leere}\text{ Menge“}$$

Beispiel 1.1.4. (a) $\{1, 2, 3, 4\} = \{3, 4, 2, 1\} = \{1, 1, 2, 3, 3, 4\}$ hat genau 4 Elemente.

(b) $\{\emptyset, 1, \{2, 3\}\}$ hat 3 Elemente, nämlich die leere Menge, die Zahl 1 und die zweielementige Menge $\{2, 3\}$. Man beachte, dass 3 kein Element von $\{\emptyset, 1, \{2, 3\}\}$ ist.

(c) $\{\{\{\{\{\{\}\}\}\}\}\}$ ist eine einelementige Menge, deren einziges Element die einelementige Menge $\{\{\{\{\{\}\}\}\}\} = \{\{\{\{\emptyset\}\}\}\}$ ist.

Notation 1.1.5. Manchmal verwendet man „. . .“, um große endliche oder unendliche Mengen zu schreiben:

$\mathbb{N} := \{1, 2, 3, 4, 5, \ldots\}$ Menge der *natürlichen Zahlen*
$\mathbb{N}_0 := \{0, 1, 2, 3, \ldots\}$
$\mathbb{Z} := \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ Menge der *ganzen Zahlen*
$\{1, \ldots, n\}$ Menge der natürlichen Zahlen $\leq n$

Warnung 1.1.6. Notationen wie in 1.1.5 sind oft missverständlich. Wie alle geübten Mathematiker verstehen wir:

- $\{1, \ldots, n\} = \{1, 2\}$ für $n = 2$
- $\{1, \ldots, n\} = \{1\}$ für $n = 1$
- $\{1, \ldots, n\} = \emptyset$ für $n = 0$.

Ein Neuling hingegen würde $\{1, \ldots, n\}$ für $n = 0$ vielleicht als $\{1, 0\} = \{0, 1\}$ auffassen.

Notation 1.1.7. $\{\mathscr{O} \mid \mathscr{E}\}$ steht für die „Menge aller Objekte $\mathscr{O}$ mit der Eigenschaft $\mathscr{E}$“.

Beispiel 1.1.8. (a) $\{x \mid x \in \mathbb{N}, 1 \leq x \leq n\} = \{1, \ldots, n\}$

(b) $\{x^2 \mid x \in \mathbb{N}\} = \{y \mid \text{es gibt ein } x \in \mathbb{N} \text{ mit } y = x^2\}$ ist die Menge der Quadratzahlen.

(c) $\mathbb{Q} := \left\{\frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N}\right\}$ ist die Menge der *rationalen Zahlen*.

Notation 1.1.9. (a) $x \left\{\begin{matrix}\in \\ \notin\end{matrix}\right\} A$ steht für „x ist $\left\{\begin{matrix}\text{ein} \\ \text{kein}\end{matrix}\right\}$ Element von A“

(b) $\{x \in A \mid \mathscr{E}\} := \{x \mid x \in A, \mathscr{E}\}$ steht für die Menge aller x in/aus A (d.h. für die Elemente x von A) mit der Eigenschaft $\mathscr{E}$.

Beispiel 1.1.10. (a) $\{x \in \mathbb{Z} \mid x^2 \leq 7\} = \{-2, -1, 0, 1, 2\}$

(b) $3 \notin \{x^2 \mid x \in \mathbb{N}\}$

(c) $4 \in \{x^2 \mid x \in \mathbb{N}\}$

(d) $\{2, 3, \{4, 5\}\} \in \{\{1, 2, \{8\}\}, \{2, 3, \{4, 5\}\}\} =: M$
$8 \in \{8\} \in \{1, 2, \{8\}\} \in M$
$8 \notin \{1, 2, \{8\}\}, \{8\} \notin M$

Bemerkung und Notation 1.1.11. Wir formulieren mathematische Aussagen meist in natürlicher Sprache. Manchmal ist es prägnanter, formale Notation zu benutzen:

$\forall x : \mathscr{E}$	„für alle x gilt $\mathscr{E}$“
$\exists x : \mathscr{E}$	„es gibt/existiert ein x mit $\mathscr{E}$“
$\forall x \in A : \mathscr{E}$	„für alle x aus A gilt $\mathscr{E}$“
$\exists x \in A : \mathscr{E}$	„es gibt ein x aus A mit Eigenschaft $\mathscr{E}$“
$\mathscr{E} \iff \mathscr{F}$	„$\mathscr{E}$ $\underbrace{\text{genau dann, wenn}}_{\text{gdw.}}$ $\mathscr{F}$“
	„$\mathscr{E}$ äquivalent $\mathscr{F}$“
$\mathscr{E} \implies \mathscr{F}$	„$\mathscr{E}$ impliziert $\mathscr{F}$“
	„wenn $\mathscr{E}$, dann $\mathscr{F}$“
	„$\mathscr{E}$ ist hinreichend für $\mathscr{F}$“
$\mathscr{E} \impliedby \mathscr{F}$	„$\mathscr{E}$ wird von $\mathscr{F}$ impliziert“
	„$\mathscr{F}$ nur dann, wenn $\mathscr{E}$“
	„$\mathscr{E}$ ist notwendig für $\mathscr{F}$“
$\mathscr{E}$ & $\mathscr{F}$	„$\mathscr{E}$ und $\mathscr{F}$“

Beachte: $\forall x \in \emptyset : \mathscr{E}$ ist immer wahr und $\exists x \in \emptyset : \mathscr{E}$ ist immer falsch.

Definition 1.1.12. Eine Menge A heißt *Teilmenge* (oder *Untermenge*) der Menge B und man schreibt $A \subseteq B$, wenn $\forall x \in A : x \in B$. („$A$ ist in B enthalten“, „B enthält A“).

Man bezeichnet dann B als *Obermenge* von A und schreibt $B \supseteq A$.

Bemerkung 1.1.13. Dass zwei Mengen A und B gleich sind, genau dann, wenn sie dieselben Elemente enthalten, kann man auch so ausdrücken:

$$A = B \iff (A \subseteq B \;\&\; B \subseteq A).$$

Fast immer ist es ratsam, die Gleichheit zweier Mengen zu zeigen, indem man die beiden Inklusionen (Teilmengenbeziehungen) *getrennt* zeigt.

Definition 1.1.14. (a) Ist M eine Menge von Mengen, so ist

$$\bigcup M := \{x \mid \exists A \in M : x \in A\}$$

die *Vereinigungsmenge* von M und für $M \underbrace{\neq}_{\text{„ungleich“}} \emptyset$ ist

$$\bigcap M := \{x \mid \forall A \in M : x \in A\}$$

die *Schnittmenge* von M. Beachte, dass $\bigcap \emptyset$ nicht generell definiert ist wegen 1.1.2. Ist M eine Menge von Teilmengen einer festen Menge A_0, so definiert man oft $\bigcap \emptyset := A_0$, denn dann gilt $\bigcap M = \{x \in A_0 \mid \forall A \in M : x \in A\}$ sowohl für $M \neq \emptyset$ als auch für $M = \emptyset$ (beachte, dass $\forall A \in \emptyset : \ldots$ wahr ist!).

(b) Für $n \in \mathbb{N}$ und Mengen $A_1, \ldots, A_n$ definiert man die *Vereinigung* $A_1 \cup \ldots \cup A_n := \bigcup \{A_1, \ldots, A_n\}$ und den *Schnitt* $A_1 \cap \ldots \cap A_n := \bigcap \{A_1, \ldots, A_n\}$.

(c) Für Mengen A und B heißt $\underbrace{A \backslash B}_{\text{„ohne“}} := \{x \in A \mid x \notin B\}$ die *Mengendifferenz*.

(d) Für jede Menge A nennt man $\mathscr{P}(A) := \{B \mid B \subseteq A\}$ ihre *Potenzmenge*.

Beispiel 1.1.15. $\bigcup \emptyset = \emptyset$, $\bigcap \emptyset$ nicht immer definiert
$\bigcup \{\emptyset\} = \emptyset, \bigcap \{\emptyset\} = \emptyset$
$\bigcup \{\{1,4,6\}, \{\{5\}\}, \emptyset\} = \{1,4,6,\{5\}\}$
$\{1,2,3\} \cup \{3,4,5\} = \{1,2,3,4,5\}$
$\{1,2,3\} \cap \{3,4,5\} = \{3\}$
$\{1,2,3\} \setminus \{3,4,5\} = \{1,2\}$
$\mathscr{P}(\emptyset) = \{\emptyset\}$
$\mathscr{P}(\{1\}) = \{\emptyset, \{1\}\}$
$\mathscr{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$

Definition 1.1.16. Eine *Abbildung* f besteht aus folgenden Angaben:

- einer Menge A, genannt *Definitionsmenge* von f,
- einer Menge B, genannt *Zielmenge* von f und
- einer Vorschrift, die jedem Element a von A genau ein Element $f(a)$ von B (das sogenannte Bild von a unter f) zuordnet.

Notation: $f\colon A \underbrace{\to}_{\text{„nach“}} B, \quad a \underbrace{\mapsto}_{\text{„wird abgebildet auf“}} f(a)$ [Bild: Veranschaulichung mit Pfeilen]
Ist f eine Abbildung mit Definitionsmenge A und Zielmenge B, so sagt man f ist eine *Abbildung von A nach B* und schreibt $f : A \to B$.

Bemerkung 1.1.17. Sind $f : A \to B$ und $g : C \to D$ Abbildungen, so

$$f = g \iff (A = C \,\&\, B = D \,\&\, \forall a \in A : f(a) = g(a)).$$

Beispiel 1.1.18.

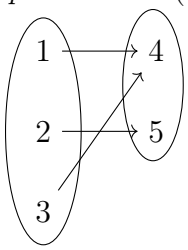
Für $f : \{0,1\} \to \{0,1\}, x \mapsto x$
$g : \{0,1\} \to \{0,1\}, x \mapsto x^2$ und
$h : \{0,1\} \to \{0,1\}, 0 \mapsto 0, 1 \mapsto 1$ gilt
$f = g = h$, aber $f \neq p$ für $p : \{0,1\} \to \{0,1,2\}, 0 \mapsto 0, 1 \mapsto 1$

Definition 1.1.19. Eine Abbildung $f : A \to B$ heißt $\left\{\begin{array}{c} \textit{injektiv} \\ \textit{surjektiv} \\ \textit{bijektiv} \end{array}\right\}$, wenn es zu jedem $b \in B \left\{\begin{array}{c} \textit{höchstens} \\ \textit{mindestens} \\ \textit{genau} \end{array}\right\}$ ein $\underbrace{a \in A}_{\text{„Urbild von } b\text{“}}$ gibt mit $f(a) = b$.

Mit anderen Worten gilt:

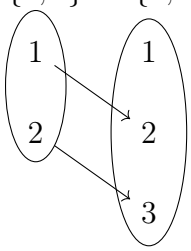
$$\begin{aligned} f \text{ injektiv} &\iff \forall a_1, a_2 \in A : (f(a_1) = f(a_2) \Longrightarrow a_1 = a_2) \\ f \text{ surjektiv} &\iff \forall b \in B : \exists a \in A : f(a) = b \text{ und} \\ f \text{ bijektiv} &\iff (f \text{ injektiv } \& \ f \text{ surjektiv}) \end{aligned}$$

Beispiel 1.1.20. (a) $\{1,2,3\} \to \{4,5\}, 1 \mapsto 4, 2 \mapsto 5, 3 \mapsto 4$



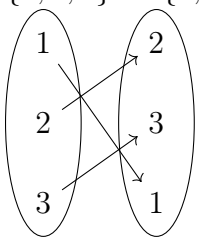
nicht injektiv, surjektiv, nicht bijektiv

(b) $\{1,2\} \to \{1,2,3\}, 1 \mapsto 2, 2 \mapsto 3$



injektiv, nicht surjektiv, nicht bijektiv

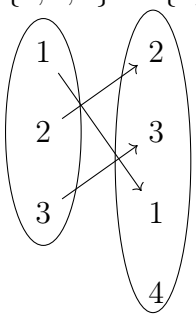
(c) $\{1,2,3\} \to \{1,2,3\}, x \mapsto x$



injektiv, surjektiv, bijektiv

(d) $\emptyset \to \emptyset$ injektiv, surjektiv, bijektiv [Bild: Zwei leere Kreise]

(e) $\{1,2,3\} \to \{1,2,3,4\}, x \mapsto x$



injektiv, nicht surjektiv, nicht bijektiv.

(f) $\mathbb{Z} \to \mathbb{N}_0, x \mapsto |x|$ nicht injektiv, surjektiv, nicht bijektiv

(g) $\mathbb{Z} \to \mathbb{Z}, x \mapsto |x|$ nicht injektiv, nicht surjektiv, nicht bijektiv

(h) $\mathbb{Z} \to \mathbb{Z}, x \mapsto -x$ injektiv, surjektiv, bijektiv

(i) ~~$\{1,2,3\} \to \{4,5\}, 2 \mapsto 4, 3 \mapsto 5$~~ keine Abbildung!

(j) ~~$\{1,2,3\} \to \{4,5\}, 1 \mapsto 4, 1 \mapsto 5, 2 \mapsto 4, 3 \mapsto 5$~~ keine Abbildung!

(k) $\mathbb{N} \to \mathbb{N}, x \mapsto x+1$ injektiv, nicht surjektiv, nicht bijektiv

Definition 1.1.21. Eine Menge A heißt *endlich*, wenn sie nur endlich viele Elemente hat. Die Anzahl ihrer Elemente nennt man dann *Mächtigkeit* (auch *Kardinalität*) $\#A$ von A. Ist A *unendlich* (d.h. nicht endlich), so setzen wir

$$\#A = \underbrace{\infty}_{\text{„unendlich“}}.$$

Wir nennen A *abzählbar* unendlich, wenn es eine bijektive Abbildung $f : \mathbb{N} \to A$ gibt, und *überabzählbar*, wenn A weder endlich noch abzählbar unendlich ist.

Satz 1.1.22 (Satz von Cantor (1891))**.** *Ist A eine Menge, so gibt es keine surjektive Abbildung von A nach $\mathscr{P}(A)$.*

Beweis. Zu zeigen ist, dass keine Abbildung von A nach $\mathscr{P}(A)$ surjektiv ist. Sei hierzu $f : A \to \mathscr{P}(A)$ eine (beliebige, aber feste) Abbildung. Wir setzen $B := \{a \in A \mid a \notin f(a)\}$ und zeigen, dass es kein $a \in A$ gibt mit $f(a) = B$ (denn dann ist insbesondere f nicht surjektiv). Dies zeigen wir durch Widerspruch: Wir nehmen an, wir haben $a \in A$ mit $f(a) = B$ und führen dies zu einem logischen Widerspruch.

Fall 1: $a \in f(a)$.

> Damit ist einerseits $a \notin B$ nach Definition von B und andererseits $a \in B$ wegen $f(a) = b$. $\lightning_{\text{„Widerspruch“}}$

Fall 2: $a \notin f(a)$.

> Dann einerseits $a \in B$ nach Definition von B und andererseits $a \notin B$ wegen $f(a) = B$. $\lightning$

□
„quod erat demonstrandum“

Veranschaulichung im Fall von $A = \mathbb{N}$:

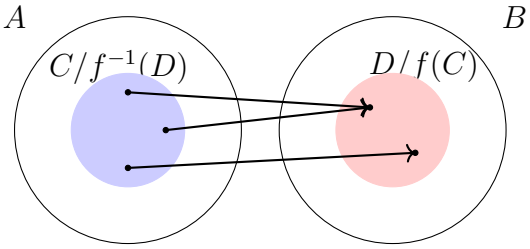
$f(1) = \{\ \underline{1},\ 3,\ 4,\ 7,\ 9,\ 10,\ \ldots\}$
$f(2) = \{\ \underline{2},\ 4,\ 5,\ 8,\ 10,\ \ldots\}$
$f(3) = \{\ 2,\ \_,\ 4,\ 6,\ 7,\ 8,\ 9,\ \ldots\}$
$\_ = \mathbb{N} \setminus B$ „Cantors Diagonalargument“

Bemerkung 1.1.23. Für endliche Mengen A folgt 1.1.22 auch aus $\#\mathscr{P}(A) = 2^{\#A} > \#A$.

Korollar 1.1.24. *$\mathscr{P}(\mathbb{N})$ ist überabzählbar.*

Beweis. Offenbar ist $\mathscr{P}(\mathbb{N})$ nicht endlich. Wäre $\mathscr{P}(\mathbb{N})$ *abzählbar unendlich*, so gäbe es gemäß Definition 1.1.21 eine bijektive Abbildung $\mathbb{N} \to \mathscr{P}(\mathbb{N})$. Dies ist nach dem Satz von Cantor 1.1.22 unmöglich. □

Definition 1.1.25. Ist $f : A \to B$ eine Abbildung, $C \subseteq A$ und $D \subseteq B$, so nennt man $f(C) := \{f(a) \mid a \in C\}$ *das Bild von C unter f* und $f^{-1}(D) := \{a \in A \mid f(a) \in D\}$ das *Urbild von D unter f*.



Beispiel 1.1.26. Für $f : \mathbb{Z} \to \mathbb{Z}, x \mapsto |x|$ gilt:

$$f(\{-3,-5,4\}) = \{3,4,5\} \text{ und}$$
$$f^{-1}(\{-3,-5,4\}) = \{-4,4\}\,.$$

Definition 1.1.27. Seien A und B Mengen. Dann bezeichnet $B^A := \{f \mid f : A \to B\}$ die Menge aller Abbildungen von A nach B. Für $n \in \mathbb{N}_0$ schreibt man oft B^n statt $B^{\{1,\ldots,n\}}$ und $(b_1,\ldots,b_n)$ statt $\{1,\ldots,n\} \to B, 1 \mapsto b_1, \ldots, n \mapsto b_n$ („n-Tupel“).
Insbesondere ist B^0 einelementig: $B^0 = \{\underbrace{()}_{\text{„leeres Tupel“}}\}$.

Definition 1.1.28. Seien A und B_a für $a \in A$ Mengen.
Setze $B := \bigcup\{B_a \mid a \in A\}$. Dann nennt man

$$\prod_{a\in A} B_a := \{f \mid f : A \to B, \forall a \in A : f(a) \in B_a\}$$

das *kartesische* [René Descartes, *1596, † 1650] *Produkt* der B_a $(a \in A)$. Für $n \in \mathbb{N}_0$ schreibt man oft $B_1 \times \ldots \times B_n$ statt $\prod_{a\in\{1,\ldots,n\}} B_a$ und $(b_1,\ldots,b_n)$ statt

$$\{1,\ldots,n\} \to B, 1 \mapsto b_1, \ldots, n \mapsto b_n.$$

Sprechweise 1.1.29. (a) Eine Abbildung, deren Definitions- und Zielmenge übereinstimmen, nennt man *Selbstabbildung*.

(b) Eine bijektive Selbstabbildung nennt man auch *Permutation*.

(c) Synonym sind jeweils:

Abbildung	$\underbrace{\text{Funktion}}_{\substack{\text{Zielmenge wenig abstrakt,}\\ \text{meist bestehend aus Zahlen}}}$	Mengenhomomorphismus
injektive Abbildung	Injektion	Mengeneinbetting/ Mengenmonomorphismus
surjektive Abbildung	Surjektion	Mengenepimorphismus
bijektive Abbildung	Bijektion	Mengenisomorphismus
Selbstabbildung		Mengenendomorphismus
bijektive Selbstabbildung	Permutation	Mengenautomorphismus
Definitionsmenge	Definitionsbereich	Quellbereich
Zielmenge	Wertevorrat	

(d) Ist $f : A \to B$ eine Abbildung, so nennt man das Bild $f(A)$ von A unter f auch das *Bild von* f. Es gilt f surjektiv $\iff f(A) = B$. Manche Leute nennen $f(A)$ die Wertemenge oder den Wertebereich von f, andere nennen B so. Daher vermeiden wir diese beiden Begriffe.

Bemerkung 1.1.30. (a) Sind $f : A \to B$ und $g : A \to C$ Abbildungen, so kann $f = g$ nur gelten, wenn $B = C$. In der Praxis wird aber dann in der Literatur mit $f = g$ oft nur $\forall a \in A : f(a) = g(a)$ gemeint (d.h. es ist gemeint $f_0 = g_0$, wobei $f_0 : A \to B \cap C, a \mapsto f(a)$ und $g_0 : A \to B \cap C, a \mapsto g(a)$).

(b) Wenn im Fall $A = \{1, \ldots, n\}$ die Abbildungen f und g aus (a) wie in 1.1.27 als Tupel geschrieben werden, dann wird diese Praxis immer angewandt, da die Zielmengen B und C in Tupelschreibweise ja gar nicht mehr spezifiziert sind. Es gilt also stets $(b_1, \ldots, b_n) = (c_1, \ldots, c_n) \iff (b_1 = c_1 \,\&\, \ldots \,\&\, b_n = c_n)$.

(c) Bemerkung (b) gilt auch für folgende *Varianten der Verallgemeinerungen* der Tupelschreibweise:
Matrizen:

$$f : \underbrace{\{1, \ldots, m\} \times \{1, \ldots, n\}}_{=\{(1,1),(1,2),\ldots,(1,n),\ldots,(m,1),\ldots,(m,n)\}} \to Z$$

$$\begin{pmatrix} f(1,1) & \ldots & f(1,n) \\ f(2,1) & \ldots & f(2,n) \\ \vdots & & \vdots \\ f(m,1) & \ldots & f(m,n) \end{pmatrix}$$

Folgen:

$$f : \mathbb{N} \to \mathbb{Z} \qquad (f(1), f(2), f(3), \ldots)$$

Familien:

$$f : \underbrace{I}_{\text{„Indexmenge“}} \to Z \qquad (f(a))_{a \in I}$$

(manchmal auch $\{f(a)\}_{a \in A} \rightsquigarrow$ schlecht wegen Verwechslungsgefahr mit der Menge $\{f(a) \mid a \in A\}$)

Definition 1.1.31. Sei $f : A \to B$ eine Abbildung und $C \subseteq A$. Dann heißt

$$f|_C : C \to B, a \mapsto f(a)$$

die *Einschränkung* (oder *Restriktion*) von f auf C.

Notation 1.1.32 (Diagramme)**.** Statt $f : A \to B$ schreibt man auch $A \xrightarrow{f} B$. Zum Beispiel steht „Gelte $A \xrightarrow{f} B \xrightarrow{g} C$“ für „Seien $f : A \to B$ und $g : B \to C$ Abbildungen“.

Definition 1.1.33. Für $f : A \to B$ heißt

$$\Gamma_f := \{(x, f(x)) \mid x \in A\} \subseteq A \times B$$

der *Graph* von f. Aus Γ_f kann man die Definitionsmenge und die Abbildungsvorschrift [→ 1.1.16] und auch das Bild [→ 1.1.29 (d)], nicht aber die Zielmenge von f zurückgewinnen.

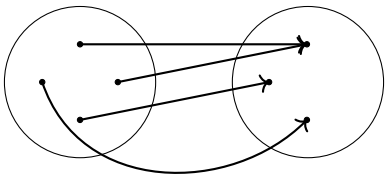
$$A = \{a \mid \exists b : (a, b) \in \Gamma_f\}$$

$$a \mapsto b \text{ falls } (a, b) \in \Gamma_f$$

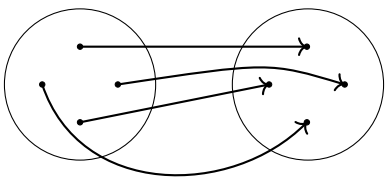
$$f(A) = \{b \mid \exists a : (a, b) \in \Gamma_f\} \subseteq B$$

§1.2 Hintereinanderschaltung und Umkehrung von Abbildungen

Erinnerung 1.2.1. Eine Abbildung $f : A \to B$ ordnet jedem $a \in A$ genau ein $b \in B$ zu.



Jedes $a \in A$ hat also genau ein Bild unter f [→ 1.1.16].
f heißt bijektiv, wenn zusätzlich jedes $b \in B$ genau ein Urbild unter f hat.



Vertauschen von Bild und Urbild („Umdrehen der Pfeile“) liefert für bijektives f eine Umkehrabbildung.

Definition 1.2.2. Für bijektives $f : A \to B$ definieren wir die *Umkehrabbildung* von f (oder zu f inverse Abbildung)

$$f^{-1} : B \to A, b \mapsto \text{das eindeutige } a \text{ mit } f(a) = b.$$

Bemerkung 1.2.3. Während f^{-1} nur für bijektive f existiert, war $f^{-1}(C)$ in 1.1.25 für jedes $f : A \to B$ und jedes $C \subseteq B$ definiert als $f^{-1}(C) = \{a \in A \mid f(a) \in C\}$. Ist $f : A \to B$ bijektiv und $C \subseteq B$, so notieren wir mit $f^{-1}(C)$ sowohl das Urbild von C unter f als auch das Bild von C unter f^{-1}, was aber konsistent ist, denn die beiden sind gleich.

Am 2. November sind wir bis hierher gekommen.

Definition 1.2.4. Für $A \xrightarrow{f} B \xrightarrow{g} C$ heißt

$$g \circ f : A \to C, a \mapsto g(f(a))$$

die *Hintereinander* $\left\{\begin{matrix} \text{schaltung} \\ \text{ausführung} \end{matrix}\right\}$ (auch *Verkettung* oder *Komposition*) von f und g. Für jede Menge A heißt

$$\mathrm{id}_A : A \to A, a \mapsto a$$

die *Identität* (oder *identische* Abbildung) auf A.

Proposition 1.2.5. (a) *Für* $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$ *gilt*

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

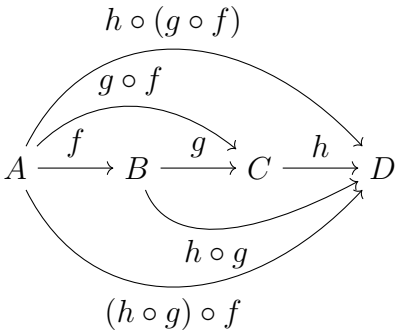
(„$\circ$ ist assoziativ“).

(b) *Für* $f : A \to B$ *gilt* $f \circ \mathrm{id}_A = f = \mathrm{id}_B \circ f$.

(c) *Für bijektive* $f : A \to B$ *gilt* $f^{-1} \circ f = \mathrm{id}_A$ *und* $f \circ f^{-1} = \mathrm{id}_B$

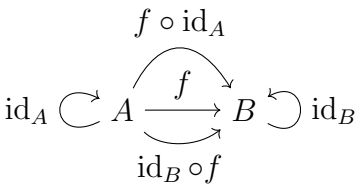
(d) *Für bijektive* $f : A \to B$ *ist auch* f^{-1} *bijektiv und es gilt* $(f^{-1})^{-1} = f$.

Beweis. (a) Gelte $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$. Dann



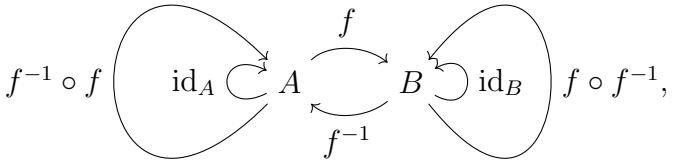
und$(h \circ (g \circ f))(a) = h((g \circ f)(a)) = h(g(f(a))) = (h \circ g)(f(a)) = ((h \circ g) \circ f)(a)$ für alle $a \in A$. Nach 1.1.16 gilt also $h \circ (g \circ f) = (h \circ g) \circ f$.

(b) Gelte $A \xrightarrow{f} B$. Dann



und $(f \circ \mathrm{id}_A)(a) = f(\mathrm{id}_A(a)) = f(a) = \mathrm{id}_B(f(a)) = (\mathrm{id}_B \circ f)(a)$ für alle $a \in A$. Nach 1.1.16 gilt also $f \circ \mathrm{id}_A = f = \mathrm{id}_B \circ f$.

(c) Sei $f : A \to B$ bijektiv. Dann



$$(f^{-1} \circ f)(a) = f^{-1}(f(a)) \overset{\substack{f(a)=f(a) \\ =\\ 1.2.2}}{} a \text{ für alle } a \in A \text{ und}$$
$$(f \circ f^{-1})(b) = f(f^{-1}(b)) \overset{1.2.2}{=} b \text{ für alle } b \in B.$$

Nach 1.1.16 gilt also $f^{-1} \circ f = \mathrm{id}_A$ und $f \circ f^{-1} = \mathrm{id}_B$.

(d) Sei $f : A \to B$ bijektiv. Dann ist auch $f^{-1} : B \to A$ bijektiv, denn ist $a \in A$, so $\{b \in B \mid f^{-1}(b) = a\} \overset{1.2.2}{=} \{b \in B \mid f(a) = b\} = \{f(a)\}$, d.h. jedes Element von A hat genau ein Urbild unter f^{-1}. Weiter gilt $(f^{-1})^{-1} : A \to B$ und

$$(f^{-1})^{-1}(a) \overset{\substack{f^{-1}(f(a)) \overset{(c)}{=} a \\ =\\ 1.2.2}}{} f(a) \text{ für alle } a \in A.$$

Nach 1.1.16 gilt also $f = (f^{-1})^{-1}$.

□

Bis hierher hätten wir am 30. Oktober kommen sollen.

Satz 1.2.6. *Seien $f : A \to B$ und $g : B \to A$ Abbildungen mit $g \circ f = \mathrm{id}_A$ und $f \circ g = \mathrm{id}_B$. Dann sind f und g bijektiv und es gilt $g = f^{-1}$ und $f = g^{-1}$.*

„f und g sind invers zueinander."

Beweis. Es ist f injektiv, denn sind $a_1, a_2 \in A$ mit $f(a_1) = f(a_2)$, so gilt

$$a_1 = \mathrm{id}_A(a_1) = (g \circ f)(a_1) = g(f(a_1)) = g(f(a_2)) = (g \circ f)(a_2) = \mathrm{id}_A(a_2) = a_2.$$

Es ist f auch surjektiv, denn ist $b \in B$, so gilt für $a := g(b)$, dass

$$f(a) = f(g(b)) = (f \circ g)(b) = \mathrm{id}_B(b) = b$$

Also ist f bijektiv.
Analog zeigt man, dass g bijektiv ist. Aus $(g \circ f) = \mathrm{id}_A$ folgt

$$g = g \circ \mathrm{id}_B \overset{1.2.5(c)}{=} g \circ (f \circ f^{-1}) \overset{1.2.5\,(a)}{=} (g \circ f) \circ f^{-1} \overset{1.2.5\,(c)}{=} \mathrm{id}_A \circ f^{-1} \overset{1.2.5\,(b)}{=} f^{-1}$$

Analog folgt $f = g^{-1}$. □

Sprechweise und Notation 1.2.7. Die Situation von 1.2.6 drücken wir sprachlich oft so aus: „Die Zuordnungen

$$\begin{aligned} a &\mapsto f(a) \\ g(b) &\mapsfrom b \end{aligned}$$

vermitteln eine Bijektion zwischen A und B."
In Zeichen:

$$\begin{aligned} A &\leftrightarrow B \\ a &\mapsto f(a) \\ g(b) &\mapsfrom b \end{aligned}$$

§1.3 Äquivalenzrelationen und Zerlegungen

Idee: Grobe Sichtweise auf eine Menge einnehmen.

Definition 1.3.1. Sei A eine Menge.

(a) Eine (zweistellige) *Relation* auf A ist eine Teilmenge von $A \times A$. Ist R eine Relation auf A, so schreibt man auch aRb statt $(a, b) \in R$.

(b) Eine *Äquivalenzrelation* auf A ist eine Relation $\sim$ auf A, für die gilt:

- $\forall a \in A : a \sim a$ „reflexiv"
- $\forall a, b \in A : (a \sim b \implies b \sim a)$ „symmetrisch"
- $\forall a, b, c \in A : ((a \sim b \ \&\ b \sim c) \implies a \sim c)$ „transitiv"

Ist $\sim$ eine Äquivalenzrelation auf A und $a \in A$, so heißt $\widetilde{a} := \{b \in A \mid a \sim b\}$ die *Äquivalenzklasse* von a bezüglich $\sim$.

Beispiel 1.3.2. Sei A eine Menge.

(a) Durch
$$a \sim b :\iff a = b \qquad (a, b \in A)$$
(das heißt durch $\sim := \{(a, b) \in A \times A \mid a = b\}$) ist eine Äquivalenzrelation definiert, deren Äquivalenzklassen alle einelementig sind („keine Vergröberung“).

(b) Durch $a \sim b$ für alle $a, b \in A$ (das heißt durch $\sim := A \times A$) ist eine Äquivalenrelation definiert, die nur eine Äquivalenzklasse besitzt („totale Vergröberung“).

Definition 1.3.3. Sei A eine Menge. Eine Menge $\mathscr{Z} \subseteq \mathscr{P}(A) \setminus \{\emptyset\}$ heißt *Zerlegung* von A, wenn $\bigcup \mathscr{Z} = A$ und $\forall B, C \in \mathscr{Z} : (B = C \text{ oder } B \cap C = \emptyset)$. Mit anderen Worten: Eine Zerlegung von A ist eine Menge von nichtleeren paarweise disjunkten Teilmengen von A, deren Vereinigung ganz A ist.

Beispiel 1.3.4. Sei A eine Menge.

(a) $\{\{a\} \mid a \in A\}$ ist eine Zerlegung von A („keine Vergröberung“).

(b) $\{A\}$ ist eine Zerlegung von A („totale Vergröberung“).

Definition 1.3.5. Sei A eine Menge.

(a) Zu jeder Äquivalenzrelation $\sim$ auf A definieren wir die zugehörige *Quotientenmenge*
$$A \underbrace{/}_{\text{„modulo“}} \sim$$
als die Menge der Äquivalenzklassen von $\sim$:
$$A/\sim := \{\widetilde{a} \mid a \in A\}$$

(b) Zu jeder Zerlegung $\mathscr{Z}$ von A definieren wir eine Relation $\sim_{\mathscr{Z}}$ auf A durch
$$a \sim_{\mathscr{Z}} b :\iff \exists Z \in \mathscr{Z} : \{a, b\} \subseteq Z$$

Satz 1.3.6. [→1.2.7] *Sei A eine Menge. Die Zuordnungen*
$$\sim \mapsto A/\sim$$
$$\sim_{\mathscr{Z}} \mapsfrom \mathscr{Z}$$
vermitteln eine Bijektion zwischen der Menge der Äquivalenzrelationen auf A und der Menge der Zerlegungen von A.

Beweis. Zu zeigen ist:

(a) Ist $\sim$ eine Äquivalenzrelation auf A, so ist $A/\sim$ eine Zerlegung von A.

(b) Ist $\mathscr{Z}$ eine Zerlegung von A, so ist $\sim_{\mathscr{Z}}$ eine Äquivalenzrelation auf A.

(c) Ist $\sim$ eine Äquivalenzrelation auf A, so ist $\sim_{A/\sim} = \sim$.

(d) Ist $\mathscr{Z}$ eine Zerlegung von A, so ist $A/\sim_{\mathscr{Z}} = \mathscr{Z}$.

Zu (a). Sei $\sim$ eine Äquivalenzrelation auf A. Zu zeigen ist:

(1) $A/\sim \subseteq \mathscr{P}(A) \setminus \{\emptyset\}$

(2) $\bigcup(A/\sim) = A$

(3) $\forall B, C \in A/\sim : (B = C \text{ oder } B \cap C = \emptyset)$

Zu (1). Sei $a \in A$. Zu zeigen ist $\widetilde{a} \in \mathscr{P}(A) \setminus \{\emptyset\}$, das heißt $\widetilde{a} \subseteq A$ und $\widetilde{a} \neq \emptyset$. Ersteres ist klar nach Definition von $\widetilde{a}$ und letzteres folgt aus $a \sim a$, denn das heißt $a \in \widetilde{a}$.

Zu (2). Es gilt $\bigcup(A/\sim) \overset{1.1.14}{=} \{a \mid \exists B \in A/\sim : a \in B\} \overset{1.3.5(a)}{=} \{a \mid \exists b \in A : a \in \widetilde{b}\}$. Wir zeigen nun die behauptete Gleichheit, indem wir beide Inklusionen getrennt zeigen:
„$\subseteq$“ Gelte $a \in \bigcup(A/\sim)$. Wähle $b \in A$ mit $a \in \widetilde{b}$. Dann $a \in \widetilde{b} \subseteq A$, also $a \in A$.
„$\supseteq$“ Gelte $a \in A$. Dann $a \in \widetilde{a}$, also $a \in \bigcup(A/\sim)$.

Zu (3). Seien $a, b \in A$. Zu zeigen: $\widetilde{a} = \widetilde{b}$ oder $\widetilde{a} \cap \widetilde{b} = \emptyset$. Gelte $\widetilde{a} \cap \widetilde{b} \neq \emptyset$. Zu zeigen ist dann $\widetilde{a} = \widetilde{b}$. Wähle $c \in \widetilde{a} \cap \widetilde{b}$. Dann $a \sim c \sim b$ und daher auch $a \sim b$. Wir zeigen nun $\widetilde{a} \subseteq \widetilde{b}$ (die andere Inklusion geht analog): Gelte $d \in \widetilde{a}$. Dann $d \sim a \sim b$, also $d \sim b$, das heißt $d \in \widetilde{b}$.

Zu (b). Sei $\mathscr{Z}$ eine Zerlegung von A. Zu zeigen ist:

(1) $\forall a \in A : a \sim_{\mathscr{Z}} a$

(2) $\forall a, b \in A : (a \sim_{\mathscr{Z}} b \implies b \sim_{\mathscr{Z}} a)$

(3) $\forall a, b, c \in A : ((a \sim_{\mathscr{Z}} b \,\&\, b \sim_{\mathscr{Z}} c) \implies a \sim_{\mathscr{Z}} c)$

Zu (1). Sei $a \in A$. Zu zeigen ist $\exists Z \in \mathscr{Z} : \{a, a\} \subseteq A$. Mit anderen Worten ist

$$\exists Z \in \mathscr{Z} : a \in A$$

zu zeigen. Dies ist aber klar, da $a \in A = \bigcup \mathscr{Z}$.
(2) ist klar nach Definition von $\sim_{\mathscr{Z}}$, da $\{a, b\} = \{b, a\}$ für alle a und b.
Zu (3). Seien $a, b, c \in A$ mit $a \sim_{\mathscr{Z}} b$ und $b \sim_{\mathscr{Z}} c$. Zu zeigen ist $a \sim_{\mathscr{Z}} c$. Wähle $Z_1, Z_2 \in \mathscr{Z}$ mit $\{a, b\} \in Z_1$ und $\{b, c\} \in Z_2$. Nun gilt $b \in Z_1 \cap Z_2$, also $Z_1 \cap Z_2 \neq \emptyset$. Nach Definition 1.3.3 folgt $Z_1 = Z_2$, also $\{a, c\} \subseteq Z_1 \cup Z_2 = Z_1 \in \mathscr{Z}$. Also $a \sim_{\mathscr{Z}} c$.

Zu (c). Seien $a, b \in A$. Zu zeigen ist $a \sim_{A/\sim} b \iff a \sim b$. Es gilt

$$\begin{aligned} a \sim_{A/\sim} b &\iff \exists Z \in A/\sim : \{a, b\} \subseteq Z \\ &\iff \exists c \in A : \{a, b\} \subseteq \widetilde{c} \\ &\iff \exists c \in A : (a \sim c \sim b) \\ &\iff a \sim b, \end{aligned}$$

wobei man für den Teil „ $\Longrightarrow$ “ der letzten Äquivalenz die Transitivität von $\sim$ benutzt und für den Teil „$\Longleftarrow$“ dieser Äquivalenz $c := a$ setzt.

Zu (d). Sei $\mathscr{Z}$ eine Zerlegung von A. Zu zeigen ist:

(1) $A/{\sim_{\mathscr{Z}}} \subseteq \mathscr{Z}$

(2) $\mathscr{Z} \subseteq A/{\sim_{\mathscr{Z}}}$

Zu (1). Sei $a \in A$. Zu zeigen ist $\widetilde{a}^{\mathscr{Z}} \in \mathscr{Z}$. Es gilt

$$\widetilde{a}^{\mathscr{Z}} = \{b \in A \mid a \sim_{\mathscr{Z}} b\} = \{b \in A \mid \exists Z \in \mathscr{Z} : \{a, b\} \subseteq Z\}.$$

Wähle $Z_0 \in \mathscr{Z}$ mit $a \in Z_0$ (dies geht, da $a \in A = \bigcup \mathscr{Z}$). Es reicht nun zu zeigen, dass

$$\{b \in A \mid \exists Z \in \mathscr{Z} : \{a, b\} \subseteq Z\} = Z_0.$$

„$\subseteq$“ Sei $b \in A$ und $Z \in \mathscr{Z}$ mit $\{a, b\} \subseteq Z$. Zu zeigen: $b \in Z_0$. Es gilt $a \in Z \cap Z_0$. Daher $Z \cap Z_0 \neq \emptyset$ und daher $Z = Z_0$. Also $b \in Z_0$ wie gewünscht.

„$\supseteq$“ Sei $b \in Z_0$. Dann gilt $\{a, b\} \subseteq Z_0 \in \mathscr{Z}$.

Zu (2). Sei $Z \in \mathscr{Z}$. Zu zeigen ist $\exists a \in A : Z = \widetilde{a}^{\mathscr{Z}}$. Wähle $a \in Z$ fest (das geht, da $Z \neq \emptyset$). Wir behaupten nun $Z = \widetilde{a}^{\mathscr{Z}}$.

„$\subseteq$“ Sei $b \in Z$. Zu zeigen ist $a \sim_{\mathscr{Z}} b$. Dies ist klar, da $\{a, b\} \subseteq Z \in \mathscr{Z}$.

„$\supseteq$“ Sei $b \in \widetilde{a}^{\mathscr{Z}}$, das heißt $b \sim_{\mathscr{Z}} a$. Also $\{a, b\} \subseteq Z'$ für ein $Z' \in \mathscr{Z}$. Zu zeigen ist $b \in Z$. Nun gilt $a \in Z \cap Z'$ und damit $Z = Z'$. Somit $b \in \{a, b\} \subseteq Z$.

□

Beispiel 1.3.7. Unter der Bijektion aus obigem Satz 1.3.6 entsprechen sich die Äquivalenzrelation $\sim$ auf $\mathbb{Z}$ definiert durch

$$a \sim b :\Longleftrightarrow a - b \text{ ist gerade Zahl} \qquad (a, b \in \mathbb{Z})$$

und die Zerlegung

$$\{\{n \in \mathbb{Z} \mid n \text{ gerade}\}, \{n \in \mathbb{Z} \mid n \text{ ungerade}\}\}.$$

Satz 1.3.8 (Homomorphiesatz für Mengen)**.** *Sei $\sim$ ein Äquivalenzrelation auf A und $f\colon A \to B$ eine Abbildung derart, dass*

$$a_1 \sim a_2 \implies f(a_1) = f(a_2)$$

für alle $a_1, a_2 \in A$.

(a) *Es gibt genau eine Abbildung* $\overline{f}\colon A/\!\sim \to B$ *mit*

$$\overline{f}(\widetilde{a}) = f(a)$$

für alle $a \in A$.

(b) $\overline{f}$ *ist injektiv* $\iff \forall a_1, a_2 \in A : (a_1 \sim a_2 \iff f(a_1) = f(a_2))$

(c) $\overline{f}$ *ist surjektiv* $\iff$ f *ist surjektiv.*

Beweis. **(a)** Klar ist, dass es höchstens eine solche Abbildung gibt, denn die Bedingung $\overline{f}(\widetilde{a}) = f(a)$ legt in eindeutiger Weise fest, was das Bild von $\widetilde{a}$ unter $\overline{f}$ sein soll (nämlich $f(a)$) und es gilt $A/\!\sim = \{\widetilde{a} \mid a \in A\}$.

Zu zeigen bleibt, dass jedem $\widetilde{a}$ nur ein Bild zugeordnet wird. Man nennt dies die *Wohldefiniertheit* von $\overline{f}$. Man muss dazu prüfen, dass für $a_1, a_2 \in A$ gilt:

$$\widetilde{a_1} = \widetilde{a_2} \implies f(a_1) = f(a_2).$$

Dies entspricht genau der vorausgesetzten Bedingung.

(c) Offensichtlich haben f und $\overline{f}$ dieselbe Zielmenge und dasselbe Bild. Benutze nun 1.1.29(d).

(b) Zieht man die Voraussetzung an f in Betracht, dann ist zu zeigen

$$\overline{f} \text{ injektiv} \iff \forall a_1, a_2 \in A : (f(a_1) = f(a_2) \implies a_1 \sim a_2).$$

Dies kann man aber umschreiben zu

$$\overline{f} \text{ injektiv} \iff \forall a_1, a_2 \in A : (\overline{f}(\widetilde{a_1}) = \overline{f}(\widetilde{a_2}) \implies \widetilde{a_1} = \widetilde{a_2}),$$

was nach Definition 1.1.19 gilt. □

[Zeichne Bild!]

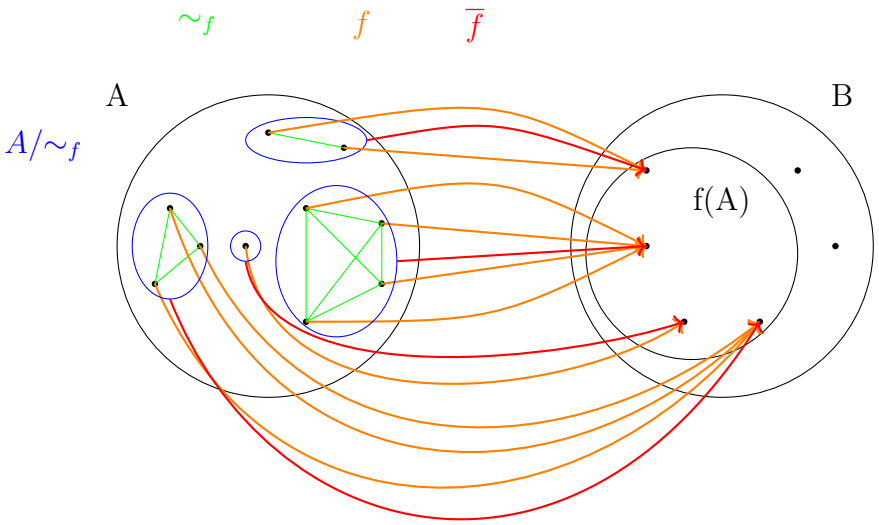
Definition und Proposition 1.3.9. Sei $f\colon A \to B$ eine Abbildung. Dann wird durch

$$a_1 \sim_f a_2 :\iff f(a_1) = f(a_2) \qquad (a_1, a_2 \in A)$$

eine Äquivalenzrelation $\sim_f$ auf A definiert, die wir die durch f *induzierte Äquivalenzrelation* nennen.

Bemerkung 1.3.10. Sei $\sim$ eine Äquivalenzrelation auf der Menge A. Dann wird $\sim$ durch eine Abbildung $f\colon A \to B$ in eine weitere Menge B induziert, nämlich durch die *kanonische Surjektion* $f\colon A \to A/\!\sim,\ a \mapsto \widetilde{a}$ (in der Tat: $a \sim b \iff \widetilde{a} = \widetilde{b} \iff f(a) = f(b)$ für alle $a, b \in A$).

Korollar 1.3.11 (Isomorphiesatz für Mengen)**.** *Sei* $f\colon A \to B$ *eine Abbildung. Dann ist* $\overline{f}\colon A/\!\sim_f \to f(A)$ *definiert durch* $\overline{f}(\widetilde{a}) = f(a)$ *für* $a \in A$ *eine Bijektion.*


$\sim_f$
f
$\overline{f}$
A
B
$A/\sim_f$
f(A)

§2 Abelsche Gruppen

[Niels Henrik Abel *1802, †1829, Abelpreis seit 2003]

§2.1 Definition und Beispiele abelscher Gruppen

Definition 2.1.1. Eine *abelsche Gruppe* ist ein geordnetes Paar (d.h. 2-Tupel) $(G,+)$, wobei G eine Menge ist und $+ : G \times G \to G$ eine Abbildung (meist infix geschrieben, d.h. man schreibt $a+b$ statt $+(a,b)$) mit folgenden Eigenschaften:

(K) $\forall a,b \in G : a+b = b+a$ „kommutativ“

(A) $\forall a,b,c \in G : a+(b+c) = (a+b)+c$ „assoziativ“

(N) $\exists e \in G : \forall a \in G : a+e = a$ „neutrales Element“

Anmerkung: sind $e, e' \in G$ neutral, d.h. $\forall a \in G : a+e = a = a+e'$, so gilt $e = e'$, denn es gilt $e' = e'+e \overset{(K)}{=} e+e' = e$. Daher gibt es *genau* ein $e \in G$ mit $\forall e \in G : a+e = a$ und man bezeichnet dieses e als das *neutrale Element* der Gruppe und schreibt dafür 0 statt e.

(I) $\forall a \in G : \exists b \in G : a+b = 0$ „inverse Elemente“

Bemerkung 2.1.2. (a) Ist $(G,+)$ eine abelsche Gruppe, so nennt man G die *zugrundeliegende* (oder *Träger-*)*Menge* und $+$ die (Gruppen-)*Addition* von $(G,+)$.

(b) Sei $(G,+)$ eine abelsche Gruppe und $a \in A$. Seien b, b' invers zu a, d.h. $a+b = 0 = a+b'$. Dann gilt $b = b'$, denn es gilt

$$b \overset{(N)}{=} b+0 = b+(a+b') \overset{(A)}{=} (b+a)+b' \overset{(K)}{=} (a+b)+b' = 0+b' \overset{(K)}{=} b'+0 \overset{(N)}{=} b'$$

Daher ist zu jedem $a \in G$ das dazu inverse Element eindeutig bestimmt und wir führen die Abbildung

$$- : G \to G, a \mapsto b \text{ falls } a+b = 0$$

ein. Statt $-(a)$ schreibt man oft $-a$ und statt $a+(-b)$ schreibt man oft $a-b$.

(c) (N) und (I) kann man nun wie folgt schreiben:

(N) $\forall a \in G : a + 0 = a$

(I) $\forall a \in G : a + (-a) = 0$

(d) Statt $+$ kann man natürlich auch andere Symbole benutzen. Zur gleichzeitigen Betrachtung mehrerer Gruppen schreibt man manchmal $(G, +_G)$, $(H, +_H)$, usw. und dann entsprechend $0_G, 0_H, -_G, -_H$. Da aus dem Kontext oft klar ist, ob $+_G$ oder $+_H$ gemeint ist, schreibt man oft schlampig $+$ sowohl für $+_G$ als auch für $+_H$. Manchmal schreibt man auch $\left\{\begin{matrix} ab = a \cdot b \\ 1 \\ a^{-1} \end{matrix}\right\}$ statt $\left\{\begin{matrix} a + b \\ 0 \\ -a \end{matrix}\right\}$, z.B. sind $(\mathbb{R} \setminus \{0\}, \cdot), (\mathbb{R}_{>0}, \cdot)$ und $(\{-1, 1\}, \cdot)$ jeweils mit der Multiplikation reeller Zahlen aus der jeweils zugrundeliegenden Menge abelsche Gruppen.

(e) Statt „$(G, +)$ ist abelsche Gruppe“ schreibt man oft auch „G ist additiv geschriebene abelsche Gruppe“ oder nur „G ist abelsche Gruppe“ (obwohl G nur die Trägermenge (siehe (a)) einer abelschen Gruppe ist). Statt „$(G, \cdot)$ ist abelsche Gruppe“ schreibt man oft auch „G ist multiplikativ geschriebene abelsche Gruppe“.

Bis hierher hätten wir am 2. November kommen sollen. Bis hierher habe ich das Skript geprüft und vorläufig für gut befunden.

Beispiel 2.1.3. (a) Ist a ein mathematisches Objekt, so ist$(\{a\}, +)$ mit

$$+ : \{a\} \times \{a\} \to \{a\}, (a, a) \mapsto a$$

eine abelsche Gruppe, in der gilt:

$$a + a = a, 0 = a, \text{ und } -a = 0$$

Dies ist die einzige abelsche Gruppe mit Trägermenge $\{a\}$.

(b) Die leere Menge ist keine Trägermenge einer abelschen Gruppe wegen (N).

(c) Ist $(G, +)$ eine zweielementige abelsche Gruppe, so gibt es $a \in G$ mit $G = \{0, a\}, a \neq 0$ und aus (I) folgt $a + 0 = 0$ oder $a + a = 0$.

Aus $a + 0 = 0$ würde aber $a \overset{(N)}{=} a + 0 = 0$ folgen im Widerspruch zu $a \neq 0$. Also gilt $a + a = 0$. Mit (K) und (N) erhält man die Addition $+$ von $(G, +)$ in Form einer *Additionstabelle*:

$+$	0	a
0	0	a
a	a	0

Vorsicht! Damit ist nicht gezeigt, dass es eine zweielementige abelsche Gruppe gibt. Es ist nur gezeigt, dass jede zweielementige abelsche Gruppe so ausschaut.
Übung: Zeige, dass durch diese Tabelle eine abelsche Gruppe definiert wird.

(d) Ist $(G,+)$ eine dreielementige abelsche Gruppe, so gibt es $a,b \in G$ mit $G = \{0,a,b\}$ und $0,a,b$ paarweise verschieden.
Wäre $a+b=a$, so folgte

$$b \overset{(N)}{=} b+0 \overset{(I)}{=} b+(a+(-a)) \overset{(A)}{=} (b+a)+(-a) \overset{(K)}{=} (a+b)+(-a) = a+(-a) \overset{(I)}{=} 0 \ \lightning$$

Also gilt $a+b \neq a$. Analog folgt $a+b \neq b$. Daher muss $a+b=0$ gelten, also ist $b=-a$. Wäre $a+a=0$, so folgte $a=-a=b$ $\lightning$. Wäre $a+a=a$, so folgte $a \overset{(N)}{=} a+0 \overset{(I)}{=} a+(a+(-a)) \overset{(A)}{=} (a+a)+(-a) = a+(-a) \overset{(I)}{=} 0$ $\lightning$. Also muss $a+a=b$ gelten. Analog zeigt man $b+b=a$. Mit (N) und (K) erhält man die Additionstabelle von $(G,+)$:

$+$	0	a	b
0	0	a	b
a	a	b	0
b	b	0	a

Vorsicht! [→ (c)]

(e) Sei A eine Menge. Dann ist $(\mathscr{P}(A),+)$ mit

$$\begin{aligned} +: \mathscr{P}(A) \times \mathscr{P}(A) &\to \mathscr{P}(A) \\ (B,C) &\mapsto B \underbrace{\Delta}_{\substack{\text{„symmetrische} \\ \text{Mengendifferenz“}}} C := (B \setminus C) \cup (C \setminus B) \end{aligned}$$

eine abelsche Gruppe mit $0 = \emptyset$ und $-B = B$ für $B \in \mathscr{P}(A)$.

(f) $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \{5n \mid n \in \mathbb{Z}\}, \left\{\frac{n}{2} \mid n \in \mathbb{Z}\right\}$ bilden zusammen mit der gewöhnlichen Addition auf ihnen jeweils eine abelsche Gruppe, nicht jedoch $\mathbb{N}$ oder $\mathbb{N}_0$.
$\{0\}, \{1\}, \{-1,1\}, \mathbb{Q} \setminus \{0\}, \mathbb{Q}_{>0}, \mathbb{R} \setminus \{0\}, \mathbb{R}_{>0}$ bilden zusammen mit der gewöhnlichen Multiplikation auf ihnen jeweils eine (multiplikativ geschriebene [→ 2.1.2 (d)]) abelsche Gruppe, nicht jedoch $\{0,1\}, \mathbb{Q}$ oder $\mathbb{R}$.

Proposition 2.1.4. *Sei G eine abelsche Gruppe* [→ 2.1.2 (e)]. *Dann gilt für alle $a,b \in G$:*

$$-(-a) = a \text{ und } -(a+b) = (-a)+(-b)$$

Beweis. Seien $a,b \in G$. Um $-(-a)=a$ zu zeigen, genügt es, $-a+a=0$ zu zeigen [→ 2.1.2 (b)], was aber sofort aus (I) und (K) folgt. Um $-(a+b)=(-a)+(-b)$ zu zeigen, ist $(a+b)+((-a)+(-b))=0$ zu zeigen.
Dies folgt aus

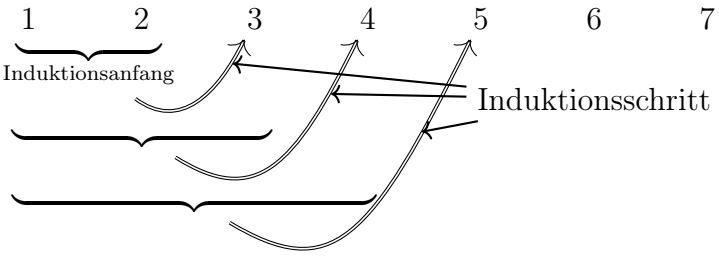
$$\begin{aligned} (a+b)+((-a)+(-b)) &\overset{(A)}{=} a+(b+((-a)+(-b))) \overset{(K)}{=} a+(b+((-b)+(-a))) \\ &\overset{(A)}{=} a+((b+(-b))+(-a)) = a+(0+(-a)) \\ &\overset{(K)}{=} a+((-a)+0) \overset{(N)}{=} a+(-a) \overset{(I)}{=} 0 \end{aligned}$$

□

Bemerkung 2.1.5. Analog zu Proposition 2.1.4 kann man bei Bedarf viele andere gewohnte Rechenregeln zeigen.

Satz 2.1.6 (Über das Weglassen von Klammern)**.** *Sei A eine Menge und $\varrho : A \times A \to A$* assoziativ*, d.h. (mit infix geschriebenem ϱ) $\forall a, b, c \in A : (a \ \varrho \ b) \ \varrho \ c = a \ \varrho \ (b \ \varrho \ c)$ (z.B. (A, ϱ) abelsche Gruppe). Dann liefert für $n \in \mathbb{N}$ und $a_1, \ldots, a_n \in A$ jede sinnvolle Klammerung von $a_1 \ \varrho \ a_2 \ \varrho \ a_3 \ \varrho \ \ldots \ \varrho \ a_n$ dasselbe Element von A.*

Beweis. durch *Induktion* nach n. Wir zeigen die Behauptung zunächst für $n = 1$ und $n = 2$ (*Induktionsanfang*) und dann für $n \in \mathbb{N}_{\geq 3}$ (*Induktionsschritt*) unter der Annahme, dass die Behauptung für $1, \ldots, n-1$ schon gezeigt wurde (*Induktionsvoraussetzung*, IV).



$\underline{n \in \{1, 2\}}$ klar.
$\underline{1, 2, \ldots, n-1 \to n \ (n \geq 3)}$: Seien zwei sinnvolle Klammerungen von $a_1 \ \varrho \ a_2 \ \varrho \ \ldots \ \varrho \ a_n$ gegeben und x und y die dadurch gegebenen Elemente von A.
Zu zeigen: $x = y$. Wähle $i, j \in \{1, \ldots, n-1\}$ mit

$$x = (a_1 \ \varrho \ \ldots \ \varrho \ a_i) \ \varrho \ (a_{i+1} \ \varrho \ \ldots \ \varrho \ a_n) \text{ und}$$

$$y = (a_1 \ \varrho \ \ldots \ \varrho \ a_j) \ \varrho \ (a_{j+1} \ \varrho \ \ldots \ \varrho \ a_n)$$

jeweils mit geeigneter Klammerung der Teilausdrücke, die nach IV aber irrelevant ist. Ist $i = j$, so sind wir fertig. Sonst können wir Œ (*ohne Einschränkung*) $i < j$ voraussetzen (sonst vertausche x und y).
Aber dann

$$\begin{aligned} x &\overset{\text{IV}}{=} (a_1 \ \varrho \ \ldots \ \varrho \ a_i) \ \varrho \ ((a_{i+1} \ \varrho \ \ldots \ \varrho \ a_j) \ \varrho \ (a_{j+1} \ \varrho \ \ldots \ \varrho \ a_n)) \\ &\overset{\varrho \text{ assoz.}}{=} ((a_1 \ \varrho \ \ldots \ \varrho \ a_i) \ \varrho \ (a_{i+1} \ \varrho \ \ldots \ \varrho \ a_j)) \ \varrho \ (a_{j+1} \ \varrho \ \ldots \ \varrho \ a_n) \\ &\overset{\text{IV}}{=} (a_1 \ \varrho \ \ldots \ \varrho \ a_j) \ \varrho \ (a_{j+1} \ \varrho \ \ldots \ \varrho \ a_n) = y \end{aligned}$$

□

Notation 2.1.7. In der Situation von 2.1.6 oder in ähnlichen Situationen, in denen der Beweis von 2.1.6 greift (etwa bei der Hintereinanderausführung von mehreren Abbildungen [→ 1.2.5 (a)]) verzichten wir oft auf Klammern oder klammern nach Belieben um.

Notation 2.1.8. $S_n := \{\sigma \mid \sigma \text{ Permutation von } \{1, \ldots, n\}\}$ für $n \in \mathbb{N}_0$ [→ 1.1.29 (b)]

Satz 2.1.9 (Über Umordnung). *Sei A eine Menge, $\varrho : A \times A \to A$ assoziativ* [→ 2.1.6] *und* kommutativ, *d.h.* $\forall a, b \in A : a \varrho b = b \varrho a$ *(z.B. (A, ϱ) abelsche Gruppe). Dann gilt für alle $n \in \mathbb{N}, a_1, \ldots, a_n \in A$ und $\sigma \in S_n$: $a_1 \varrho \ldots \varrho a_n = a_{\sigma(1)} \varrho \ldots \varrho a_{\sigma(n)}$.*

Beweis. durch Induktion nach n
$\underline{n = 1}$ klar
$\underline{n - 1 \to n \ (n \geq 2)}$ Seien $n \in \mathbb{N}, a_1, \ldots, a_n \in A$ und $\sigma \in S_n$. Wähle $i \in \{1, \ldots, n\}$ mit $\sigma(i) = n$. Es ist $\tau : \{1, \ldots, n-1\} \to \{1, \ldots, n-1\}, j \mapsto \begin{cases} \sigma(j), & \text{falls } j < i \\ \sigma(j+1), & \text{falls } j \geq i \end{cases}$ eine Bijektion, wie man sich sofort überlegt. Nach IV gilt $a_1 \varrho \ldots \varrho a_{n-1} = a_{\tau(1)} \varrho \ldots \varrho a_{\tau(n-1)}$ und daher

$$\begin{aligned} a_1 \varrho \ldots \varrho a_{n-1} \varrho a_n &= a_{\tau(1)} \varrho \ldots \varrho a_{\tau(n-1)} \varrho a_n \\ &\overset{\text{Def. von } \tau}{=} a_{\sigma(1)} \varrho \ldots \varrho a_{\sigma(i-1)} \varrho a_{\sigma(i+1)} \varrho \ldots \varrho a_{\sigma(n-1)} \varrho a_n \\ &= (a_{\sigma(1)} \varrho \ldots \varrho a_{\sigma(i-1)}) \varrho ((a_{\sigma(i+1)} \varrho \ldots \varrho a_{\sigma(n)}) \varrho a_{\sigma(i)}) \\ &\overset{\varrho \text{ komm.}}{=} (a_{\sigma(1)} \varrho \ldots \varrho a_{\sigma(i-1)}) \varrho (a_{\sigma(i)} \varrho (a_{\sigma(i+1)} \varrho \ldots \varrho a_{\sigma(n)})) \\ &= a_{\sigma(1)} \varrho \ldots \varrho a_{\sigma(n)} \end{aligned}$$

□

Bis hierher hätten wir am 6. November kommen sollen.

Notation 2.1.10. Sei G eine abelsche Gruppe, ferner $(a_i)_{i \in I}$ eine Familie in G, ferner $n = \#I \in \mathbb{N}$, so gilt $a_{\sigma(1)} + \ldots + a_{\sigma(n)} = a_{\tau(1)} + \ldots + a_{\tau(n)}$ für alle Bijektionen $\sigma, \tau : \{1, \ldots, n\} \to I$ [→ 2.1.6, 2.1.9]. (denn $\sigma^{-1} \circ \tau \in S_n$ und daher ist $b_1 + \ldots + b_n = b_{\sigma^{-1}(\tau(1))} + \ldots + b_{\sigma^{-1}(\tau(n))}$ für $b_1 = a_{\sigma(1)}, \ldots, b_n = a_{\sigma(n)}$ und wir notieren dieses Element von G mit $\sum_{i \in I} a_i$. Wir setzen $\sum_{i \in \emptyset} a_i = 0$. Statt $\sum_{i \in \{m, \ldots, n\}} a_i$ schreibt man auch $\sum_{i=m}^{n} a_i$. Beachte $\sum_{i=1}^{0} a_i \overset{1.1.6}{=} \sum_{i \in \emptyset} = 0$.

Satz und Definition 2.1.11. [→ 1.1.28, 1.1.27]. *Sei I eine Menge für jedes $i \in I$ sei $(G_i, +_i)$ eine abelsche Gruppe. Dann ist $\prod_{i \in I} G_i$ mit*

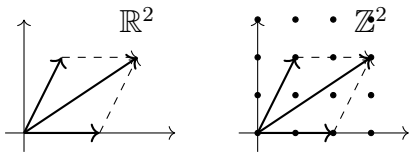
$$+ : \prod_{i \in I} G_i \times \prod_{i \in I} G_i \to \prod_{i \in I} G_i, (g, h) \mapsto (i \mapsto g(i) + h(i))$$

wieder eine abelsche Gruppe, genannt das direkte Produkt *der G_i ($i \in I$). Für alle $g, h \in \prod_{i \in I} G_i$ gilt: $(g + h)(i) = g(i) +_i h(i), 0(i) = 0_i, (-g)(i) = -_i g(i)$. („punktweise Addition").*

Beweis. Übungsaufgabe. □

Korollar 2.1.12. *Seien $n \in \mathbb{N}_0$ und $(G_1, +_1), \ldots, (G_n, +_n)$ abelsche Gruppen. Dann ist $(G_1 \times \ldots \times G_n, +)$ mit $(a_1, \ldots, a_n) + (b_1, \ldots, b_n) := (a_1 + b_1, \ldots, a_n + b_n)$ für alle $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in G_1 \times \ldots \times G_n$ wieder eine abelsche Gruppe mit $0 = (0_1, \ldots, 0_n)$, $-(a_1, \ldots, a_n) = (-_1 a_1, \ldots, -_n a_n)$ für alle $(a_1, \ldots, a_n) \in G_1 \times \ldots \times G_n$.*

Beispiel 2.1.13. $\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}}_{n \text{ mal}}$ mit „Vektoraddition“:



§2.2 Untergruppen und Gruppenhomomorphismen

Definition 2.2.1. Seien $(G, +_G)$ und $(H, +_H)$ abelsche Gruppen. Dann heißt $(H, +_H)$ eine *Untergruppe* von $(G, +_G)$, falls $H \subseteq G$ und $\forall a, b \in H : (a +_H b = a +_G b)$.

Proposition 2.2.2. *Sei $(G, +_G)$ eine abelsche Gruppe und H eine Menge. Dann ist H genau dann Trägermenge einer Untergruppe von $(G, +_G)$, wenn gilt:*

(a) $H \subseteq G$

(b) $0_G \in H$

(c) $\forall a, b \in H : (a +_G b \in H)$

(d) $\forall a \in H : (-_G a \in H)$

In diesem Fall gibt es genau eine Abbildung $+_H : H \times H \to H$, mit der $(H, +_H)$ eine Untergruppe von $(G, +_G)$ wird.
Es gilt dann:

(b') $0_H = 0_G$

(c') $\forall a, b \in H; (a +_H b = a +_G b)$

(d') $\forall a \in H : (-_G a = -_H a)$

Beweis. Wir zeigen zunächst: $(*) : (a) \,\&\, (b) \,\&\, (c) \,\&\, (d) \Longrightarrow H$ ist Trägermenge einer Untergruppe von $(G, +_G)$. Seien (a), (b), (c), (d). Definiere eine Abbildung $+_H : H \times H \to H, (a, b) \mapsto a +_G b$ unter Ausnutzung von (a) und (c). Sicherlich gelten wegen (a) auch (K) und (A) aus 2.1.1 in $(H, +_H)$. Gleiches gilt für (N) und (I) wegen (b) bzw. (d). Damit ist $(*)$ gezeigt.
Sei nun H Trägermenge einer Untergruppe von $(G, +_G)$. Dann existiert eine Gruppenaddition $+_H : H \times H \to H$ derart, dass $(H, +_H)$ eine Untergruppe von $(G, +_G)$ ist. Nach 2.2.1 gelten (a) und (c'). Aus (c') folgt weiter, dass es *genau ein* solches $+_H$ gibt. Offenbar gilt (b') $\Longrightarrow$ (b), (c') $\Longrightarrow$ (c), (d') $\Longrightarrow$ (d).
z.Z also noch: (b') und (d').

Zu (b'): Es gilt: $0_H +_G 0_H \overset{(c')}{=} 0_H +_H 0_H \overset{(N) \text{ in } H}{=} 0_H$.
Daraus folgt:

$$\begin{aligned} 0_H &\overset{(N) \text{ in } G}{=} 0_H +_G 0_G \\ &\overset{(I)}{=} 0_H +_G (0_H +_G (-_G 0_H)) \\ &\overset{(A) \text{ in } G}{=} (0_H +_G 0_H) +_G (-_G 0_H) \\ &\overset{\text{s.o.}}{=} 0_H +_G (-_G 0_H) = 0_G \end{aligned}$$

Zu (d'): Sei $a \in H$. Dann ist $a +_G (-_H a) \overset{(c')}{=} a +_H (-_H a) \overset{(I)}{=} 0_H = 0_G$. Daher $-_H a = -_G a$ für alle $a \in H$. □

Bemerkung 2.2.3. Ist $(H, +_H)$ eine Untergruppe der abelschen Gruppe $(G, +_G)$, so schreibt man wegen (b') - (d') fast immer $+, -, 0$ statt $+_H, -_H, 0_H$. Daher erwähnt man die Gruppenaddition oft nicht mehr explizit und spricht z.B. von einer „Untergruppe H der abelschen Gruppe G".

Beispiel 2.2.4. (a) Für jede abelsche Gruppe G sind $\{0\}$ und G (Trägermengen von) Untergruppen von G.
Für $\#G \leq 3$ besitzt G keine weiteren Untergruppen.

(b) Gelte $X \subseteq A$. Betrachte wieder die abelsche Gruppe $\mathscr{P}(A)$ mit $B + C = B \Delta C = (B \setminus C) \cup (C \setminus B)$. Es ist $H := \{B \in \mathscr{P}(A) \mid B \cap X = \emptyset\}$ eine Untergruppe von G. (Übungsaufgabe).

(c) Folgende Inklusionen sind Untergruppenbeziehungen:

$$\{10n \mid n \in \mathbb{Z}\} \begin{Bmatrix} \subseteq \{5n \mid n \in \mathbb{Z}\} \\ \subseteq \{2n \mid n \in \mathbb{Z}\} \end{Bmatrix} \subseteq \mathbb{Z} \begin{Bmatrix} \subseteq \left\{\frac{n}{3} \mid n \in \mathbb{Z}\right\} \\ \subseteq \left\{\frac{n}{2} \mid n \in \mathbb{Z}\right\} \end{Bmatrix} \subseteq \left\{\frac{n}{6} \mid n \in \mathbb{Z}\right\} \subseteq \mathbb{Q}$$
$$\subseteq \left\{p + q\sqrt{2} \mid p, q \in \mathbb{Q}\right\} \subseteq \mathbb{R}$$

$$\mathbb{Z} \times \mathbb{Z} \begin{Bmatrix} \subseteq \mathbb{Q} \times \mathbb{Z} \\ \subseteq \mathbb{Z} \times \mathbb{Q} \end{Bmatrix} \subseteq \mathbb{Q} \times \mathbb{Q} \subseteq \mathbb{Q} \times \mathbb{R} \subseteq \mathbb{R} \times \mathbb{R}.$$

(d) $\mathbb{N}_0$ ist *keine* Untergruppe von $\mathbb{Z}$ bezüglich der gewöhnlichen Addition.
$(\mathbb{Q} \setminus \{0\}, \cdot)$ ist *keine* Untergruppe von $(\mathbb{Q}, +)$.

Proposition 2.2.5. *Sei G eine abelsche Gruppe und M eine Menge von Untergruppen von G. Dann ist auch $\bigcap M$ eine Untergruppe von G ($\bigcap \emptyset = G$).*

Beweis. z.Z.:

(a) $\bigcap M \subseteq G$

(b) $0 \in \bigcap M$

(c) $\forall a, b \in \bigcap M : (a + b \in \bigcap M)$

(d) $\forall a \in \bigcap M : (-a \in \bigcap M)$

(a) Œ: $M \neq \emptyset$. Sei $a \in \bigcap M$. Wähle $H \in M$. Dann $a \in H$. Also $a \in H \subseteq G$, d.h. $a \in G$.

(b) Sei $H \in M$. z.Z. $0 \in H$. Klar, da H Untergruppe von G und also (b) *für* H gilt.

(c) Seien $a, b \in \bigcap M$. z.Z.: $a+b \in \bigcap M$. Sei $H \in M$. Dann $a, b \in H$. Da H Untergruppe, folgt $a + b \in H$.

(d) Sei $a \in \bigcap M$. z.Z.: $-a \in \bigcap M$. Sei $H \in M$. Dann $a \in H$. Da H Untergruppe ist $-a \in H$. □

Satz und Definition 2.2.6. *Sei G eine abelsche Gruppe und $E \subseteq G$. Dann existiert* eine eindeutig bestimmte *kleinste Untergruppe H von G mit $E \subseteq H$. (d.h. Ist H' eine Untergruppe von G mit $E \subseteq H'$, so folgt $H \subseteq H'$.) Dieses H nennt man die von E* erzeugte Untergruppe *von G und notiert es mit $\langle E \rangle_G$.*

Beweis. $M := \bigcap\{H \mid H$ Untergruppe von G und $E \subseteq H\}$. M ist Untergruppe von G nach 2.2.5. Offenbar ist $E \subseteq M$. $M \subseteq H'$ für alle Untergruppen H' mit $E \subseteq H'$ ist trivial nach Definition von M. □

Satz 2.2.7. *Sei G eine abelsche Gruppe und $E \subseteq G$. Dann*

$$\langle E \rangle_G = \left\{ \sum_{i=1}^{m} a_i - \sum_{i=1}^{n} b_i \;\middle|\; m, n \in \mathbb{N}, a_1, \ldots, a_m, b_1, \ldots, b_n \in E \right\}$$

Beweis. Übungsaufgabe. □

Beispiel 2.2.8. $\langle \{3, 2\} \rangle_{\mathbb{Z}} = \mathbb{Z}$

Definition 2.2.9. Seien G und H abelsche Gruppen. Dann heißt $f : G \to H$ ein *Gruppenhomomorphismus*, falls $\forall a, b \in G : f(a +_G b) = f(a) +_H f(b)$.

Gedanke 2.2.10. Idee eines Homomorphismus (Hom.): „erst rechnen, dann abbilden“ ist dasselbe wie „erst abbilden, dann rechnen“.

Proposition 2.2.11. *Seien $(G, +_G)$ und $(H, +_H)$ abelsche Gruppen, ferner $f : G \to H$ ein Gruppenhomomorphismus. Dann gilt $f(0_G) = 0_H$ und $\forall a \in G(f(-_G a) = -_H f(a))$.*

Beweis. Aus $0_G = 0_G +_G 0_G$ folgt $f(0_G) = f(0_G +_G 0_G) \overset{f \text{ Hom.}}{=} f(0_G) +_H f(0_G)$ und daher $0_H = f(0_G) -_H f(0_G) = f(0_G) +_H f(0_G) -_H f(0_G) = f(0_G)$.
Sei nun $a \in G$. Um $f(-_G a) = -_H f(a)$ zu zeigen, müssen wir zeigen, dass $f(a) +_H f(-_G a) = 0_H$.
Es gilt aber $f(a) +_H f(-_G a) \overset{f \text{ Hom.}}{=} f(a +_G (-_G a)) = f(0_G) = 0_H$. □

Notation 2.2.12. Ein Gruppenhomomorphismus $f : G \to H$ heißt (Gruppen-) $\left\{\begin{array}{c}\textit{Einbettung oder Mono-}\\ \textit{Epi-}\\ \textit{Iso-}\end{array}\right\}$ *morphismus*, falls f $\left\{\begin{array}{c}\text{injektiv}\\ \text{surjektiv}\\ \text{bijektiv}\end{array}\right\}$ ist.
ein Gruppenhomomorphismus $f : G \to G$ heißt auch Endomorphismus, ein Isomorphismus $f : G \to G$ heißt Automorphismus von G.

Beispiel 2.2.13.

	Hom.	Einb.	Epi	Iso	Endo	Auto
$\mathbb{Z} \to \mathbb{Z}, a \mapsto 2a$	✓	✓	✗	✗	✓	✗
$\mathbb{Z} \to \mathbb{Q}, a \mapsto 2a$	✓	✓	✗	✗	✗	✗
$\mathbb{Q} \to \mathbb{Q}, a \mapsto 2a$	✓	✓	✓	✓	✓	✓
$\mathbb{R} \setminus \{0\} \to \mathbb{R}_{>0}$ (mit der Multiplikation), $a \mapsto \lvert a\rvert$	✓	✗	✓	✗	✗	✗

Proposition 2.2.14. (a) *Seien G, H, I abelsche Gruppen, ferner $G \xrightarrow{f} H \xrightarrow{g} I$ Gruppenhomomorphismus. Dann ist auch $g \circ f$ Gruppenhomomorphismus.*

(b) *Ist $f : G \to H$ ein Gruppenisomorphismus, so auch f^{-1}.*

Beweis. (a) Zu zeigen: $(g \circ f)(a + b) = (g \circ g)(a) + (g \circ f)(b)$ für alle $a, b \in G$. Es ist aber

$$\begin{aligned}(g \circ f)(a + b) = g(f(a + b)) &\overset{f \text{ Hom.}}{=} g(f(a) + f(b))\\ &\overset{g \text{ Hom.}}{=} g(f(a)) + g(f(b)) = (g \circ f)(a) + (g \circ f)(b).\end{aligned}$$

(b) Sei $f : G \to H$ ein Gruppenisomorphismus. Dann ist f^{-1} bijektiv. Es ist also noch zu zeigen, dass f^{-1} Homomorphismus, d.h. $f^{-1}(a + b) = f^{-1}(a) + f^{-1}(b)$ für alle $a, b \in H$. Da f injektiv, reicht es, zu zeigen, dass $f(f^{-1}(a+b)) = f(f^{-1}(a)+f^{-1}(b))$. Das folgt aber wegen

$$f(f^{-1}(a + b)) = a + b = f(f^{-1}(a)) + f(f^{-1}(b)) \overset{f \text{ Hom.}}{=} f(f^{-1}(a) + f^{-1}(b)).$$

□

Sprechweise und Notation 2.2.15. Zwei abelsche Gruppen G und H heißen *isomorph*, wenn es einen Isomorphismus von G nach H gibt, in Zeichen $G \cong H$.
Ein Gruppenisomorphismus führt also die Additionstabelle der einen abelschen Gruppe in eine Additionstabelle der anderen abelschen Gruppe über:
Ein Isomorphismus tauscht die Elemente aus, ohne die „Gruppenstruktur“ zu verändern. Alle „strukturellen“, d.h. nicht auf die Natur der Elemente bezogenen Eigenschaften einer abelschen Gruppe übertragen sich daher unter Isomorphismen.

$+$	$\cdots$	a	$\cdots$
$\vdots$		$\vert$	
b	$-$	$a+b$	$-$
$\vdots$		$\vert$	

$+$	$\cdots$	$f(a)$	$\cdots$
$\vdots$		$\vert$	
$f(b)$	$-$	$f(a+b)$	$-$
$\vdots$		$\vert$	

$+_G$	0_G	a	b
0_G	0_G	a	b
a	a	b	0_G
b	b	0_G	a

$+_H$	0_H	c	d
0_H	0_H	c	d
c	c	d	0_H
d	d	0_H	c

Beispiel 2.2.16. 1. Seien $(G, +_G)$ und $(H, +_H)$ abelsche Gruppen mit $\#G = \#H = 3$. Dann kann man [→ 2.1.3 d)] a, b, c, d finden, so dass $G = \{0_G, a, b\}$ und $H = \{0_H, c, d\}$ und die Additionstabellen von G und H wie folgt lauten:
Dann ist $f : G \to H, 0_G \mapsto 0_H, a \mapsto c, b \mapsto d$ ein Isomorphismus zwischen G und H.
Man sagt, dass es „bis auf Isomorphie“ genau eine dreielementige abelsche Gruppe gibt.

2. Frage: Wie viele Automorphismen hat eine dreielementige abelsche Gruppe?
Antwort: Genau 2. Die Identität und diejenige, die die beiden von 0 verschiedenen Elemente vertauscht.

Sprechweise 2.2.17. „Isomorphismus = Umbenennung der Elemente“

§2.3 Quotientengruppen[→ §1.3]

Idee: Grobe Sichtweise auf eine abelsche Gruppe einnehmen.

Definition 2.3.1. Sei G eine abelsche Gruppe. Eine *Kongruenzrelation* auf G ist eine Äquivalenzrelation $\equiv$ auf G, für die gilt:

$$(*) \qquad \forall a, a', b, b' \in G : ((a \equiv a' \ \& \ b \equiv b') \implies a + b \equiv a' + b')$$

Für $a \in G$ nennt man $\overline{a} := \overline{\overline{a}}$ („ein Strich gleich drei Strich“-Regel zur Vereinfachung der Notation!) statt Äquivalenz- auch *Kongruenzklasse* von a bezüglich $\equiv$.

Bemerkung 2.3.2. In Definition 2.3.1 drückt Bedingung $(*)$ gerade folgendes aus:

$$(**) \qquad \begin{array}{ccc} G/\equiv \ \times \ G/\equiv & \to & G/\equiv \\ (\overline{a}, \overline{b}) & \mapsto & \overline{a+b} \quad (a, b \in G) \end{array} \qquad \text{ist wohldefiniert.}$$

In der Tat: $(**) \iff (\forall a, b, a', b' \in G : ((\overline{a}, \overline{b}) = (\overline{a'}, \overline{b'}) \implies \overline{a+b} = \overline{a'+b'})) \iff (*)$.

Satz und Definition 2.3.3. *Sei G eine abelsche Gruppe und $\equiv$ eine Kongruenzrelation auf G. Dann wird die Quotientenmenge $A/\equiv$ vermöge der durch*

$$\overline{a} + \overline{b} := \overline{a+b} \qquad (a, b \in G)$$

festgelegten („vertreterweisen") Addition zu einer abelschen Gruppe, die man die zu $\equiv$ gehörige Quotientengruppe *von G nennt (auch „G nach $\equiv$" oder „G modulo $\equiv$"). In ihr gilt $0 = \overline{0}$ und $-\overline{a} = \overline{-a}$ für alle $a \in G$.*

Beweis. Die Wohldefiniertheit der Addition auf $A/\equiv$ als Abbildung haben wir schon in Bemerkung 2.3.2 geklärt. Wir prüfen die Axiome (K), (A), (N) und (I) aus 2.1.1 nach:

(K) $\overline{a} + \overline{b} = \overline{a+b} = \overline{b+a} = \overline{b} + \overline{a}$ für alle $a, b \in G$.

(A) $(\overline{a}+\overline{b})+\overline{c} = \overline{a+b}+\overline{c} = \overline{(a+b)+c} = \overline{a+(b+c)} = \overline{a}+\overline{b+c} = \overline{a}+(\overline{b}+\overline{c})$ für alle $a, b, c \in G$.

(N) Für $a \in G$ gilt $\overline{a} + \overline{0} = \overline{a+0} = \overline{a}$. Daher ist $0 = \overline{0}$.

(I) Für $a \in G$ gilt $\overline{a} + \overline{-a} = \overline{a+(-a)} = \overline{0} = 0$. Daher ist $-\overline{a} = \overline{-a}$ für $a \in G$.

□

Proposition 2.3.4. *Sei G eine abelsche Gruppe und $\equiv$ eine Kongruenzrelation auf G. Dann ist $H := \overline{0}$ eine Untergruppe von G, für die gilt:*

(a) $\forall a, b \in G : (a \equiv b \iff a - b \in H)$

(b) $\forall a \in G : \overline{a} = \{a + b \mid b \in H\}$

(c) *Für jedes $a \in G$ ist $H \to \overline{a},\ b \mapsto a + b$ bijektiv.*

Beweis. Um zu zeigen, dass H eine Untergruppe von G ist, sind nach Proposition 2.2.2 zeigen:

(1) $H \subseteq G$

(2) $0 \in H$

(3) $\forall a, b \in H : a + b \in H$

(4) $\forall a \in H : -a \in H$

(1) und (2) sind trivial. Um (3) zu sehen, beobachten wir, dass für alle $a, b \in H$ wegen $a \equiv 0$ und $b \equiv 0$ aus $(*)$ folgt $a+b \equiv 0+0 = 0$ und daher $a+b \in H$. Schließlich erhält man (4) daraus, dass für alle $a \in H$ aus $a \equiv 0$ und $-a \equiv -a$ gemäß $(*)$ $0 = a-a \equiv 0-a = -a$ und damit $-a \equiv 0$ folgt.

(a) Seien $a, b \in G$. Gilt $a \equiv b$, so wegen $-b \equiv -b$ gemäß $(*)$ auch $a - b \equiv b - b = 0$ und daher $a - b \in \overline{0} = H$. Gilt umgekehrt $a - b \in H$, so $a - b \equiv 0$ und wegen $b \equiv b$ gemäß $(*)$ auch $a = a - b + b \equiv b$.

(b) Sei $a \in G$. Wir behaupten $\overline{a} = \{a + b \mid b \in H\}$.
„$\subseteq$“ Ist $c \in \overline{a}$, so gilt $c \equiv a$, woraus mit $(*)$ folgt $b := c - a \equiv a - a = 0$ und $c = a + b$.
„$\supseteq$“ Ist umgekehrt $b \in \overline{0}$, so folgt mit $(*)$, dass $a + b \equiv a + 0 = a$ und damit $a + b \in \overline{a}$.
(c) Die Surjektivität ist gerade (b), die Injektivität ist leicht zu zeigen. □

Definition 2.3.5. [→ 1.3.5(b)] Sei G eine abelsche Gruppe. Zu jeder Untergruppe H von G definieren wir eine Relation $\equiv_H$ auf G durch

$$a \equiv_H b : \iff a - b \in H.$$

Satz 2.3.6. [→1.3.6] *Sei G eine abelsche Gruppe. Die Zuordnungen*

$$\equiv \mapsto \overline{0}$$
$$\equiv_H \mapsfrom H$$

vermitteln eine Bijektion zwischen der Menge der Kongruenzrelationen auf G und der Menge der Untergruppen von G.

Beweis. Zu zeigen ist:

(a) Ist $\equiv$ eine Kongruenzrelation auf G, so ist $\overline{0}$ eine Untergruppe von G.

(b) Ist H eine Untergruppe von G, so ist $\equiv_H$ eine Kongruenzrelation auf G.

(c) Ist $\equiv$ eine Kongruenzrelation auf G, so $\equiv_{\overline{0}} = \equiv$.

(d) Ist H eine Untergruppe von G, so $\overline{0}^{H} = H$.

(a) wurde bereits in Proposition 2.3.4 gezeigt.

Zu (b). Seien H eine Untergruppe von G und $a, a', b, b' \in G$ mit $a \equiv_H a'$ und $b \equiv_H b'$. Wir behaupten $a + b \equiv_H a' + b'$. Nach Definition 2.3.5 gilt $a - a' \in H$ und $b - b' \in H$. Es folgt $a + b - (a' + b') \overset{2.1.9}{=} (a - a') + (b - b') \in H$. Wieder wegen Definition 2.3.5 entspricht dies der Behauptung.

Zu (c). Seien $\equiv$ eine Kongruenzrelation auf G und $a, b \in G$. Zu zeigen ist $a \equiv_{\overline{0}} b \iff a \equiv b$. Dies sieht man leicht:

$$a \equiv_{\overline{0}} b \overset{2.3.5}{\iff} a - b \in \overline{0} \overset{1.3.1(b)}{\iff} a - b \equiv 0 \overset{(*)}{\iff} a \equiv b.$$

Zu (d). Ist H eine Untergruppe von G, so

$$\overline{0}^{H} \overset{1.3.1(b)}{=} \{a \in G \mid a \equiv_H 0\} \overset{2.3.5}{=} \{a \in G \mid a - 0 \in H\} = H.$$

□

Notation, Sprechweise und Proposition 2.3.7. *Sei H eine Untergruppe der abelschen Gruppe G. Dann nennt man $G/H := G/\equiv_H$ die Quotientengruppe von G nach (oder modulo) H. Die Kongruenzklassen $\overline{a}^H$ $(a \in G)$ [→ 2.3.1] von $\equiv_H$ nennen wir auch die* Nebenklassen *von H (in G). Wegen 2.3.4*(c) *haben alle Nebenklassen von H dieselbe Mächtigkeit* [→ 1.1.21] *und da sie eine Zerlegung* [→ 1.3.3] *von G bilden, gilt*

$$\#G = (\#(G/H))(\#H),$$

falls G endlich ist, denn $\#(G/H)$ ist dann die Anzahl der Nebenklassen von H und jede Nebenklasse von H hat $\#H$ viele Elemente.

Beispiel 2.3.8. Sei G eine abelsche Gruppe.

(a) $G/G = \{G\} = \{0\}$ ist einelementig.

(b) Es gilt $G/\{0\} = \{\{a\} \mid a \in G\}$, wobei $\{a\} + \{b\} = \{a + b\}$ für alle $a, b \in G$ gilt. Man sieht sofort, dass $G \to G/\{0\},\ a \mapsto \{a\}$ ein Isomorphismus ist. Insbesondere gilt $G/\{0\} \cong G$.

Beispiel 2.3.9. Sei $n \in \mathbb{Z}$. Die von $\{n\}$ erzeugte Untergruppe von $\mathbb{Z}$ ist $\langle n\rangle := \langle\{n\}\rangle_{\mathbb{Z}} = \{cn \mid c \in \mathbb{Z}\}$. Es gilt

$$\mathbb{Z}/\langle n\rangle = \left\{\overline{a}^{\langle n\rangle} \mid a \in \mathbb{Z}\right\}, \qquad \text{wobei}$$

$$\overline{a}^{\langle n\rangle} = \overline{b}^{\langle n\rangle} \iff a \equiv_{\langle n\rangle} b \iff a - b \in \langle n\rangle \iff \exists c \in \mathbb{Z} : a - b = cn.$$

Man überlegt sich sofort: Ist $n \in \mathbb{N}$, so hat $\mathbb{Z}/\langle n\rangle$ genau n Elemente und es gilt

$$\mathbb{Z}/\langle n\rangle = \left\{\overline{0}^{\langle n\rangle}, \overline{1}^{\langle n\rangle}, \ldots, \overline{n-1}^{\langle n\rangle}\right\}.$$

Ist $n = 0$, so ist $Z/\langle 0\rangle = \{\{a\} \mid a \in \mathbb{Z}\} \cong \mathbb{Z}$. Ist $-n \in \mathbb{N}$, so gilt $\langle n\rangle = \langle -n\rangle$ und daher $\mathbb{Z}/\langle n\rangle = \mathbb{Z}/\langle -n\rangle$.

Definition und Proposition 2.3.10. Seien G und H abelsche Gruppen und $f\colon G \to H$ ein Homomorphismus.

(a) Es ist $\equiv_f := \sim_f$ [→1.3.9] eine Kongruenzrelation auf G.

(b) Es ist der *Kern* $\ker f := f^{-1}(\{0\})$ [→1.1.25] von f eine Untergruppe von G.

(c) Unter der Bijektion aus Satz 2.3.6 entsprechen sich $\equiv_f$ und $\ker f$, das heißt

$$\ker f = \overline{0}^f \qquad \text{und} \qquad \equiv_f = \equiv_{\ker f}.$$

Beweis. **(a)** Seien $a, a', b, b' \in G$ mit $a \equiv_f a'$ und $b \equiv_f b'$. Zu zeigen ist $a + b \equiv_f a' + b'$. Nach Definition von $\equiv_f = \sim_f$ in 1.3.9 gilt $f(a) = f(a')$ und $f(b) = f(b')$ und es ist $f(a + b) = f(a' + b')$ zu zeigen. Dies ist aber klar, denn

$$f(a+b) \overset{2.2.9}{=} f(a) + f(b) = f(a') + f(b') \overset{2.2.9}{=} f(a'+b').$$

(b) Nach 2.2.2 ist zu zeigen:

$$\ker f \subseteq G, \quad 0 \in \ker f, \quad \forall a, b \in \ker f : a + b \in \ker f \quad \text{und} \quad \forall a \in \ker f : -a \in \ker f.$$

Trivial ist $\ker f \subseteq G$, da $f^{-1}(\{0\})$ natürlich in der Definitionsmenge G von f enthalten ist. Die Bedingung $0 \in \ker f$ entspricht genau der Beobachtung $f(0) = 0$ aus 2.2.11. Sind $a, b \in \ker f$, dann gilt $f(a) = 0 = f(b)$ und daher $f(a + b) = f(a) + f(b) = 0 + 0 = 0$, also $a + b \in \ker f$. Ist schließlich $a \in \ker f$, also $f(a) = 0$, so gilt $f(-a) \overset{2.2.11}{=} -f(a) = -0 \overset{0+0=0}{=} 0$ und daher $-a \in \ker f$.

(c) Die erste Gleichheit ergibt sich wie folgt:

$$\ker f \overset{(b)}{=} f^{-1}(\{0\}) \overset{1.1.25}{=} \{a \in G \mid f(a) = 0\} \overset{2.2.11}{=} \{a \in G \mid f(a) = f(0)\} \overset{1.3.9}{=} \{a \in G \mid a \equiv_f 0\} \overset{1.3.1(b)}{=} \overset{-f}{0}.$$

Um die zweite Gleichheit zu zeigen, seien $a, b \in G$. Wir zeigen $a \equiv_f b \iff a \equiv_{\ker f} b$. Es gilt

$$a \equiv_f b \overset{1.3.9}{\iff} f(a) = f(b) \overset{(I)}{\iff} f(a) - f(b) = 0 \overset{2.2.11}{\iff} f(a) + f(-b) = 0 \overset{2.2.9}{\iff} f(a-b) = 0 \overset{1.1.25}{\iff} a - b \in f^{-1}(\{0\}) \overset{(b)}{\iff} a - b \in \ker f \overset{2.3.5}{\iff} a \equiv_{\ker f} b.$$

□

Satz 2.3.11 (Homomorphiesatz für abelsche Gruppen)**.** [→1.3.8] *Seien G und H abelsche Gruppen, I eine Untergruppe von G und $f\colon G \to H$ ein Homomorphismus mit $I \subseteq \ker f$.*

(a) *Es gibt genau eine Abbildung $\overline{f}\colon G/I \to H$ mit $\overline{f}(\overset{-}{a}^I) = f(a)$ für alle $a \in G$. Diese Abbildung $\overline{f}$ ist ein Homomorphismus.*

(b) *$\overline{f}$ ist injektiv $\iff$ $I = \ker f$*

(c) *$\overline{f}$ ist surjektiv $\iff$ f ist surjektiv.*

Beweis. Nach Satz 2.3.6 (genauer der Wohldefiniertheit der dortigen Abbildung von rechts nach links) ist $\equiv_I$ eine Kongruenzrelation auf G. Aus der Voraussetzung $I \subseteq \ker f$ erhält man

$$a \equiv_I b \implies f(a) = f(b)$$

für alle $a, b \in G$, denn sind $a, b \in G$ mit $a \equiv_I b$, so $a - b \in I \subseteq \ker f$ nach Definition 2.3.5 und damit $f(a) - f(b) \overset{f \text{ Hom.}}{=} f(a-b) = 0$. Unter Beachtung von $G/I \overset{2.3.7}{=} G/\equiv_I$ erhält man die in (a) behauptete Existenz und Eindeutigkeit der Abbildung $\overline{f}$ daher aus dem Homomorphiesatz für Mengen 1.3.8. Dass $\overline{f}$ ein Homomorphismus ist, rechnet man sofort nach:

$$\overline{f}(\bar{a}^I) + \overline{f}(\bar{b}^I) \overset{(a)}{=} f(a) + f(b) \overset{f \text{ Hom.}}{=} f(a+b) = \overline{f}(\overline{a+b}^I) \overset{2.3.3}{=} \overline{f}(\bar{a}^I + \bar{b}^I)$$

für alle $a, b \in G$. Damit ist (a) gezeigt. Die Aussage (c) folgt direkt aus 1.3.8(c). Schließlich ist es eine leichte Übung zu zeigen, dass die Aussage $I = \ker f$ äquivalent ist zu $\forall a, b \in G : (a \equiv_I b \iff f(a) = f(b))$, womit (b) nichts anderes als 1.3.8(b) ist. □

Bemerkung 2.3.12. [→1.3.10] Sei I eine Untergruppe einer abelschen Gruppe G. Dann wird I durch einen Gruppenhomomorphismus $f\colon G \to H$ in eine weitere abelsche Gruppe H induziert, nämlich durch den *kanonischen Epimorphismus* [→2.2.12]

$$f\colon G \to G/I,\ a \mapsto \bar{a}^I.$$

In der Tat: $\ker f = \{a \in G \mid f(a) = 0\} = \{a \in G \mid \bar{a}^I = 0\} = \{a \in G \mid a \in I\} = I$.

Notation und Proposition 2.3.13. *Seien G und H abelsche Gruppen und $f\colon G \to H$ ein Homomorphismus. Dann schreiben wir meist* $\operatorname{im} f := f(G) \overset{1.1.25}{=} \{f(a) \mid a \in G\}$ *für das in* 1.1.29(d) *eingeführte Bild von f. Es ist* $\operatorname{im} f$ *eine Untergruppe von H.*

Beweis. Zu zeigen sind gemäß 2.2.2:

(a) $\operatorname{im} f \subseteq H$

(b) $0 \in \operatorname{im} f$

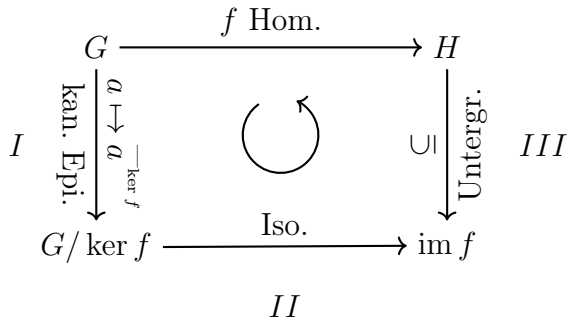
(c) $\forall a, b \in \operatorname{im} f : a + b \in \operatorname{im} f$

(d) $\forall a \in \operatorname{im} f : -a \in \operatorname{im} f$

(a) ist trivial und (b) folgt aus $f(0) \overset{2.2.11}{=} 0$. Sind $a, b \in \operatorname{im} f$, so gibt es $c, d \in G$ mit $f(c) = a$ und $f(d) = b$, woraus $f(c+d) \overset{f \text{ Hom.}}{=} f(c) + f(d) = a + b$ und damit $a + b \in \operatorname{im} f$ folgt. Dies zeigt (c). Ist schließlich $a \in \operatorname{im} f$, so gibt es $c \in G$ mit $f(c) = a$, woraus $f(-c) \overset{2.2.11}{=} -f(c) = -a$ und damit $-a \in \operatorname{im} f$ folgt. Damit haben wir auch (d) gezeigt und sind fertig. □

Korollar 2.3.14 (Isomorphiesatz für abelsche Gruppen)**.** [→1.3.11] *Seien G und H abelsche Gruppen und $f\colon G \to H$ ein Homomorphismus. Dann ist $\overline{f}\colon G/\ker f \to \operatorname{im} f$ definiert durch $\overline{f}(\bar{a}^{\ker f}) = f(a)$ für $a \in G$ ein Isomorphismus* [→2.2.12]*. Insbesondere* $G/\ker f \cong \operatorname{im} f$ [→2.2.15].

Bemerkung 2.3.15. Der Isomorphiesatz klärt uns über die Natur von Homomorphismen auf:



Drei Phasen:

$$G \xrightarrow[\text{„vergröbern“}]{I} G/\ker f \xrightarrow[\text{„umbenennnen“}]{II} \operatorname{im} f \xrightarrow[\text{„neue Elemente dazufügen“}]{III} H$$

§3 Kommutative Ringe

[Julius Wilhhelm Richard Dedekind *1831, †1916]

§3.1 Definition und Beispiele kommutativer Ringe

Definition 3.1.1. Ein *kommutativer Ring* ist ein Tripel (d.h. 3-Tupel) $(A, +, \cdot)$, wobei $(A, +)$ eine abelsche Gruppe ist und $\cdot : A \times A \to A$ eine (meist unsichtbar oder infix geschriebene, d.h. man schreibt ab oder $a \cdot b$ statt $\cdot(a, b)$) Abbildung mit folgenden Eigenschaften:

($\dot{\text{K}}$) $\forall a, b \in A : ab = ba$

($\dot{\text{A}}$) $\forall a, b, c \in A : (ab)c = a(bc)$

($\dot{\text{N}}$) $\exists e \in A : \forall a \in A : ae = a$

(D) $\forall a, b, c \in A : a(b + c) = (ab) + (ac)$ „distributiv“

Bemerkung 3.1.2. (a) Sind $e, e' \in A$ mit $\forall a \in A : ae = a = ae'$, so $e' = e'e \overset{(\dot{\text{K}})}{=} ee' = e$. Daher ist e wie in (N) eindeutig bestimmt und man schreibt dafür 1 statt e (die „Eins“ oder das „Einselement“ des kommutativen Ringes).

(b) Manchmal lässt man $\left\{\begin{matrix}(\dot{\text{A}}) \\ (\dot{\text{N}})\end{matrix}\right\}$ in der Definition 3.1.1 weg und bezeichnet einen kommutativen Ring mit $\left\{\begin{matrix}(\dot{\text{A}}) \\ (\dot{\text{N}})\end{matrix}\right\}$ als $\left\{\begin{matrix}\text{assoziativen} \\ \text{unitären}\end{matrix}\right\}$ kommutativen Ring. Statt „unitärer Ring“ sagt man auch „kommutativer Ring mit Eins“.

(c) Man nennt A bzw. $(A, +)$ die *zugrundeliegende Menge bzw. abelsche Gruppe*, $+$ die Addition und $\cdot$ die Multiplikation von $(A, +, \cdot)$. Es heißt A auch die *Trägermenge* und $(A, +)$ die *additive Gruppe* von $(A, +, \cdot)$.

(d) Wie bei abelschen Gruppen ist auch bei kommutativen Ringen ein schlampiger Sprachgebrauch üblich, z.B. „Sei A ein kommutativer Ring“ statt „Sei $(A, +, \cdot)$ ein kommutativer Ring.“ [→ 2.1.2 (e)]

(e) Wegen (Ȧ) kann man beim Multiplizieren mehrerer Elemente eines kommutativen Ringes beliebig umklammern [→ 2.1.6] und damit auf Klammern verzichten [→ 2.1.7]. Weiter kann man die Elemente auch in beliebiger Reihenfolge multiplizieren [→ 2.1.9].

(f) Es gilt die Konvention „Punkt vor Strich“, d.h. $\cdot$ bindet stärker als $+$ und $-$: $a \cdot b + c \cdot d$ und $ab - cd$ steht für $(ab) + (cd)$ und $(ab) - (cd)$.

(g) (D) sagt nichts anderes, als dass für jedes $a \in A$ die Abbildung $A \to A, x \mapsto ax$ ein Gruppenendomorphismus von $f_a : (A, +)$ ist [→ 2.2.9]. Insbesondere gilt $a \cdot 0 = 0$ für alle $a \in A$ und $a(-b) = -(ab)$ für alle $a, b \in A$ [→ 2.2.11].

Proposition 3.1.3. *Sei A ein kommutativer Ring. Dann $\#A = 1 \iff 0 = 1$ in A.*

Beweis. „$\Longrightarrow$“: Ist $A = \{a\}$, so gilt $0 = a = 1$.

„$\Longleftarrow$“: Gelte $0_A = 1_A$. Dann gilt für jedes $a \in A$ $a \overset{(\dot{\mathrm{N}})}{=} a \cdot 1_A = a \cdot 0_A \overset{3.1.2\,(\mathrm{g})}{=} 0_A$, also $A = \{0_A\} = \{1_A\}$.

□

Beispiel 3.1.4. [→ 2.1.3]

(a) $(\{a\}, +, \cdot)$ mit $+, \cdot : \{a\} \times \{a\} \to \{a\}, (a, a) \mapsto a$ ist ein kommutativer Ring mit $0 = a = 1$.

(b) $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ (mit gewöhnlicher Addition und Multiplikation) sind kommutative Ringe.

(c) Sei A eine Menge. Dann ist $(\mathscr{P}(A), +, \cdot)$ mit $+, \cdot : \mathscr{P}(A) \times \mathscr{P}(A) \to \mathscr{P}(A)$ definiert durch $B + C := B \Delta C = (B \setminus C) \cup (C \setminus B)$ und $B \cdot C := B \cap C$ für $B, C \in \mathscr{P}(A)$ ein kommutativer Ring mit 1. Es gilt $1 = A$.

(d) Genauso wie man in 2.1.11 das direkte Produkt von abelschen Gruppen eingeführt hat, kann man auch das direkte Produkt von kommutativen Ringen über punktweise Addition und Multiplikation einführen (Übung).
Auf diese Weise ist insbesondere $A^{\mathbb{N}_0}$ ein kommutativer Ring. Man kann die abelsche Gruppe $A^{\mathbb{N}_0}$ aber auch mit einer anderen Multiplikation, der sogenannten *Faltung*, zu einem kommutativen Ring machen. Da dies für uns wichtig sein wird, formulieren wir es in einem Satz.

Satz 3.1.5. *Sei A ein kommutativer Ring. Dann ist $(A^{\mathbb{N}_0}, +, *)$ mit*

$$f + g : \mathbb{N}_0 \to A, \qquad k \mapsto f(k) + g(k) \qquad \text{und}$$

$$\underbrace{f * g}_{\text{„gefaltet“}} : \mathbb{N}_0 \to A, \qquad k \mapsto \sum_{i=0}^{k} f(i) \cdot g(k - i) \qquad \text{für alle } f, g \in A^{\mathbb{N}_0}$$

ein kommutativer Ring mit $1 : \mathbb{N}_0 \to A, 0 \mapsto 1, k \mapsto 0$ für $k \in \mathbb{N}$.

Beweis. Wir wissen aus 2.1.11 schon, dass $(A^{\mathbb{N}_0}, +)$ eine abelsche Gruppe ist. Man rechnet nun als Übung (K̇), (Ȧ),(Ṅ) und (D) nach. □

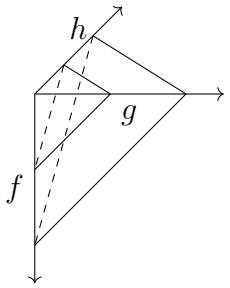
	$g(0)$	$g(1)$	$g(2)$	$\cdots$
$f(0)$	$f(0)g(0)$	$f(0)g(1)$	$f(0)g(2)$	$\cdots$
$f(1)$	$f(1)g(0)$	$f(1)g(1)$	$\ldots$	$\cdots$
$f(2)$	$f(2)g(0)$	$\ldots$	$\ldots$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Faltung ist Summe über die Diagonale.

(Ṅ):

	1	0	0	$\cdots$
$f(0)$	$f(0)$	0	0	$\cdots$
$f(1)$	$f(1)$	0	0	$\cdots$
$f(2)$	$f(2)$	0	0	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$

(Ȧ):



Faltung ist Summe über Raumdiagonalen.

Notation 3.1.6. [→ 2.1.10] Sei A ein kommutativer Ring [→ 3.1.1]. Ist $(a_i)_{i\in I}$ eine Familie in A und $n := \#I \in \mathbb{N}$, so gilt $a_{\sigma(1)} \cdot \ldots \cdot a_{\sigma(n)} = a_{\tau(1)} \cdot \ldots \cdot a_{\tau(n)}$ für alle Bijektionen $\sigma, \tau : \{1, \ldots, n\} \to I$ [→ 2.1.6, 2.1.9] und wir notieren dieses Element von A mit $\prod_{i\in I} a_i$. Wir setzen $\prod_{i\in\emptyset} a_i = 1$. Satt $\prod_{i\in\{m,\ldots,n\}} a_i$ schreibt man auch $\prod_{i=m}^{n} a_i$. Beachte $\prod_{i=1}^{0} a_i \overset{1.1.6}{=} \prod_{i\in\emptyset} a_i = 1$. Für $n \in \mathbb{N}_0$ setzen wir weiter $a^n := \prod_{i=1}^{n} a = \underbrace{a \cdot \ldots \cdot a}_{n\text{-mal}}$ (insbesondere $a^0 = 1$).

§3.2 Unterringe, Ringhomomorphismen und Polynome

Definition 3.2.1. [→ 2.2.1] Seien $(A, +_A, \cdot_A)$ und $(B, +_B, \cdot_B)$ kommutative Ringe. Dann heißt $(B, +_B, \cdot_B)$ ein *Unterring* von $(A, +_A, \cdot_A)$, wenn $B \subseteq A, 1_A \in B, \forall a, b \in B : a +_B b = a +_A b$ und $\forall a, b \in B : a \cdot_B b = a \cdot_A b$.

Proposition 3.2.2. [→ 2.2.2] *Sei $(A, +_A, \cdot_A)$ ein kommutativer Ring und B eine Menge. Genau dann ist B Trägermenge* [→ 3.1.2 (c)] *eines Unterrings von $(A, +_A, \cdot_A)$, wenn B Trägermenge einer Untergruppe von $(A, +_A)$ ist* [→ 2.2.2] *und $1_A \in B$ sowie $\forall a, b \in B : a \cdot_A b \in B$ gelten. In diesem Fall gibt es* genau *ein Paar $(+_B, \cdot_B)$, mit dem $(B, +_B, \cdot_B)$ ein Unterring von $(A, +_A, \cdot_A)$ wird.*
Es gilt dann $1_B = 1_A, \forall a, b \in B : a \dot{+}_B b = a \dot{+}_A b$ und $\forall a \in B : -_B a = -_A a$.

Beweis. Einfach mit 2.2.2. □

Beispiel 3.2.3. (a) $A := \{0_{\mathbb{Z}}\}$ mit der gewöhnlichen Addition und Multiplikation ist ein kommutativer Ring (in dem $1_A = 0_A = 0_{\mathbb{Z}}$ gilt), aber kein Unterring von $\mathbb{Z}$, denn $1_{\mathbb{Z}} \notin A$.

(b) Folgende Inklusionen sind Unterringbeziehungen: $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \left\{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\right\} \subseteq \mathbb{R}$ (beachte $(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ab + bc)\sqrt{2}$ für alle $a, b, c, d \in \mathbb{Q}$.)

Notation und Satz 3.2.4. *Sei A ein kommutativer Ring, B ein Unterring von A und $x \in A$. Dann ist*

$$\underbrace{B[x]}_{\text{„}B \text{ adjungiert } x\text{“}} := \left\{ \sum_{k=0}^{n} a_k x^k \mid n \in \mathbb{N}_0, a_0, \ldots, a_n \in B \right\}$$

der kleinste Unterring C von A mit $B \cup \{x\} \subseteq C$.

Beweis. Zu zeigen:

(a) $B[x]$ ist Unterring von A mit $x \in B[x]$.

(b) Ist C Unterring von A mit $B \cup \{x\} \subseteq C$, so gilt $B[x] \subseteq C$.

(b) ist klar.
Für (a) ist zu zeigen:

(1) $\{0, 1\} \subseteq B[x] \subseteq A$

(2) $\forall a, b \in B[x] : (a + b \in B[x] \,\&\, ab \in B[x])$

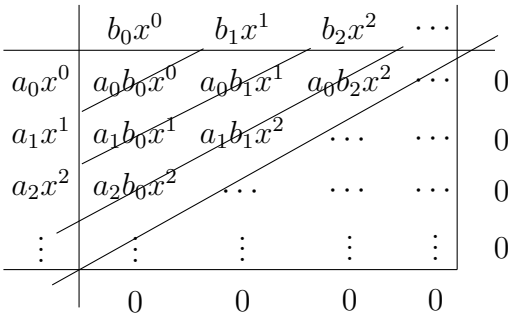
(3) $\forall a \in B[x] : -a \in B[x]$

Zu (1). $0 = 0 \cdot x^0 \in B[x]$
$1 = 1 \cdot x^0 \in B[x]$
$B[x] \subseteq A$ ist klar.

Zu (2). Seien $a, b \in B[x]$, etwa $a = \sum_{k=0}^{m} a_k x^k$ und $b = \sum_{k=0}^{n} b_k x^k$ mit $m, n \in \mathbb{N}_0, a_0, \ldots, a_m, b_0, \ldots, b_n \in B$. Setze $a_k := 0$ für $k \geq m + 1$ und $b_k := 0$ für $k \geq n + 1$. Dann gilt mit $\max\{m, n\} :=$ größte Element der Menge $\{m, n\}$:

$$a + b \overset{2.1.9}{=} \sum_{k=0}^{\max\{m,n\}} (a_k x^k + b_k x^k) \overset{(D)}{=} \sum_{k=0}^{\max\{m,n\}} (\underbrace{a_k + b_k}_{\in B}) x^k \in B[x] \text{ und}$$

$$ab \overset{(D)}{=} \sum_{k=0}^{m+n} \sum_{i=0}^{k} (a_i b_{k-i}) x^k \in B[x]$$



Diagonalen betrachten. $(a_0x^0 + \ldots) + \_)(b_0x^0 + \ldots) = \ldots$

Zu (3). ist klar.

□

Beispiel 3.2.5. $\mathbb{Q}[\sqrt{2}] = \left\{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\right\}$, denn „$\supseteq$“ ist klar und „$\subseteq$“ gilt, da auf der rechten Seite ein Unterring von $\mathbb{R}$ steht [$\to$ 3.2.3 (b)], der $\mathbb{Q} \cup \left\{\sqrt{2}\right\}$ enthält und $\mathbb{Q}[\sqrt{2}]$ der kleinste solche ist [$\to$ 3.2.4].

Definition 3.2.6. Sei A ein kommutativer Ring. Ein kommutativer Ring B heißt *Polynomring* über A in x, wenn $B = A[x]$ und

$$\forall n \in \mathbb{N}_0 : \forall a_0, \ldots, a_n \in A : \left(\sum_{k=0}^{n} a_k x^k = 0 \Longrightarrow a_0 = \ldots = a_n = 0\right).$$

Man nennt dann die Elemente von B *Polynome* der *Unbestimmten oder Variablen* x. Ist $p = \sum_{k=0}^{n} a_k x^k$ ($a_k \in A$) ein Polynom, so ist der *k-te Koeffizient* a_k von p eindeutig bestimmt. (denn $\sum_k a_k x^k = \sum_k b_k x^k$ mit $a_k, b_k \in A$ impliziert $\sum_k (a_k - b_k)x^k = 0$ und damit $a_k - b_k = 0$ für alle k, d.h. $a_k = b_k$ für alle k).
Ist $p = \sum_{k=0}^{n} a_k x^k$ ($a_k \in A$) mit $a_n \neq 0$, so heißt a_n der *Leitkoeffizient* oder *höchste Koeffizient* von p und $\underbrace{\deg}_{\text{„degree“}}\ p := n$ der *Grad* von p. Wir setzen $\deg(0) := -\infty$.

Beispiel 3.2.7. Ist $A = \{0\}$ ein einelementiger Ring [$\to$ 3.1.3, 3.1.4 (a)], so ist A ein Polynomring über sich selber in 0.

Definition 3.2.8. [$\to$ 2.2.9] Seien $(A, +_A, \cdot_A)$ und $(B, +_B, \cdot_B)$ kommutative Ringe. Dann heißt f ein (Ring-)*Homomorphismus* von $(A, +_A, \cdot_A)$ nach $(B, +_B, \cdot_B)$, wenn f ein Gruppenhomomorphismus von $(A, +_A)$ nach $(B, +_B)$ ist mit $\underline{f(1_A) = 1_B}$ und $\forall a, b \in A$: $f(a \cdot_A b) = f(a) \cdot_B f(b)$.

Beispiel 3.2.9. (a) $f : \mathbb{R} \to \mathbb{R}, m \mapsto 0$ ist kein Ringhomomorphismus, da $f(1) = 0 \neq 1$.

(b) $f : \mathbb{Q} \to \mathbb{R}, x \mapsto x$ ist ein Ringhomomorphismus.

(c) Sei A eine Menge und $B \subseteq A$. Wie in 3.1.4 (c) machen wir $\mathscr{P}(A)$ und $\mathscr{P}(B)$ zu einem kommutativen Ring vermöge der symmetrischen Mengendifferenz als Addition und dem Schnitt als Multiplikation. Dann ist $\mathscr{P}(A) \to \mathscr{P}(B), C \mapsto C \cap B$ ein Ringhomomorphismus (nachrechnen!).

Definition 3.2.10. [→ 2.2.12] Ein Ringhomomorphismus $f : A \to B$ heißt (Ring-) $\left\{\begin{matrix}\textit{Einbettung oder Mono-}\\ \textit{Epi-}\\ \textit{Iso-}\end{matrix}\right\}$*morphismus*, wenn f $\left\{\begin{matrix}\text{injektiv}\\ \text{surjektiv}\\ \text{bijektiv}\end{matrix}\right\}$ ist. Einen Ringhomomorphismus $f : A \to A$ nennt man ach einen (Ring-)*Endomorphismus* von A und, falls er bijektiv ist, (Ring-)*Automorphismus* von A.

Satz 3.2.11. *Sei A ein kommutativer Ring mit $0 \neq 1$ und $x \notin A$. Dann gibt es einen Polynomring über A in x.*

Beweis. Betrachte $(A^{\mathbb{N}_0}, +, *)$ wie in 3.1.5. Es ist $\varphi : A \to A^{\mathbb{N}_0}, a \mapsto (a, 0, 0, 0, \ldots)$ [→ 1.1.30 (c)] eine Ringeinbettung. Daher ist $\widehat{\varphi} : A \to \operatorname{im}\varphi$ ein Ringisomorphismus und $A' := \operatorname{im}\varphi = \{(a, 0, 0, 0, \ldots) \mid a \in A\}$ ein Unterring von $A^{\mathbb{N}_0}$. Da die Behauptung eine *strukturelle* Aussage über A macht (2.2.15 gilt sinngemäß natürlich auch für kommutative Ringe statt abelsche Gruppen), reicht es, die Behauptung für A' statt A zu zeigen, denn $A \cong A'$. Aus ähnlichen Gründen kann man nach Definition 3.2.6 x durch irgendein festes Element außerhalb von A' austauschen. Wegen $0 \neq 1$ gilt $(0, 1, 0, 0, \ldots) \notin A'$ und wir können daher $x = (0, 1, 0, 0, \ldots)$ annehmen. Wir behaupten, dass nun der Unterring $A'[x]$ [→ 3.2.4] von $A^{\mathbb{N}_0}$ ein Polynomring über A' in x ist [→ 3.2.6]. Seien hierzu $n \in \mathbb{N}_0$ und $a_0, \ldots, a_n \in A$ mit $\sum_{k=0}^{n} \widehat{\varphi}(a_k)x^k = 0$. Z.z.: $a_0 = \ldots = a_n = 0$. Durch Induktion nach $k \in \mathbb{N}_0$ zeigt man $x^k = (0, 0, \ldots, 0, 1, 0, 0, \ldots)$. Daher $(a_0, \ldots, a_n, 0, 0, \ldots) = \sum_{k=0}^{n} \widehat{\varphi}(a_k)x^k = 0 = (0, 0, 0, \ldots)$. Also $a_0 = \ldots = a_n = 0$. □

Bemerkung 3.2.12. Schreibt man $A[X]$ mit großem X, so meint man meist stillschweigend, dass $A[X]$ ein Polynomring über A in X ist.

Satz 3.2.13. *Seien A und B kommutative Ringe und $\varphi : A \to B$ ein Ringhomomorphismus. Sei $x \in B$. Dann gibt es genau einen Ringhomomorphismus $\psi : A[X] \to B$ mit $\psi|_A = \varphi$ und $\psi(X) = x$. Für jedes Polynom $p = \sum_k a_k X^k$ $(a_k \in A)$ gilt $\psi(p) = \sum_k \varphi(a_k)x^k$.*

Beweis. Offensichtlich muss ψ so definiert werden, wenn es ein Ringhomomorphismus mit $\psi|_A = \varphi$ und $\psi(X) = x$ sein soll. Es bleibt nur zu zeigen, dass ψ ein solcher ist. Zu zeigen:

(a) $\psi|_A = \varphi$,

(b) $\psi(X) = x$,

(c) $\psi(1) = 1$,

(d) $\forall p, q \in A[X] : \psi(p + q) = \psi(p) + \psi(q)$,

(e) $\forall p, q \in A[X] : \psi(pq) = \psi(p)\psi(q)$

(a), (b), (c) sind klar nach Definition von ψ.
Für (d) und (e) seien $a_k, b_k \in A$ beliebig.

Zu (c): $\psi(1) = \psi(1 \cdot X^0) \overset{\text{Def. v. } \psi}{=} \varphi(1) \cdot x^0 \overset{\varphi(1)=1}{=} 1 \cdot x^0 \overset{x^0=1}{=} 1 \cdot 1 = 1$

Zu (d):

$$\begin{aligned}\psi\left(\sum_k a_k X^k + \sum_k b_k X^k\right) &\overset{2.1.10,\ (D)}{=} \psi\left(\sum_k (a_k + b_k) X^k\right) \overset{\text{Def. v. } \psi}{=} \sum_k \varphi(a_k + b_k) x^k \\ &\overset{\varphi \text{ Hom.}}{=} \sum_k (\varphi(a_k) + \varphi(b_k)) x^k \\ &\overset{(D),\ 2.1.10}{=} \sum_k \varphi(a_k) x^k + \sum_k \varphi(b_k) x^k \\ &\overset{\text{Def. v. } \psi}{=} \psi\left(\sum_k a_k X^k\right) + \psi\left(\sum_k b_k X^k\right)\end{aligned}$$

Zu (e):

$$\begin{aligned}\psi\left(\left(\sum_k a_k X^k\right)\left(\sum_k b_k X^k\right)\right) &\overset{(D)}{=} \psi\left(\sum_k \left(\sum_{i=0}^{k} a_i b_{k-i}\right) X^k\right) \\ &\overset{\text{Def. v. } \psi}{=} \sum_k \varphi\left(\sum_{i=0}^{k} a_i b_{k-i}\right) x^k \\ &= \sum_k \left(\sum_{i=0}^{k} \varphi(a_i)\varphi(b_{k-i})\right) x^k \\ &\overset{(D)}{=} \left(\sum_k \varphi(a_k) x^k\right)\left(\sum_k \varphi(b_k) x^k\right)\end{aligned}$$

□

Korollar 3.2.14. *Sei A ein Unterring des kommutativen Ringes B und $x \in B$. Dann gibt es genau einen Ringhomomorphismus $\psi : A[X] \to B$ mit $\psi(a) = a$ für $a \in A$ und $\psi(X) = x$. Für jedes Polynom $p = \sum_k a_k X^k$ ($a_k \in A$) gilt $\psi(p) = \sum_k a_k x^k$. Insbesondere gilt* $\operatorname{im} \psi = A[x]$.

Notation 3.2.15. In der Situation von 3.2.14 schreibt man auch $p(x)$ statt $\psi(p)$ (obwohl p keine Funktion, sondern ein Polynom ist).
Da ψ ein Ringhomomorphismus ist, gilt dann $(p+q)(x) = p(x)+q(x)$, $(pq)(x) = p(x)q(x)$ und $1(x) = 1$ für $p, q \in A[X]$.

Satz 3.2.16. *Seien $A[X]$ und $A[Y]$ Polynomringe über dem kommutativen Ring A in X bzw. Y. Dann gibt es genau einen Ringisomorphismus $\psi : A[X] \to A[Y]$ mit $\psi(a) = a$ für $a \in A$ und $\psi(X) = Y$.*

Beweis. Nach 3.2.14 gibt es genau einen Ringhomomorphismus $\psi : A[X] \to A[Y]$ mit $\psi(a) = a$ für $a \in A$ und $\psi(X) = Y$. Offensichtlich gilt $\operatorname{im}\psi = A[Y]$, d.h. ψ ist surjektiv. Noch zu zeigen: ψ injektiv.
Es reicht $\ker\psi = \{0\}$ zu zeigen. Gelte hierzu $\psi\left(\sum_k a_k X^k\right) = 0$ $(a \in A)$. Dann $\sum_k a_k Y^k = 0$, also $a_k = 0$ für alle k, da $A[Y]$ Polynomring über A in Y. Daher $\sum_k a_k X^k = 0$ wie gewünscht. □

Sprechweise 3.2.17. Wegen 3.2.16 spricht man oft auch von <u>dem</u> Polynomring $A[X]$ über A in X.

§3.3 Ideale und Quotientenringe [→ §1.3, §2.3]

Idee: Grobe Sichtweise auf kommutative Ringe einnehmen.

Definition 3.3.1. Sei A ein kommutativer Ring. Eine *Kongruenzrelation* auf A ist eine Kongruenzrelation $\equiv$ auf der additiven Gruppe von A [→2.3.1], für die gilt:

$$(*) \qquad \forall a, a', b, b' \in A : ((a \equiv a' \ \&\ b \equiv b') \implies ab \equiv a'b')$$

Bemerkung 3.3.2. In Definition 3.3.1 drückt Bedingung $(*)$ gerade aus, dass

$$\begin{aligned} A/\!\equiv \ \times \ A/\!\equiv \ &\to \ A/\!\equiv \\ (\overline{a}, \overline{b}) &\mapsto \overline{ab} \quad (a, b \in A) \end{aligned}$$

wohldefiniert ist.

Satz und Definition 3.3.3. *Sei A ein kommutativer Ring und $\equiv$ eine Kongruenzrelation auf A. Dann wird die Quotientengruppe $A/\!\equiv$ [→2.3.3] vermöge der durch*

$$\overline{a}\,\overline{b} := \overline{ab} \qquad (a, b \in G)$$

festgelegten („vertreterweisen“) Multiplikation zu einem kommutativen Ring, den man den zu $\equiv$ gehörigen Quotientenring *von A nennt (auch „A nach $\equiv$“ oder „A modulo $\equiv$“). In ihm gilt $1 = \overline{1}$.*

Beweis. Aus 2.3.3 wissen wir schon, dass A bezüglich der Addition eine abelsche Gruppe bildet. Es sind daher nur noch (K̇), (Ȧ), (Ṅ) und (D) aus Definition 3.1.1 nachzurechnen:

(K̇) $\overline{a}\,\overline{b} = \overline{ab} = \overline{ba} = \overline{b}\,\overline{a}$ für alle $a, b \in A$.

(Ȧ) $(\overline{a}\,\overline{b})\,\overline{c} = \overline{ab}\,\overline{c} = \overline{(ab)c} = \overline{a(bc)} = \overline{a}\,\overline{bc} = \overline{a}\,(\overline{b}\,\overline{c})$ für alle $a, b, c \in A$.

(Ṅ) Für $a \in A$ gilt $\overline{a}\,\overline{1} = \overline{a1} = \overline{a}$. Daher ist $1 = \overline{1}$.

(D) $\overline{a}\,(\overline{b} + \overline{c}) = \overline{a}\,\overline{b + c} = \overline{a(b + c)} = \overline{ab + ac} = \overline{ab} + \overline{ac} = \overline{a}\,\overline{b} + \overline{a}\,\overline{c}$ für alle $a, b, c \in A$.

□

Definition 3.3.4. Sei A ein kommutativer Ring. Eine Untergruppe I der additiven Gruppe von A [→3.1.2(c)] heißt *Ideal* von A, wenn $\forall a \in A : \forall b \in I : ab \in I$.

Proposition 3.3.5. *Sei A ein kommutativer Ring und $I \subseteq A$. Genau dann ist I ein Ideal von A, wenn folgende Bedingungen gelten:*

(a) $0 \in I$

(b) $\forall a, b \in I : a + b \in I$

(c) $\forall a \in A : \forall b \in I : ab \in I$

Beweis. Dies folgt direkt aus 2.2.2 und 3.3.4, denn wenn (c) gilt, so ist Bedingung 2.2.2(d) automatisch erfüllt, da dann

$$-a = -(a1) \overset{3.1.2(g)}{=} a(-1) = (-1)a \overset{(c)}{\in} I$$

für alle $a \in I$. □

Satz 3.3.6. [→2.3.6] *Sei A ein kommutativer Ring. Die Zuordnungen*

$$\equiv \mapsto \overline{0}$$
$$\equiv_I \hookleftarrow I$$

vermitteln eine Bijektion zwischen der Menge der Kongruenzrelationen auf A und der Menge der Ideale von A.

Beweis. Zu zeigen ist:

(a) Ist $\equiv$ eine Kongruenzrelation auf A, so ist $\overline{0}$ ein Ideal von A.

(b) Ist I ein Ideal von A, so ist $\equiv_I$ eine Kongruenzrelation auf A.

(c) Ist $\equiv$ eine Kongruenzrelation auf A, so $\equiv_{\overline{0}} = \equiv$.

(d) Ist I ein Ideal von A, so $\overline{0}^{-I} = I$.

Zu (a). Sei $\equiv$ eine Kongruenzrelation auf A. Wir wissen aus 2.3.4 (oder 2.3.6) schon, dass $\overline{0}$ eine Untergruppe von A ist. Gemäß Definition 3.3.4 bleibt $\forall a \in A : \forall b \in \overline{0} : ab \in \overline{0}$ zu zeigen. Seien hierzu $a \in A$ und $b \in \overline{0}$. Dann $a \equiv a$ und $b \equiv 0$, woraus mit 3.3.1$(*)$ folgt

$$ab \equiv a0 \overset{3.1.2(g)}{=} 0,$$

das heißt $ab \in \overline{0}$.

Zu (b). Sei I eine Ideal von A. Wir wissen aus 2.3.6 schon, dass $\equiv_I$ ein Kongruenzrelation der additiven Gruppe von A ist. Es bleibt 3.3.1$(*)$ zu zeigen. Seien hierzu $a, a', b, b' \in A$ mit $a \equiv_I a'$ und $b \equiv_I b'$. Zu zeigen ist $ab \equiv_I a'b'$. Nun gilt $b - b' \in I$ und daher auch $ab - ab' = a(b - b') \in I$. Genauso gilt $a - a' \in I$ und damit auch $b'a - b'a' = b'(a - a') \in I$. Hiermit $ab \equiv_I ab' = b'a \equiv_I b'a' = a'b'$ und daher $ab \equiv_I a'b'$.

(c) und (d) folgen aus Satz 2.3.6. □

Definition 3.3.7. [→2.3.7] Sei I ein Ideal des kommutativen Ringes A. Dann nennt man $A/I := A/\equiv_I$ den Quotientenring (oder *Restklassenring*) von A nach (oder modulo) I. Die Kongruenzklassen $\bar{a}^A$ ($a \in A$) [→2.3.1] von $\equiv_I$ nennt man manchmal auch die *Restklassen* von I (in A).

Proposition 3.3.8. [→2.2.5] *Sei A ein kommutativer Ring und M eine Menge von Idealen von A. Dann ist auch $\bigcap M$ ein Ideal von A (mit $\bigcap \emptyset := A$).*

Beweis. Nach 2.2.5 wissen wir schon, dass $\bigcap M$ eine Untergruppe der additiven Gruppe von A ist. Nach Definition 3.3.4 bleibt $\forall a \in A : \forall b \in \bigcap M : ab \in \bigcap M$ zu zeigen. Seien hierzu $a \in A$ und $b \in \bigcap M$. Dann gilt für jedes $I \in M$, dass $b \in I$ und damit $ab \in I$, weil I ein Ideal ist. Also gilt $ab \in \bigcap M$. □

Satz und Definition 3.3.9. [→2.2.6] *Sei A ein kommutativer Ring und $E \subseteq A$. Dann gibt es* das kleinste *Ideal I von A mit $E \subseteq I$. Man nennt es das von E in A* erzeugte Ideal *und notiert es mit $(E)_A$ oder (etwas schlampig) mit (E).*

Beweis. Völlig analog zum Beweis von 2.2.6 (benutze 3.3.8 statt 2.2.5). □

Satz 3.3.10. [→2.2.7] *Sei A ein kommutativer Ring und $E \subseteq A$. Dann gilt*

$$(E)_A = \left\{ \sum_{i=1}^{m} a_i b_i \mid m \in \mathbb{N}_0, a_1, \ldots, a_m \in A, b_1, \ldots, b_m \in E \right\}.$$

Beweis. Der Beweis ist völlig analog zum Beweis von Satz 2.2.7 und stellt gleichzeitig einen neuen Beweis für 3.3.9 dar! □

Definition 3.3.11. Sei A ein kommutativer Ring. Dann nennt man Ideale von A der Form

$$(a_1, \ldots, a_n) := (a_1, \ldots, a_n)_A := (\{a_1, \ldots, a_n\})_A \overset{(D)}{=} \left\{ \sum_{i=1}^{m} b_i a_i \mid b_1, \ldots, b_m \in A \right\}$$

mit $m \in \mathbb{N}_0$ und $a_1, \ldots, a_m \in A$ *endlich erzeugt* (e.e.) und Ideale der Form

$$(a) = \{ba \mid b \in A\}$$

mit $a \in A$ *Hauptideale* von A.

Beispiel 3.3.12. (a) Für $n \in \mathbb{Z}$ gilt $(n) = \{bn \mid b \in \mathbb{Z}\} = \langle n \rangle$. Ist $n > 0$, so gilt

$$\mathbb{Z}/(n) = \left\{ \bar{a}^{(n)} \mid a \in \mathbb{Z} \right\}$$

und $\mathbb{Z}/(n)$ hat genau n Elemente. Die additive Gruppe des Ringes $\mathbb{Z}/(n)$ ist die abelsche Gruppe $\mathbb{Z}/\langle n \rangle$.

(b) Betrachte $\mathbb{Z}/(9)$ und schreibe kurz $\overline{a}$ statt $\overline{a}^{(9)} = \overline{a}^{\equiv_{(9)}}$. Dann $\mathbb{Z}/(9) = \{\overline{0}, \ldots, \overline{8}\}$, $\overline{10} = 1$, $\overline{100} = \overline{10}\ \overline{10} = \overline{1}\ \overline{1} = \overline{1} = 1$, $\overline{1000} = 1$ und so weiter. Es gilt $\overline{17368} = \overline{1} + \overline{7} + \overline{3} + \overline{6} + \overline{8} = \overline{1} + \overline{8} + \overline{3} + \overline{6} + \overline{7} = \overline{9} + \overline{9} + \overline{7} = \overline{7}$. Daher gilt $17368 \equiv_{(9)} 7$. Der Rest von 17368 bei Division durch 9 ist also 7.

Satz 3.3.13. *Jedes Ideal von $\mathbb{Z}$ ist ein Hauptideal von $\mathbb{Z}$.*

Beweis. Sei I ein Ideal von $\mathbb{Z}$. Ist $I = \{0\}$, so ist $I = (0)$. Also bleibt nur der Fall zu betrachten, dass es ein $n \in \mathbb{N}$ gibt mit $n \in I$. Wähle dann das kleinste solche n. Wir behaupten nun $I = (n)$. Die Inklusion $I \supseteq (n)$ ist klar. Um $I \subseteq (n)$ zu beweisen, sei $a \in I$. Zu zeigen ist $a \in (n)$. Da $\mathbb{Z}/(n) = \left\{\overline{0}^{(n)}, \ldots, \overline{n-1}^{(n)}\right\}$, gibt es $b \in \{0, \ldots, n-1\}$ mit $a \equiv_{(n)} b$, das heißt $a - b \in (n) \subseteq I$. Wir wissen $b \in I$ (denn $b \overset{a-b\in I}{\equiv_I} a \overset{a\in I}{\equiv_I} 0$) und damit sogar $b = 0$, da n kleinstmöglich gewählt wurde. Es folgt $a \equiv_{(n)} b = 0$ und daher $a \in (n)$. □

Korollar 3.3.14. *Sei H eine Untergruppe von $\mathbb{Z}$. Dann gilt $H = \langle n \rangle$ für ein $n \in \mathbb{Z}$.*

Beweis. Jede Untergruppe der abelschen Gruppe $(\mathbb{Z}, +)$ ist ein Ideal des kommutativen Ringes $(\mathbb{Z}, +, \cdot)$, denn Multiplizieren mit einer ganzen Zahl lässt sich durch iteriertes Addieren oder Subtrahieren ausdrücken. □

Definition und Proposition 3.3.15. [→2.3.10] Seien A und B kommutative Ringe und $f\colon A \to B$ ein Homomorphismus. Dann ist $\equiv_f$ eine Kongruenzrelation auf A und $\ker f$ ein Ideal von A.

Beweis. Da sich $\equiv_f$ und $\ker f$ nach 2.3.10 unter der Bijektion aus Satz 3.3.6 entsprechen, reicht es eine der beiden Behauptungen zu zeigen. Wir entscheiden uns für die zweite. Dass $\ker f$ eine Untergruppe der additiven Gruppe von A ist, wurde schon in 2.3.10(b) gezeigt. Nach Definition 3.3.4 reicht es daher $\forall a \in A : \forall b \in \ker f : ab \in \ker f$ zu zeigen. Sind nun $a \in A$ und $b \in \ker f$, so gilt $f(b) = 0$ und daher

$$f(ab) = f(a)f(b) = f(a)\ 0 \overset{3.1.2(g)}{=} 0,$$

also $ab \in \ker f$. □

Satz 3.3.16 (Homomorphiesatz für kommutative Ringe)**.** [→2.3.11] *Seien A und B kommutative Ringe, I ein Ideal von A und $f\colon A \to B$ ein Homomorphismus mit $I \subseteq \ker f$.*

(a) *Es gibt genau eine Abbildung $\overline{f}\colon A/I \to B$ mit $\overline{f}(\overline{a}^I) = f(a)$ für alle $a \in A$. Diese Abbildung $\overline{f}$ ist ein Homomorphismus.*

(b) *$\overline{f}$ ist injektiv $\iff$ $I = \ker f$*

(c) *$\overline{f}$ ist surjektiv $\iff$ f ist surjektiv.*

Beweis. Existenz und Eindeutigkeit der Abbildung $\overline{f}$ erhält man dem Homomorphiesatz für abelsche Gruppen 2.3.11. Dass $\overline{f}$ ein Gruppenhomomorphismus ist, wissen wir ebenfalls aus 2.3.11. Wir rechnen nach, dass $\overline{f}$ ein Ringhomomorphismus [→3.2.8] ist, wobei wir $\equiv := \equiv_I$ und damit $\overline{a} = \overline{a}^I = \overline{a}^{\equiv_I}$ für alle $a \in A$ schreiben:

$$\overline{f}(1) \overset{3.3.3}{=} \overline{f}(\overline{1}) \overset{\text{Def. von } \overline{f}}{=} f(1) \underset{3.2.8}{\overset{f \text{ Hom.}}{=}} 1 \qquad \text{und}$$

und

$$\overline{f}(\overline{a}\,\overline{b}) \overset{3.3.3}{=} \overline{f}(\overline{ab}) \overset{\text{Def. von } \overline{f}}{=} f(ab) \underset{3.2.8}{\overset{f \text{ Hom.}}{=}} f(a)f(b) \overset{\text{Def. von } \overline{f}}{=} \overline{f}(\overline{a})\,\overline{f}(\overline{b})$$

für alle $a, b \in A$. Teil (b) und (c) der Behauptung folgen unmittelbar aus 2.3.11. □

Bemerkung 3.3.17. [→2.3.12] Sei I ein Ideal des kommutativen Ringes A. Dann wird I durch einen Ringhomomorphismus $f\colon A \to B$ in einen weiteren kommutativen Ring B induziert, nämlich durch den *kanonischen Epimorphismus*

$$f\colon A \to A/I,\ a \mapsto \overline{a}^I.$$

In der Tat gilt $\ker f = I$ nach 2.3.12.

Proposition 3.3.18. [→2.3.13] *Seien A und B kommutative Ringe und $f\colon A \to B$ ein Ringhomomorphismus. Dann ist* $\operatorname{im} f$ *ein Unterring von B.*

Beweis. Aus 2.3.13 wissen wir schon, dass $\operatorname{im} f$ eine Untergruppe der additiven Gruppe von B ist. Zu zeigen sind dann gemäß 3.2.2 noch:

(a) $1 \in \operatorname{im} f$

(b) $\forall a, b \in \operatorname{im} f : ab \in \operatorname{im} f$

(a) folgt daraus, dass nach der Definition 3.2.8 eines Ringhomomorphismus gilt $f(1) = 1$. Um (b) zu zeigen, seien $a, b \in \operatorname{im} f$. Wähle dann $c, d \in G$ mit $f(c) = a$ und $f(d) = b$. Es folgt $f(cd) \overset{f \text{ Hom.}}{=} f(c)f(d) = ab$ und damit $ab \in \operatorname{im} f$. □

Korollar 3.3.19 (Isomorphiesatz für kommutative Ringe)**.** [→2.3.14] *Seien A und B kommutative Ringe und $f\colon A \to B$ ein Homomorphismus. Dann ist $\overline{f}\colon A/\ker f \to \operatorname{im} f$ definiert durch $\overline{f}(\overline{a}^{\ker f}) = f(a)$ für $a \in A$ ein Isomorphismus* [→3.2.10]*. Insbesondere $A/\ker f \cong \operatorname{im} f$ (in der zu 2.2.15 analogen Bedeutung).*

§4 Körper

§4.1 Definitionen und Beispiele von Körpern

Definition und Proposition 4.1.1. Sei $(A, +, \cdot)$ ein kommutativer Ring [$\to$ 3.1.1]. Die Elemente von $A^\times := \{a \in A \mid \exists b \in A : ab = 1\}$ nennt man *Einheiten* oder *invertierbare Elemente* von A. Es ist $(A^\times, \cdot)$ mit $\cdot : A^\times \times A^\times \to A^\times, (a, b) \mapsto ab$ eine abelsche Gruppe.

Beweis. $\cdot : A^\times \times A^\times \to A^\times$ ist wohldefiniert, denn sind $a, b \in A^\times$, so auch $ab \in A^\times$. In der Tat: Seien $a, b \in A^\times$. Dann gibt es $a', b' \in A$ mit $aa' = 1 = bb'$. Es folgt $(ab)(a'b') \overset{3.1.2\,(e)}{=} (aa')(bb') = 1 \cdot 1 \overset{(\dot{\mathrm{N}})}{=} 1$, also $ab \in A^\times$. Die Axiome (K), (A), (N) für $(A^\times, \cdot)$ [$\to$ 2.1.1] folgen aus den Axiomen ($\dot{\mathrm{K}}$), ($\dot{\mathrm{A}}$), ($\dot{\mathrm{N}}$) für $(A, +, \cdot)$ [$\to$ 3.1.1], wobei $1 \in A^\times$ zu beachten ist. Um schließlich (I) für $(A^\times, \cdot)$ zu zeigen, sei $a \in A^\times$. Dann gibt es $b \in A$ mit $ab = 1$. Es gilt aber $ba \overset{(\dot{\mathrm{K}})}{=} ab = 1$, also $b \in A^\times$. Also haben wir $b \in A^\times$ gefunden mit $ab = 1$. □

Bemerkung 4.1.2. In jedem kommutativen Ring $(A, +, \cdot)$ „steckt“ also nicht nur die additive abelsche Gruppe $(A, +)$, sondern auch die multiplikativ geschriebene Einheitengruppe $(A^\times, \cdot)$. Oft ist $A^\times$ viel kleiner als A.

Beispiel 4.1.3. (a) $\mathbb{Z}^\times = \{-1, 1\}, \mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}, \mathbb{R}^\times = \mathbb{R} \setminus \{0\}$

(b) $\overline{3} \cdot \overline{7} = \overline{1} = 1$ in $\mathbb{Z}/(10)$, denn $3 \cdot 7 \equiv_{(10)} 1$, also $\overline{3}, \overline{7} \in (\mathbb{Z}/(10))^\times$
$\overline{2} \cdot \overline{5} = \overline{0} = 0$ in $\mathbb{Z}/(10)$, also $\overline{2}, \overline{5} \notin (\mathbb{Z}/(10))^\times$
$\overline{1} \cdot \overline{1} = 1$ in $\mathbb{Z}/(10)$, also $\overline{1} \in (\mathbb{Z}/(10))^\times$
$\overline{4} \cdot \overline{5} = \overline{6} \cdot \overline{5} = \overline{8} \cdot \overline{5} = 0$, also $\overline{4}, \overline{6}, \overline{8} \notin (\mathbb{Z}/(10))^\times$
$\overline{9} = \overline{-1} \in (\mathbb{Z}/(10))^\times$, da $\overline{-1} \cdot \overline{-1} = \overline{1} = 1$.
Insgesamt $(\mathbb{Z}/(10))^\times = \{\overline{1}, \overline{3}, \overline{7}, \overline{9}\} \subseteq \mathbb{Z}/(10) = \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{9}\}$.

Definition 4.1.4. Ein kommutativer Ring A heißt *Körper*, wenn $A^\times = A \setminus \{0\}$.

Beispiel 4.1.5. (a) Ein einelementiger Ring $A = \{0\} = \{1\}$ [$\to$ 3.1.3, 3.1.4 (a)] ist kein Körper, denn $A^\times = A$.

(b) $\mathbb{Z}/(2)$ und $\mathbb{Z}/(3)$ sind Körper.

(c) $\mathbb{Z}/(4)$ ist kein Körper, denn $\overline{2} \cdot \overline{2} = 0$ und daher $\overline{2} \notin (\mathbb{Z}/(4))^\times$.

(d) $\mathbb{Z}$ ist kein Körper.

(e) $\mathbb{Q}$ und $\mathbb{R}$ sind Körper.

Definition 4.1.6. Wir nennen $n \in \mathbb{N}$ mit $n \geq 2$ eine *Primzahl*, wenn es keine $s, t \in \mathbb{N}$ mit $s, t \geq 2$ und $n = st$ gibt. Wir schreiben $\mathbb{P} = \{2, 3, 5, 7, 11, 13, 17, \ldots\}$ für die Menge der Primzahlen.

Satz 4.1.7. *Sei $n \in \mathbb{N}_0$. Dann $\mathbb{Z}/(n)$ Körper $\iff$ $n \in \mathbb{P}$.*

Beweis. Fall 1: $n = 0$.

$$\mathbb{Z}/(n) = \mathbb{Z}/(0) = \{\{m\} \mid m \in \mathbb{Z}\} \cong \mathbb{Z} \text{ (vgl. 2.3.8 (b))}$$

Da $\mathbb{Z}$ kein Körper ist und $\mathbb{Z}/(n) \cong \mathbb{Z}$, ist $\mathbb{Z}/(n)$ auch kein Körper (vgl. 2.2.15). Gleichzeitig ist auch $n = 0 \notin \mathbb{P}$. Also sind beide Aussagen falsch und damit äquivalent.

Fall 2: $n = 1$. $\mathbb{Z}/(1) = \{0\}$ (vgl. 2.3.8 (a)) ist kein Körper [$\to$ 4.1.5 (a)]. Gleichzeitig $n = 1 \notin \mathbb{P}$.

Fall 3: $n \in \mathbb{N}, n \geq 2$.

„$\Longrightarrow$“: Sei $\mathbb{Z}/(n)$ ein Körper. Zu zeigen: $n \in \mathbb{P}$.
Seien $s, t \in \mathbb{N}$ mit $n = st$. Zu zeigen: $s = 1$ oder $t = 1$. Beachte $s = 1 \iff t = n$ und $t = 1 \iff s = n$.
Annahme: Weder $t = n$ noch $s = n$.
Dann gilt $1 \leq t \leq n - 1$ und $1 \leq s \leq n - 1$.
Also $\overline{s}, \overline{t} \in \left\{\overline{1}, \ldots, \overline{n-1}\right\} = (\mathbb{Z}/(n))^\times$.
$0 = \overline{n} = \overline{st} = \overline{s}\overline{t} \in (\mathbb{Z}/(n))^\times$, da $(\mathbb{Z}/(n))^\times$ abelsche Gruppe $\lightning$.

„$\Longleftarrow$“: Sei $n \in \mathbb{P}$. Zu zeigen: $A := \mathbb{Z}/(n)$ Körper. Zu zeigen $A^\times = A \setminus \{0\}$.
„$\subseteq$“ klar, da $0a = 0 \neq 1$ für $a \in A$, also $0 \notin A^\times$.
„$\supseteq$“: Sei $a \in A \setminus \{0\}$. Zu zeigen: $a \in A^\times$.

$$a \in A^\times \iff 1 \in (a) \iff A = (a) \iff \#(a) = \#A \iff \#(a) = n$$

Da (a) eine Untergruppe der additiven Gruppe von A ist, gilt nach 2.3.7

$$(*) \qquad \underbrace{(\#(a))}_{\substack{\geq 2 \\ \text{wegen } \{0,a\} \subseteq (a) \\ \text{und } 0 \neq a}} \quad (\#(A/(a))) = \#A = n \in \mathbb{P}$$

Aus $(*)$ folgt daher $\#(a) = n$ wie gewünscht.

□

Korollar 4.1.8 („Lemma von Euklid“)**.** [Euklid von Alexandria ≈ -300] *Sei $n \in \mathbb{N}$ mit $n \geq 2$. Dann*

$$n \in \mathbb{P} \iff \forall a, b \in \mathbb{N} : (ab \in (n) \Longrightarrow (a \in (n) \text{ oder } b \in (n)))$$

Beweis. „$\Longleftarrow$“: Gelte die Bedingung rechts und seien $s,t \in \mathbb{N}$ mit $n = st$. Zu zeigen: $s = 1$ oder $t = 1$. Nun $st = n \in (n)$ und daher $s \in (n)$ oder $t \in (n)$. Wäre $s \neq 1$ und $t \neq 1$, so $s < n$ und $t < n$ und damit $s = t = 0$ $\lightning$.

„$\Longrightarrow$“: Gelte $n \in \mathbb{P}$ und seien $a, b \in \mathbb{N}$ mit $a \notin (n)$ und $b \notin (n)$. Zu zeigen: $a, b \notin (n)$. Wegen $\overline{a} \neq 0$ und $\overline{b} \neq 0$ in $\mathbb{Z}/(n)$ gilt nach 4.1.7 $\overline{a}, \overline{b} \in (\mathbb{Z}/(n))^\times$ und daher nach 4.1.1 $\overline{ab} = \overline{a}\overline{b} \in (\mathbb{Z}/(n))^\times$. Insbesondere $\overline{ab} \neq 0$ in $\mathbb{Z}/(n)$.

□

Bemerkung 4.1.9. Mit dem „Wissen“ über „Primfaktorzerlegungen“ aus der Schule ist 4.1.8 auch klar, aber wurde dort dieses „Wissen“ begründet? Mit 4.1.8 kann man es begründen!
Wir werden dies aber später viel allgemeiner machen!

Notation 4.1.10. $\mathbb{F}_p := \mathbb{Z}/(p)$ für $p \in \mathbb{P}$.

Notation 4.1.11. Sei A ein kommutativer Ring, $a \in A$ und $b \in A^\times$. $\frac{a}{b} := ab^{-1}$ [→ 2.1.2 (d)]

Beispiel 4.1.12. $\frac{\overline{3}}{\overline{4}} = \overline{3} \cdot \overline{2} = \overline{6}$ in $\mathbb{F}_7$, da $\overline{4}^{-1} = \overline{2}$, denn $\overline{2} \cdot \overline{4} = \overline{8} = \overline{1} = 1$.

§4.2 Die komplexen Zahlen

[Leonhard Euler *1707 †1783]

Definition 4.2.1. Sei A ein kommutativer Ring. Ein Element $a \in A$ heißt eine *imaginäre Einheit* oder *Wurzel aus* -1 in A, wenn $a^2 = -1$.

Bemerkung 4.2.2. Schreiben wir $\mathring{i}$, so meinen wir meist stillschweigend, dass $\mathring{i}$ eine imaginäre Einheit ist (genauso für $\mathring{j}$).

Satz 4.2.3. *Sei K ein Körper, der keine imaginäre Einheit besitzt.*

(a) *Es gibt einen kommutativen Ring C und (eine imaginäre Einheit) $\mathring{i} \in C$ mit* [→3.2.4]

$$C = K[\mathring{i}].$$

(b) *Gilt $C = K[\mathring{i}]$, so $C = \{a + b\mathring{i} \mid a, b \in K\}$ und für alle $a, b \in K$ gilt:*

$$a + b\mathring{i} = 0 \iff a = b = 0.$$

(c) *Gilt $C = K[\mathring{i}]$ und $D = K[\mathring{j}]$, so gibt es genau einen Ringisomorphismus $\varphi\colon C \to D$ mit $\varphi(a) = a$ für alle $a \in K$ und $\varphi(\mathring{i}) = \mathring{j}$.*

Beweis. **(a).** Es ist $\varphi\colon K \to K[X]/(X^2+1),\ a \mapsto \overline{a}^{(X^2+1)}$ eine Ringeinbettung [→3.2.10]. In der Tat: φ ist ein Ringhomomorphismus [→ 3.2.8] (nämlich die Einschränkung [→ 1.1.31] des kanonischen Epimorphismus [→ 3.3.17] von $K[X]$ nach $K[X]/(X^2+1)$) und es gilt $\ker\varphi = \{0\}$, denn ist $a \in K$ mit $\varphi(a) = 0$, so gilt $a \in (X^2+1)$ und damit $a = 0$, denn außer dem Nullpolynom hat jedes Polynom im Hauptideal [→3.3.11] (X^2+1) einen Grad [→3.2.6] ≥ 2. Nun ist $\hat{\varphi}\colon K \to \operatorname{im}\varphi$ ein Ringisomorphismus und $K' := \operatorname{im}\varphi$ ein Unterring von $K[X]/(X^2+1)$ [→3.3.18]. Es reicht, die Behauptung für K' statt K zu zeigen, denn $K \cong K'$ (vergleiche 2.2.15 und Beweis von 3.2.11). Setze

$$\mathring{i} := \overline{X}^{(X^2+1)} \qquad \text{und} \qquad C := K'[\mathring{i}] \subseteq K[X]/(X^2+1).$$

Dann $\mathring{i}^2 = \overline{X}^2 = \overline{X^2} = \overline{-1} = -1$ in C.

(b). Sei $C = K[\mathring{i}]$. Wie in 3.2.5 zeigt man sofort $K[\mathring{i}] = \{a + b\mathring{i} \mid a, b \in K\}$. Seien $a, b \in K$. Zu zeigen ist $a + b\mathring{i} = 0 \iff a = b = 0$. Hier ist „$\Longleftarrow$“ klar. Um „$\Longrightarrow$“ zu zeigen, gelte $a + b\mathring{i} = 0$. Wäre $b \neq 0$, dann wäre $\mathring{i} = -\frac{a}{b} \in K$ im Widerspruch zur Voraussetzung, dass K keine imaginäre Einheit besitzt. Also $b = 0$ und genauso sieht man $a = 0$.

(c). Eindeutigkeit: Es gilt sogar mehr: Gilt $C = K[\mathring{i}]$ und ist D irgendein kommutativer Oberring von C und $j \in D$, so gilt für jeden Ringhomomorphismus $\varphi\colon C \to D$ mit $\varphi(a) = a$ für alle $a \in K$ und $\varphi(\mathring{i}) = j$, dass $\varphi(a + b\mathring{i}) = \varphi(a) + \varphi(b)j$ für alle $a, b \in K$ und hierdurch ist φ eindeutig festgelegt, denn $C \overset{(b)}{=} \{a + b\mathring{i} \mid a, b \in K\}$.

Existenz: Da Umkehrabbildungen und Hintereinanderschaltungen von Ringisomorphismen offensichtlich wieder Ringisomorphismen sind, reicht es einen kommutativen Ring A zu finden so, dass jeder kommutative Ring C von der Form $C = K[\mathring{i}]$ isomorph zu A ist. Wir setzen $A := K[X]/(X^2+1)$. Nach 3.2.14 gibt es einen Ringhomomorphismus $\psi\colon K[X] \to C$ mit $\psi(a) = a$ für alle $a \in K$ und $\psi(X) = \mathring{i}$ (also $\psi\colon K[X] \to C,\ p \mapsto p(\mathring{i})$ in der Notation von 3.2.15). Nun ist $\operatorname{im}\psi$ ein Unterring von C [→3.3.18], der $K \cup \{\mathring{i}\}$ enthält, womit $C = K[\mathring{i}] \subseteq \operatorname{im}\psi \subseteq C$ und damit $\operatorname{im}\psi = C$ gilt. Aus dem Isomorphiesatz für kommutative Ringe 3.3.19 folgt nun $K[X]/\ker\psi \cong \operatorname{im}\psi = C$. Um $A \cong C$ zu zeigen, reicht es also $\ker\psi = (X^2+1)$ zu zeigen. Hier ist „$\supseteq$“ trivial. Um „$\subseteq$“ zu zeigen, sei $p \in \ker\psi$. Wir zeigen $p \equiv_{(X^2+1)} 0$. Offensichtlich gibt es $a, b \in K$ mit $p \equiv_{(X^2+1)} a + bX$ und damit erst recht $p \equiv_{\ker\psi} a + bX$, also $0 = \psi(p) = \psi(a + bX) = a + b\mathring{i}$. Wegen (b) gilt $a = b = 0$ und daher $p \equiv_{(X^2+1)} 0 + 0X = 0$.

Alternativer weniger abstrakter Beweis für (c). Die Eindeutigkeit ist wieder klar nach (b).

Existenz: Gelte $C = K[\mathring{i}]$ und $D = K[\mathring{j}]$. Es ist $\varphi\colon C \to D,\ a + b\mathring{i} \mapsto a + b\mathring{j}$ nach (b) wohldefiniert. Es ist φ ein Homomorphismus [→3.2.8], denn

$$\begin{aligned}\varphi((a + b\mathring{i}) + (c + d\mathring{i})) &= \varphi((a+c) + (b+d)\mathring{i}) = (a+c) + (b+d)\mathring{j} \\ &= (a + b\mathring{j}) + (c + d\mathring{j}) = \varphi(a + b\mathring{i}) + \varphi(c + d\mathring{i}) \text{ und}\end{aligned}$$

$$\varphi((a+b\mathring{i})(c+d\mathring{i})) = \varphi(ac+bd\mathring{i}^2+(ad+bc)\mathring{i}) = \varphi((ac-bd)+(ad+bc)\mathring{i})$$
$$= (ac-bd)+(ad+bc)\mathring{j} = (a+b\mathring{j})(c+d\mathring{j}) \text{ für alle } a,b,c,d \in K$$

und $\varphi(1) = \varphi(1+0\mathring{i}) = 1+0\mathring{j} = 1$. Es ist φ injektiv, denn sind $a, b \in K$ mit $\varphi(a+b\mathring{i}) = 0$, so gilt $a + b\mathring{j} = 0$ und daher nach (b) (angewandt auf $D = K[\mathring{j}]$) $a = b = 0$ und damit $a + b\mathring{i} = 0$. Schließlich ist wieder mit (b) angewandt auf $D = K[\mathring{j}]$ die Abbildung φ auch surjektiv. □

Notation 4.2.4. Ist K ein Körper mit imaginärer Einheit, so setzen wir $K[\mathring{i}] := K$. Ist K ein Körper ohne imaginäre Einheit, so bezeichne $K[\mathring{i}]$ ab jetzt einen fest gewählten kommutativen Ring C mit $C = K[\mathring{i}]$, in dem $\mathring{i}$ eine imaginäre Einheit ist [→4.2.3(a)]. Wegen 4.2.3(c) ist $K[\mathring{i}]$ im Wesentlichen (bis auf Isomorphie) eindeutig bestimmt.

Satz 4.2.5. *Sei K ein Körper. Dann ist auch $K[\mathring{i}]$ ein Körper.*

Beweis. Der Fall, dass K eine imaginäre Einheit besitzt, ist trivial, da dann $K[\mathring{i}] = K$. Besitze also K keine imaginäre Einheit und wende Teil (b) des Satzes an. Seien also $a, b \in K$ mit $a + b\mathring{i} \neq 0$. Dann $(a+b\mathring{i})(a-b\mathring{i}) = a^2+b^2$. Es reicht zu zeigen $a^2+b^2 \neq 0$, denn dann

$$(a+b\mathring{i})\frac{a-b\mathring{i}}{a^2+b^2} = 1$$

und daher $a + b\mathring{i} \in K[\mathring{i}]^\times$ wie gewünscht. Wir nehmen an, dass $a^2+b^2 = 0$ und suchen einen Widerspruch. Wäre $a \neq 0$, so $1 + (\frac{b}{a})^2 = \frac{a^2+b^2}{a^2} = 0$ und damit $-1 = (\frac{b}{a})^2$ im Widerspruch dazu, dass K keine imaginäre Einheit besitzt. Also $a = 0$. Analog zeigt man $b = 0$. Dann aber $a = b = 0$ im Widerspruch zu $a + b\mathring{i} \neq 0$. □

Definition 4.2.6. $\mathbb{C} := \mathbb{R}[\mathring{i}]$ nennt man den Körper der *komplexen Zahlen*.

Bemerkung 4.2.7. Ist $\mathring{i}$ eine imaginäre Einheit in einem kommutativen Ring A, so auch $\mathring{j} := -\mathring{i}$, denn $\mathring{j}^2 = \mathring{j}\mathring{j} = (-\mathring{i})(-\mathring{i}) = -\mathring{i}(-\mathring{i}) = -(-\mathring{i}\mathring{i}) = \mathring{i}\mathring{i} = \mathring{i}^2 = -1$. Ist nun K ein Körper ohne imaginäre Einheit, so ist $K[\mathring{i}] = K[-\mathring{i}]$ und nach 4.2.3(c) gibt es einen Ringautomorphismus [→3.2.10] φ auf $K[\mathring{i}]$ mit $\varphi(a) = a$ für alle $a \in K$ und $\varphi(\mathring{i}) = -\mathring{i}$. Auf $\mathbb{C}$ bezeichnet man diesen Ringautomorphismus

$$\mathbb{C} \to \mathbb{C},\ a + b\mathring{i} \mapsto a - b\mathring{i} \qquad (a, b \in \mathbb{R})$$

als *komplexe Konjugation* und $a-b\mathring{i}$ als das *komplex Konjugierte* zu $a+b\mathring{i}$. Wir schreiben auch z^* für das komplex Konjugierte von $z \in \mathbb{C}$. Andere Autoren schreiben dafür meist $\overline{z}$, aber dies könnte nicht nur zur Verwechslung mit unserer Notation für Kongruenzklassen führen, sondern ist auch aus anderen Gründen weniger modern.

Definition 4.2.8. Sei $z \in \mathbb{C}$. Wegen können wir 4.2.3(b) $z = a + b\mathring{i}$ mit eindeutig bestimmten $a, b \in \mathbb{R}$ schreiben. Wir definieren den *Realteil* von z durch

$$\mathrm{Re}(z) := \frac{1}{2}(z + z^*) = a \in \mathbb{R},$$

den *Imaginärteil* von z durch

$$\mathrm{Im}(z) := \frac{1}{2\overset{\circ}{i}}(z - z^*) = b \in \mathbb{R}$$

und den *Betrag* von z durch

$$|z| := \sqrt{a^2 + b^2} = \sqrt{z^* z} \in \mathbb{R}_{\geq 0}.$$

Definition 4.2.9. Sei K ein Körper und $p \in K[X]$ ein Polynom. Ein Element $a \in K$ heißt *Nullstelle* von p, wenn $p(a) = 0$ [→3.2.15].

Proposition 4.2.10 („Abspalten von Nullstellen“)**.** *Ist K ein Körper, $p \in K[X]$ und $a \in K$ eine Nullstelle von p, so gibt es $q \in K[X]$ mit $p = (X - a)q$.*

Beweis. Zu zeigen ist, dass p im Hauptideal [→3.3.11] $(X - a)$ von $K[X]$ liegt. Dazu äquivalent ist, dass $\overline{p} = 0$ in $K[X]/(X - a)$. Ist $p = \sum_{k=0}^{n} a_k X^k$ mit $n \in \mathbb{N}_0$ und $a_0, \ldots, a_n \in K$, so gilt $\overline{p} \overset{3.3.3}{=} \sum_{k=0}^{n} \overline{a_k}\,\overline{X}^k \overset{\overline{X}=\overline{a}}{=} \sum_{k=0}^{n} \overline{a_k}\,\overline{a}^k \overset{3.3.3}{=} \overline{p(a)} = \overline{0} = 0$. □

Korollar 4.2.11. *Sei K ein Körper. Ist dann $p \in K[X]$ und $\deg(p) = n \in \mathbb{N}_0$ [→3.2.6], so hat p höchstens n Nullstellen in K.*

Beweis. Wir zeigen durch Induktion nach $n \in \mathbb{N}_0$, dass jedes $p \in K[X]$ vom Grad n höchstens n Nullstellen in K hat.

$\underline{n = 0}$ Sei $p \in K[X]$ vom Grad 0. Dann gilt $p \in K^\times = K \setminus \{0\}$. Dann hat p offensichtlich keine Nullstelle in K.

$\underline{n - 1 \to n \quad (n \geq 1)}$ Sei $p \in K[X]$ vom Grad n. Hat p keine Nullstelle, so sind wir fertig. Sonst wählen wir eine Nullstelle $a \in K$ von p. Nach 4.2.10 gibt es $q \in K[X]$ mit $p = (X - a)q$. Offensichtlich gilt $\deg(q) = \deg(p) - 1 = n - 1$, weswegen nach Induktionsvoraussetzung q höchstens $n-1$ Nullstellen in K hat. Da in einem Körper ein Produkt zweier Elemente offensichtlich nur dann null sein kann, wenn schon einer der beiden Faktoren null war, ist die einzige Nullstelle, die p zusätzlich noch haben kann, offenbar a. Also hat p höchstens n Nullstellen in K. □

Satz 4.2.12 (Fundamentalsatz der Algebra)**.** [Jean-Robert Argand *1768 †1822] *Jedes Polynom $p \in \mathbb{C}[X]$ vom Grad ≥ 1 hat eine Nullstelle in $\mathbb{C}$.*

Bemerkung 4.2.13. (a) Durch sukzessives Abspalten von Nullstellen mit 4.2.10 kann man den Fundamentalsatz der Algebra auch wie folgt formulieren: Für jedes $p \in \mathbb{C}[X]$ vom Grad $n \in \mathbb{N}_0$ gibt es komplexe Zahlen $a_1, \ldots, a_n \in \mathbb{C}$ und ein $c \in \mathbb{C}^\times$ mit

$$p = c(X - a_1) \ldots (X - a_n).$$

(b) Der Fundamentalsatz der Algebra ist so erstaunlich, da wir durch das Adjungieren einer imaginären Einheit zu $\mathbb{R}$ a priori nur sicherstellen, dass das Polynom $X^2 + 1$ eine Nullstelle bekommt. Der Satz besagt, dass damit alle anderen Polynome vom Grad ≥ 1 automatisch auch eine Nullstelle erhalten.

(c) Im 17. Jahrhundert gab es von verschiedenen Mathematikern Äußerungen, die man als eine Vermutung der Gültigkeit des Fundamentalsatz deuten könnte, auch wenn der Begriff der komplexen Zahlen noch nicht auf soliden Grundlagen stand.

(d) Lückenhafte Beweisversuche mit wertvollen Ideen gab es seit 1746 [Jean-Baptiste le Rond d'Alembert *1717 †1783]. Mehrere wertvolle Versuche stammen von Carl Friedrich Gauss [*1777 †1855], unter anderem der erste algebraische Beweis aus dem Jahr 1816, der im Kern völlig richtig ist aber allerdings erst später auf solide Grundlagen gestellt wurde.

(e) Entgegen dem, was mancherorts geschrieben wird, dürfte der erste (lediglich modulo den damals noch etwas wackligen Grundlagen der Analysis) als richtig geltende Beweis des Fundamentalsatzes von Jean-Robert Argand [*1768 †1822] im Jahr 1814 geführt worden sein. Wir geben unten eine sehr grobe Skizze, die der Leser mit etwas Anfängeranalysis zu einem Beweis ausbauen können sollte.

(f) Der für Anfänger am leichtesten zu verstehende *algebraische* Beweis ist der Beweis von Gauss aus dem Jahr 1816. Leider würde er an dieser Stelle zuviel Zeit in Anspruch nehmen. Der Leser kann ihn aber in der Literatur nachlesen (siehe Theorem 2.11 im Buch von Basu, Pollack und Roy: Algorithms in Real Algebraic Geometry, `http://perso.univ-rennes1.fr/marie-francoise.roy/bpr-ed2-posted1.pdf`).

(g) In der einführenden Algebra-Vorlesung im dritten Semester geben wir einen sehr schönen algebraischen Beweis mit Hilfe von Galoistheorie [Évariste Galois *1811 †1832]. Dieser Beweis wird in Wirklichkeit sogar sehr viel mehr zeigen als jeder analytische Beweis. Wir sollten dabei natürlich kein Ergebnis benutzen, was schon auf dem Fundamentalsatz fußt. Um dies leichter überprüfen zu können, werden wir alle Resultate, in deren Beweis wir den Fundamentalsatz benutzen, in dieser Vorlesung entsprechend kennzeichnen.

Beweisskizze für den Fundamentalsatz der Algebra 4.2.12. Wir folgen der Beweisidee von Argand. Wir benutzen dabei die aus der Analysis bekannte Geometrie der Multiplikation von komplexen Zahlen und die Konzepte der Stetigkeit und der Kompaktheit. Sei $p = a_n X^n + a_{n-1}X^{n-1} + \cdots + a_0$ mit $n \in \mathbb{N}$, $n \geq 1$, $a_0, \ldots, a_n \in \mathbb{C}$ und $a_n \neq 0$. Dann gilt für alle $z \in \mathbb{C}$

$$|p(z)| \geq |a_n||z|^n - |a_0| - \cdots - |a_{n-1}||z|^{n-1}$$

und daher $\lim_{|z|\to\infty} |p(z)| = \infty$. Daraus folgt die Existenz eines globalen Minimalpunkts $z_0 \in \mathbb{C}$ von $\mathbb{C} \to \mathbb{R}_{\geq 0}$, $z \mapsto |p(z)|$, das heißt es gibt $z_0 \in \mathbb{C}$ mit $|p(z_0)| \leq |p(z)|$ für alle $z \in \mathbb{C}$ (dieser Punkt schien Argand intuitiv klar zu sein, erfordert aber heutzutage eine Begründung, die auch ein Anfänger geben kann). Œ $z_0 = 0$. Dann gilt mit $S := \{\zeta \in \mathbb{C} \mid |\zeta| = 1\}$ für alle $\zeta \in S$ und $r \in \mathbb{R}_{\geq 0}$

$$|p(r\zeta)|^2 - |p(0)|^2 \geq 0.$$

Schreibe $p = p(0) + X^k q$ mit einem $k \in \mathbb{N}$ und einem $q \in \mathbb{C}[X]$ mit $q(0) \neq 0$. Die Ungleichung lautet dann

$$|p(0) + r^k \zeta^k q(r\zeta)|^2 - |p(0)|^2 \geq 0$$

für alle $\zeta \in S$ und $r \in \mathbb{R}_{\geq 0}$. Unter Beachtung von

$$|z_1 + z_2|^2 = (z_1 + z_2)^*(z_1 + z_2) = z_1^* z_1 + z_1^* z_2 + z_1 z_2^* + z_2^* z_2 = \\ |z_1|^2 + z_1^* z_2 + (z_1^* z_2)^* + |z_2|^2 \overset{4.2.8}{=} |z_1|^2 + 2\,\mathrm{Re}(z_1^* z_2) + |z_2|^2$$

für alle $z_1, z_2 \in \mathbb{C}$ folgt daraus

$$2r^k \,\mathrm{Re}(p(0)^* \zeta^k q(r\zeta)) + r^{2k} |q(r\zeta)|^2 \geq 0$$

für alle $\zeta \in S$ und $r \in \mathbb{R}_{\geq 0}$. Daraus folgt

$$2\,\mathrm{Re}(p(0)^* \zeta^k q(r\zeta)) + r^k |q(r\zeta)|^2 \geq 0$$

für alle $\zeta \in S$ und $r \in \mathbb{R}_{>0}$. Indem man für festes $\zeta \in S$ nun den Grenzwert für $r \to 0$ betrachtet, erhält man

$$2\,\mathrm{Re}(p(0)^* \zeta^k q(0)) \geq 0$$

für alle $\zeta \in S$. Man muss also nur noch zeigen, dass es für festes $z \in \mathbb{C}^\times$ nicht vorkommen kann, dass $\mathrm{Re}(z\zeta^k) \geq 0$ für alle $\zeta \in S$ gilt. Dies kann man auf verschiedene Weisen schließen. Es ist aber klar, wenn man die Geometrie der Multiplikation von komplexen Zahlen verstanden hat. □

Beispiel 4.2.14. Weitere Beispiele zu imaginären Einheiten:

(a) $\mathbb{F}_3$ hat keine imaginäre Einheit, da $\mathbb{F}_3 \overset{4.1.10}{=} \mathbb{Z}/(3) = \{\overline{0}, \overline{1}, \overline{2}\}$ und $\overline{0}^2 = 0 \neq \overline{2} = -1$, $\overline{1}^2 = 1 \neq \overline{2} = -1$ und $\overline{2}^2 = \overline{4} = 1 \neq \overline{2} = -1$ in $\mathbb{F}_3$. Wegen $\#\mathbb{F}_3 = 3$ folgt $\#\mathbb{F}_3[\mathring{i}] = 9$ nach 4.2.3(b). Es ist also $\mathbb{F}_9 := \mathbb{F}_3[\mathring{i}]$ ein neunelementiger Körper.

(b) $\mathbb{F}_5$ hat eine imaginäre Einheit, denn $\overline{2}\,\overline{2} = \overline{4} = -1$ in $\mathbb{F}_5$. Es gilt also $\mathbb{F}_5[\mathring{i}] = \mathbb{F}_5$.

(c) In $\mathbb{F}_7$ gilt $0^2 = 0$, $1^2 = 1$, $\overline{2}^2 = \overline{4}$, $\overline{3}^2 = \overline{2}$, $\overline{4}^2 = \overline{2}$, $\overline{5}^2 = \overline{4}$ und $\overline{6}^2 = \overline{1}$. Also hat $\mathbb{F}_7$ keine imaginäre Einheit und $\mathbb{F}_{49} := \mathbb{F}_7[\mathring{i}]$ ist ein Körper mit 49 Elementen.

§5 Homogene lineare Gleichungssysteme

In diesem Kapitel sei stets K ein Körper. [→ 4.1.4]

§5.1 Matrizen in Stufenform

Sprechweise 5.1.1. Ein *homogenes lineares Gleichungssystem* über K ist (ggf. nach Umstellen) von der Form

$$(*) \qquad \begin{array}{ccccccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & 0 & & \\ \vdots & & & & \vdots & & & & (x \in K^n) \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n & = & 0 & & \end{array}$$

wobei die *Koeffizienten* $a_{ij} \in K$ $(1 \leq i \leq m, 1 \leq j \leq n)$ vorgegeben sind und die *Unbekannten* x_j $(1 \leq j \leq n)$ gesucht sind (m Gleichungen in n Unbekannten).
einzelne Zeilen: „homogene lineare Gleichungen“
„homogen“: rechte Seite gleich Null
„linear“: keine Produkte der Unbekannten
Eine *Lösung* von $(*)$ ist ein n-Tupel $(x_1, \dots, x_n) \in K^n$, welches alle Gleichungen gleichzeitig erfüllt. Es wird sich als praktisch herausstellen, solche Lösungen als *Spaltenvektor* zu schreiben, das heißt, man schreibt $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ statt $(x_1, \dots, x_n)$.

Beispiel 5.1.2. (a) Sei $K := \mathbb{F}_2$. $0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ sind Lösungen des homogenen Gleichungssystems

$$\begin{aligned} x_1 + x_3 &= 0 \\ x_2 &= 0 \end{aligned}$$

(b) Sei $K := \mathbb{C}$ und das lineare Gleichungssystem

$$\begin{aligned} (1 + \mathring{i})x_1 - 2x_2 + x_3 &= 0 \\ \mathring{i}x_2 - x_3 &= 0 \end{aligned}$$

Es sind $0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 1-3\mathring{i} \\ 2 \\ 2\mathring{i} \end{pmatrix}$ Lösungen, denn $(1+\mathring{i})(1-3\mathring{i}) - 2 \cdot 2 + 2\mathring{i}$ und $\mathring{i}2 - 2\mathring{i} = 0$. Für jedes $\lambda \in \mathbb{C}$ ist auch $\begin{pmatrix} \lambda(1-3\mathring{i}) \\ \lambda 2 \\ \lambda 2\mathring{i} \end{pmatrix}$ eine Lösung. Es gibt also unendlich viele Lösungen.

Proposition 5.1.3. *Ist U die* Lösungsmenge *von* 5.1.1 $(*)$ *(das heißt, die Menge aller Lösungen $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ von $(*)$), so gilt*

(a) $0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in U$

(b) *Sind* $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in U$ *und* $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in U$, *so* $x + y \overset{2.1.12}{=} \begin{pmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{pmatrix} \in U$

(c) *Sind* $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in U$ *und* $\underbrace{\lambda}_{\text{„lambda“}} \in K$, *so* $\lambda x := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix} \in U$.

Beweis. (a) klar mit 3.1.2 (g)

(b) Sind $x, y \in U$, so gilt für alle $i \in \{1, \ldots, n\}$

$$\begin{aligned} a_{i1}(x_1 + y_1) + \ldots + a_{in}(x_n + y_n) &\overset{\substack{3.1.1\,(D) \\ 2.1.9}}{=} (a_{i1}x_1 + \ldots + a_{in}x_n) + (a_{i1}y_1 + \ldots + a_{in}y_n) \\ &= 0 + 0 = 0 \end{aligned}$$

(c) Sind $x \in U$ und $\lambda \in K$, so gilt für alle $i \in \{1, \ldots, m\}$

$$a_{i1}(\lambda x_1) + \ldots + a_{in}(\lambda xn) \overset{\substack{3.1.2\,(e) \\ 3.1.1(D)}}{=} \lambda(a_{i1}x_1 + \ldots + a_{in}x_n) = \lambda \cdot 0 \overset{3.1.2\,(g)}{=} 0.$$

□

Bemerkung 5.1.4. In der Situation von 5.1.3 kann man (a), (b) und (c) wie folgt zusammenfassen:
Sind $r \in \mathbb{N}_0$ und $x^{(1)}, \ldots, x^{(r)} \in U$, so ist für alle $\lambda_1, \ldots, \lambda_r \in K$ auch deren *Linearkombination* $\sum_{i=1}^{r} \lambda_i x^{(i)} = \lambda_1 x^{(1)} + \ldots + \lambda_r x^{(r)}$ ein Element von U.
[(a) $\leftrightsquigarrow r = 0$, (b) $\leftrightsquigarrow (r = 2 \,\&\, \lambda_1 = \lambda_2 = 1)$, (c) $\leftrightsquigarrow r = 1$, (b) & (c) $\rightsquigarrow r \geq 2$]

Sprechweise und Notation 5.1.5. Für $r \in \mathbb{N}_0$ und $x^{(1)}, \ldots, x^{(r)} \in K^n$ bezeichnen wir die Menge aller Linearkombinationen

$$\operatorname{span}\left(x^{(1)}, \ldots, x^{(r)}\right) := \left\{ \sum_{i=1}^{r} \lambda_i x^{(i)} \;\middle|\; \lambda_1, \ldots, \lambda_r \in K \right\}$$

als *Spann* von $x^{(1)}, \ldots, x^{(r)}$.

Beispiel 5.1.6. (a) $\operatorname{span}() = \{0\} \subseteq K^n$

(b) $\operatorname{span}\begin{pmatrix}1\\1\\0\end{pmatrix} = \left\{ \begin{pmatrix}\lambda\\\lambda\\0\end{pmatrix} \;\middle|\; \lambda \in K \right\} \subseteq K^3$

(c) $\operatorname{span}\left(\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\-1\end{pmatrix}\right) = \left\{ \begin{pmatrix}\lambda+\mu\\\lambda\\-\mu\end{pmatrix} \;\middle|\; \lambda, \underbrace{\mu}_{\text{„my“}} \in K \right\} \subseteq K^3$

(d) Die Lösungsmenge des linearen Gleichungssystems

$$\begin{aligned} x_1 - 2x_2 + 0 - x_4 &= 0 \\ x_3 - 2x_4 &= 0 \qquad\qquad (\text{mit } 2 := 1+1 \in K) \end{aligned}$$

ist

$$\begin{aligned} \left\{ \begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} \;\middle|\; x_1 = 2x_2 + x_4, x_3 = 2x_4, x_2, x_4 \in K \right\} &= \left\{ \begin{pmatrix}2x_2+x_4\\x_2\\2x_4\\x_4\end{pmatrix} \;\middle|\; x_2, x_4 \in K \right\} \\ &= \left\{ \begin{pmatrix}2x_2\\x_2\\0\\0\end{pmatrix} + \begin{pmatrix}x_4\\0\\2x_4\\x_4\end{pmatrix} \;\middle|\; x_2, x_4 \in K \right\} \\ &= \left\{ x_2 \begin{pmatrix}2\\1\\0\\0\end{pmatrix} + x_4 \begin{pmatrix}1\\0\\2\\1\end{pmatrix} \;\middle|\; x_2, x_4 \in K \right\} \\ &= \left\{ \lambda \begin{pmatrix}2\\1\\0\\0\end{pmatrix} + \mu \begin{pmatrix}1\\0\\2\\1\end{pmatrix} \;\middle|\; \lambda, \mu \in K \right\} \\ &= \operatorname{span}\left(\begin{pmatrix}2\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\2\\1\end{pmatrix} \right) \end{aligned}$$

Bemerkung 5.1.7. Da ein lineares Gleichungssystem unendlich viele Lösungen haben kann, versucht man endlich viele „Basislösungen“ $x^{(1)}, \ldots, x^{(r)}$ zu berechnen derart, dass die Lösungsmenge genau $\operatorname{span}\left(x^{(1)}, \ldots, x^{(r)}\right)$ ist (Existenz noch unklar!). Zusätzlich will man, dass dabei keine der berechneten „Basislösungen“ überflüssig ist. Das in §5.2 beschriebene *Gauß-Verfahren* (lange vor Gauß bekannt, z.B. um -100 in China) wird dies leisten.

Erinnerung und Definition 5.1.8. [→ 1.1.30 (c)] Sei Z eine Menge und $m, n \in \mathbb{N}_0$. Eine $m \times n$-Matrix über Z ist eine Abbildung $A : \{1. \ldots, m\} \times \{1, \ldots, n\} \to Z$, die man meist in der Form $(A(i,j))_{(i,j)\in\{1,\ldots,m\}\times\{1,\ldots,n\}} = (A(i,j))_{1\le i\le m, 1\le j\le n}$ schreibt.
Die Menge aller $m \times n$-Matrizen über Z bezeichnet man mit $Z^{m\times n}$.

Sprechweise und Notation 5.1.9. Ist $A = (a_{ij})_{1\le i\le m, 1\le j\le n} \in K^{m\times n}$ und $x \in K^n$, so setzen wir

$$Ax := A \cdot x := \begin{pmatrix} a_{11}x_1 & + & \ldots & + & a_{1n}x_n \\ \vdots & & & & \vdots \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n \end{pmatrix} \in K^n.$$

Damit können wir 5.1.1 $(*)$ kompakt schreiben als

$$(*) \qquad Ax = 0 \qquad (x \in K^n).$$

Man nennt A die *Koeffizientenmatrix* von $(*)$.

Definition 5.1.10. Eine Matrix $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \in K^{m \times n}$ heißt in (Zeilen-) *Stufenform*, wenn sie von der Gestalt

$$A = \begin{pmatrix} 0 & \dots & a_{1j_1} & (*) & \dots & (*) & \dots & (*) & \dots & (*) & \dots & \dots & (*) \\ 0 & \dots & 0 & \dots & a_{2j_2} & (*) & \dots & (*) & \dots & (*) & \dots & \dots & (*) \\ 0 & \dots & \dots & \dots & 0 & \dots & a_{3j_3} & (*) & \dots & (*) & \dots & \dots & (*) \\ \vdots & & \vdots & & \vdots & & \vdots & \ddots & & \vdots & & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & \dots & a_{rj_r} & (*) & \dots & (*) \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ \vdots & & & & & & & & & & & \ddots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix}$$

ist mit $r \in \{0, \dots, m\}$, $j_1, \dots, j_r \in \{1, \dots, n\}$, $j_1 < j_2 < \dots < j_r$ und $a_{ij} \neq 0$ für alle $i \in \{1, \dots, r\}$, wobei „0" für Nulleinträge und „$(*)$" für beliebige Einträge steht. Dabei sind die *Anzahl der Stufen* r und die *Stufenposition* $j_1, \dots, j_r$ eindeutig bestimmt. Wir nennen A in *reduzierter Stufenform*, wenn zusätzlich $a_{ij_i} = 1$ und alle anderen Einträge der j_i-ten Spalte gleich null sind.

$$A = \begin{pmatrix} 0 & \dots & 1 & (*) & 0 & (*) & 0 & (*) & (*) & 0 & (*) & \dots & (*) \\ 0 & \dots & 0 & \dots & 1 & (*) & 0 & (*) & (*) & 0 & (*) & \dots & (*) \\ 0 & \dots & \dots & \dots & 0 & \dots & 1 & (*) & (*) & 0 & (*) & \dots & (*) \\ \vdots & & \vdots & & \vdots & & \vdots & \ddots & & \vdots & & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & \dots & 1 & (*) & \dots & (*) \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ \vdots & & & & & & & & & & & \ddots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix}$$

(d.h. die Einträge oberhalb von a_{ij_i} denn die unterhalb sind sowieso gleich 0).

Beispiel 5.1.11. (a) () ist 0×0-Matrix in reduzierter Stufenform mit 0 Stufen. Gleichzeitig ist () eine $m \times 0$- und $0 \times n$-Matrix für alle $m, n \in \mathbb{N}_0$.

(b) $(\overline{0})$ ist eine 1×1-Matrix in reduzierter Stufenform mit 0 Stufen.

(c) $(\underline{|1})$ ist eine 1×1-Matrix in reduzierter Stufenform mit 1 Stufe.

(d) $(\underline{|1, 0})$; ist 1×2-Matrix in reduzierter Stufenform mit 1 Stufe und Stufenposition 1.

(e) $\begin{pmatrix} 0 \\ \underline{1} \end{pmatrix}$ ist eine 2×1-Matrix *nicht in Stufenform.*

(f) $\begin{pmatrix} 0 & \boxed{1} & 0 & 3 & 0 & 1 \\ 0 & 0 & \boxed{1} & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ ist 4×6-Matrix in Stufenform mit 3 Stufen und Stufenpositionen 2,3,5 (hierbei $2 := 1 + 1$ und $3 := 1 + 1 + 1$).

Sprechweise und Bemerkung 5.1.12. Ist ein lineares Gleichungssystem [→ 5.1.9] $(*)$ $Ax = 0$ $(x \in \mathbb{R}^n)$ mit $A \in K^{m \times n}$ in Stufenform gegeben, so nennen wir für $j \in \{1, \dots, n\}$ die Unbekannte $\left\{ \begin{matrix} \textit{abhängig} \\ \textit{frei} \end{matrix} \right\}$ in $(*)$, wenn j $\left\{ \begin{matrix} \text{eine} \\ \text{keine} \end{matrix} \right\}$ Stufenposition [→ 5.1.10] von A ist. Hat also A genau r Stufen, so gibt es r abhängige und $n - r$ freie Unbekannte. Ist A sogar in reduzierter Stufenform [→ 5.1.10], also

$$A = \begin{pmatrix} 0 & \dots & \boxed{1} & (*) & 0 & (*) & 0 & (*) & (*) & 0 & (*) & \dots & (*) \\ 0 & \dots & 0 & \dots & \boxed{1} & (*) & 0 & (*) & (*) & 0 & (*) & \dots & (*) \\ 0 & \dots & \dots & \dots & 0 & \dots & \boxed{1} & (*) & (*) & 0 & (*) & \dots & (*) \\ \vdots & & \vdots & & \vdots & & \vdots & & \ddots & \vdots & & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & \dots & \boxed{1} & (*) & \dots & (*) \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & & & & & & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \} r \\ \\ \\ \end{matrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \end{matrix} \left. \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \end{matrix} \right\} m$$
$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{n}$$

so kann man offensichtlich $Ax = 0$ als ein System von r homogenen linearen Gleichungen schreiben, auf deren rechten Seiten nur freie Unbekannte auftauchen. Es folgt, dass in $(*)$ für jede Festlegung der freien Unbekannten genau eine Wahl der abhängigen Unbekannten existiert derart, dass $Ax = 0$ gilt.
Damit kann man dann unmittelbar $x^{(1)}, \dots, x^{(n-r)} \in K^n$ bestimmen mit

$$\{x \in K^n \mid Ax = 0\} = \operatorname{span}\left(x^{(1)}, \dots, x^{(n-r)}\right).$$

Beispiel 5.1.13. $K = \mathbb{Q}, n = 7$

$$(*) \qquad \begin{array}{rrrrl} x_1 & -3x_4 & +x_5 & +x_7 & = 0 \\ x_2 & -x_3 & +x_5 & & = 0 \\ x_6 & -x_7 & & & = 0 \end{array} \qquad (x_1, \dots, x_7 \in \mathbb{Q})$$

kann man schreiben als $(*)$ $Ax = 0$ $(x \in \mathbb{Q}^7)$ mit

$$A := \begin{pmatrix} 1 & 0 & 0 & -3 & 1 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \in \mathbb{Q}^{3 \times 7}.$$

Es ist A in reduzierter Stufenform gemäß 5.1.10 und in $(*)$ sind x_1, x_2, x_6 abhängig und

x_3, x_4, x_5, x_7 frei gemäß 5.1.12. Die Lösungsmenge von $(*)$ ist

$$\left\{ x \in \mathbb{Q}^7 \;\middle|\; x_1 = 3x_4 - x_5 - x_7, x_2 = x_3 - x_5, x_6 = x_7, x_3, x_4, x_5, x_7 \in \mathbb{Q} \right\}$$
$$= \left\{ \begin{pmatrix} 3x_4 - x_5 - x_7 \\ x_3 - x_5 \\ x_3 \\ x_4 \\ x_5 \\ x_7 \\ x_7 \end{pmatrix} \;\middle|\; x_3, x_4, x_5, x_7 \in \mathbb{Q} \right\}$$
$$= \left\{ x_3 \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_7 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \;\middle|\; x_3, x_4, x_5, x_7 \in \mathbb{Q} \right\}$$
$$\overset{5.1.5}{=} \operatorname{span}\left(\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right).$$

Offensichtlich kann man keine der so berechneten Basislösungen streichen, ohne den Spann zu verändern, denn ist x_j frei in $(*)$, so gibt es eine Lösung $x \in K^n$ von $(*)$ mit $x_j = 1$ und $x_k = 0$ für alle anderen freien Unbekannten x_k (es ist aber die j-te Komponente der zu diesen x_k gehörenden Basisvektor gleich 0).

Bemerkung 5.1.14. (a) Da man stets wie in 5.1.13 vorgehen kann, ist geklärt, wie man homogene Lineare Gleichungssysteme $Ax = 0$ $(x \in K^n)$ mit Koeffizientenmatrix A *in reduzierter Stufenform* löst (das heißt, deren Lösungsmenge als Spann endlich vieler Lösungstupel darstellt, von denen keines überflüssig ist).

(b) Im allgemeinen Fall versucht man eine homogenes lineares Gleichungssystem so umzuschreiben, dass dessen Koeffizientenmatrix schließlich in reduzierter Stufenform vorliegt. Folgende Operationen ändern die Lösungsmenge nicht:

(1) Addieren des λ-fachen einer Gleichung zu einer anderen $(\lambda \in K)$.

(2) Multiplizieren einer Gleichung mit λ $(\lambda \in K^\times)$

In der Tat: (1) kann rückgängig gemacht werden durch Addieren des $(-\lambda)$-fachen (d.h. Subtrahieren des λ-fachen) der einen Gleichung zu der anderen und (2) durch Multiplikation derselben Gleichung mit λ^{-1}. Das Gauß-Verfahren führt diese Operationen gleich auf den Zeilen der Koeffizientenmatrix A durch.

§5.2 Gauß-Verfahren

Notation, Sprechweise und Bemerkung 5.2.1. *Sei $A \in K^{m \times n}$. Wir betrachten folgende* elementare Zeilenoperationen [→ 5.1.14 (b) (1),(2)]*:*

$$\underbrace{Z_i}_{\text{„Zeile } i\text{“}} \underbrace{\leftarrow}_{\text{„wird“}} Z_i + \lambda Z_j \ (i, j \in \{1, \ldots, m\}, i \neq j, \lambda \in K)$$

(Addieren des λ-fachen einer Zeile zu einer anderen)

$$Z_i \leftarrow \lambda Z_i \ (i \in \{1, \ldots, m\}, \lambda \in K^\times)$$

(Multiplizieren einer Zeile mit einem $\lambda \neq 0$).

Wir werden A durch sukzessive Anwendung endlich vieler dieser Operationen in reduzierte Stufenform überführen. Man kann dabei auch Zeilenoperationen *erlauben, die man durch endlich viele Operationen simulieren kann, z.B. die folgenden:*

$$Z_i \underbrace{\leftrightarrow}_{\text{„wird vertauscht“}} Z_j \ (i,j \in \{1,\dots,m\})$$

(Vertauschen zweier Zeilen)

$$\textit{Simulation durch} \begin{cases} \textit{Nichtstun}, & \textit{falls } i = j. \\ \left.\begin{array}{l} Z_i \leftarrow Z_i + Z_j, \\ Z_j \leftarrow Z_j - Z_i, \\ Z_i \leftarrow Z_i + Z_j, \\ Z_j \leftarrow -Z_j, \end{array}\right\} & \textit{falls } i \neq j. \end{cases}$$

$$\begin{pmatrix} a \\ b \end{pmatrix} \overset{Z_1 \leftarrow Z_1 + Z_2}{\sim} \begin{pmatrix} a+b \\ a \end{pmatrix} \overset{Z_2 \leftarrow Z_2 - Z_1}{\sim} \begin{pmatrix} a+b \\ -a \end{pmatrix} \overset{Z_1 \leftarrow Z_1 + Z_2}{\sim} \begin{pmatrix} b \\ -a \end{pmatrix} \overset{Z_2 \leftarrow -Z_2}{\sim} \begin{pmatrix} b \\ a \end{pmatrix}$$

Definition und Proposition 5.2.2. Für $A, B \in K^{m\times n}$ sagen wir, B *geht aus A durch Zeilenoperationen hervor*, wenn man B aus A durch eine endliche Abfolge von Zeilenoperationen gewinnen kann. (vgl. 5.1.14 (b)).
Geht B aus A durch Zeilenoperationen hervor, so auch A aus B (vgl. 5.1.14 (b)).
Durch $A \sim B :\iff B$ geht aus A durch Zeilenoperation hervor, $(A, B \in K^{m\times n})$ wird also eine Äquivalenzrelation auf $K^{m\times n}$ definiert [→ 1.3.1 (b)].

Algorithmus 5.2.3. Man kann *zum Beispiel* wie folgt eine Matrix $A \in K^{m\times n}$ durch Zeilenoperationen in eine Matrix $B \in K^{m\times n}$ in Stufenform überführen [→ 5.1.10]:

$$A = \begin{pmatrix} 0 & \dots & 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & * & * & \dots & * \\ 0 & \dots & 0 & a_{ij} & * & \dots & * \\ 0 & \dots & 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & * & * & \dots & * \end{pmatrix} \overset{Z_i \leftrightarrow Z_1}{\sim} \begin{pmatrix} 0 & \dots & 0 & a_{ij} & * & \dots & * \\ 0 & \dots & 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & * & * & \dots & * \end{pmatrix}$$

$$\overset{Z_1 \leftarrow \frac{1}{a_{ij}} Z_1}{\sim} \begin{pmatrix} 0 & \dots & 0 & 1 & * & \dots & * \\ 0 & \dots & 0 & a_{i2} & * & \dots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & a_{im} & * & \dots & * \end{pmatrix} \overset{\substack{Z_2 \leftarrow Z_2 - a_{i2} Z_1 \\ Z_3 \leftarrow Z_3 - a_{i3} Z_1 \\ \vdots \\ Z_m \leftarrow Z_m - a_{im} Z_1}}{\sim} \left(\begin{array}{cccc|c} 0 & \dots & 0 & 1 & (*)\dots(*) \\ 0 & \dots & 0 & 0 & \text{rekursive} \\ \vdots & \vdots & \vdots & \vdots & \text{Anwendung} \\ \vdots & \vdots & \vdots & \vdots & \text{des} \\ \vdots & \vdots & \vdots & \vdots & \text{Verfahrens} \\ 0 & \dots & 0 & 0 & (*)\dots(*) \end{array}\right)$$

$\uparrow$ erste Spalte $\neq 0$, $a_{ij} \neq 0$

$\sim \dots \sim B$

Algorithmus 5.2.4. Eine Matrix $B \in K^{m\times n}$ in Stufenform kann man wie folgt in eine Matrix $C \in K^{m\times n}$ in *reduzierter* Stufenform überführen:

$$B = \begin{pmatrix} 0 & \dots & \boxed{b_{1j_1}} & (*) & \dots & (*) & \dots & (*) & \dots & (*) & \dots & \dots & (*) \\ 0 & \dots & \dots & 0 & \boxed{b_{2j_2}} & (*) & \dots & (*) & \dots & (*) & \dots & \dots & (*) \\ 0 & \dots & \dots & \dots & \dots & 0 & \boxed{b_{3j_3}} & (*) & \dots & (*) & \dots & \dots & (*) \\ \vdots & & & \vdots & & \vdots & & & & \vdots & & & \vdots \\ 0 & \dots & \dots & 0 & \dots & 0 & \dots & & 0 & \boxed{b_{rj_r}} & (*) & \dots & (*) \end{pmatrix}$$

$$\underset{i\in\{1,\dots,r\}}{\overset{Z_i \leftarrow \frac{1}{b_{ij_i}} Z_i,}{\sim}} \begin{pmatrix} 0 & \dots & \boxed{1} & (*) & b_{1j_2} & (*) & b_{1j_3} & (*) & b_{1j_r} & \dots & \dots & (*) \\ 0 & \dots & \dots & 0 & \boxed{1} & (*) & b_{2j_3} & (*) & b_{2j_r} & \dots & \dots & (*) \\ 0 & \dots & \dots & \dots & \dots & 0 & \boxed{1} & (*) & b_{3j_r} & \dots & \dots & (*) \\ \vdots & & & \vdots & & \vdots & & & \vdots & & & \vdots \\ 0 & \dots & \dots & 0 & \dots & 0 & \dots & 0 & \boxed{1} & (*) & \dots & (*) \end{pmatrix}$$

$$\underset{i\in\{1,\dots,r\}}{\overset{Z_i \leftarrow Z_i - b_{ij_i} Z_r,}{\sim}} \begin{pmatrix} 0 & \dots & \boxed{1} & (*) & b_{1j_2} & (*) & b_{1j_3} & (*) & 0 & \dots & \dots & (*) \\ 0 & \dots & \dots & 0 & \boxed{1} & (*) & b_{2j_3} & (*) & b_{2j_r} & \dots & \dots & (*) \\ 0 & \dots & \dots & \dots & \dots & 0 & \boxed{1} & (*) & b_{3j_r} & \dots & \dots & (*) \\ \vdots & & & \vdots & & \vdots & & & \vdots & & & \vdots \\ 0 & \dots & \dots & 0 & \dots & 0 & \dots & 0 & \boxed{1} & (*) & \dots & (*) \end{pmatrix}$$

$$\underset{i\in\{1,\dots,r-2\}}{\overset{Z_i \leftarrow Z_i - b_{ij_{r-1}} Z_{r-1},}{\sim}} \quad \dots \quad \underset{i\in\{1,\dots,1\}}{\overset{Z_i \leftarrow Z_i - b_{ij_2} Z_2,}{\sim}}$$

$$\begin{pmatrix} 0 & \dots & \boxed{1} & (*) & 0 & (*) & 0 & (*) & 0 & (*) & \dots & (*) \\ 0 & \dots & \dots & 0 & \boxed{1} & (*) & 0 & (*) & 0 & (*) & \dots & (*) \\ 0 & \dots & \dots & \dots & \dots & 0 & \boxed{1} & (*) & 0 & (*) & \dots & (*) \\ \vdots & & & \vdots & & \vdots & & & \vdots & & & \vdots \\ 0 & \dots & \dots & 0 & \dots & 0 & \dots & 0 & \boxed{1} & (*) & \dots & (*) \end{pmatrix}$$

Bemerkung 5.2.5. Es ist nun geklärt, wie man homogene lineare Gleichungssysteme löst:

(a) Bringe Koeffizientematrix auf Stufenform [→ 5.2.3]

(b) Bringe sie sogar in *reduzierte* Stufenform [→ 5.2.4]

(c) Schreibe die Lösungsmenge als Spann [→ 5.1.13]

Beispiel 5.2.6. $K = \mathbb{F}_5, n = 4$

$$(*) \qquad \begin{array}{llll} \overline{4}x_2 & +x_3 & +\overline{3}x_4 & = 0 \\ \overline{2}x_1 & +\overline{3}x_2 & +x_4 & = -\overline{2}x_3 \\ x_1 & +\overline{2}x_2 & +\overline{3}x_3 & +\overline{4}x_4 = 0 \\ \overline{2}x_1 & +\overline{4}x_2 & +x_3 & +\overline{3}x_4 = 0 \end{array}$$

kann man schreiben als $(*) \quad Ax = 0 \quad (x \in {\mathbb{F}_5}^4)$ mit

$$A := \begin{pmatrix} \overline{0} & \overline{4} & \overline{1} & \overline{3} \\ \overline{2} & \overline{3} & \overline{2} & \overline{1} \\ \overline{1} & \overline{2} & \overline{3} & \overline{4} \\ \overline{2} & \overline{4} & \overline{1} & \overline{3} \end{pmatrix} \overset{\substack{Z_1 \leftrightarrow Z_3 \\ Z_2 \leftarrow Z_2 - 2Z_1 \\ Z_4 \leftarrow Z_4 - \overline{2}Z_1}}{\sim} \begin{pmatrix} \overline{1} & \overline{2} & \overline{3} & \overline{4} \\ \overline{0} & \overline{4} & \overline{1} & \overline{3} \\ \overline{0} & \overline{4} & \overline{1} & \overline{3} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \end{pmatrix} \overset{Z_3 \leftarrow Z_3 - Z_2}{\sim} \begin{pmatrix} \overline{1} & \overline{2} & \overline{3} & \overline{4} \\ \overline{0} & \overline{4} & \overline{1} & \overline{3} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \end{pmatrix}$$

$$\overset{Z_2 \leftarrow \frac{1}{4} Z_2}{\sim} \begin{pmatrix} \overline{1} & \overline{2} & \overline{3} & \overline{4} \\ \overline{0} & \overline{1} & \overline{4} & \overline{2} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \end{pmatrix} \overset{Z_1 \leftarrow Z_1 - 2Z_2}{\sim} \begin{pmatrix} \boxed{\overline{1}} & \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \boxed{\overline{1}} & \overline{4} & \overline{2} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \end{pmatrix}$$

reduzierte Stufenform. Stufenpositionen 1, 2, abhängig: x_1, x_2, frei: x_3, x_4.

Die Lösungsmenge von $(*)$ ist

$$\begin{aligned} \left\{ x \in {\mathbb{F}_5}^4 \;\middle|\; Ax = 0 \right\} &= \left\{ x \in {\mathbb{F}_5}^4 \;\middle|\; x_1 = 0, x_2 = -\overline{4}x_3 - \overline{2}x_4 = x_3 + \overline{3}x_4 \right\} \\ &= \left\{ \begin{pmatrix} 0 \\ x_3 + 3x_4 \\ x_3 \\ x_4 \end{pmatrix} \;\middle|\; x_3, x_4 \in \mathbb{F}_5 \right\} \\ &= \left\{ x_3 \begin{pmatrix} \overline{0} \\ \overline{1} \\ \overline{1} \\ \overline{0} \end{pmatrix} + x_4 \begin{pmatrix} \overline{0} \\ \overline{3} \\ \overline{0} \\ \overline{1} \end{pmatrix} \;\middle|\; x_3, x_4 \in \mathbb{F}_5 \right\} \\ &= \operatorname{span}\left(\begin{pmatrix} \overline{0} \\ \overline{1} \\ \overline{1} \\ \overline{0} \end{pmatrix}, \begin{pmatrix} \overline{0} \\ \overline{3} \\ \overline{0} \\ \overline{1} \end{pmatrix} \right) \end{aligned}$$

Bemerkung 5.2.7. Ein homogenes lineares Gleichungssystem mit weniger Gleichungen als unbekannten hat immer eine nichttriviale Lösung (eine Lösung $\neq 0$), denn mit 5.2.3 und 5.2.4 kann man seine Koeffizientenmatrix in *reduzierter* Stufenform annehmen und da diese Matrix breiter als hoch ist, hat das Gleichungssystem dann eine freie Unbekannte [→ 5.1.12].

Bemerkung 5.2.8. Beim Lösen von homogenen linearen Gleichungssystemen kann es manchmal sinnvoll sein, die Unbekannten $x_1, \ldots, x_n$ anders zu nummerieren, um schneller zu einer Koeffizientenmatrix in reduzierter Stufenform zu gelangen. Da man dies manchmal erst im Laufe der Berechnung bemerkt, kann man auch *Spalten* vertauschen, wenn man sich merkt, welche Spalte zu welcher Unbekannten gehört.

Beispiel 5.2.9. $Ax = 0, (x \in \mathbb{R}^5)$ mit $A = \begin{pmatrix} 1 & 0 & 0 & 3 & 1 \\ 2 & 0 & 1 & 3 & 1 \\ 3 & 1 & 0 & 3 & 1 \end{pmatrix}$.

$$A \overset{\substack{Z_2 \leftarrow Z_2 - Z_1 \\ Z_3 \leftarrow Z_3 - Z_1}}{\sim} \begin{pmatrix} 1 & 0 & 0 & 3 & 1 \\ 1 & 0 & \boxed{1} & 0 & 0 \\ 2 & \boxed{1} & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{matrix} x_5 & x_4 & x_3 & x_2 & x_1 \end{matrix}$$

$$\left[A \underbrace{\quad \not\sim \quad}_{\text{wahrscheinlich}} \right] \begin{pmatrix} \boxed{1} & 3 & 0 & 0 & 1 \\ 0 & 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & 0 & \boxed{1} & 2 \end{pmatrix}$$

in reduzierter Stufenform, 3 Stufen, Stufenpositionen 1, 3, 4, abhängig: x_5, x_3, x_2, frei: x_4, x_1.

$$\left\{ x \in \mathbb{R}^5 \;\middle|\; Ax = 0 \right\} = \left\{ \begin{pmatrix} x_1 \\ -2x_2 \\ -1x_1 \\ x_4 \\ -x_1 - 3x_4 \end{pmatrix} \;\middle|\; x_1, x_4 \in \mathbb{R} \right\} = \operatorname{span}\left(\begin{pmatrix} 1 \\ -2 \\ -1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -3 \end{pmatrix} \right)$$

§5.3 Dualität

Definition 5.3.1. Ist $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in K^{m \times n}$, so nennt man

$$\underbrace{\operatorname{row}}_{\text{„row space“}} A := \operatorname{span}\left(\begin{pmatrix} a_{11} \\ \vdots \\ a_{1n} \end{pmatrix}, \dots, \begin{pmatrix} a_{m1} \\ \vdots \\ a_{mn} \end{pmatrix} \right) \subseteq K^n$$

den *Zeilenraum* und $\ker A := \left\{ x \in K^n \;\middle|\; \underbrace{Ax}_{[\to 5.1.9]} = 0 \right\}$ den *Kern* von A.

Satz 5.3.2. *Sind $A, B \in K^{m \times n}$ in reduzierter Stufenform mit $\ker A = \ker B$, so gilt $A = B$.*

Beweis. Wir zeigen $\forall n \in \mathbb{N}_0 : \forall m \in \mathbb{N}_0; \forall A, B \in K^{m \times n}$:

$$(A, B \text{ in reduzierter Stufenform } \& \ker A = \ker B) \Longrightarrow A = B$$

durch Induktion nach n.

$\underline{n = 0}$ Sind $m \in \mathbb{N}_0$ und $A, B \in K^{m \times 0}$, so gilt $A = () = B$.

$\underline{n - 1 \to n \ (n \in \mathbb{N})}$ Seien $m \in \mathbb{N}_0$ und $A, B \in K^{m \times n}$ in reduzierter Stufenform mit $\ker A = \ker B$. Zu zeigen: $A = B$.

Fall 1: $A = \begin{pmatrix} 0 & \boxed{\quad (*) \quad} \\ \vdots & \\ 0 & \end{pmatrix}$. Dann $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \ker A = \ker B$ und

$B = \begin{pmatrix} 0 & \boxed{\quad (*) \quad} \\ \vdots & \\ 0 & \end{pmatrix}$. Schreibe $A = \begin{pmatrix} 0 & \boxed{\quad A' \quad} \\ \vdots & \\ 0 & \end{pmatrix}$ und

$B = \begin{pmatrix} 0 & \boxed{\quad B' \quad} \\ \vdots & \\ 0 & \end{pmatrix}$ mit $A', B' \in K^{m \times (n-1)}$. Dann sind A' und B' in

reduzierter Stufenform und es gilt

$$\ker A' = \left\{ \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} \in K^{n-1} \;\middle|\; \begin{pmatrix} 0 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \ker A \right\}$$
$$= \left\{ \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} \in K^{n-1} \;\middle|\; \begin{pmatrix} 0 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \ker B \right\} = \ker B'$$

und daher $A' = B'$ nach IV. Also $A = B$.

Fall 2: $A = \begin{pmatrix} 1 & * & \dots & * \\ 0 & & & \\ \vdots & & (*) & \\ 0 & & & \end{pmatrix}$. Dann $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \notin \ker A = \ker B$ und daher

$B = \begin{pmatrix} 1 & * & \dots & * \\ 0 & & & \\ \vdots & & (*) & \\ 0 & & & \end{pmatrix}$. Schreibe $A = \begin{pmatrix} 1 & * & \dots & * \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$ und

$B = \begin{pmatrix} 1 & * & \dots & * \\ 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{pmatrix}$ mit $A', B' \in K^{(m-1)\times(n-1)}$ und daher $A' = B'$ nach IV. Dann sind A' und B' in reduzierter Stufen fom und es gilt

$$\ker A' = \left\{ \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} \in K^{n-1} \;\middle|\; \exists x_1 \in K : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \ker A \right\}$$
$$= \left\{ \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} \;\middle|\; \exists x_1 \in K : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \ker B \right\} = \ker B'$$

Behauptung 1: A und B haben dieselben Stufenpositionen.

Begründung: 1 ist Stufenposition von A und von B. Sei $j \in \{2, \dots, n\}$. Dann j Stufenposition von $A \iff j-1$ Stufenpostion von $A' = B' \iff j$ Stufenposition von B.

Behauptung 2: Die Gleichungssysteme $Ax = 0$ und $Bx = 0$ ($x \in K^n$) haben dieselben abhängigen und freien Unbekannten.

Begründung: folgt sofort aus Behauptung 1.

Sei nun $j \in \{2, \dots, n\}$, a der j-te Eintrag von A und b der j-te Eintrag von B in der ersten Zeile. Zu zeigen: $a = b$. Ist j eine Stufenposition von A, so nach Behauptung 1 auch von B und daher $a = 0 = b$.
Sei also nun j keine Stufenposition von A. Dann ist x_j frei (für beide Gleichungssysteme, siehe Behauptung 2) und man findet $x \in \ker A = \ker B$ mit $x_j = 1$ und $x_k = 0$ für alle anderen freien Unbekannten x_k. Es folgt $1 \cdot x_1 + a x_j = 0 = 1 \cdot x_1 + b x_j$ und damit $a = -x_1 = b$. □

Bemerkung 5.3.3. Ist $A \in K^{m\times n}$, so

$$\ker A = \{x \in K^n \mid \forall a \in \operatorname{row} A : a_1x_1 + \cdots + a_nx_n = 0\} \ [\to 5.1.9,\ 5.3.1].$$

Satz 5.3.4. *Seien $A, B \in K^{m\times n}$. Dann*

$$A \sim B \iff \ker A = \ker B \iff \operatorname{row} A = \operatorname{row} B$$

Beweis. Zeilenoperationen auf einer Matrix $K^{m\times n}$ ändern weder ihre Äquivalenzklasse [→ 5.2.2], noch ihren Kern [→ 5.1.14 (b)] noch ihren Zeilenraum (sieht man leicht). Also können wir nach 5.2.3 und 5.2.4 A und B in reduzierter Stufenform annehmen. Dann

$$\ker A = \ker B \overset{5.3.2}{\Longrightarrow} A = B \Longrightarrow A \sim B \Longrightarrow \operatorname{row} A = \operatorname{row} B \overset{5.3.3}{\Longrightarrow} \ker A = \ker B$$

□

Korollar 5.3.5. [→ 5.3.3] *Ist $A \in K^{m\times n}$, so*

$$\operatorname{row} A = \{a \in K^n \mid \forall x \in \ker A : a_1x_1 + \ldots + a_nx_n = 0\}$$

Beweis. Sei $A \in K^{m\times n}$.

„$\subseteq$“ ist klar nach Def. 5.3.1.

„$\supseteq$“ Sei $a \in K^n$ mit $\forall x \in \ker A : a_1x_1 + \ldots + a_nx_n = 0$.
Dann $\ker \left(\begin{smallmatrix} a_1 & \overset{A}{\cdots} & a_n \end{smallmatrix}\right) = \ker A = \ker \left(\begin{smallmatrix} 0 & \overset{A}{\cdots} & 0 \end{smallmatrix}\right)$ und daher

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \operatorname{row} \left(\begin{smallmatrix} a_1 & \overset{A}{\cdots} & a_n \end{smallmatrix}\right) \overset{5.3.4}{=} \operatorname{row} \left(\begin{smallmatrix} 0 & \overset{A}{\cdots} & 0 \end{smallmatrix}\right) = \operatorname{row} A.$$

□

Bemerkung 5.3.6. Gegeben seien $x^{(1)}, \ldots, x^{(m)} \in K^n$. Wir wollen ein homogenes lineares Gleichungssystem ohne redundante Gleichungen finden, dessen Lösungsmenge genau $\operatorname{span}\left(x^{(1)}, \ldots, x^{(m)}\right)$ ist. Diese sogenannte *duale Aufgabe* kann man wegen 5.3.5 wie folgt lösen:

Schreibe $x^{(1)}, \ldots, x^{(m)}$ als Zeilen der Matrix $X \in K^{m\times n}$. Löse das Gleichungssystem $X \cdot a = 0$ ($a \in K^n$) und fasse die gefundenen Basislösungen $a^{(1)}, \ldots, a^{(k)} \in K^n$ als Zeilen der Matrix $A \in K^{k\times n}$ auf. Es ist $A \cdot x = 0$ ($x \in K^n$) ein Gleichungssystem wie gewünscht.

Beispiel 5.3.7. Wir suchen ein Gleichungssystem, dessen Lösungsmenge $\operatorname{span}\left(\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right)$ ist ($K = \mathbb{R}$).

$$X := \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \overset{Z_1 \leftarrow \frac{1}{2}Z_1}{\sim} \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 1 & 1 \end{pmatrix} \overset{Z_1 \leftarrow Z_1 - \frac{1}{2}Z_2}{\sim} \begin{pmatrix} \boxed{1} & 0 & -\frac{1}{2} \\ 0 & \boxed{1} & 1 \end{pmatrix}$$

$$\left\{ a \in \mathbb{R}^3 \;\middle|\; X \cdot a = 0 \right\} = \left\{ a \in \mathbb{R}^3 \;\middle|\; a_1 = \frac{1}{2}a_3, a_2 = -a_3 \right\} = \left\{ \begin{pmatrix} \frac{1}{2}a_3 \\ -a_3 \\ a_3 \end{pmatrix} \;\middle|\; a_3 \in \mathbb{R} \right\}$$
$$= \operatorname{span}\left(\begin{pmatrix} \frac{1}{2} \\ -1 \\ 1 \end{pmatrix}\right) = \operatorname{span}\left(\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}\right)$$

$x_1 - 2x_2 + 2x_3 = 0$ ($x \in \mathbb{R}^3$) ist ein Gleichungssystem wie gewünscht.

§6 Vektorräume

§6.1 Definitionen und Beispiele von Vektorräumen, Untervektorräume

Definition 6.1.1. [→ 3.1.1] Ein *Vektorraum* (*VR*) ist ein Tupel $(K, +_K, \cdot_K, V, +, \cdot)$, wobei $(K, +_K, \cdot_K)$ ein Körper, $(V, +)$ eine abelsche Gruppe [→ 2.1.1] und $\cdot : K \times V \to V$ eine (meist unsichtbar oder infix geschriebene) Abbildung ist mit folgenden Eigenschaften:

(V) $\forall a, b \in K : \forall v \in V : (a \cdot_K b)v = a(bv)$ „verträglich“

($\overrightarrow{\text{N}}$) $\forall v \in V : 1_K v = v$

($\overrightarrow{\text{D}}$) $\forall a \in K : \forall v, w \in V : a(v + w) = (av) + (aw)$

(D') $\forall a, b \in K : \forall v \in V : (a +_K b)v = (av) + (bv)$

Bemerkung 6.1.2. Sei $(K, +_K, \cdot_K, V, +, \cdot)$ ein Vektorraum.

(a) Die Elemente von V nennt man oft *Vektoren*. Wir bezeichnen sie oft mit $v, w, u, x, y, \ldots$. Früher bezeichnete man sie mit $\overrightarrow{v}, \overrightarrow{w}, \ldots$

(b) Die Elemente von K nennt man oft *Skalare*. Wir bezeichnen sie oft mit $a, b, c, d, \lambda, \mu, \alpha, \beta, \gamma, \delta, \ldots$

(c) Sprechweisen:

- $(K, +_K, \cdot_K)$ Grund-/Skalarkörper
- V zugrundeliegende (oder Träger-)Menge
- $(V, +)$ zugrundeliegende abelsche Gruppe oder additive Gruppe.
- $+$ (Vektor-)Addition
- $\cdot$ Skalarmultiplikation (nicht verwechseln mit Skalarprodukt!)

(d) „Punkt vor Strich“ [→ 3.1.2 (f)]: $av + bw = (av) + (bw)$ für $a, b \in K, v, w \in V$.

(e) $(\overrightarrow{\mathrm{D}})$ besagt, dass für jedes $a \in K$ die Abbildung $V \to V, v \mapsto av$ ein Gruppenendomorphismus von $(V, +)$ ist. Insbesondere $a \cdot 0 = 0$ [→ 3.1.2 (g)] und $a(-v) = -av$ für alle $a \in K$.

(f) Für alle $v \in V$ gilt $0_K \cdot v = 0$, denn $0_K v + 0_K v \overset{(\mathrm{D'})}{=} (0_K + 0_K)v \overset{(\mathrm{N})}{=} 0_K v$.

(g) Für alle $a \in K$ gilt $(-a)v = -av$, denn $av + (-a)v \overset{(\mathrm{D'})}{=} (a - a)v = 0_K v \overset{(\mathrm{f})}{=} 0$.

Sprechweise und Notation 6.1.3. [→ 2.1.2 (e), 3.1.2 (d)] Statt von einem Vektorraum $(K, +_K, \cdot_K, V, +, \cdot)$ spricht man auch von einem K-Vektorraum V oder von einem Vektorraum V (über K). Statt $+_K$ und $\cdot_K$ schreibt man + und $\cdot$.

Beispiel 6.1.4. (a) Jede einelementige abelsche Gruppe $V = \{0\}$ kann man für jeden Körper K auf genau eine Weise zu einem K-Vektorraum machen (nämlich indem man die Skalarmultiplikation durch $\lambda \cdot 0 := 0$ für $\lambda \in K$ definiert). Mann nennt dann V einen *Nullvektorraum.*

(b) Sei der Körper K ein Unterring [→ 3.2.1] des kommutativen Ringes A [→ 3.1.1]. Dann ist A ein K-Vektorraum vermöge der Skalarmultiplikation $\cdot : K \times A \to A, (\lambda, a) \mapsto \lambda \cdot_A a$.

Zum Beispiel wird in dieser Weise der Polynomring $K[X]$ ein K-Vektorraum und die komplexen Zahlen bilden einen reellen Vektorraum (d.h. $\mathbb{C}$ ist ein $\mathbb{R}$-Vektorraum). Es ist $\mathbb{C}$ aber auch ein $\mathbb{C}$-Vektorraum. Allgemeiner ist jeder Körper ein Vektorraum über sich selbst, d.h. K ist ein K-Vektorraum vermöge der Körpermultiplikation als Skalarmultiplikation.

(c) Ist A eine Menge, so wird die abelsche Gruppe $\mathscr{P}(A)$ aus 2.1.3 (e) ein $\mathbb{F}_2$-Vektorraum vermöge

$$0_{\mathbb{F}_2} B := \emptyset \text{ und } 1_{\mathbb{F}_2} B := B \text{ für } B \in \mathscr{P}(A).$$

Satz und Definition 6.1.5. [→ 2.1.11] *Sei K ein Körper, I eine Menge und für jedes $i \in I$ ein K-Vektorraum V_i gegeben. Das in 2.1.11 definierte direkte Produkt $\prod_{i \in I} V_i$ der additiven Gruppe der V_i wird dann vermöge der „punktweisen" Skalarmultiplikation*

$$K \times \prod_{i \in I} V_i \to \prod_{i \in I} V_i, (a, f) \mapsto (i \mapsto af(i))$$

zu einem K-Vektorraum, dem direkten Produkt *der K-Vektorräume V_i $(i \in I)$. Für $a \in K$ und $f \in \prod_{i \in I} V_i$ gilt $(af)(i) = af(i)$ für alle $i \in I$.*

Beweis. Übung. □

Korollar 6.1.6. [→ 2.1.12] *Sei K ein Körper. Sind $n \in \mathbb{N}_0$ und $V_1, \ldots, V_n$ K- Vektorräume, so ist die abelsche Gruppe $V_1 \times \ldots \times V_n$ zusammen mit der durch*

$$a \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} := \begin{pmatrix} av_1 \\ \vdots \\ av_n \end{pmatrix} \quad (a \in K, v_1 \in V_1, \ldots, v_n \in V_n)$$

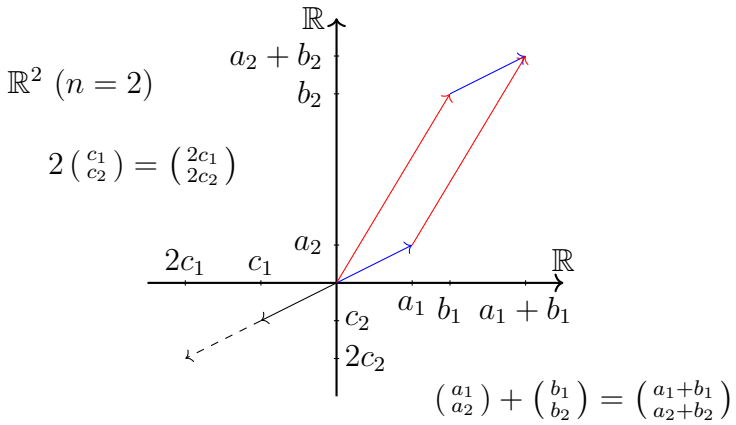
definierte Skalarmultiplikation ein K-Vektorraum (wir schreiben hier n-Tupel wieder als „Spaltenvektoren“). Insbesondere ist für jeden K-Vektorraum V und jedes $n \in \mathbb{N}_0$ die abelsche Gruppe $V^n = \underbrace{V \times \ldots \times V}_{n\text{-mal}}$ ein K-Vektorraum.

Bemerkung 6.1.7. Sei K ein Körper und $n \in \mathbb{N}_0$. Als sehr wichtiger Spezialfall von 6.1.6 ergibt sich K^n ist mit der durch

$$\lambda \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} := \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix} \quad (\lambda \in K, a_1, \ldots, a_n \in K) [\rightarrow \text{ 5.1.3 (c)}]$$

definierten Skalarmultipliaktion ein K-Vektorraum (beachte $K^0 = \{()\}$ ist ein Nullvektorraum [→ 6.1.4 (a)]).

Beispiel 6.1.8. $\mathbb{R}^n = \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n-\text{mal}}$



Vektoraddition: Aneinanderfügen von Pfeilen, Skalarmultiplikation: Strecken.

Definition 6.1.9. [→ 2.2.1, 3.2.1] Seien U und V K-Vektorräume. Dann heißt U ein *Untervektorraum* oder *(linearer) Unterraum* (*UR*) von V, wenn die additive Gruppe [→ 6.1.2] von U eine Untergruppe der additiven Gruppe von V ist und $\forall a \in K : \forall u \in U : a \cdot_U u = a \cdot_V u$.

Proposition 6.1.10. [→ 2.2.2, 3.2.2] *Sei V ein K-Vektorraum und U eine Menge. Dann ist U genau dann Trägermenge eines Unterraums von V, wenn gilt:*

(a) $U \subseteq V$,

(b) $0_V \in U$,

(c) $\forall u, v \in U : u +_V v \in U$ *und*

(d) $\forall \lambda \in K : \forall u \in U : \lambda \cdot_V u \in U$.

In diesem Fall gibt es genau *ein Paar $(+_U, \cdot_U)$, mit dem $(K, +_K, \cdot_K, U, +_U, \cdot_U)$ ein Unterraum von $(K, +_K, \cdot_K, V, +_V, \cdot_V)$ wird. Es gilt dann*

(a') $0_U = 0_V$,

(b') $\forall u, v \in U : u +_U v = u +_V v$ *und*

(c') $\forall \lambda \in K : \forall u \in U : \lambda \cdot_U = \lambda \cdot_V v$.

Beweis. Einfach mit 2.2.2 unter Verwendung von $-_V u \overset{6.1.2\,(g)}{=} (-1) \cdot_V u$ für alle $u \in V$. □

Beispiel 6.1.11. (a) Ist K ein Körper, so ist die Lösungsmenge eines *homogenen* linearen Gleichungssystems über K [→ 5.1.1] in n Unbekannten ein Untervektorraum des Vektorraums K^n. Dies besagt gerade 5.1.3.

(b) Die Mengen $\mathbb{R}$ der reellen und $\left\{ a\mathring{\imath} \mid a \in \mathbb{R} \right\}$ der *rein imaginären* Zahlen sind jeweils Unterräume des $\mathbb{R}$-Vektorraums $\mathbb{C}$, aber nicht des $\mathbb{C}$-Vektorraums $\mathbb{C}$.

(c) Ist K ein Körper, so ist für jedes $d \in \{-\infty\} \cup \mathbb{N}_0$ $K[X]_d := \{p \in K[X] \mid \deg p \leq d\}$ ein Unterraum des K-Vektorraums $K[X]$. (schlampig gesagt: ein K-Unterraum von $K[X]$).

Proposition 6.1.12. [→ 2.2.5, 3.3.8] *Sei V ein K-Vektorraum und M eine Menge von Unterräumen von V. Dann ist $\bigcap M$ ein Unterraum von V (mit $\bigcap \emptyset := V$).*

Beweis. Einfach mit 2.2.5. □

Satz und Definition 6.1.13. [→ 2.2.6, 3.3.9] *Sei V ein K-Vektorraum und $E \subseteq V$. Dann gibt es* den kleinsten *Unterraum U von V mit $E \subseteq U$. Man nennt ihn den von E in V* $\left\{ \begin{matrix} \text{erzeugten} \\ \text{aufgespannten} \end{matrix} \right\}$ *Unterraum oder* $\left\{ \begin{matrix} \textit{die lineare Hülle} \\ \textit{den Spann} \end{matrix} \right\}$ *von E (in V) und notiert in mit* $\operatorname{span}(E)$.

Beweis. Völlig analog zum Beweis von 2.2.6 (benutze 6.1.12 statt 2.2.5). □

Satz 6.1.14. [→ 2.2.7, 3.3.10] *Sei V ein K-Vektorraum und $E \subseteq V$. Dann*

$$\operatorname{span}(E) = \left\{ \underbrace{\sum_{i=1}^{n} \lambda_i v_i}_{\text{„Linearkombinationen“}} \;\middle|\; n \in \mathbb{N}_0, \lambda_1, \ldots, \lambda_n \in K, v_1, \ldots, v_n \in E \right\}.$$

Beweis. Der Beweis ist analog zum Beweis von Satz 2.2.7 und stellt gleichzeitig einen neuen Beweis für 6.1.13 dar. □

§6.2 Basen

Definition 6.2.1. Seien V ein K-Vektorraum, $n \in \mathbb{N}_0$ und $v_1, \ldots, v_n \in V$.

(a) Man nennt

$$\operatorname{span}(v_1,\dots,v_n) := \operatorname{span}(\{v_1,\dots,v_n\}) \overset{6.1.1(\text{D'})}{=} \left\{ \sum_{i=1}^{n} \lambda_i v_i \;\middle|\; \lambda_1,\dots,\lambda_n \in K \right\}$$

den Spann von $v_1,\dots,v_n$ [→5.1.5, 6.1.13, 6.1.14]. Man sagt, $v_1,\dots,v_n$ *erzeugen* V (oder *spannen* V auf oder bilden ein *Erzeugendensystem* von V), wenn

$$V = \operatorname{span}(v_1,\dots,v_n).$$

(b) Man nennt $v_1,\dots,v_n$ *linear unabhängig*, wenn

$$\operatorname{span}(v_1,\dots,v_{i-1},v_{i+1},\dots,v_n) \neq \operatorname{span}(v_1,\dots,v_n)$$

für alle $i \in \{1,\dots,n\}$. Andernfalls nennt man $v_1,\dots,v_n$ *linear abhängig*

(c) Es heißt $(v_1,\dots,v_n)$ eine *(geordnete) Basis* von V, falls $v_1,\dots,v_n$ linear unabhängig sind und V erzeugen. Man sagt dann auch, $v_1,\dots,v_n$ *bilden* eine Basis von V.

Beispiel 6.2.2. Sei K ein Körper und $n \in \mathbb{N}_0$. Ist für $i \in \{1,\dots,n\}$

$$e_i := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \leftarrow i\text{-te Stelle} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in K^n$$

der i-te *Standardvektor* (oder *kanonische Einheitsvektor*) im K^n, so ist

$$\underline{e} := (e_1,\dots,e_n)$$

eine Basis des K^n. Man nennt sie die *Standardbasis* (auch natürliche Basis oder kanonische Basis) des K^n.

Proposition 6.2.3. *Seien V ein K-Vektorraum, $n \in \mathbb{N}_0$ und $v_1,\dots,v_n \in V$. Dann sind folgende Aussagen äquivalent:*

(a) *$v_1,\dots,v_n$ sind linear unabhängig.*

(b) *$\forall \lambda_1,\dots,\lambda_n \in K : (\lambda_1 v_1 + \dots + \lambda_n v_n = 0 \implies \lambda_1 = \dots = \lambda_n = 0)$, das heißt „der Nullvektor lässt sich aus den v_i nur trivial kombinieren“.*

Beweis. Die logischen Negationen von (a) und (b) sind:

(ā) Es gibt ein $i \in \{1,\dots,n\}$ mit $\operatorname{span}(v_1,\dots,v_{i-1},v_{i+1},\dots,v_n) = \operatorname{span}(v_1,\dots,v_n)$.

(b̄) Es gibt $i \in \{1,\dots,n\}$ und $\lambda_1,\dots,\lambda_n \in K$ mit $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ und $\lambda_i \neq 0$.

Wir zeigen, dass für alle $i \in \{1,\dots,n\}$ äquivalent sind:

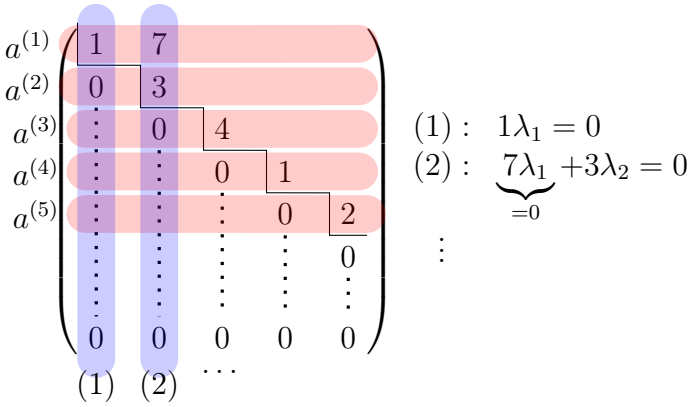
(ã) $\operatorname{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) = \operatorname{span}(v_1, \ldots, v_n)$.

(b̃) Es gibt $\lambda_1, \ldots, \lambda_n \in K$ mit $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ und $\lambda_i \neq 0$.

Dies ist klar, da (ã) und (b̃) beide äquivalent sind zu $v_i \in \operatorname{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$. □

Bemerkung 6.2.4. (a) In §5.1 hatten wir gesehen, wie man aus einer gegebenen Matrix in reduzierter Stufenform eine Basis ihres Kerns ablesen kann [→5.1.14(a)].

(b) Die Nichtnullzeilen einer Matrix in Stufenform [→5.1.10] bilden eine Basis ihres Zeilenraums. In der Tat: Sie spannen den Zeilenraum gemäß seiner Definition 5.3.1 natürlich auf. Um zu zeigen, dass sie linear unabhängig sind, bezeichne r die Anzahl der Stufen und $a^{(i)}$ die i-te Zeile für jedes $i \in \{1, \ldots, r\}$. Seien $\lambda_1, \ldots, \lambda_r \in K$ mit $\sum_{i=1}^{r} \lambda_i a^{(i)} = 0$. Angenommen es gilt nicht $\lambda_1 = \cdots = \lambda_r = 0$. Dann gibt es ein $k \in \{1, \ldots, r\}$ mit $\lambda_1 = \cdots = \lambda_{k-1} = 0 \neq \lambda_k$. Ist j_k die k-te Stufenposition der Matrix, so folgt nun $\lambda_k a^{(k)}_{j_k} = 0$ und daher $\lambda_k = 0$. Widerspruch!



(c) Bemerkungen (a) und (b) zeigen, wie man mit dem Gauß-Verfahren aus §5.2 zu einer gegebenen Matrix $A \in K^{m \times n}$ (K ein Körper) Basen der Untervektorräume $\ker A$ und $\operatorname{row} A$ des K-Vektorraums K^n berechnen kann.

Lemma 6.2.5 (Austauschlemma von Steinitz)**.** [Ernst Steinitz *1871 †1928]
Sei $(v_1, \ldots, v_n)$ eine Basis des K-Vektorraums V. Seien $\lambda_1, \ldots, \lambda_n \in K$, $w = \sum_{i=1}^{n} \lambda_i v_i$ und $j \in \{1, \ldots, n\}$ mit $\lambda_j \neq 0$. Dann ist auch $(v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n)$ eine Basis von V.

Beweis. Wegen

$$v_j = \frac{1}{\lambda_j} \left(\sum_{\substack{i=1 \\ i \neq j}}^{n} (-\lambda_i) v_i + 1w \right)$$

gilt $v_i \in \operatorname{span}(v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n)$ für alle $i \in \{1, \ldots, n\}$ und daher

$$\operatorname{span}(v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n) = V.$$

Es ist noch zu zeigen, dass $v_1, \dots, v_{j-1}, w, v_{j+1}, \dots, v_n$ linear unabhängig sind. Seien hierzu $\mu_1, \dots, \mu_n \in K$ mit

$$\sum_{\substack{i=1 \\ i \neq j}}^{n} \mu_i v_i + \mu_j w = 0.$$

Zu zeigen ist $\mu_i = 0$ für alle $i \in \{1, \dots, n\}$. Es gilt

$$0 = \sum_{\substack{i=1 \\ i \neq j}}^{n} \mu_i v_i + \mu_j \sum_{\substack{i=1 \\ i \neq j}}^{n} \lambda_i v_i + \mu_j \lambda_j v_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} (\mu_i + \mu_j \lambda_i) v_i + \mu_j \lambda_j v_j.$$

Da $v_1, \dots, v_n$ linear unabhängig sind, folgt $\mu_j = 0$ (denn $\mu_j \lambda_j = 0$ und $\lambda_j \neq 0$) und somit $\mu_i = \mu_i + \mu_j \lambda_i = 0$ für alle $i \in \{1, \dots, n\} \setminus \{j\}$. □

Satz 6.2.6 (Austauschsatz von Steinitz)**.** *Sei $(v_1, \dots, v_n)$ eine Basis des Vektorraums V und seien $w_1, \dots, w_m \in V$ linear unabhängig. Dann gibt es paarweise verschiedene $i_1, \dots, i_m \in \{1, \dots, n\}$ so, dass $v_1, \dots, v_n$ nach Ersetzen von v_{i_j} durch w_j ($j \in \{1, \dots, m\}$) immer noch eine Basis von V ist. Insbesondere gilt $m \leq n$.*

Beweis. Induktion nach m.
$\underline{n = 0}$ nichts zu zeigen
$\underline{m - 1 \to m \quad (m \in \mathbb{N})}$ Nach Induktionsvoraussetzung und Umnummerierung bilden

$$w_1, \dots, w_{m-1}, v_m, \dots, v_n$$

eine Basis von V (insbesondere $m - 1 \leq n$). Wähle $\lambda_i, \mu_j \in K$ mit

$$w_m = \sum_{i=1}^{m-1} \lambda_i w_i + \sum_{j=m}^{n} \mu_j v_j.$$

Da $w_1, \dots, w_m$ linear unabhängig sind, gibt es ein $\mu_j \neq 0$ (insbesondere $m \leq n$). Wende nun das Austauschlemma 6.2.5 an. □

Definition 6.2.7. [→6.2.1] Sei V ein Vektorraum und $G \subseteq V$.

(a) Es heißt G ein *Erzeugendensystem* von V (auch G *erzeugt* V oder *spannt* V *auf*), wenn $\operatorname{span}(G) = V$ [→6.1.13].

(b) Es heißt G *linear unabhängig* in V, wenn $\operatorname{span}(G \setminus \{v\}) \neq \operatorname{span}(G)$ für alle $v \in G$. Andernfalls heißt G *linear abhängig* in V.

(c) Es heißt G eine (ungeordnete) *Basis* von V, wenn G ein linear unabhängiges Erzeugendensystem von V ist.

Beispiel 6.2.8. Sei K ein Körper.

(a) Ist $V = \{0\}$ ein Nullvektorraum [→6.1.4(a)], so ist $\emptyset$ die einzige Basis von V.

(b) Die Basen des K- Vektorraums K [→6.1.4(b)] sind genau die einelementigen Mengen $\{a\}$ mit $a \in K^\times$.

(c) Es ist $\{1, \mathring{i}\}$ eine Basis des $\mathbb{R}$-Vektorraums $\mathbb{C}$ [→4.2.3(b)].

(d) Die Menge der *Monome* $\{1, X, X^2, X^3, \dots\}$ ist eine Basis des K-Vektorraums $K[X]$ [→3.2.4, 3.2.6].

(e) Die Menge $\{e_1, \dots, e_n\}$ der Standardvektoren [→6.2.2] ist eine Basis des K- Vektorraums K^n.

(f) $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \dots \right\}$ ist keine Basis des K-Vektorraums

$$K^{\mathbb{N}} = \{f \mid f \colon \mathbb{N} \to K\} = \{(a_i)_{i \in \mathbb{N}} \mid \forall i \in \mathbb{N} : a_i \in K\}$$

aller Folgen in K [→6.1.5], denn der Spann ist nur der Untervektorraum

$$\{f \mid f : \mathbb{N} \to K, \underbrace{\{i \in \mathbb{N} \mid f(i) \neq 0\}}_{=:\mathrm{supp}(f)} \text{ endlich}\}$$

aller Folgen mit endlichem *Träger* (*support* auf Englisch).

(g) Ist V ein K-Vektorraum und $G \subseteq V$, so ist G ein Erzeugendensystem des Vektorraums $\mathrm{span}(G)$ und es ist G genau dann eine Basis von $\mathrm{span}(G)$, wenn G linear unabhängig ist.

Warnung 6.2.9. [→6.2.1, 6.2.7] Seien V ein Vektorraum und $v_1, \dots, v_n \in V$. Dann sind $v_1, \dots, v_n$ linear unabhängig in V genau dann, wenn die Menge $\{v_1, \dots, v_n\}$ linear unabhängig in V ist und $v_1, \dots, v_n$ *paarweise verschieden* sind. Dementsprechend bilden $v_1, \dots, v_n$ eine (geordnete) Basis von V genau dann, wenn die Menge $\{v_1, \dots, v_n\}$ eine (ungeordnete) Basis von V ist und $v_1, \dots, v_n$ *paarweise verschieden* sind.

Beispiel 6.2.10. $\{\binom{0}{1}, \binom{1}{1}, \binom{1}{1}\} \overset{\substack{1.1.1 \\ 1.1.3}}{=} \{\binom{0}{1}, \binom{1}{1}\}$ ist eine Basis von $\mathbb{R}^2$, aber $\binom{0}{1}, \binom{1}{1}, \binom{1}{1}$ bilden keine Basis von $\mathbb{R}^2$.

Proposition 6.2.11. *Sei V ein K-Vektorraum und $F \subseteq V$. Dann sind äquivalent:*

(a) *F ist linear unabhängig.*

(b) *Jede Teilmenge von F ist linear unabhängig.*

(c) *Jede endliche Teilmenge von F ist linear unabhängig.*

(d) *Alle paarweise verschiedenen $v_1, \dots, v_n \in F$ sind linear unabhängig.*

Beweis. **(a) $\Longrightarrow$ (b)** Gelte (a) und sei $E \subseteq F$. Sei $v \in E$. Zu zeigen ist

$$\operatorname{span}(E \setminus \{v\}) \neq \operatorname{span}(E),$$

was äquivalent zu $v \notin \operatorname{span}(E \setminus \{v\})$ ist. Es gilt aber sogar $v \notin \operatorname{span}(F \setminus \{v\})$.

(b) $\Longrightarrow$ (c) ist trivial.

(c) $\Longrightarrow$ (d) ist klar (vergleiche Warnung 6.2.9)

(d) $\Longrightarrow$ (a) zeigen wir durch Kontraposition: Gelte (a) nicht, das heißt es gebe ein $v \in F$ mit $\operatorname{span}(F \setminus \{v\}) = \operatorname{span}(F)$. Wir zeigen, dass (d) nicht gilt. Wegen $v \in \operatorname{span}(F) = \operatorname{span}(F \setminus \{v\})$ und 6.1.14 gibt es paarweise verschiedene $v_1, \ldots, v_n \in F \setminus \{v\}$ mit $v \in \operatorname{span}(v_1, \ldots, v_n)$. Dann sind $v_1, \ldots, v_n, v$ paarweise verschieden und linear abhängig, denn $\operatorname{span}(v_1, \ldots, v_n) = \operatorname{span}(v_1, \ldots, v_n, v)$. □

Notation 6.2.12. [→ 1.1.12] Eine Menge A heißt *echte* Teilmenge der Menge B, und wir schreiben $A \subset B$, wenn $A \subseteq B$ und $A \neq B$. Man bezeichnet dann B auch als *echte* Obermenge von A, wofür wir $B \supset A$ schreiben.

Bemerkung 6.2.13. Um den folgenden Satz 6.2.14 besser zu verstehen, sollte man zunächst sein Korollar 6.2.15 betrachten und beim ersten Lesen des Beweises $F := \emptyset$ und $G := V$ setzen. Beachte auch, dass im Beweises des Satzes die folgenden Inklusionen gelten:

$$F \subseteq C \subseteq B \subseteq D \subseteq G$$

Satz 6.2.14. *Sei V ein Vektorraum und $F \subseteq B \subseteq G \subseteq V$. Sei F linear unabhängig und G ein Erzeugendensystem von V. Dann sind äquivalent:*

(a) *B ist eine Basis von V.*

(b) *B ist ein Erzeugendensystem von V aber kein C mit $F \subseteq C \subset B$ ist ein Erzeugendensystem von V.*

(c) *B ist linear unabhängig aber kein D mit $B \subset D \subseteq G$ ist linear unabhängig.*

Beweis. Bezeichne K den Grundkörper von V [→6.1.2(c)].

(a) $\Longrightarrow$ (b) Sei B eine Basis von V. Gelte $C \subset B$. Wir zeigen $\operatorname{span}(C) \neq V$. Wähle hierzu $v \in B \setminus C$. Dies folgt aus

$$\operatorname{span}(C) \overset{C \subseteq B \setminus \{v\}}{\subseteq} \operatorname{span}(B \setminus \{v\}) \subset \operatorname{span}(B).$$

(b) $\Longrightarrow$ (c) Gelte (b) und $B \subset D \subseteq G$. Zu zeigen ist, dass B linear unabhängig und D linear abhängig ist. Um letzteres zu zeigen, reicht es $v \in D \setminus B$ zu wählen, denn für dieses gilt $B \subseteq D \setminus \{v\} \subseteq D \subseteq V$ und daher

$$V \overset{(b)}{=} \operatorname{span}(B) \subseteq \operatorname{span}(D \setminus \{v\}) \subseteq \operatorname{span}(D) \subseteq V,$$

also $\operatorname{span}(D \setminus \{v\}) = \operatorname{span}(D)$. Um zu zeigen, dass B linear unabhängig ist, wenden wir 6.2.11(d) zusammen mit 6.2.3 an und beachten dabei, dass $B = F \cup (B \setminus F)$. Seien also $v_1, \ldots, v_m \in F$ und $w_1, \ldots, w_n \in B \setminus F$ jeweils paarweise verschieden und $\lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_n \in K$ mit $\sum_{i=1}^m \lambda_i v_i + \sum_{j=1}^n \mu_j w_j = 0$. Zu zeigen ist $\lambda_i = 0$ und $\mu_j = 0$ für alle i und j. Wäre ein $\mu_j \neq 0$, so

$$w_j = \frac{1}{\mu_j} \left(\sum_{i=1}^m \lambda_i v_i + \sum_{\substack{\ell=1 \\ \ell \neq j}}^n \mu_\ell w_\ell \right) \in \operatorname{span}(B \setminus \{w_j\}),$$

also $V \overset{(b)}{=} \operatorname{span}(B) = \operatorname{span}(B \setminus \{w_j\})$ im Widerspruch zu (b). Also $\mu_j = 0$ für alle j. Da F linear unabhängig ist, gilt auch $\lambda_i = 0$ für alle i.

(c) $\Longrightarrow$ (a) Gelte (c). Zu zeigen ist $\operatorname{span}(B) = V$. Hierzu genügt es $G \subseteq \operatorname{span}(B)$ zu zeigen (denn dann $V = \operatorname{span}(G) \subseteq \operatorname{span}(B) = V$, also $\operatorname{span}(B) = V$). Sei also $v \in G$. Zu zeigen ist $v \in \operatorname{span}(B)$. Œ $v \notin B$. Dann $B \subset D := B \cup \{v\} \subseteq G$. Nach (c) ist D linear abhängig, das heißt es gibt nach 6.2.11 in Kombination mit 6.2.3 paarweise verschiedene $v_1, \ldots, v_n \in B$ und es gibt $\lambda_1, \ldots, \lambda_n, \lambda \in K$, die nicht alle null sind, mit $\lambda_1 v_1 + \cdots + \lambda_n v_n + \lambda v = 0$. Da B linear unabhängig ist, gilt $\lambda \neq 0$ und es gilt $v \in \operatorname{span}(v_1, \ldots, v_n) \subseteq \operatorname{span}(B)$ wie gewünscht. □

Korollar 6.2.15. *Sei V ein Vektorraum und $B \subseteq V$. Dann sind äquivalent:*

(a) *B ist eine Basis von V.*

(b) *B ist ein minimales Erzeugendensystem von V.*

(c) *B ist eine maximale linear unabhängige Teilmenge von V.*

Beweis. Nehme $F := \emptyset$ und $G := V$ im 6.2.14. □

Definition 6.2.16. Ein Vektorraum heißt *endlich erzeugt* (abgekürzt: e.e.), falls er ein endliches Erzeugendensystem besitzt.

Beispiel 6.2.17. Für jedes $d \in \{-\infty\} \cup \mathbb{N}_0$ ist der K-Vektorraum $K[X]_d$ [→6.1.11(c)] endlich erzeugt, denn es ist zum Beispiel $\{1, X, X^2, \ldots, X^d\}$ ein endliches Erzeugendensystem (sogar eine Basis). Dagegen ist der K-Vektorraum $K[X]$ nicht endlich erzeugt, denn jedes endliche $G \subseteq K[X]$ ist in einem der Unterräume $K[X]_d$ ($d \in \{-\infty\} \cup \mathbb{N}_0$) enthalten.

Korollar 6.2.18. *Jeder endlich erzeugte Vektorraum besitzt eine endliche Basis.*

Beweis. Entferne aus einem endlichen Erzeugendensystem sukzessive überflüssige Elemente bis ein minimales Erzeugendensystem vorliegt. Wende nun 6.2.15(b) an. □

Satz 6.2.19. *Sei V ein Vektorraum und G ein endliches Erzeugendensystem von V. Dann gilt für jede linear unabhängige Menge F von V, dass $\#F \leq \#G$.*

Beweis. Analog zum Beweis von 6.2.18 kann man Œ G als Basis voraussetzen. Dann folgt aber die Behauptung aus dem „insbesondere“ in der Aussage des Austauschsatzes von Steinitz 6.2.6. □

Korollar 6.2.20. *Sei V ein endlich erzeugter Vektorraum, $F \subseteq G \subseteq V$, F linear unabhängig und G ein Erzeugendensystem von V. Dann gibt es eine endliche Basis B von V mit $F \subseteq B \subseteq G$.*

Beweis. Starte mit $B := F$ und vergrößere solange es möglich ist B durch Hinzunahme eines Elements aus $G \setminus B$ so, dass B dabei linear unabhängig bleibt. Dieser Prozess muss wegen 6.2.19 nach endlich vielen Schritten abbrechen. Falls dann $B = G$ gilt, so ist B linear unabhängig und ein Erzeugendensystem, also eine Basis. Gilt aber dann $B \neq G$, so ist Bedingung 6.2.14(c) erfüllt und B ist wieder eine Basis. □

Bemerkung 6.2.21. Mit „transfiniten“ Beweistechniken (zum Beispiel dem „Zornschen Lemma“) werden wir in der Linearen Algebra II zeigen, dass Korollar 6.2.20 richtig bleibt mit „Vektorraum“ statt „endlich erzeugter Vektorraum“ und „Basis“ statt „endlicher Basis“. Insbesondere besitzt *jeder* Vektorraum eine Basis. In der Linearen Algebra I werden wir dies weder benutzen noch beweisen.

Satz 6.2.22. *Sei V ein Vektorraum. Es sind äquivalent:*

(a) *V ist endlich erzeugt.*

(b) *V hat eine endliche Basis.*

(c) *Jede linear unabhängige Teilmenge von V ist endlich.*

Beweis. (a) $\Longrightarrow$ (b) ist 6.2.18 und (b) $\Longrightarrow$ (c) folgt aus 6.2.19. Schließlich zeigen wir (c) $\Longrightarrow$ (a) durch Kontraposition: Es gelte (a) nicht. Dann finden wir rekursiv eine Folge $(v_n)_{n\in\mathbb{N}}$ mit $v_n \notin \operatorname{span}(v_1, \ldots, v_n)$ für alle $n \in \mathbb{N}$. Es ist dann $\{v_n \mid n \in \mathbb{N}\}$ unendlich und linear unabhängig, denn wären $\lambda_1, \ldots, \lambda_n \in K$ mit $\sum_{i=1}^{n} \lambda_i v_i = 0$ und $\lambda_n \neq 0$, so folgte $v_n = -\frac{1}{\lambda_n} \sum_{i=1}^{n-1} \lambda_i v_i \in \operatorname{span}(v_1, \ldots, v_{n-1})$, was der Wahl der Folge widerspricht. □

Satz 6.2.23. *Seien B und C Basen eines endlich erzeugten Vektorraums. Dann sind B und C endlich und es gilt $\#B = \#C$.*

Beweis. Wegen 6.2.22(c) sind B und C endlich. Wegen 6.2.19 gilt $\#B \leq \#C$ und $\#C \leq \#B$. □

Definition 6.2.24. Sei V ein Vektorraum. Die *Dimension* von V ist definiert durch

$$\dim(V) := \begin{cases} \#B & \text{falls } V \text{ endlich erzeugt ist und } B \text{ eine Basis von } V \text{ ist} \\ \infty & \text{falls } V \text{ nicht endlich erzeugt ist.} \end{cases}$$

Diese Definition ist wegen 6.2.23 sinnvoll.

Bemerkung 6.2.25. Sei V ein K-Vektorraum.

(a) Wenn man weiß, dass jeder Vektorraum eine Basis hat [→ 6.2.21], kann man äquivalent definieren $\dim(V) := \#B$ [→ 1.1.21], falls B eine Basis von V ist.

(b) $\dim(V) = 0 \iff V = \{0\}$

(c) $\dim(K^n) = n$ für alle $n \in \mathbb{N}_0$ [→6.2.8(e)]

(d) Es gilt V endlich erzeugt $\iff \dim(V) < \infty$. Statt von einem „endlich erzeugten" spricht man daher meist von einem „endlichdimensionalen" Vektorraum.

(e) Manchmal schreibt man $\dim_K(V)$ statt „Dimension von V als K-Vektorraum". Zum Beispiel gilt $\dim_{\mathbb{R}} \mathbb{C} = 2$ [→6.1.4(b), 6.2.8(c)] und $\dim_{\mathbb{C}} \mathbb{C} = 1$ [→6.2.8(b)].

Satz 6.2.26. [→6.2.1] *Sei V ein endlichdimensionaler Vektorraum der Dimension n. Seien $v_1, \ldots, v_n \in V$. Dann sind äquivalent:*

(a) *$v_1, \ldots, v_n$ erzeugen V.*

(b) *$v_1, \ldots, v_n$ sind linear unabhängig in V.*

(c) *$v_1, \ldots, v_n$ bilden eine Basis von V.*

Beweis. Es reicht (a) $\implies$ (c) und (b) $\implies$ (c) zu zeigen.

(a) $\implies$ (c) Wende Korollar 6.2.20 an mit $F = \emptyset$ und $G = \{v_1, \ldots, v_n\}$, um eine Basis B von V mit $B \subseteq G$ zu erhalten. Es muss nach 6.2.23 $\#B = n$ und daher $B = G$ gelten. Da insbesondere $v_1, \ldots, v_n$ paarweise verschieden sind, bilden sie eine Basis von V.

(b) $\implies$ (c) Wende Korollar 6.2.20 an mit $F = \{v_1, \ldots, v_n\}$ und $G = V$, um eine Basis B von V mit $F \subseteq B$ zu erhalten. Es muss nach 6.2.23 $\#B = n$ und daher $B = G$ gelten, denn $v_1, \ldots, v_n$ sind paarweise verschieden. Also bilden $v_1, \ldots, v_n$ eine Basis von V. □

Korollar 6.2.27. *Sei V ein endlichdimensionaler Vektorraum und U ein Untervektorraum von V* [→6.1.9]. *Dann*

$$U = V \iff \dim U = \dim V.$$

Beweis. Für die nichttriviale Richtung sei $n := \dim U = \dim V$. Zu zeigen ist $U = V$. Wähle eine Basis $v_1, \ldots, v_n$ von U. Da $v_1, \ldots, v_n$ linear unabhängig im Vektorraum V der Dimension n sind, bilden sie nach Satz 6.2.26 eine Basis von V, insbesondere ein Erzeugendensystem von V. Also gilt $U = \operatorname{span}(v_1, \ldots, v_n) = V$. □

Proposition 6.2.28. *Sei U ein Untervektorraum des Vektorraums V, so gilt* $\dim U \leq \dim V$.

Beweis. Œ $\dim V < \infty$. Ergänze eine Basis von U zu einer Basis von V. □

Bemerkung 6.2.29. Sei K ein Körper.

(a) Die Dimension des Kerns einer Matrix $A \in K^{m\times n}$ in reduzierter Stufenform mit r Stufen ist die Anzahl der freien Unbekannten des Gleichungssystems $Ax = 0$ ($x \in K^n$), also $n - r$ [→6.2.4(a)]. Dasselbe gilt für eine Matrix in Stufenform, da sich die Anzahl der Stufen bei Überführung in reduzierte Stufenform nicht ändert.

(b) Die Dimension des Zeilenraums einer Matrix $A \in K^{m\times n}$ in Stufenform ist die Anzahl der Stufen [→6.2.4(b)].

(c) Da sich Kern und Zeilenraum bei elementaren Zeilenoperationen nicht ändern, ist mit (a) und (b) klar, wie man deren Dimension berechnet.

Beispiel 6.2.30. $\begin{pmatrix}1\\3\\0\end{pmatrix}, \begin{pmatrix}2\\1\\3\end{pmatrix}, \begin{pmatrix}-1\\1\\0\end{pmatrix}$ spannen $\mathbb{R}^3$ auf, denn der Zeilenraum von

$$\begin{pmatrix}1&3&0\\2&1&3\\-1&1&0\end{pmatrix} \overset{\substack{Z_2\leftarrow Z_2-2Z_1\\ Z_3\leftarrow Z_3+Z_1}}{\sim} \begin{pmatrix}1&3&0\\0&-5&3\\0&4&0\end{pmatrix} \overset{\substack{Z_2\leftarrow Z_2+Z_3\\ Z_2\leftarrow -Z_2\\ Z_3\leftarrow \frac{1}{4}Z_3}}{\sim} \begin{pmatrix}1&3&0\\0&1&-3\\0&1&0\end{pmatrix} \overset{\substack{Z_2\leftarrow Z_2-Z_3\\ Z_2\leftrightarrow Z_3}}{\sim} \begin{pmatrix}1&3&0\\0&1&0\\0&0&-3\end{pmatrix}$$

ist dreidimensional.

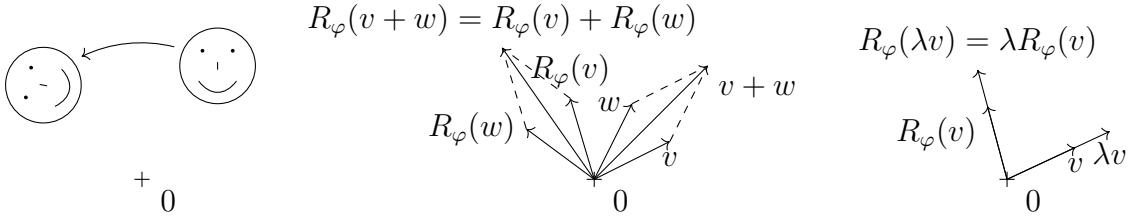
§6.3 Lineare Abbildungen

In diesem Abschnitt sei K stets ein Körper.

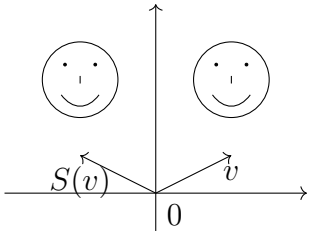
Definition 6.3.1. [→ 2.2.9, 2.2.12, 3.2.8, 3.2.10] Seien V und W K-Vektorräume. Dann heißt f ein *((K-)Vektorraum-)Homomorphismus* oder eine *(K-)lineare Abbildung* von V nach W, wenn $f : V \to W$ ein Gruppenhomomorphismus der additiven Gruppe von V und W ist („Additivität“) mit $\forall v \in V : \forall \lambda \in K : f(\lambda v) = \lambda f(v)$ („Homogenität“). Eine lineare Abbildung $f : V \to W$ heißt ((K-)Vektorraum-) $\left\{\begin{matrix}\textit{Einbettung} \text{ oder } \textit{Mono-}\\ \textit{Epi-}\\ \textit{Iso-}\end{matrix}\right\}\textit{morphismus}$, wenn f $\left\{\begin{matrix}\text{injektiv}\\ \text{surjektiv}\\ \text{bijektiv}\end{matrix}\right\}$ ist. Einen Vektorraumhomomorphismus $f : V \to V$ nennt man auch einen ((K-)Vektorraum-) *Endomorphismus* von V und, falls er bijektiv ist, ((K-)Vektorraum-)*Automorphismus* von V.

Beispiel 6.3.2. Die folgenden Abbildungen sind linear:

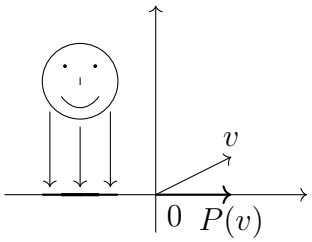
(a) $R_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ Drehung um den Winkel $\varphi \in \mathbb{R}$ (gegen den Uhrzeigersinn am Ursprung).



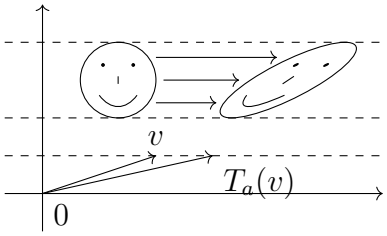
(b) $S : \mathbb{R}^2 \to \mathbb{R}^2, \binom{x}{y} \mapsto \binom{-x}{y}$ Spiegelung an der zweiten Koordinatenachse.



(c) $P : \mathbb{R}^2 \to \mathbb{R}^2, \binom{x}{y} \mapsto \binom{x}{0}$ Projektion auf die erste Koordinatenachse.



(d) $T_a : \mathbb{R}^2 \to \mathbb{R}^2, \binom{x}{y} \mapsto \binom{x+ay}{y}$ Scherung an der ersten Koordinatenachse um $a \in \mathbb{R}$.



(e) Sei $A \in K^{m\times n}$. Dann ist $f_A : K^n \to K^m, x \mapsto Ax$ [→ 5.1.9] linear (nachrechnen!). Es gilt $\ker f_A = \ker A$ [→ 2.3.10 (b), 5.3.1]. Weiter ist $\operatorname{im} A := \operatorname{im} f_A$ [→ 2.3.13] der von den Spalten von A aufgespannte Unterraum von K^m, den wir auch den *Spaltenraum* von A (vgl. auch 5.3.1) oder das *Bild* von A nennen.

(f) $D : K[X] \to K[X], \sum_{k=0}^n a_k X^k \mapsto \sum_{k=1}^n k a_k X^{k-1}$ $(n \in \mathbb{N}_0, a_0, \ldots, a_n \in K)$ mit $k = \underbrace{1 + \ldots + 1}_{k\text{-mal}} \in K$.
$D^{(d)} : K[X]_d \to K[X]_d, p \mapsto D_p$ $[d \in \mathbb{N}_0]$ formale Ableitung.

(g) $E_{a_1,\ldots,a_n} : K[X] \to K^n, p \mapsto \begin{pmatrix} p(a_1) \\ \vdots \\ p(a_n) \end{pmatrix}$ $[n \in \mathbb{N}_0, a_1, \ldots, a_n \in K]$.
$E^{(d)}_{a_1,\ldots,a_n} : K[X]_d \to K^n, p \mapsto \begin{pmatrix} p(a_1) \\ \vdots \\ p(a_n) \end{pmatrix}$ $[d \in \mathbb{N}_0]$

(h) Die komplexe Konjugation $C : \mathbb{C} \to \mathbb{C}, z \mapsto z^*$ ist $\mathbb{R}$-linear, aber nicht $\mathbb{C}$-linear. In der Tat: C ist nach 4.2.7 ein Automorphismus des kommutativen Ringes $\mathbb{C}$, insbesondere der additiven Gruppe von $\mathbb{C}$. Es gilt $C(\lambda z) = (\lambda z)^* = \lambda^* z^* = \lambda z^* = \lambda C(z)$ für alle $\lambda \in \mathbb{R}$ und $z \in \mathbb{C}$, aber $C(\mathring{i}1) = -\mathring{i} \neq \mathring{i} = \mathring{i}C(1)$.

Proposition 6.3.3. *Seien U, V, W K-Vektorräume.*

(a) *Sind $U \xrightarrow{f} V \xrightarrow{g} W$ Vektorraumhomomorphismen, so auch $g \circ f$.*

(b) *Ist $f : V \to W$ ein Vektorraumisomorphismus, so auch f^{-1}.*

Beweis. Nach 2.2.14 muss jeweils nur noch die Homogenität [→ 6.3.1] nachgerechnet werden.

(a) Sind $U \xrightarrow{f} V \xrightarrow{g} W$ Vektorraumhomomorphismen, $u \in U$ und $\lambda \in K$, so

$$(g \circ f)(\lambda u) = g(f(\lambda u)) = g(\lambda f(u)) = \lambda g(f(u)) = \lambda(g \circ f)(u).$$

(b) Sind $f : V \to W$ ein Vektorraumisomorphismus, $w \in W$ und $\lambda \in K$, so

$$\begin{aligned} f(f^{-1}(\lambda w)) = (f \circ f^{-1})(\lambda w) = \mathrm{id}_W(\lambda w) = \lambda w = \lambda \,\mathrm{id}_W(w) = \lambda(f \circ f^{-1})(w) \\ = \lambda(f(f^{-1}(w))) \overset{f \text{ Hom.}}{=} f(\lambda f^{-1}(w)) \end{aligned}$$

und daher $f^{-1}(\lambda w) = \lambda f^{-1}(w)$, da f injektiv ist.

□

Proposition 6.3.4. *Seien V und W K-Vektorräume, $B \subseteq V$ und $g : B \to W$. Ist B $\left\{\begin{array}{c}\text{eine Basis}\\ \text{ein Erzeugendensystem}\end{array}\right\}$ von V, so gibt es $\left\{\begin{array}{c}\text{genau}\\ \text{höchstens}\end{array}\right\}$ eine lineare Abbildung $f : V \to W$ mit $f|_B = g$.*

Beweis. Seien zunächst B ein Erzeugendensystem von V und $f_1, f_2 : V \to W$ linear mit $f_1|_B = g = f_2|_B$. Zu zeigen: $f_1 = f_2$. Sei $v \in V$. Wegen $V = \operatorname{span} B$ gibt es $n \in \mathbb{N}_0, v_1, \ldots, v_n \in B$ und $\lambda_1, \ldots, \lambda_n \in B$ mit $v = \sum_{i=1}^n \lambda_i v_i$. Dann gilt

$$f(v) = \sum_{i=1}^n \lambda_i f_1(v_i) = \sum_{i=1}^n \lambda_1 g(v_i) = \sum_{i=1}^n \lambda_i f_2(v_i) = f_2(v).$$

Sei nun B sogar eine Basis von V. Dann gibt es für jedes $v \in V$ eine eindeutig bestimmte Familie $(\lambda u)_{u \in B}$ in K mit endlichem Träger $\{u \in B \mid \lambda u \neq 0\}$ und $v = \sum_{u \in B} \lambda_u u := \sum_{\substack{u \in B \\ \lambda_u \neq 0}} \lambda_u u$. Daher ist

$$f : V \to W, \sum_{u \in B} \lambda_u u \mapsto \sum_{u \in B} \lambda_u g(u) \qquad ((\lambda_u)_{u \in B} \in K^B \text{ mit endlichem Träger})$$

eine wohldefinierte Abbildung.
Zu zeigen: f linear. Seien $(\lambda_u)_{u\in B}$ und $(\mu_u)_{u\in B}$ Familien in K mit endlichen Trägern und $\lambda \in K$.
Dann

$$f\left(\sum_{u\in B}\lambda_u u+\sum_{u\in B}\mu_u u\right) \overset{6.1.1\ (D')}{=} f\left(\sum_{u\in B}(\lambda_u+\mu_u)u\right)$$
$$=\sum_{u\in B}(\lambda_u+\mu_u)g(u) \overset{(D')}{=} \sum_{u\in B}\lambda_u g(u)+\sum_{u\in B}\mu_u g(u)$$
$$= f\left(\sum_{u\in B}\lambda_u u\right)+f\left(\sum_{u\in B}\mu_u u\right) \quad \text{und}$$

$$f\left(\lambda\sum_{u\in B}\lambda_u u\right) \overset{(\vec{D})}{=} f\left(\sum_{u\in B}\lambda(\lambda_u u)\right) \overset{(V)}{=} f\left(\sum_{u\in B}(\lambda\lambda_u)u\right)=\sum_{u\in B}(\lambda\lambda_u)g(u)$$
$$\overset{(V)}{=}\sum_{u\in B}\lambda(\lambda_u g(u)) \overset{(\vec{D})}{=} \lambda\sum_{u\in B}\lambda_u g(u)=\lambda f\left(\sum_{u\in B}\lambda_u u\right).$$

□

Beispiel 6.3.5. (a) Um zu zeigen, dass für alle $p\in\mathbb{R}[X]_2$ gilt:

$$(*)\qquad \int_{-1}^{1} p(x)\,dx = p\left(\frac{1}{\sqrt{3}}\right)+p\left(-\frac{1}{\sqrt{3}}\right),$$

reicht es wegen der Linearität der beiden Abbildungen $\mathbb{R}[X]_2\to\mathbb{R}, p\mapsto\int_{-1}^{1}p(x)\,dx$ und $\mathbb{R}[X]_2\to\mathbb{R}, p\mapsto p\left(\frac{1}{\sqrt{3}}\right)+p\left(-\frac{1}{\sqrt{3}}\right)$ zu zeigen, dass $(*)$ für alle $p\in\{1,X,X^2\}$ gilt, denn $\{1,X,X^2\}$ erzeugt den $\mathbb{R}$-Vektorraum $\mathbb{R}[X]_2$.

$$\int_{-1}^{1}1\,dx = 2 = 1+1$$

$$\int_{-1}^{1}x\,dx = 0 = \frac{1}{\sqrt{3}}-\frac{1}{\sqrt{3}}$$

$$\int_{-1}^{1}x^2\,dx = \left[\frac{x^3}{3}\right]_{x=-1}^{1} = \frac{1}{3}+\frac{1}{3} = \left(\frac{1}{\sqrt{3}}\right)^2+\left(-\frac{1}{\sqrt{3}}\right)^2$$

(b) Die formale Ableitung $D: K[X]\to K[X]$ aus 6.3.2 (f) hätte man wegen 6.2.8 (d) auch wie folgt definieren können: „Sei $D: K[X]\to K[X], 1\mapsto 0, X^k\mapsto kX^{k-1}$ $(k\in\mathbb{N})$ linear."

Proposition 6.3.6. *Seien V ein K-Vektorraum, $v_1, \dots, v_n \in V$ und $\underline{v} = (v_1, \dots, v_n)$. Die Abbildung* $\mathrm{vec}_{\underline{v}} : K^n \to V, \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \mapsto \sum_{i=1}^n \lambda_i v_i$ *ist linear. Sie ist* $\left\{ \begin{array}{c} \text{injektiv} \\ \text{bijektiv} \\ \text{surjektiv} \end{array} \right\}$ *genau dann, wenn* $v_1, \dots, v_n$ $\left\{ \begin{array}{c} \text{linear unabhängig in } V \text{ sind} \\ \text{eine Basis von } V \text{ bilden} \\ V \text{ erzeugen} \end{array} \right\}$.

Beweis. $\mathrm{vec}_{\underline{v}}$ ist offensichtlich linear, da es die eindeutig bestimmte Abbildung $K^n \to V$ ist, die e_i [→ 6.2.2] auf v_i abbildet. Es gilt $\mathrm{vec}_{\underline{v}}$ injektiv $\iff \ker \mathrm{vec}_{\underline{v}} = \{0\} \overset{6.2.3\,(a)}{\iff}$ $v_1, \dots, v_n$ linear unabhängig in V und $\mathrm{vec}_{\underline{v}}$ surjektiv $\iff \mathrm{im}\,\mathrm{vec}_{\underline{v}} \overset{6.2.1\,(a)}{\iff} v_1, \dots, v_n$ erzeugen V. □

Notation und Proposition 6.3.7. *Sei $\underline{v} = (v_1, \dots, v_n)$ eine Basis des K-Vektorraums V. Dann sind*

$$\mathrm{vec}_{\underline{v}} : K^n \to V, \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \mapsto \sum_{i=1}^n \lambda_i v_i \quad \text{und}$$

$$\mathrm{coord}_{\underline{v}} := \mathrm{vec}_{\underline{v}}^{-1} : V \to K^n, \sum_{i=1}^n \lambda_i v_i \mapsto \underbrace{\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}}_{\text{„Koordination“}} \qquad (\lambda_1, \dots, \lambda_n \in K)$$

nach 6.3.6 und 6.3.3 Vektorraumisomorphismen.

Satz 6.3.8. *Sei V ein K-Vektorraum mit Basis $\underline{v} = (v_1, \dots, v_n)$. Sei W ein weiterer K-Vektorraum und $f : V \to W$ linear. Genau dann ist f* $\left\{ \begin{array}{c} \text{injektiv} \\ \text{bijektiv} \\ \text{surjektiv} \end{array} \right\}$, *wenn* $f(v_1), \dots, f(v_n)$ $\left\{ \begin{array}{c} \text{linear unabhängig in } W \text{ sind} \\ \text{eine Basis von } W \text{ bilden} \\ W \text{ erzeugen} \end{array} \right\}$.

Beweis. $K^n \xrightarrow{\mathrm{vec}_{\underline{v}}} V \xrightarrow{f} W$
Da $\mathrm{vec}_{\underline{v}}$ bijektiv ist, kann man f durch $f \circ \mathrm{vec}_{\underline{v}}$, V durch K^n und $\underline{v}$ durch $\underline{e}$ ersetzen. Wende nun 6.3.6 an unter Beachtung von $f \circ \mathrm{vec}_{\underline{v}} = \mathrm{vec}_{(f(v_1),\dots,f(v_n))}$. □

Definition 6.3.9. [→ 2.2.15] Zwei K-Vektorräume V und W heißen *isomorph*, wenn es einen Isomorphismus von V nach W gibt, in Zeichen $V \cong W$.

Satz 6.3.10. *Seien V und W K-Vektorräume, von denen mindestens einer endlich dimensioniert ist. Dann $V \cong W \iff \dim V = \dim W$.*

Beweis. Œ V endlich dimensioniert [→ 6.2.25 (d)], etwa mit Basis $(v_1, \dots, v_n)$ [→ 6.2.22].

„$\Longrightarrow$“ Gelte $V \cong W$. Wähle Isomorphismus $f : V \to W$. Dann $(f(v_1), \dots, f(v_n))$ Basis von W [→ 6.3.8]. Also $\dim V = \dim W$ [→ 6.2.24].

„$\Longleftarrow$“ Gelte $\dim V = \dim W$. Dann $\dim W = n$, d.h. es gibt eine Basis $(w_1, \ldots, w_n)$ von W. Definiere lineare Abbildung $f : V \to W, v_i \mapsto w_i$ [$\to$ 6.3.4]. Nach 6.3.8 ist f bijektiv, also ein Isomorphismus.

□

Korollar 6.3.11. *Jeder Vektorraum der Dimension $n \in \mathbb{N}_0$ ist isomorph zu K^n.*

Bemerkung 6.3.12. „Bis auf Isomorphie“ gibt es also keine anderen endlich dimensionalen K-Vektorräume als die K^n ($n \in \mathbb{N}_0$). Dabei entspricht die Wahl einer geordneten Basis [$\to$ 6.2.1 (c)] der Wahl eines Isomorphismus:

Satz 6.3.13. *Sei V ein K-Vektorraum mit $n := \dim V < \infty$. Die Zuordnungen*

$$\begin{aligned} \underline{v} &\mapsto \mathrm{vec}_{\underline{v}} \\ (f(e_1), \ldots, f(e_n)) &\mapsfrom f \end{aligned}$$

vermitteln eine Bijektion [$\to$ 1.2.7] zwischen der Menge der geordneten Basen von V und der Menge der Vektorraumisomorphismen $K^n \to V$.

Beweis. Zu zeigen:

(a) Ist $\underline{v}$ Basis von V, so ist $\mathrm{vec}_{\underline{v}}$ ein Isomorphismus.

(b) Ist $f : K^n \to V$ ein Isomorphismus, so $(f(e_1), \ldots, f(e_n))$ Basis von V.

(c) Ist $\underline{v} = (v_1, \ldots, v_n)$ Basis von V, so gilt $\underline{v} = (\mathrm{vec}_{\underline{v}}(e_1), \ldots, \mathrm{vec}_{\underline{v}}(e_n))$.

(d) Ist $f : K^n \to V$ ein Isomorphismus, so $f = \mathrm{vec}_{(f(e_1),\ldots,f(e_n))}$.

(a) folgt aus 6.3.5, (b) aus 6.3.8, (c) aus 6.3.7 und (d) aus 6.3.4. □

§7 Matrizen

[Arthur Cayley *1821, †1895]

In diesem Kapitel sei stets K ein Körper.

§7.1 Matrixdarstellungen von linearen Abbildungen

Definition 7.1.1. Seien V und W K-Vektorräume mit Basen [→ 6.2.1 (c), 6.3.7] $\underline{v} = (v_1, \ldots, v_n)$ und $\underline{w} = (w_1, \ldots, w_m)$. Eine Matrix $A \in K^{m\times n}$ heißt *Darstellungsmatrix* einer linearen Abbildung $f : V \to W$ bezüglich der Basen $\underline{v}$ und $\underline{w}$, falls [→ 6.3.2 (e)]

$$(*) \qquad f = \mathrm{vec}_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}}$$

$$\begin{array}{ccccc} \sum_{j=1}^n \lambda_j v_j & & V & \xrightarrow{\quad f \quad} & W & & \sum_{i=1}^n \mu_i w_i \\ \Big\updownarrow & & \mathrm{coord}_{\underline{v}}\Big\downarrow\cong & & \cong\Big\uparrow \mathrm{vec}_{\underline{w}} & & \Big\updownarrow \\ \begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_n\end{pmatrix} & & K^n & \xrightarrow[x \mapsto Ax]{} & K^m & & \begin{pmatrix}\mu_1\\ \vdots \\ \mu_n\end{pmatrix} \end{array}$$

Bemerkung 7.1.2. In der Situation von 7.1.1 gilt

$$\begin{aligned} (*) &\overset{\mathrm{vec}_{\underline{w}}^{-1}=\mathrm{coord}_{\underline{w}}}{\iff} \mathrm{coord}_{\underline{w}} \circ f = f_A \circ \mathrm{coord}_{\underline{v}} \\ &\iff \forall j \in \{1,\ldots,n\} : \mathrm{coord}_{\underline{w}}(f(v_j)) = \underbrace{f_A(\underbrace{\mathrm{coord}_{\underline{v}}(v_j)}_{e_j})}_{=Ae_j} \\ &\iff \forall j \in \{1,\ldots,n\} : Ae_j = \mathrm{coord}_{\underline{w}}(f(v_j)) \end{aligned}$$

„In den Spalten stehen die *Koordinaten* der Bilder der Basisvektoren."

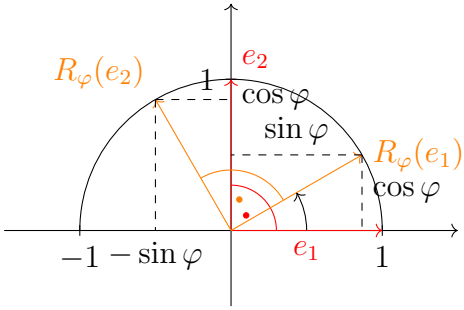
Zu jeder linearen Abbildung zwischen endlichdimensionalen Vektorräumen gibt es also bezüglich gegebener Basen jeweils *genau eine* Darstellungsmatrix.

Notation 7.1.3. $M(f, \underline{v}, \underline{w})$ steht für das eindeutig bestimmte A aus 7.1.1.

Beispiel 7.1.4. [→ 6.3.2] $\underline{e} = (e_1, e_2) = (\binom{1}{0}, \binom{0}{1})$ Standardbasis des $\mathbb{R}^2$ [→ 6.2.2].

(a) $R_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ Drehung um φ

$$R_\varphi(e_1) = \begin{pmatrix}\cos\varphi\\ \sin\varphi\end{pmatrix}$$
$$R_\varphi(e_2) = \begin{pmatrix}-\sin\varphi\\ \cos\varphi\end{pmatrix}$$
$$R_\varphi(e_1) = \underline{(\cos\varphi)}e_1 + \underline{(\sin\varphi)}e_2$$
$$R_\varphi(e_2) = \underline{(-\sin\varphi)}e_1 + \underline{(\cos\varphi)}e_2$$
$$M(R_\varphi, \underline{e}, \underline{e}) = \begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}$$



(b) $S(e_1) = \underline{(-1)}e_1 + \underline{0}\cdot e_2$

$S(e_2) = \underline{0}e_1 + \underline{1}e_2$

$M(S, \underline{e}, \underline{e}) = \begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}$

$\underline{v} := (\binom{0}{1}, \binom{1}{1})$ und $\underline{w} := (\binom{2}{1}, \binom{1}{-1})$ sind Basen des $\mathbb{R}^2$.

$$S(v_1) = S\binom{0}{1} = \binom{0}{1} = \frac{1}{3}\binom{2}{1} + \left(-\frac{2}{3}\binom{1}{-1}\right) = \frac{1}{3}w_1 + \left(-\frac{2}{3}\right)w_2$$

$$S(v_2) = S\binom{1}{1} = \binom{-1}{1} = 0\binom{2}{1} + (-1)\binom{1}{-1} = 0w_1 + (-1)w_2$$

$$M(S, \underline{v}, \underline{w}) = \begin{pmatrix}\frac{1}{3} & 0\\ -\frac{2}{3} & -1\end{pmatrix}$$

(c) $M(P, \underline{e}, \underline{e}) = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$, denn $P(e_1) = 1e_1 + 0e_2$ und $P(e_2) = 0e_1 + 0e_2$.

(d) $M(T_a, \underline{e}, \underline{e}) = \begin{pmatrix}1 & a\\ 0 & 1\end{pmatrix}$, denn $T_a(e_1) = 1e_1 + 0e_2$ und $T_a(e_2) = ae_1 + e_2$.

(e) Schreibe $\underline{v}$ und $\underline{w}$ für die Standardbasen [→ 6.2.2] des K^n und K^m. Es gilt $\mathrm{vec}_{\underline{v}} = \mathrm{id}_{K^n}$ und $\mathrm{vec}_{\underline{w}} = \mathrm{id}_{K^m}$. Daher $\mathrm{coord}_{\underline{v}} = \mathrm{id}_{K^n}^{-1} = \mathrm{id}_{K^n}$ und somit $f_A = \mathrm{id}_{K^m} \circ f_A \circ \mathrm{id}_{K^n} = \mathrm{vec}_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}}$, das heißt $M(f_A, \underline{v}, \underline{w}) = A$.

(f) $\underline{v} := (1, X, \ldots, X^d)$ ist Basis von $K[X]_d$ [→ 6.2.8 (d)]

$$D^{(d)}(X^k) = \begin{cases}0 & \text{falls } k = 0\\ kX^{k-1} & \text{falls } k \in \{1, \ldots, d\}\,.\end{cases}$$

$$\text{Also } M(D^{(d)}, \underline{v}, \underline{v}) = \begin{pmatrix}0 & 1 & 0 & \cdots & 0\\ \vdots & 0 & 2 & \ddots & \vdots\\ \vdots & \vdots & 0 & \ddots & 0\\ \vdots & \vdots & \vdots & \ddots & d\\ 0 & 0 & 0 & \cdots & 0\end{pmatrix},\ D^{(d)}(X^2) = 2X = \underline{0}\cdot 1 + \underline{2}\cdot X + \underline{0}\cdot X^2 + \ldots$$

(g) $\underline{v}$ wie eben, $\underline{w} :=$ Standardbasis des K^n.

$$E^{(d)}_{a_1,\dots,a_n}(X^k) = \begin{pmatrix} a_1^k \\ \vdots \\ a_n^k \end{pmatrix}$$

$$M(E_{a_1,\dots,a_n}, \underline{v}, \underline{w}) = \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^d \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^d \end{pmatrix}$$

Vandermonde-Matrix
[Alexandre-Théophile Vandermonde, *1735, †1796]

(h) $\underline{v} := (1, \mathring{i})$ ist Basis des $\mathbb{R}$-Vektorraums $\mathbb{C}$ [→ 6.2.8 (c)].

$$C(1) = \underline{1} \cdot 1 + \underline{0} \cdot \mathring{i}, \quad C(\mathring{i}) = \underline{0} \cdot 1 + (\underline{-1})\mathring{i}$$

$$M(C, \underline{v}, \underline{v}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$w := (1 + \mathring{i}, 1 - \mathring{i})$ ist auch Basis des $\mathbb{R}$-Vektorraums $\mathbb{C}$.

$$C(1 + \mathring{i}) = 1 - \mathring{i}, \quad C(1 - \mathring{i}) = 1 + \mathring{i}$$

$$M(C, \underline{w}, \underline{w}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$C(1) = 1 = \frac{1}{\underline{2}}(1 + \mathring{i}) + \frac{1}{\underline{2}}(1 - \mathring{i}), \quad C(\mathring{i}) = -\mathring{i} = \underline{\left(-\frac{1}{2}\right)}(1 + \mathring{i}) + \frac{1}{\underline{2}}(1 - \mathring{i})$$

$$M(C, \underline{v}, \underline{w}) = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$C(1 + \mathring{i}) = \underline{1} \cdot 1 + \underline{(-1)}\mathring{i}, \quad C(1 - \mathring{i}) = \underline{1} \cdot 1 + \underline{1} \cdot \mathring{i}$$

$$M(C, \underline{w}, \underline{v}) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

Erinnerung 7.1.5 (Spezialfall von 6.1.5). Ist I eine Menge und V ein K-Vektorraum, so ist auch $V^I \overset{1.1.27}{=} \{f \mid f : I \to V\}$. ein K-Vektorraum vermöge $(f + g)(i) = f(i) + g(i)$ und $(\lambda f)(i) = \lambda(f(i))$ für alle $f, g \in V^I$ und $\lambda \in K$.

Der Spezialfall $I = \{1, \dots, m\} \times \{1, \dots, n\}$ [→ 5.1.8] und $V = K$ liefert den K-Vektorraum $K^{m \times n}$ der $m \times n$-Matrizen über K:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

$$\lambda \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix} \qquad (a_{ij}, b_{ij}, \lambda \in K)$$

Notation und Proposition 7.1.6. *Sind V und W K-Vektorräume, so ist*

$$\mathrm{Hom}(V,W) := \{f \mid f : V \to W \text{ linear}\}$$

ein Unterraum des K-Vektorraums W^V.

Beweis. Nach 6.1.10 ist zu zeigen:

(a) $0 \in \mathrm{Hom}(V,W)$ (wobei $0 : V \to W, v \mapsto 0_W$)

(b) $\forall f, g \in \mathrm{Hom}(V,W) : f + g \in \mathrm{Hom}(V,W)$

(c) $\forall f \in \mathrm{Hom}(V,W) : \forall \mu \in K : \mu f \in \mathrm{Hom}(V,W)$

Zu (a). $0(v_1 + v_2) = 0_W = 0_W + 0_W = 0(v_1) + 0(v_2)$ für $v_1, v_2 \in K$
$0(\lambda v) = 0w = \lambda \cdot 0w = \lambda 0(v)$ für $v \in V$ und $\lambda \in K$

Zu (b). Seien $f, g : V \to W$ linear. Zu zeigen: $f + g$ linear.

$$\begin{aligned}(f+g)(v_1+v_2) &= f(v_1+v_2) + g(v_1+v_2) = f(v_1) + f(v_2) + g(v_1) + g(v_2)\\ &= f(v_1) + g(v1) + f(v_2) + g(v_2) = (f+g)(v_1) + (f+g)(v_2)\end{aligned}$$

für alle $v_1, v_2 \in V$.

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

für alle $v \in V$ und $\lambda \in K$.

Zu (c). Sei $f : V \to W$ linear und $\mu \in K$. Zu zeigen: μf linear.

$$\begin{aligned}(\mu f)(v_1+v_2) &= \mu(f(v_1+v_2)) = \mu(f(v_1) + f(v_2)) = \mu(f(v_1)) + \mu(f(v_2))\\ &= (\mu f)(v_1) + (\mu f)(v_2)\end{aligned}$$

für alle $v_1, v_2 \in V$.

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

für alle $v \in V$ und $\lambda \in K$.

□

Übung 7.1.7. (a) Sind V und W K-Vektorräume, $U \overset{f}{\underset{g}{\rightrightarrows}} V \xrightarrow{h} W$, h *linear* und $\lambda \in K$, so

$$h \circ (f+g) = h \circ f + h \circ g \qquad \text{und} \qquad h \circ (\lambda f) = \lambda(h \circ f)$$

(b) Sind W ein K-Vektorraum, $U \xrightarrow{f} V \overset{g}{\underset{h}{\rightrightarrows}} W$ Abbildungen und $\lambda \in K$, so

$$(g+h) \circ f = g \circ f + h \circ f \qquad \text{und} \qquad (\lambda g) \circ f = \lambda(g \circ f)$$

Satz 7.1.8. *Seien V und W K-Vektorräume, $\underline{v} = (v_1, \ldots, v_n)$ eine Basis von V und $\underline{w} = (w_1, \ldots, w_m)$ eine Basis von W.*

Dann sind $\Phi : \begin{cases} \mathrm{Hom}(V,W) \to K^{m\times n} \\ f \mapsto M(f, \underline{v}, \underline{w}) \end{cases}$ *und* $\Psi : \begin{cases} K^{m\times n} \to \mathrm{Hom}(V,W) \\ A \mapsto \mathrm{vec}_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}} \end{cases}$ *zueinander inverse Vektorraumisomorphismen.*

Beweis. Nach 1.2.6 und 6.3.3 (b) ist zu zeigen:

(a) $\Phi \circ \Psi = \mathrm{id}_{K^{m\times n}}$

(b) $\Psi \circ \Phi = \mathrm{id}_{\mathrm{Hom}(V,W)}$

(c) Ψ ist linear.

Zu (a). Für $A \in K^{m\times n}$ gilt $(\Phi \circ \Psi)(A) = \Phi(\Psi(A)) = M(vec_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}}, v, w) \overset{7.1.1}{\underset{7.1.3}{=}} A$.

Zu (b). Für $f \in \mathrm{Hom}(V,W)$ gilt $(\Psi \circ \Phi)(f) = \mathrm{vec}_{\underline{w}} \circ f_{M(f,\underline{v},\underline{w})} \circ \mathrm{coord}_{\underline{v}} \overset{7.1.1}{\underset{7.1.3}{=}} f$.

Zu (c). Seien $A, B \in K^{m\times n}$ und $\lambda \in K$.

Zu zeigen: $\mathrm{vec}_{\underline{w}} \circ f_{A+B} \circ \mathrm{coord}_{\underline{v}} = \mathrm{vec}_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}} + \mathrm{vec}_{\underline{w}} \circ f_B \circ \mathrm{coord}_{\underline{v}}$
und $\mathrm{vec}_{\underline{w}} \circ f_{\lambda A} \circ \mathrm{coord}_{\underline{v}} = \lambda(\mathrm{vec}_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}})$

Die rechten Seiten sind nach 7.1.7 gleich

$$\mathrm{vec}_{\underline{w}} \circ (f_A + f_B) \circ \mathrm{coord}_{\underline{v}} \qquad \text{und} \qquad \mathrm{vec}_{\underline{w}} \circ (\lambda f_A) \circ \mathrm{coord}_{\underline{v}}\,.$$

Es reicht also $f_{A+B} = f_A + f_B$ und $f_{\lambda A} = \lambda f_A$ zu zeigen. Sei hierzu $x \in K^n$. Zu zeigen: $(A+B)x = Ax + Bx$ und $(\lambda A)x = \lambda(Ax)$. Dies rechnet man sofort nach.

□

Korollar 7.1.9. *Sind V und W endlichdimensionale K-Vektorräume mit $n = \dim V$ und $m = \dim W$, so $\mathrm{Hom}(V,W) \cong K^{m\times n}$. Insbesondere gilt $\dim \mathrm{Hom}(V,W) = mn$.*

Definition und Bemerkung 7.1.10. Sei V ein K-Vektorraum mit Basen $\underline{v} = (v_1, \ldots, v_n)$ und $\underline{w} = (w_1, \ldots, w_n)$. Dann heißt $M(\underline{v}, \underline{w}) := M(\mathrm{id}_V, \underline{v}, \underline{w}) \in K^{n\times n}$ die *Matrix des Basiswechsels* von $\underline{v}$ nach $\underline{w}$. Dies ist nach 7.1.1 die eindeutig bestimmte Matrix $A \in K^{n\times n}$ mit $\mathrm{coord}_{\underline{w}} = f_A \circ \mathrm{coord}_{\underline{v}}$. Sind also $\lambda_1, \ldots, \lambda_n$ die Koordinaten eines Vektors bezüglich $\underline{v}$, so sind $\mu_1, \ldots, \mu_n$ mit $\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} := M(\underline{v}, \underline{w}) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ die Koordinaten desselben Vektors bezüglich $\underline{w}$.

Definition 7.1.11. Ist V ein K-Vektorraum mit Basis $\underline{v} = (v_1, \ldots, v_n)$ und $f : V \to V$ linear, so heißt $M(f, \underline{v}) := M(f, \underline{v}, \underline{v}) \in K^{n\times n}$ *Darstellungsmatrix* von f bezüglich $\underline{v}$.

§7.2 Matrizenkalkül

Definition 7.2.1. (auch gültig, wenn K nur ein kommutativer Ring statt einem Körper ist!) Seien $m, n, r \in \mathbb{N}_0$, $A = (a_{ij})_{1\le i\le m, 1\le j\le n} \in K^{m\times n}$, $B = (b_{jk})_{1\le j\le n, 1\le k\le r} \in K^{n\times r}$. Dann ist das *Matrizenprodukt* $AB = A \cdot B \in K^{m\times r}$ definiert durch

$$AB = \left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)_{1\le i\le m, 1\le k\le r}.$$

Veranschaulichung:

$$\begin{pmatrix} 1 & 3 & 4 \\ -1 & 2 & 3 \\ 0 & 1 & 5 \\ 4 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 5 & 1 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 16 & 15 \\ 9 & 11 \\ 5 & 16 \\ 9 & 1 \end{pmatrix}$$

Bemerkung 7.2.2. (a) Ist $n = 0$, so ist $AB = 0 \in K^{m\times r}$ die *Nullmatrix*, aber $m, r \in \mathbb{N}_0$ können beliebig gewählt werden. Nur in diesem Ausnahmefall müsste man in die Notation $A \cdot B$ eigentlich m und r aufnehmen, aber aus dem Zusammenhang ist ohnehin meist klar, was m und r sein sollen.

(b) Damit das Matrixprodukt zweier Matrizen definiert ist, muss die erste Matrix genau so viele Spalten haben, wie die zweite Zeilen hat. Mit anderen Worten: Die Zeilen der ersten Matrix müssen genauso lang sein, wie die Spalten der zweiten Matrix. Der Eintrag in der i-ten Zeile und k-ten Spalte von AB ist dann das innere Produkt der i-ten Zeile von A mit der k-ten Spalte von B („Zeile mal Spalte“). Dabei nennt man $\sum_{i=1}^{n} x_i y_i$ für $x, y \in K^n$ das *innere Produkt* von x und y.

(c) Sind $x^{(1)}, \dots, x^{(r)}$ die Spalten von B, so sind $Ax^{(1)}, \dots, Ax^{(r)}$ die Spalten von AB:

$$A(x^{(1)} \dots x^{(r)}) = (Ax^{(1)} \dots Ax^{(r)}).$$

Matrizenmultiplikation ist also „simultanes Multiplizieren mit Spaltenvektoren“.

(d) Sind $A \in K^{m\times n}$ und $x_1, \dots, x_n \in K$, so ist

$$A \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{\in K^n} = A \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{\in K^{n\times 1}}.$$

Beispiel 7.2.3.

(a) $$\begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 3 & 3 & 0 & 2 \end{pmatrix}$$

$3 \times \underline{2}$ $\quad$ $\underline{2} \times 4$ $\quad$ 3×4

(b) $\begin{pmatrix}1 & 1\\ 2 & -1\end{pmatrix}\ \begin{pmatrix}1 & 0\\ 1 & 1\\ 1 & 0\end{pmatrix}$ ist nicht definiert.

$2 \times \underline{2} \qquad \underline{3} \times 2$

(c) $\begin{pmatrix}1 & 3 & 0 & 1\end{pmatrix}\ \begin{pmatrix}0 & 1\\ 1 & 1\\ -1 & 0\\ 0 & 2\end{pmatrix} = \begin{pmatrix}3 & 6\end{pmatrix}$

$1 \times \underline{4} \qquad \underline{4} \times 2 \qquad 1 \times 2$

Lemma 7.2.4. *Seien $m, n, r \in \mathbb{N}_0$, $A \in K^{m\times n}$ und $B \in K^{n\times r}$. Dann gilt $f_{AB} = f_A \circ f_B$.*

Beweis. Wegen $AB \in K^{m\times r}$ haben wir [→ 6.3.2 (e)]

$$K^r \xrightarrow{f_B} K^n \xrightarrow{f_A} K^m, \qquad K^r \xrightarrow{f_{AB}} K^m$$

so dass Definitions- und Zielmengen von f_{AB} und $f_A \circ f_B$ übereinstimmen. Da beide Abbildungen linear sind, reicht es nach 6.3.4, die Gleichheit auf den Standardvektoren [→ 6.2.2] $e_k \in K^r$ zu zeigen: Sei $k \in \{1, \ldots, r\}$. Zu zeigen ist $(AB)e_k = A(Be_k)$. Da Be_k die k-te Spalte von B ist, ist $A(Be_k)$ nach 7.2.2(c) die k-te Spalte von AB, welche natürlich $(AB)e_k$ ist. □

Satz 7.2.5. („Matrizenprodukt entspricht Hintereinanderschaltung von linearen Abbildungen") *Seien U, V, W K-Vektorräume der Dimensionen r, n, m mit geordneten Basen $\underline{u}, \underline{v}, \underline{w}$. Seien $U \xrightarrow{g} V \xrightarrow{f} W$ linear. Dann gilt $M(f \circ g, \underline{u}, \underline{w}) = M(f, \underline{v}, \underline{w})M(g, \underline{u}, \underline{v})$.*

Beweis. Setzt man $A := M(f, \underline{v}, \underline{w}) \in K^{m\times n}$, $B := M(g, \underline{u}, \underline{v}) \in K^{n\times r}$, so ist $AB = M(f \circ g, \underline{u}, \underline{w})$ zu zeigen, das heißt $f \circ g = \mathrm{vec}_{\underline{w}} \circ f_{AB} \circ \mathrm{coord}_{\underline{u}}$ [→7.1.1]. Nun gilt aber:

$$\begin{aligned} f \circ g &= (\mathrm{vec}_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}}) \circ (\mathrm{vec}_{\underline{v}} \circ f_B \circ \mathrm{coord}_{\underline{u}}) \\ &= \mathrm{vec}_{\underline{w}} \circ f_A \circ \underbrace{(\mathrm{coord}_{\underline{v}} \circ \mathrm{vec}_{\underline{v}})}_{=\mathrm{id}_{K^n}} \circ f_B \circ \mathrm{coord}_{\underline{u}} \\ &= \mathrm{vec}_{\underline{w}} \circ (f_A \circ f_B) \circ \mathrm{coord}_{\underline{u}} \overset{7.2.4}{=} \mathrm{vec}_{\underline{w}} \circ f_{AB} \circ \mathrm{coord}_{\underline{u}} \end{aligned}$$

□

Korollar 7.2.6. („Matrizenmultiplikation ist assoziativ") *Seien $m, n, r, s \in \mathbb{N}_0$, $A \in K^{m\times n}$, $B \in K^{n\times r}$ und $C \in K^{r\times s}$. Dann gilt $(AB)C = A(BC)$.*

Beweis. Bezeichne $\underline{e}^{(\ell)}$ die Standardbasis von K^ℓ für $\ell \in \mathbb{N}_0$. Dann gilt

$$\begin{aligned}(AB)C &= (M(f_A, \underline{e}^{(n)}, \underline{e}^{(m)})M(f_B, \underline{e}^{(r)}, \underline{e}^{(n)}))M(f_C, \underline{e}^{(s)}, \underline{e}^{(r)}) \\ &= M(f_A \circ f_B, \underline{e}^{(r)}, \underline{e}^{(m)})M(f_C, \underline{e}^{(s)}, \underline{e}^{(r)})) \\ &= M((f_A \circ f_B) \circ f_C, \underline{e}^{(s)}, \underline{e}^{(m)}) \overset{1.2.5(a)}{=} M(f_A \circ (f_B \circ f_C), \underline{e}^{(s)}, \underline{e}^{(m)}) \\ &= M(f_A, \underline{e}^{(n)}, \underline{e}^{(m)})M(f_B \circ f_C, \underline{e}^{(s)}, \underline{e}^{(n)}) \\ &= M(f_A, \underline{e}^{(n)}, \underline{e}^{(m)})(M(f_B, \underline{e}^{(r)}, \underline{e}^{(n)})M(f_C, \underline{e}^{(s)}, \underline{e}^{(r)})) = A(BC)\end{aligned}$$

□

Bemerkung 7.2.7. (a) Man kann für Korollar 7.2.6 auch den folgenden direkten Beweis geben, welcher zeigt, dass es auch richtig bleibt, wenn K nur ein kommutativer Ring statt ein Körper ist: Für $i \in \{1, \ldots, m\}$ und $\ell \in \{1, \ldots, s\}$ gilt

$$\begin{aligned}((AB)C)_{i\ell} &= \sum_{k=1}^{r} (AB)_{ik} C_{k\ell} \\ &= \sum_{k=1}^{r} \left(\sum_{j=1}^{n} A_{ij} B_{jk} \right) C_{k\ell} \\ &= \sum_{k=1}^{r} \sum_{j=1}^{n} A_{ij} B_{jk} C_{k\ell} \\ &= \sum_{j=1}^{n} \sum_{k=1}^{r} A_{ij} B_{jk} C_{k\ell} \\ &= \sum_{j=1}^{n} A_{ij} \sum_{k=1}^{r} B_{jk} C_{k\ell} \\ &= \sum_{j=1}^{n} A_{ij} (BC)_{j\ell} = (A(BC))_{i\ell}.\end{aligned}$$

(b) Aus 7.1.7 und 7.1.8 folgt ähnlich wie im Beweis von 7.2.6, dass für alle $m, n, r \in \mathbb{N}_0$ gilt

$$\begin{aligned}&\forall A, B \in K^{m\times n} : \forall C \in K^{n\times r} : (A+B)C = AC + BC, \\ &\forall A \in K^{m\times n} : \forall B, C \in K^{n\times r} : A(B+C) = AB + AC \qquad \text{und} \\ &\forall \lambda \in K : \forall A \in K^{m\times n} : \forall B \in K^{n\times r} : (\lambda A)B = \lambda(AB) = A(\lambda B).\end{aligned}$$

Dies kann man aber auch direkt nachrechnen und zwar sogar dann, wenn K nur ein kommutativer Ring statt ein Körper ist, wobei man dann λA für $\lambda \in K$ und $A \in K^{m\times n}$ analog zu 7.1.5 „eintragweise“ definiert (und $A + B$ für $A, B \in K^{m\times n}$ schon durch 2.1.11 genauso wie in 7.1.5 „eintragweise“ definiert ist).

(c) Wegen 7.2.6 können wir beim Multiplizieren von mehreren Matrizen auf Klammern verzichten [→2.1.7].

Beispiel 7.2.8. Seien $\varphi, \psi \in \mathbb{R}$. Dann $R_{\varphi+\psi} = R_\varphi \circ R_\psi$ aus geometrischen Gründen und mit 7.2.5 daher $M(R_{\varphi+\psi}, \underline{e}) = M(R_\varphi, \underline{e})M(R_\psi, \underline{e})$, was mit 7.1.4(a) heißt

$$\begin{pmatrix} \cos(\varphi+\psi) & -\sin(\varphi+\psi) \\ \sin(\varphi+\psi) & \cos(\varphi+\psi) \end{pmatrix} = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \begin{pmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{pmatrix}$$
$$\overset{7.2.5}{=} \begin{pmatrix} (\cos\varphi)(\cos\psi) - (\sin\varphi)(\sin\psi) & -(\cos\varphi)(\sin\psi) - (\sin\varphi)(\cos\psi) \\ (\sin\varphi)(\cos\psi) + (\cos\varphi)(\sin\psi) & -(\sin\varphi)(\sin\psi) + (\cos\varphi)(\cos\psi) \end{pmatrix}.$$

Es folgen die *Additionstheoreme*

$$\cos(\varphi+\psi) = (\cos\varphi)(\cos\psi) - (\sin\varphi)(\sin\psi) \text{ und}$$
$$\sin(\varphi+\psi) = (\sin\varphi)(\cos\psi) + (\cos\varphi)(\sin\psi).$$

Definition und Proposition 7.2.9. (auch falls K nur ein kommutativer Ring statt einem Körper) Für $n \in \mathbb{N}_0$ heißt

$$I_n := \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in K^{n\times n}$$

die *Einheitsmatrix* der Größe n. Falls n aus dem Zusammenhang klar ist, schreibt man oft I statt I_n. Man überprüft sofort $\forall A \in K^{m\times n} : AI_n = A$ und $\forall B \in K^{n\times r} : I_n B = B$. Eine Matrix $A \in K^{n\times n}$ heißt *invertierbar* (falls K ein Körper auch *regulär*), wenn es ein $B \in K^{n\times n}$ gibt mit

$$AB = I_n = BA.$$

In diesem Fall ist B eindeutig bestimmt (hat B' dieselben Eigenschaften, so $B' = B'I_n = B'AB = I_nB = B$) und heißt die zu A *inverse* Matrix, in Zeichen A^{-1}.

Proposition 7.2.10. *Seien V ein Vektorraum mit Basis $\underline{v} = (v_1, \ldots, v_n)$ und $f\colon V \to V$ linear. Dann $M(f, \underline{v}) = I_n \iff f = \mathrm{id}_V$.*

Beweis. $M(f,\underline{v}) = I_n \overset{7.1.1}{\iff} f = \mathrm{vec}_{\underline{v}} \circ \underbrace{f_{I_n}}_{=\mathrm{id}_{K^n}} \circ \mathrm{coord}_{\underline{v}} \iff f = \underbrace{\mathrm{vec}_{\underline{v}} \circ \mathrm{coord}_{\underline{v}}}_{\mathrm{id}_V}$ □

Proposition 7.2.11. *Seien V und W K-Vektorräume mit Basen $\underline{v} = (v_1, \ldots, v_n)$ und $\underline{w} = (w_1, \ldots, w_n)$. Sei $f\colon V \to W$ linear. Dann ist $M(f, \underline{v}, \underline{w})$ invertierbar genau dann, wenn f bijektiv ist, und in diesem Fall gilt*

$$M(f, \underline{v}, \underline{w})^{-1} = M(f^{-1}, \underline{w}, \underline{v}).$$

Beweis. Ist f bijektiv, so gilt $f \circ f^{-1} = \mathrm{id}_W$ und $f^{-1} \circ f = \mathrm{id}_V$ nach 1.2.5(c), also

$$I_n \overset{7.2.10}{=} M(f \circ f^{-1}, \underline{w}, \underline{w}) \overset{7.2.5}{=} M(f, \underline{v}, \underline{w}) \cdot M(f^{-1}, \underline{w}, \underline{v}) \text{ und}$$
$$I_n \overset{7.2.10}{=} M(f^{-1} \circ f, \underline{v}, \underline{v}) \overset{7.2.5}{=} M(f^{-1}, \underline{w}, \underline{v}) \cdot M(f, \underline{v}, \underline{w}),$$

das heißt $M(f, \underline{v}, \underline{w})$ ist invertierbar mit $M(f, \underline{v}, \underline{w})^{-1} = M(f^{-1}, \underline{w}, \underline{v})$. Sei nun umgekehrt $A := M(f, \underline{v}, \underline{w})$ invertierbar, etwa $B \in K^{n\times n}$ mit $AB = I_n = BA$. Dann gilt für $g := \mathrm{vec}_{\underline{v}} \circ f_B \circ \mathrm{coord}_{\underline{w}} \colon W \to V$ unter Beachtung von $f \overset{7.1.1}{=} \mathrm{vec}_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}}$:

$$g \circ f = \mathrm{vec}_{\underline{v}} \circ f_B \circ f_A \circ \mathrm{coord}_{\underline{v}} = \mathrm{id}_V \text{ und } f \circ g = \mathrm{vec}_{\underline{w}} \circ f_A \circ f_B \circ \mathrm{coord}_{\underline{w}} = \mathrm{id}_W,$$

da $f_B \circ f_A \overset{7.2.4}{=} f_{BA} = f_{I_n} \overset{7.2.10}{=} \mathrm{id}_{K^n}$ und $f_A \circ f_B \overset{7.2.4}{=} f_{AB} = f_{I_n} \overset{7.2.10}{=} \mathrm{id}_{K^n}$. Aus 1.2.6 folgt, dass dann f bijektiv ist. □

Proposition 7.2.12. *Seien V und W endlichdimensionale K-Vektorräume derselben Dimension* [→6.2.24] *und $f \colon V \to W$ linear. Dann gilt*

$$f \text{ injektiv} \iff f \text{ bijektiv} \iff f \text{ surjektiv.}$$

Beweis. Wähle mit 6.2.18 eine Basis $\underline{v} = (v_1, \ldots, v_n)$ von V. Nach 6.3.8 gilt:

$$\begin{aligned} f \text{ injektiv} &\iff f(v_1), \ldots, f(v_n) \text{ linear unabhängig in } W, \\ f \text{ bijektiv} &\iff f(v_1), \ldots, f(v_n) \text{ bilden Basis von } W, \\ f \text{ surjektiv} &\iff f(v_1), \ldots, f(v_n) \text{ spannen } W \text{ auf.} \end{aligned}$$

Wegen $\dim W = n$ sind nach 6.2.26 die rechts stehenden Bedingungen aber äquivalent. □

Satz 7.2.13. *Seien $A, B \in K^{n\times n}$. Dann $AB = I_n \iff BA = I_n$.*

Beweis. Wegen Symmetrie reicht es zu „ $\Longrightarrow$ “ zu zeigen. Gelte hierzu $AB = I_n$. Dann $f_A \circ f_B \overset{7.2.4}{=} f_{AB} = f_{I_n} \overset{7.2.10}{=} \mathrm{id}_{K^n}$, woraus folgt, dass f_B injektiv ist. Aus 7.2.12 folgt, dass f_B bijektiv ist, woraus man mit 7.2.11 die Invertierbarkeit von B erhält. Es folgt

$$BA = BA(BB^{-1}) = B(AB)B^{-1} = BB^{-1} = I_n.$$

□

Korollar 7.2.14. *Sei $A \in K^{n\times n}$. Dann ist A invertierbar genau dann, wenn es*

$$x^{(1)}, \ldots, x^{(n)} \in K^n$$

gibt mit $Ax^{(j)} = e_j$ für alle $j \in \{1, \ldots, n\}$. In diesem Fall sind $x^{(1)}, \ldots, x^{(n)}$ eindeutig bestimmt und $x^{(j)}$ ist die j-te Spalte von A^{-1}.

Beweis. Folgt direkt aus 7.2.13 und 7.2.2(c). □

Zur Berechnung von Matrixinversen, muss man also sogenannte *inhomogene* lineare Gleichungssysteme lösen, was Gegenstand des nächsten Abschnitts ist.

§7.3 Inhomogene lineare Gleichungssysteme [→ §5]

Sprechweise und Bemerkung 7.3.1. [→ 5.1.1, 5.1.12] Ein *lineares Gleichungssystem* über K ist (ggf. nach Umstellen) von der Form $(*)\quad Ax = b \quad (x \in K^n)$, wobei $A \in K^{m\times n}$ und $b \in K^m$ vorgegeben sind und $x \in K^n$ gesucht ist (m Gleichungen in n Unbekannten). Ist $b = 0$, so heißt $(*)$ *homogen*, ansonsten *inhomogen*. Der homogene Fall ist hier zugelassen, wurde aber schon in §5 behandelt. Ist A in Stufenform [→ 5.1.10], so nennen wir wieder für $j \in \{1, \dots, n\}$ die Unbekannte x_j $\left\{\begin{matrix} abhängig \\ frei \end{matrix}\right\}$ in $(*)$, wenn j $\left\{\begin{matrix} \text{eine} \\ \text{keine} \end{matrix}\right\}$ Stufenposition [→ 5.1.10] von A ist. Ist A sogar in *reduzierter Stufenform* mit r Stufen und Stufenpostionen $j_1, \dots, j_r$, so kann man offensichtlich $(*)$ als ein System von m linearen Gleichungen

$$\begin{array}{rcl} x_{j_1} & = & b_1 + \dots \\ \vdots & & \vdots \\ x_{j_r} & = & b_r + \dots \\ 0 & = & b_{r+1} \\ 0 & = & b_m \end{array} \qquad \begin{array}{c} \begin{array}{cccccccc} & j_1 & & j_2 & & j_3 & & \\ \end{array} \\ \begin{pmatrix} 1 & * & 0 & * & 0 & \dots \\ & & 1 & * & 0 & \dots \\ & & & & 1 & \dots \end{pmatrix} x = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \end{array}$$

schreiben, auf deren rechten Seiten nur freie Unbekannte auftauchen. Es sind nun zwei Fälle zu unterscheiden:

Fall 1. nicht $b_{r+1} = \dots = b_m = 0$. Dann ist $(*)$ unlösbar (leere Lösungsmenge).

Fall 2. $b_{r+1} = \dots = b_m = 0$. Dann existiert für jede Festlegung der freien Unbekannten wieder genau eine Wahl der abhängigen Unbekannten derart, dass $Ax = b$ gilt. Damit kann man dann unmittelbar $x^{(0)}, \xi^{(1)}, \dots, \xi^{(n-r)} \in K^n$ bestimmen mit

$$\{x \in K^n \mid Ax = b\} = \left\{x^{(0)} + \xi \mid \xi \in \operatorname{span}\left(\xi^{(1)}, \dots, \xi^{(n-r)}\right)\right\}$$

Beispiel 7.3.2. [→ 5.1.13] $K = \mathbb{Q}$.

$$(*) \qquad \begin{array}{rrrrl} x_1 & -3x_4 & +x_5 & +x_7 & = 1 \\ & x_2 & -x_3 & +x_5 & = 2 \\ & & x_6 & -x_7 & = -1 \end{array} \qquad (x_1, \dots, x_7 \in \mathbb{Q})$$

x_1, x_2, x_6 abhängig, x_3, x_4, x_5, x_7 frei.

$$\left\{ x \in \mathbb{Q}^7 \;\middle|\; \begin{array}{rrrrrr} x_1 = & 1 & +3x_4 & -x_5 & -x_7 & \\ x_2 = & 2 & +x_3 & -x_5 & & , x_3, x_4, x_5, x_7 \in \mathbb{Q} \\ x_6 = & -1 & +x_7 & & & \end{array} \right\}$$

$$= \left\{ \begin{pmatrix} 1+3x_4-x_5-x_7 \\ 2+x_3-x_5 \\ x_3 \\ x_4 \\ x_5 \\ -1+x_7 \\ x_7 \end{pmatrix} \;\middle|\; x_3, x_4, x_5, x_7 \in \mathbb{Q} \right\}$$

$$= \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_7 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \;\middle|\; x_3, x_4, x_5, x_7 \in \mathbb{Q} \right\}$$

$$= \left\{ \underbrace{\begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}}_{x_{(0)}} + \xi \;\middle|\; \xi \in \operatorname{span} \left(\underbrace{\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\xi^{(1)}}, \underbrace{\begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\xi^{(2)}}, \underbrace{\begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}}_{\xi^{(3)}}, \underbrace{\begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}}_{\xi^{(4)}} \right) \right\}$$

Keines der so berechneten $\xi^{(\ell)}$ ist überflüssig, denn ist x_j eine freie Unbekannte, so gibt es im Fall 2 von 7.3.1 eine Lösung von $(*)$ mit $x_j = 1$ (und $x_k = 0$ für alle anderen freien Unbekannten x_k) (es sind aber die i-ten Komponenten von $x^{(0)}$ und von zu diesen x_k gehörenden $\xi^{(\ell)}$ gleich 0).

Bemerkung 7.3.3. [→ 5.1.14]

(a) Da man stets wie in 7.3.2 vorgehen kann, ist geklärt, wie man lineare Gleichungssysteme $Ax = b$ mit *Koeffizientenmatrix* $A \in K^{m\times n}$ in *reduzierter Stufenform* und *rechter Seite* $b \in K^m$ löst.

(b) Im allgemeinen Fall schreibt man ein lineares Gleichungssystem so um, dass dessen Koeffizientenmatrix schließlich in *reduzierter Stufenform* vorliegt. Dazu bildet man die *erweiterte Koeffizientenmatrix* $(Ab) \in K^{m\times(n+1)}$ und führt sie durch erlaubte Zeilenoperationen [→ §5.2] in eine Matrix $(A'b') \in K^{m\times(n+1)}$ mit A' in Stufenform über. Es gilt dann

$$\begin{aligned} \{x \in K^n \mid Ax = b\} &= \{x \in K^n \mid \left(\begin{smallmatrix} x \\ -1 \end{smallmatrix}\right) \in \ker(Ab)\} \\ &\underset{5.3.4}{\overset{(Ab)\sim(A'b')}{=}} \{x \in K^n \mid \left(\begin{smallmatrix} x \\ -1 \end{smallmatrix}\right) \in \ker(A'b')\} \\ &= \{x \in K^n \mid A'x = b'\}. \end{aligned}$$

Will man gleich mehrere lineare Gleichungssysteme mit derselben Koeffizientenmatrix aber verschiedenen rechten Seiten lösen, so kann man die Koeffizientenmatrix um mehrere rechte Seiten erweitern.

Beispiel 7.3.4. [→ 5.2.6] $K = \mathbb{F}_5, 2 := 1+1, 3 := 1+1+1$ usw.

$$A := \begin{pmatrix} 0 & 4 & 1 & 3 \\ 2 & 3 & 2 & 1 \\ 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} \in \mathbb{F}_5^{4\times 4}, \quad b_1 := \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Um $\{x \in \mathbb{F}_5^4 \mid Ax = b_i\}$ für $i \in \{1,2\}$ zu berechnen, bilden wir die erweiterte Koeffizientenmatrix $(A, b_1, b_2) \in \mathbb{F}_5^{4\times 6}$ und berechnen $A' \in \mathbb{F}_5^{4\times 4}$ in reduzierter Stufenform und $b_1', b_2' \in \mathbb{F}_5^4$ mit $(A, b_1, b_2) \sim (A', b_1', b_2')$.

$$\left(\begin{array}{cccc|cc} 0 & 4 & 1 & 3 & 1 & 1 \\ 2 & 3 & 2 & 1 & 1 & 0 \\ 1 & 2 & 3 & 4 & 0 & 0 \\ 2 & 4 & 1 & 3 & 0 & 0 \end{array}\right) \overset{\substack{Z_1 \leftrightarrow Z_3 \\ Z_2 \leftarrow Z_2 - 2Z_1 \\ Z_4 \leftarrow Z_4 - 2Z_1}}{\sim} \left(\begin{array}{cccc|cc} 1 & 2 & 3 & 4 & 0 & 0 \\ 0 & 4 & 1 & 3 & 1 & 0 \\ 0 & 4 & 1 & 3 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\overset{Z_3 \leftarrow Z_3 - 2Z_2}{\sim} \left(\begin{array}{cccc|cc} 1 & 2 & 3 & 4 & 0 & 0 \\ 0 & 4 & 1 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \overset{Z_2 \leftarrow \frac{1}{4} Z_2}{\sim} \left(\begin{array}{cccc|cc} 1 & 2 & 3 & 4 & 0 & 0 \\ 0 & 1 & 4 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\overset{Z_1 \leftarrow Z_1 - 2Z_2}{\sim} \left(\begin{array}{cccc|cc} \boxed{1} & 0 & 0 & 0 & 2 & 0 \\ 0 & \boxed{1} & 4 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

nicht beachtet in red. Stufenform

reduzierte Stufenform, 2 Stufen, Stufenpositionen $1, 2$, abhängig: x_1, x_2, frei: x_3, x_4.

$$\left\{x \in \mathbb{F}_5^4 \mid Ax = b_2\right\} = \emptyset$$

$$\begin{aligned} \left\{x \in \mathbb{F}_5^4 \mid Ax = b_1\right\} &= \left\{ \begin{pmatrix} 2 \\ 4-4x_3-2x_4 \\ x_3 \\ x_4 \end{pmatrix} \;\middle|\; x_3, x_4 \in \mathbb{F}_5 \right\} \\ &= \left\{ \begin{pmatrix} 2 \\ 4 \\ 0 \\ 0 \end{pmatrix} + \xi \;\middle|\; \xi \in \operatorname{span}\left(\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right) \right\} \end{aligned}$$

Mit 7.2.14 kann man so auch Matrixinverse berechnen.

Satz 7.3.5. *Seien $A, B, S \in K^{n\times n}$ mit $(A\ I_n) \sim (S\ B)$ und S in reduzierter Stufenform. Dann ist A invertierbar* [→ 7.2.9] *genau dann, wenn $S = I_n$ gilt. In diesem Fall gilt $B = A^{-1}$.*

Beweis. Zu zeigen:

(a) $S \neq I_n \Longrightarrow A$ nicht invertierbar

(b) $S = I_n \Longrightarrow (A$ invertierbar $\ \& \ B = A^{-1})$

Bezeichne b_j die j-te Spalte von B für $j \in \{1, \ldots, n\}$. Aus $(A\ I_n) \sim (S\ B)$ folgt $(A\ e_j) \sim (S\ b_j)$ für $j \in \{1, \ldots, n\}$.

Zu (a). Gelte $S \neq I_n$. Dann gilt $n \geq 1$. Wegen $(A\ e_1) \sim (S\ b_1)$ gilt $\{x \in K^n \mid Ax = e_1\} = \{x \in K^n \mid Sx = b_1\}$. Da S in reduzierter Stufenform ist, hat $S \in K^{n \times n}$ höchstens $n-1$ Stufen und das Gleichungssystem $Sx = b_1 \quad (x \in K^n)$ mindestens eine freie Variable [→ 7.3.1]. Daher gibt es *kein* oder *mehrere* $x^{(1)} \in K^n$ mit $Ax^{(1)} = e_1$, je nachdem ob in 7.3.1 Fall 1 oder 2 eintritt. Wäre A invertierbar, so stünde dies im Widerspruch zu 7.2.14.

Zu (b). Gelte $S = I_n$. Dann gilt $\{x \in K^n \mid Ax = e_j\} = \{x \in K^n \mid I_n x = b_j\} = \{b_j\}$ und insbesondere $Ab_j = e_j$ für $j \in \{1, \ldots, n\}$. Nach 7.2.14 ist A invertierbar und b_j die j-te Spalte von A^{-1}, das heißt $A^{-1} = B$.

□

Beispiel 7.3.6. $K = \mathbb{R}$.

Ist $\left(\begin{smallmatrix} 1 & 2 \\ -1 & 3 \end{smallmatrix}\right)$ invertierbar? Wenn ja, was ist das Inverse?

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{array}\right) \overset{Z_2 \leftarrow Z_2 + Z_1}{\sim} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 5 & 1 & 1 \end{array}\right) \overset{Z_2 \leftarrow \frac{1}{5} Z_2}{\sim} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{1}{5} & \frac{1}{5} \end{array}\right)$$
$$\overset{Z_1 \leftarrow Z_1 - 2Z_2}{\sim} \left(\begin{array}{cc|cc} 1 & 0 & \frac{3}{5} & -\frac{2}{5} \\ 0 & 1 & \frac{1}{5} & \frac{1}{5} \end{array}\right).$$

Also ist $\left(\begin{smallmatrix} 1 & 2 \\ -1 & 3 \end{smallmatrix}\right)$ invertierbar und $\left(\begin{smallmatrix} 1 & 2 \\ -1 & 3 \end{smallmatrix}\right)^{-1} = \frac{1}{5}\left(\begin{smallmatrix} 3 & -2 \\ 1 & 1 \end{smallmatrix}\right)$.

§8 Quotienten und direkte Summen

Sei wieder K ein Körper.

§8.1 Quotientenvektorräume [→ §1.3, §2.3, §3.3]

Definition 8.1.1. [→ 2.3.1, 3.3.1] Sei V ein K-Vektorraum. Eine *Kongruenzrelation* auf V ist eine Kongruenzrelation $\equiv$ auf der additiven Gruppe von V, für die gilt: $\forall v, w \in V : \forall \lambda \in K : (v \equiv w \Longrightarrow \lambda v \equiv \lambda w)$.

Bemerkung 8.1.2. [→ 2.3.2, 3.3.2] Definition 8.1.1 wurde gerade so definiert, dass

$$\begin{aligned} K \times V/\equiv &\to V/\equiv \\ (\lambda, \overline{v}) &\mapsto \overline{\lambda v} \qquad\qquad (\lambda \in K, v \in V) \end{aligned}$$

wohldefiniert ist.

Satz und Definition 8.1.3. [→ 2.3.3,3.3.3] *Ist V ein K-Vektorraum und $\equiv$ eine Kongruenzrelation auf V, so wird die Quotientengruppe $V/\equiv$ vermöge der Skalarmultiplikation definiert durch*

$$\lambda\overline{v} := \overline{\lambda v} \quad (\lambda \in K, v \in V)$$

zu einem K-Vektorraum („Quotientenvektorraum“).

Beweis. direktes Nachrechnen, von (V), ($\overrightarrow{\text{N}}$), ($\overrightarrow{\text{D}}$), (D') aus 6.1.1. □

Satz 8.1.4. [→ 2.3.6, 3.3.6] *Sei V ein Vektorraum. Die Zuordnungen*

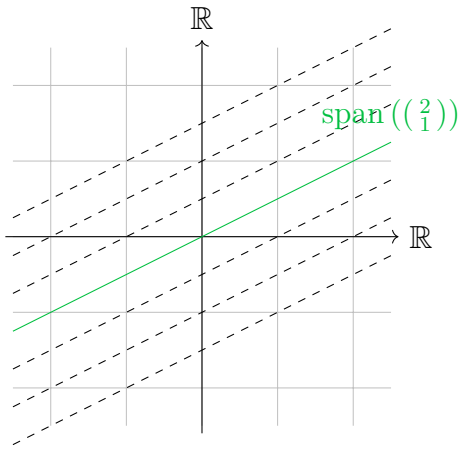
$$\begin{aligned} \equiv &\mapsto \overline{0} \\ \equiv_U &\leftarrow\!\shortmid U \end{aligned}$$

vermitteln eine Bijektion zwischen der Menge der Kongruenzrelationen auf V und der Menge der Unterräume auf V.

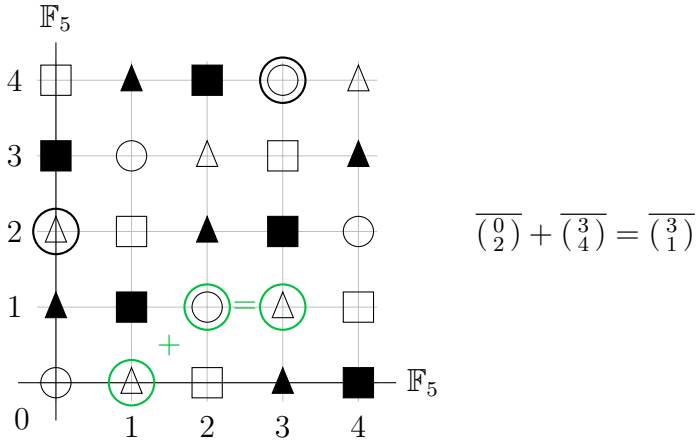
Beweis. Übung (vgl. 3.3.6). □

Notation 8.1.5. [→ 2.3.7,3.3.7] Sei V ein Vektorraum und U ein Unterraum von V. $V/U := V/\equiv_U$ („V modulo U“).

Beispiel 8.1.6. (a) $\mathbb{R}^2/\operatorname{span}\left(\binom{2}{1}\right)$ besteht aus allen Geraden in der Ebene mit Steigung $\frac{1}{2}$.



(b) $\mathbb{F}_5^2/\operatorname{span}\left(\binom{2}{1}\right) = \left\{\underbrace{\overline{0}}_{\bigcirc}, \underbrace{\overline{\binom{1}{0}}}_{\triangle}, \underbrace{\overline{\binom{2}{0}}}_{\square}, \underbrace{\overline{\binom{3}{0}}}_{\blacktriangle}, \underbrace{\overline{\binom{4}{0}}}_{\blacksquare}\right\}$



$\overline{\binom{0}{2}} + \overline{\binom{3}{4}} = \overline{\binom{3}{1}}$

Proposition 8.1.7. [→ 2.3.10, 2.3.13, 3.3.15, 3.3.18] *Seien V und W K-Vektorräume und $f : V \to W$ linear. Dann ist $\equiv_f$ eine Kongruenzrelation auf V und $\ker f$ ein Unterraum von V. Weiter ist $\operatorname{im} f$ ein Unterraum von W.*

Beweis. Übung. □

Proposition 8.1.8. *Sei V ein K-Vektorraum und U ein Unterraum von V. Dann ist die kanonische Surjektion $V \to V/U$ linear.*

Beweis. Übung. □

Satz 8.1.9 (Homomorphiesatz für Vektorräume)**.** [→ 2.3.11, 3.3.16] *Seien V und W K-Vektorräume, U ein Unterraum von V und $f : V \to W$ linear mit $U \subseteq \ker f$.*

(a) *Es gibt genau eine Abbildung* $\overline{f} : V/U \to W$ *mit* $\overline{f}(\overline{v}) = f(v)$ *für alle* $v \in V$. *Diese Abbildung ist linear.*

(b) $\overline{f}$ *injektiv* $\iff U = \ker f$

(c) $\overline{f}$ *surjektiv* $\iff f$ *surjektiv.*

Beweis. folgt fast alles aus 2.3.11. Nur noch zu zeigen [→ 6.3.1]: $\forall v \in V : \forall \lambda \in K : \overline{f}(\lambda\overline{v}) = \lambda\overline{f}(\overline{v})$. Sei also $v \in V$ und $\lambda \in K$. Dann $\overline{f}(\lambda\overline{v}) = \overline{f}(\overline{\lambda v}) = f(\lambda v) = \lambda f(v) = \lambda\overline{f}(\overline{v})$. □

Korollar 8.1.10 (Isomorphiesatz für Vektorräume)**.** [→ 2.3.14,3.3.19] *Seien* V *und* W K*-Vektorräume und* $f : V \to W$ *linear. Dann ist* $\overline{f} : V/\ker f \to \operatorname{im} f$ *definiert durch* $\overline{f}(\overline{v}) = f(v)$ *für* $v \in V$ *ein* K*-Vektorraumisomorphismus. Insbesondere* $V/\ker f \cong \operatorname{im} f$.

Lemma 8.1.11. *Sei* V *ein Vektorraum und* U *ein Unterraum von* V *mit Basis* $(u_1, \ldots, u_m)$. *Seien* $v_1, \ldots, v_n \in V$. *Dann*

$$(\overline{v_1}, \ldots, \overline{v_n}) \text{ Basis von } V/U \iff (u_1, \ldots, u_m, v_1, \ldots, v_n) \text{ Basis von } V.$$

Insbesondere $\dim(U) + \dim(V/U) = \dim(V)$ *falls* $\dim V < \infty$.

Beweis. $K :=$ Grundkörper von V [→ 6.1.2 (c)].

„$\Longrightarrow$“ Sei $(\overline{v_1}, \ldots, \overline{v_n})$ Basis von V/U. Zu zeigen:

(a) $u_1, \ldots, u_m, v_1, \ldots, v_n$ linear unabhängig in V.

(b) $V = \operatorname{span}(u_1, \ldots, u_m, v_1, \ldots, v_n)$.

Zu (a). Seien $\lambda_i, \mu_i \in K$ mit $\sum_{i=1}^m \lambda_i u_i + \sum_{j=1}^n \mu_j v_j = 0$. Zu zeigen: $\lambda_i = \mu_i = 0$.

Aus $\sum_{i=1}^m \mu_j \overline{v_j} = \overline{\sum_{i=1}^m \lambda_i u_i + \sum_{j=1}^n \mu_j v_j} = 0$ folgt $\mu_j = 0$ für alle j, da $\overline{v_1}, \ldots, \overline{v_n}$ linear unabhängig.

Da $u_1, \ldots, u_m$ auch linear unabhängig, folgt $\lambda_i = 0$ für alle i.

Zu (b). Sei $v \in V$. Zu zeigen: $\exists \lambda_i, \mu_i \in K : v = \sum_{i=1}^m \lambda_i u_i + \sum_{j=1}^n \mu_j v_j$. Da $\overline{v_1}, \ldots, \overline{v_n}$ den Vektorraum V/U aufspannen, gibt es $\mu_j \in K$ mit $\overline{v} = \sum_{j=1}^n \mu_i, \overline{v_j}$. Es folgt $v - \sum_{j=1}^n \mu_j v_j \in U$. Da $u_1, \ldots, u_m$ den Vektorraum U aufspannen, gibt es $\lambda_i \in K$ mit $v - \sum_{j=1}^n \mu_j v_j = \sum_{i=1}^m \lambda_i v_i$.

„$\Longleftarrow$“ Sei $(u_1, \ldots, u_m, v_1, \ldots, v_n)$ Basis von V. Zu zeigen:

(a) $\overline{v_1}, \ldots, \overline{v_n}$ sind linear unabhängig in V/U.

(b) $V/U = \operatorname{span}(\overline{v_1}, \ldots, \overline{v_n})$

Zu (a). Seien $\mu_j \in K$ mit $\sum_{j=1}^n \mu_j \overline{v_j} = 0$. Zu zeigen: $\mu_j = 0$.

Aus $\sum_{j=1}^n \mu_j v_j \in U$ folgt, dass es $\lambda_i \in K$ gibt mit $\sum_{j=1}^n \mu_j v_j = \sum_{i=1}^m \lambda_i u_i$. Es folgt $\sum_{i=1}^m (-\lambda_i) u_i + \sum_{j=1}^n \mu_j v_j = 0$ und daher $-\lambda_i = \mu_j = 0$ für alle i, j.

Zu (b). Sei $v \in V$. Zu zeigen: $\exists \mu_j \in K : \overline{v} = \sum_{j=1}^{n} \mu_j \overline{v_j}$.
Wähle $\lambda_i, \mu_j \in K$ mit $v = \sum_{i=1}^{m} \lambda_i u_i + \sum_{j=1}^{n} \mu_j v_j$. Dann $\overline{v} = \sum_{j=1}^{n} \mu_j \overline{v_j}$.

□

Satz 8.1.12 (Dimensionsformel für lineare Abbildungen)**.** *Seien V und W K- Vektorräume mit $\dim V < \infty$ und $f : V \to W$ linear. Dann $\dim \ker f + \dim \operatorname{im} f = \dim V$.*

Beweis.

$$\dim \ker f + \dim V/\ker f \overset{8.1.11}{=} \dim V$$
$$V/\ker f \overset{8.1.10}{\cong} \operatorname{im} f$$

□

Korollar 8.1.13. *Zeilen und Spaltenraum* [→ 5.3.1, 6.3.2 (e)] *einer Matrix über einem Körper haben dieselbe Dimension.*

Beweis. Sei $A \in K^{m \times n}$. Die Dimensionsformel 8.1.12 für $f_A : K^n \to K^m$ besagt wegen $\ker f_A = \ker A$ und $\operatorname{im} f_A = \operatorname{im} A$, dass $\dim \ker A + \dim \operatorname{im} A = \dim (K^n) = n$. Also können wir die Behauptung schreiben als

$$\dim \ker A + \dim \operatorname{row} A = n.$$

Wähle B in reduzierter Stufenform mit $A \sim B$ [→ 5.2.3]. Wegen $\ker B = \ker A$ und $\operatorname{row} B = \operatorname{row} A$ [→ 5.3.4] reicht es zu zeigen, dass $\dim \ker B + \dim \operatorname{row} B = n$. Benutze nun 6.2.29. □

Definition 8.1.14. Sei $A \in K^{m \times n}$. Man nennt $\operatorname{rank} A := \dim \operatorname{im} A \overset{8.1.13}{=} \dim \operatorname{row} A$ den *Rang* von A.

Proposition 8.1.15. *Seien V und W K-Vektorräume mit Basen $\underline{v} = (v_1, \ldots, v_n)$ und $\underline{w} = (w_1, \ldots, w_m)$. Sei $f : V \to W$ linear. Dann $\operatorname{rank} M(f, \underline{v}, \underline{w}) = \dim \operatorname{im} f$.*

Beweis.

$$A := M(f, \underline{v}, \underline{w}), f \overset{7.1.1}{=} \operatorname{vec}_{\underline{w}} \circ f_A \circ \operatorname{coord}_{\underline{v}}$$
$$\dim \operatorname{im} f = \dim \operatorname{vec}_{\underline{w}}(\operatorname{im}(f_A \circ \operatorname{coord}_{\underline{v}})) \underset{6.3.7}{\overset{\text{vec}_{\underline{w}} \text{ Iso.}}{=}} \dim \operatorname{im}(f_A \circ \operatorname{coord}_{\underline{v}}) \underset{6.3.7}{\overset{\text{coord}_{\underline{v}} \text{ Iso.}}{=}} \dim \operatorname{im} f_A$$
$$\overset{6.3.2\,(e)}{=} \dim \operatorname{im} A \overset{8.1.14}{=} \operatorname{rank} A.$$

□

$$\begin{array}{ccc} V & \xrightarrow{\quad f \quad} & W \\ {\scriptstyle \text{coord}_{\underline{v}}}\uparrow\downarrow \cong & \circlearrowleft & \cong \uparrow\downarrow {\scriptstyle \text{vec}_{\underline{w}}} \\ K^n & \xrightarrow{\quad f_A \quad} & K^m \end{array} \qquad f \cong f_A$$

Proposition 8.1.16. *Sei $A \in K^{n\times n}$. Dann ist A invertierbar* [→ 7.2.9] $\iff \operatorname{rank} A = n$.

Beweis.

$$A \text{ invertierbar} \overset{7.2.11}{\iff} f_A \text{ bijektiv} \overset{7.2.12}{\iff} f_A \text{ surjektiv} \iff \operatorname{im} f_A = K^n \overset{6.2.27}{\iff} \dim \operatorname{im} f_A = n \overset{8.1.15}{\iff} \operatorname{rank} A = n.$$

□

Proposition 8.1.17. *Seien $A \in K^{m\times n}$ und $B \in K^{n\times r}$. Dann gilt*

$$\operatorname{rank}(AB) \leq \operatorname{rank} A \qquad \textit{und} \qquad \operatorname{rank}(AB) \leq \operatorname{rank}(B).$$

Beweis.

$$\operatorname{im}(AB) = \{ABx \mid x \in K^r\} \subseteq \{Ay \mid y \in K^n\} = \operatorname{im} A$$

und daher $\operatorname{rank}(AB) = \dim \operatorname{im}(AB) \overset{6.2.28}{\leq} \dim \operatorname{im} A = \operatorname{rank} A$.

$$\operatorname{row}(AB) = \{(x_1, \ldots, x_m)AB \mid x \in K^m\} \subseteq \{(y_1, \ldots, y_n)B \mid y \in K^n\} = \operatorname{row} B$$

und daher $\operatorname{rank}(AB) = \dim \operatorname{row}(AB) \overset{6.2.28}{\leq} \dim \operatorname{row} B = \operatorname{rank} B$. □

§8.2 Direkte Summen

Definition 8.2.1. Seien $n \in \mathbb{N}_0$ und $U_1, \ldots, U_n$ Unterräume des Vektorraums V. Betrachte die lineare Abbildung $f : U_1 \times \ldots \times U_n \to V, (u_1, \ldots, u_n) \mapsto u_1 + \ldots + u_n$ [→ 6.1.6]. Der Unterraum

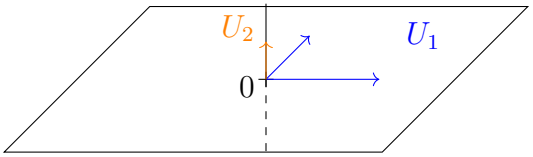
$$\sum_{i=1}^{n} U_i := U_1 + \ldots + U_n := \operatorname{im} f = \{u_1 + \ldots + u_n \mid u_1 \in U_1, \ldots, u_n \in U_n\}$$

heißt *Summe* von $U_1, \ldots, U_n$. Falls f injektiv ist, so sagt man, diese Summe ist *direkt* und schreibt dann auch $\bigoplus_{i=1}^{n} U_i := U_1 \oplus \ldots \oplus U_n := \operatorname{im} f$.

Lemma 8.2.2. *Seien $V_1, \ldots, V_n$ K-Vektorräume und sei B_i eine Basis von V_i für alle $i \in \{1, \ldots, n\}$. Dann ist $(B_1 \times \{0\} \times \{0\} \times \ldots) \cup (\{0\} \times B_2 \times \{0\} \times \ldots) \cup \ldots$ eine Basis von $V_1 \times \ldots \times V_n$. Insbesondere gilt $\dim(V_1 \times \ldots \times V_n) = \sum_{i=1}^{n} \dim(V_i)$, falls alle V_i endlichdimensional sind.*

Beweis. Übung. □

Korollar 8.2.3. *Seien V ein Vektorraum und $U_1, \ldots, U_n$ Unterräume von V mit $V = U_1 \oplus \ldots \oplus U_n$* [→ 8.2.1]. *Ferner sei B_i eine Basis von U_i für alle $i \in \{1, \ldots, n\}$. Dann ist $B_1 \cup \ldots \cup B_n$ eine Basis von V. Insbesondere gilt $\dim(V) = \sum_{i=1}^{n} \dim(U_i)$, falls alle U_i endlichdimensional sind.*



Proposition 8.2.4. *Sei V ein endlichdimensionaler Vektorraum mit Unterräumen $U_1, \ldots, U_n$. Dann $V = U_1 \oplus \ldots \oplus U_n \iff \dim V = \dim\left(\sum_{i=1}^n U_i\right) = \sum_{i=1}^n \dim U_i$.*

Beweis.

$$V = U_1 + \ldots + U_n \overset{6.2.27}{\iff} \dim V = \dim\left(\sum_{i=1}^n U_i\right)$$

Noch zu zeigen:

$$f : \begin{cases} U_1 \times \ldots \times U_n & \to U_1 + \ldots + U_n \\ (v_1, \ldots, v_n) & \mapsto v_1 + \ldots + v_n \end{cases} \text{ ist injektiv } \iff \sum_{i=1}^n \dim U_i = \dim\left(\sum_{i=1}^n U_i\right)$$

„$\Longrightarrow$“ Ist f injektiv, so ist f ein Vektorraumisomorphismus und daher $\sum_{i=1}^n \dim U_i \overset{8.2.1}{=} \dim(U_1 \times \ldots \times U_n) = \dim(U_1 + \ldots + U_n)$.

„$\Longleftarrow$“ Ist $\sum_{i=1}^n \dim U_i = \dim\left(\sum_{i=1}^n U_i\right)$, so haben nach 8.2.1 Definitions- und Zielvektorraum von f dieselbe Dimension und nach 7.2.14 ist f injektiv (da f surjektiv.).

□

Satz 8.2.5 (Dimensionsformel für Unterräume)**.** *Seien U und W Unterräume des endlichdimensionalen Vektorraums V. Dann* $\dim(U \cap W) + \dim(U + W) = (\dim U) + (\dim W)$.

Beweis. $f : \begin{cases} U \times W & \to U + W \\ (u, w) & \mapsto u + w \end{cases}$ ist Vektorraumepimorphismus [→ 6.3.1]

$$\begin{aligned} \ker f &= \{(u, w) \in U \times W \mid u + w = 0\} \\ &= \{(u, w) \in U \times W \mid w = -u\} \\ &= \{(u, -u) \mid u \in U, -u \in W\} \\ &= \{(u, -u) \mid u \in U, u \in W\} \\ &= \{(u, -u) \mid u \in U \cap W\} \end{aligned}$$

Daher ist $U \cap W \to \ker f, u \mapsto (u, -u)$ ein Vektorraumisomorphismus und somit $\dim(U \cap W) = \dim \ker f$. Die Dimensionsformel für f [→ 8.1.12] liefert $\dim \ker f + \dim \operatorname{im} f = \dim(U \times W)$. Daraus folgt mit 8.2.1 die Behauptung. □

§9 Determinanten

In diesem Kapitel sei stets K ein kommutativer Ring.

§9.1 Definition und Eigenschaften von Determinanten

Definition 9.1.1. Sei $\sigma \in S_n$ [→2.1.8]. Ein *Fehlstand* von σ ist ein Paar $(i,j) \in \{1,\dots,n\}^2$ mit $i < j$ und $\sigma(i) > \sigma(j)$. Hat σ genau m Fehlstände, so definieren wir das *Vorzeichen* (oder *Signum*) von σ durch $\operatorname{sgn}\sigma := (-1)^m \in \{-1,1\} \subseteq \mathbb{Z}$.

Beispiel 9.1.2. $\sigma\colon \{1,\dots,5\} \to \{1,\dots,5\},\ 1 \mapsto 2,\ 2 \mapsto 3,\ 3 \mapsto 1,\ 4 \mapsto 5,\ 5 \mapsto 4$ hat genau die Fehlstände $(1,3)$, $(2,3)$ und $(4,5)$ und daher Vorzeichen $(-1)^3 = -1$.

Definition 9.1.3. Die Permutationen

$$\tau_{k\ell}\colon \{1,\dots,n\} \to \{1,\dots,n\},\ i \mapsto \begin{cases} \ell & \text{falls } i = k \\ k & \text{falls } i = \ell \\ i & \text{sonst} \end{cases}$$

mit $k,\ell \in \{1,\dots,n\}$ und $k \neq \ell$ heißen *Transpositionen*.

Satz 9.1.4. $\forall \sigma,\tau \in S_n : \operatorname{sgn}(\sigma\circ\tau) = (\operatorname{sgn}\sigma)(\operatorname{sgn}\tau)$

Beweis. Ist $\varrho \in S_n$, so gilt

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

denn das Produkt auf der rechten Seite hat wegen $\prod_{i<j} |\varrho(j)-\varrho(i)| = \prod_{i<j} |j-i|$ den Betrag 1 und hat gleichzeitig dasselbe Vorzeichen wie $\operatorname{sgn}\varrho$, da der Faktor $\frac{\varrho(j)-\varrho(i)}{j-i}$ genau dann negativ ist, wenn (i,j) ein Fehlstand ist. Nun gilt

$$\begin{aligned} \operatorname{sgn}(\sigma\circ\tau) &= \prod_{i<j} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{j-i} \\ &= \prod_{i<j} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} \prod_{i<j} \frac{\tau(j)-\tau(i)}{j-i} = (\operatorname{sgn}\sigma)(\operatorname{sgn}\tau), \end{aligned}$$

wobei das Produkt $\prod_{i<j} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)}$ deswegen gleich $\operatorname{sgn}\sigma$ ist, weil in der Liste

$$(\tau(1),\tau(2)),(\tau(1),\tau(3)),(\tau(1),\tau(4)),\ldots,(\tau(1),\tau(n)),\quad \ldots\quad ,(\tau(n-1),\tau(n))$$

für jedes $(i,j)\in\{1,\ldots,n\}$ entweder genau einmal (i,j) oder genau einmal (j,i) auftaucht (ob ersteres oder letzteres ist egal wegen $\frac{\sigma(j)-\sigma(i)}{j-i}=\frac{\sigma(i)-\sigma(j)}{i-j}$). □

Korollar 9.1.5. *Jede Transposition hat Vorzeichen -1.*

Beweis. Sei $\tau_{k\ell}\in S_n$ eine Transposition. Dann kann man $\sigma\in S_n$ wählen mit $\sigma(k)=1$ und $\sigma(\ell)=2$. Es gilt dann $\sigma\circ\tau_{k\ell}=\tau_{12}\circ\sigma$. Mit 9.1.4 erhält man daraus $(\operatorname{sgn}\sigma)(\operatorname{sgn}\tau_{k\ell})=(\operatorname{sgn}\tau_{12})(\operatorname{sgn}\sigma)$ und somit $\operatorname{sgn}\tau_{k\ell}=\operatorname{sgn}\tau_{12}$. Nun hat aber τ_{12} nur den Fehlstand $(1,2)$ und daher Vorzeichen -1. □

Lemma 9.1.6. *Jede Permutation $\sigma\in S_n$ ist Hintereinanderschaltung [→2.1.7] endlich vieler Transpositionen.*

Beweis. Dies entspricht der Tatsache, dass man n nebeneinander angeordnete Objekte durch paarweise Vertauschungen in jede beliebige Reihenfolge bringen kann. □

Satz 9.1.7. *Für jedes $n\in\mathbb{N}_0$ und $e\in K$ gibt es genau eine Funktion $\delta_e^{(n)}\colon K^{n\times n}\to K$ mit folgenden Eigenschaften:*

(a) *Für alle $i\in\{1,\ldots,n\}$ und Zeilen $a_1,\ldots,a_{i-1},b,c,a_{i+1},\ldots,a_n\in K^n$ gilt*

$$\delta_e^{(n)}\begin{pmatrix}a_1\\ \vdots\\ a_{i-1}\\ b+c\\ a_{i+1}\\ \vdots\\ a_n\end{pmatrix}=\delta_e^{(n)}\begin{pmatrix}a_1\\ \vdots\\ a_{i-1}\\ b\\ a_{i+1}\\ \vdots\\ a_n\end{pmatrix}+\delta_e^{(n)}\begin{pmatrix}a_1\\ \vdots\\ a_{i-1}\\ c\\ a_{i+1}\\ \vdots\\ a_n\end{pmatrix}$$

(b) *Für alle $i\in\{1,\ldots,n\}$, $a_1,\ldots,a_n\in K^n$ und $\lambda\in K$ gilt*

$$\delta_e^{(n)}\begin{pmatrix}a_1\\ \vdots\\ a_{i-1}\\ \lambda a_i\\ a_{i+1}\\ \vdots\\ a_n\end{pmatrix}=\lambda\delta_e^{(n)}\begin{pmatrix}a_1\\ \vdots\\ a_{i-1}\\ a_i\\ a_{i+1}\\ \vdots\\ a_n\end{pmatrix}.$$

(c) *Für alle $A\in K^{n\times n}$ mit zwei identischen Zeilen gilt $\delta_e^{(n)}(A)=0$.*

(d) $\delta_e^{(n)}(I_n) = e$

Es gilt

$$(*) \qquad \delta_e^{(n)}(A) = e \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

für alle $A = (a_{ij})_{1 \le i,j \le n} \in K^{n\times n}$.

Beweis. Schreibe $e_j := (0, \ldots, 0, \overbrace{1}^{j\text{-te Stelle}}, 0 \ldots, 0) \in K^n$ für $j \in \{1, \ldots, n\}$. Wir zeigen:

(1) Für jedes $\delta_e^{(n)} \colon K^{n\times n} \to K$ mit (a)–(d) gilt $(*)$.

(2) $\delta_e^{(n)} \colon K^{n\times n} \to K$ definiert durch $(*)$ erfüllt (a)–(d).

Zu (1). Es habe $\delta := \delta_e^{(n)} \colon K^{n\times n} \to K$ die Eigenschaften (a)–(d). Geht B aus A durch Vertauschen zweier Zeilen hervor, so gilt $\delta(B) = -\delta(A)$, denn sind $a, b \in K^n$ diese zwei Zeilen, so gilt

$$0 \overset{(c)}{=} \delta \begin{pmatrix} \vdots \\ a+b \\ \vdots \\ a+b \\ \vdots \end{pmatrix} \overset{(a)}{=} \delta \underbrace{\begin{pmatrix} \vdots \\ a \\ \vdots \\ a \\ \vdots \end{pmatrix}}_{=0 \text{ nach (c)}} + \delta \begin{pmatrix} \vdots \\ a \\ \vdots \\ b \\ \vdots \end{pmatrix} + \delta \begin{pmatrix} \vdots \\ b \\ \vdots \\ a \\ \vdots \end{pmatrix} + \delta \underbrace{\begin{pmatrix} \vdots \\ b \\ \vdots \\ b \\ \vdots \end{pmatrix}}_{=0 \text{ nach (c)}} .$$

Mit 9.1.6, 9.1.5 und 9.1.4 folgt daraus für alle $\sigma \in S_n$

$$\delta \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} = (\operatorname{sgn}\sigma)\ \delta \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Sei nun $A = (a_{ij})_{1 \le i,j \le n} \in K^{n\times n}$. Dann

$$\begin{aligned} \delta(A) &= \delta \begin{pmatrix} \sum_{j=1}^n a_{1j} e_j \\ \vdots \\ \sum_{j=1}^n a_{nj} e_j \end{pmatrix} \overset{(a)}{\underset{(b)}{=}} \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n a_{1j_1} \cdots a_{nj_n} \delta \begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix} \\ &\overset{(c)}{=} \sum_{\sigma \in S_n} a_{1j_1} \cdots a_{nj_n} \delta \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} = e \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}. \end{aligned}$$

Zu (2). Für $\delta := \delta_e^{(n)}$ definiert durch $(*)$ sind (a),(b) und (d) unmittelbar einsichtig. Um (c) zu zeigen, sei $A \in K^{n\times n}$ derart, dass die k-te Zeile und ℓ-te Zeile ($k \ne \ell$) von A übereinstimmen. Definiere $A_n := \{\sigma \in S_n \mid \operatorname{sgn}\sigma = 1\}$ und beachte, dass [→9.1.4, 9.1.5]

$$\Phi \colon A_n \to S_n \setminus A_n,\ \sigma \mapsto \sigma \circ \tau_{k\ell}$$

eine Bijektion ist. Dann

$$\begin{aligned}\delta(A) &\overset{(*)}{=} e \sum_{\sigma\in A_n} (a_{1\sigma(1)}\cdots a_{n\sigma(n)} - a_{1(\sigma\circ\tau_{k\ell})(1)}\cdots a_{n(\sigma\circ\tau_{k\ell})(n)}) \\ &= e \sum_{\sigma\in A_n} \left(\prod_{i\in\{1,\dots,n\}\setminus\{k,\ell\}} a_{i\sigma(i)} \right) \underbrace{(a_{k\sigma(k)}a_{\ell\sigma(\ell)} - \underbrace{a_{k\sigma(\ell)}}_{=a_{\ell\sigma(\ell)}} \underbrace{a_{\ell\sigma(k)}}_{=a_{k\sigma(k)}})}_{=0} = 0.\end{aligned}$$

□

Bemerkung 9.1.8. (a) In 9.1.7 besagen (a) und (b), dass $\delta_e^{(n)}$ *linear in den Zeilen* ist, das heißt der Wert einer Matrix unter $\delta_e^{(n)}$ hängt linear von einer Zeile ab, wenn man alle anderen Zeilen fixiert.

(b) 9.1.7$(*)$ nennt man auch die *Leibniz-Regel* [Gottfried Wilhelm von Leibniz *1646 †1716].

Definition 9.1.9. Sei $n \in \mathbb{N}_0$ und $A = (a_{ij})_{1\le i,j\le n} \in K^{n\times n}$. Dann heißt

$$\det(A) := \delta_1^{(n)}(A) \overset{9.1.7(*)}{=} \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma) a_{1\sigma(1)}\cdots a_{n\sigma(n)}$$

die *Determinante* von A.

Beispiel 9.1.10. (a) $\det() = \delta_1^{(0)}() \overset{()=I_0}{=} \delta_1^{(0)}(I_0) \overset{9.1.7(d)}{=} 1$

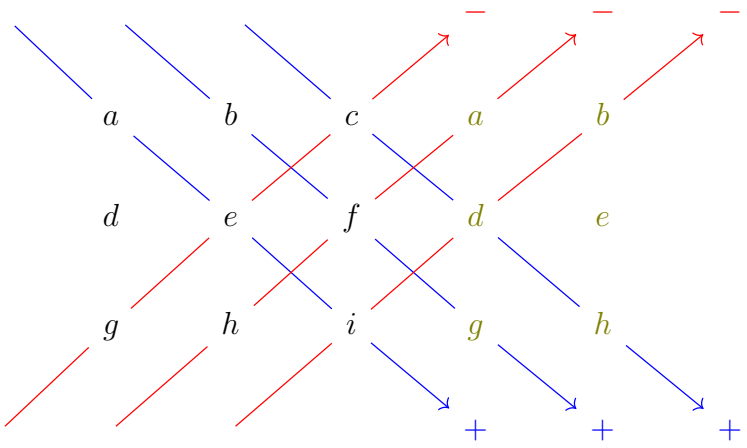
(b) Die Determinante einer Matrix, die eine Nullzeile enthält, ist null. Dies folgt zum Beispiel aus 9.1.7(b). Insbesondere gilt für die Nullmatrix $0 \in K^{n\times n}$ im Fall $n \ge 1$, dass $\det(0) = 0$ (nicht allerdings, wenn $n = 0$ und $0 \ne 1$ in K, wie in (a) gesehen).

(c) $\det(a) = a$ für alle $a \in K$

(d) Sind $a, b, c, d \in K$, so $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$.

(e) Sind $a, b, c, d, e, f, g, h, i \in K$, so gilt

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei - afh - bdi + bfg + cdh - ceg.$$



Regel von Sarrus

Satz 9.1.11. *Sei* $A \in K^{n\times n}$ *von der Gestalt*

$$A = \begin{pmatrix} \boxed{A_1} & & * \\ & \boxed{A_2} & \\ 0 & & \ddots \\ & & \boxed{A_m} \end{pmatrix} \text{ mit quadratischen Matrizen } A_i.$$

Dann gilt $\det A = \prod_{i=1}^{m} \det A_i$. *Insbesondere ist die Determinante einer Matrix* A *in* oberer Dreiecksgestalt $A = \begin{pmatrix} a_1 & & * \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$ *das Produkt ihrer* Diagonaleinträge, *das heißt* $\det A = \prod_{i=1}^{n} a_i$.

Beweis. Es reicht, für $A = \left(\begin{array}{c|c} B & * \\ \hline 0 & C \end{array}\right)$ mit $B \in K^{r\times r}$ und $C \in K^{t\times t}$ zu zeigen

$$\det A = (\det B)(\det C),$$

denn der Rest folgt dann mit Induktion. Wegen der Nulleinträge links unten sind in der Leibniz-Formel 9.1.7(∗) nur diejenigen Summanden ungleich null, die zu einem $\sigma \in S_n$ gehören, für welches es $\varrho \in S_r$ und $\tau \in S_t$ gibt mit $\sigma(i) = \varrho(i)$ für $i \in \{1, \dots, r\}$ und $\sigma(i) = \tau(i-r) + r$ für $i \in \{r+1, \dots, n\}$ (beachte $r + t = n$). Dabei ist die Anzahl der Fehlstände [→9.1.1] von σ offensichtlich die Summe der Anzahl der Fehlstände von ϱ und τ, weswegen $\operatorname{sgn}\sigma = (\operatorname{sgn}\varrho)(\operatorname{sgn}\tau)$ gilt. Es folgt

$$\det A = \sum_{\varrho\in S_r} \sum_{\tau\in S_t} (\operatorname{sgn}\varrho)(\operatorname{sgn}\tau) b_{1\varrho(1)} \cdots b_{r\varrho(r)} c_{1\tau(1)} \cdots c_{t\tau(t)},$$

wobei $B = (b_{ij})_{1\le i,j\le r}$ und $C = (c_{ij})_{1\le i,j\le t}$. Daher gilt

$$\det A = \left(\sum_{\varrho\in S_r} (\operatorname{sgn}\varrho) b_{1\varrho(1)} \cdots b_{r\varrho(r)}\right) \left(\sum_{\tau\in S_t} (\operatorname{sgn}\tau) c_{1\tau(1)} \cdots c_{t\tau(t)}\right) \overset{9.1.7(*)}{=} (\det B)(\det C).$$

□

Bemerkung 9.1.12. Sei K ein Körper. Zur Berechnung von Determinanten ist es oft am effizientesten, die Matrix durch Zeilenoperationen [→§5.2] auf obere Dreiecksgestalt (zum Beispiel Stufenform [→5.1.10]) zu bringen und den Effekt auf die Determinante dabei mitzuprotokollieren. Für die beiden *elementaren* Zeilenoperationen [→5.2.1] gilt:

(a) Sind $A, B \in K^{n\times n}$ mit $A \overset{Z_i \leftarrow Z_i + \lambda Z_j}{\sim} B$ $(i, j \in \{1, \dots, n\}, i \neq j, \lambda \in K)$, so gilt $\det A = \det B$ [→9.1.7(a)–(c)].

(b) Sind $A, B \in K^{n\times n}$ mit $A \overset{Z_i \leftarrow \lambda Z_i}{\sim} B$ $(i \in \{1, \dots, n\}, \lambda \in K^\times)$, so gilt $\det A = \frac{1}{\lambda} \det B$ [→9.1.7(b)].

Daraus ergibt sich der Effekt auf andere Zeilenoperationen:

(c) Sind $A, B \in K^{n\times n}$ mit $A \overset{Z_i \leftrightarrow Z_j}{\sim} B$ $(i, j \in \{1, \dots, n\}, i \neq j)$, so gilt $\det A = -\det B$ [durch Simulation wie in 5.2.1 aus (a) und (b) oben oder wie im Beweis von 9.1.7].

(d) Sind $A, B \in K^{n\times n}$ mit $A \overset{Z_i \leftarrow \sum_{j=1}^n \lambda_j Z_j}{\sim} B$ $(i \in \{1, \dots, n\}, \lambda_1, \dots, \lambda_n \in K, \lambda_i \neq 0)$, so gilt $\det A = \frac{1}{\lambda_i} \det B$ [durch Simulation wie in 5.2.1 aus (a) und (b) oben].

Beispiel 9.1.13. Sei K ein Körper, schreibe $2 := 1+1 \in K$ und so weiter. Gelte $3 \in K^\times$.

$$A := \begin{pmatrix} 1 & 2 & -1 & 0 \\ 2 & 3 & 0 & 1 \\ -1 & 1 & 1 & 0 \\ 0 & 9 & 3 & -3 \end{pmatrix} \overset{\substack{Z_2 \leftarrow Z_2 - 2Z_1 \\ Z_3 \leftarrow Z_3 + Z_1}}{\sim} \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 1 \\ 0 & 3 & 0 & 0 \\ 0 & 9 & 3 & -3 \end{pmatrix}$$

$$\overset{\substack{Z_4 \leftarrow \frac{1}{3} Z_4 \\ Z_4 \leftarrow Z_4 - Z_3 \\ Z_3 \leftarrow Z_3 + 3Z_2}}{\sim} \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 6 & 3 \\ 0 & 0 & 1 & -1 \end{pmatrix} \overset{\substack{Z_3 \leftrightarrow Z_4 \\ Z_4 \leftarrow Z_4 - 6Z_3}}{\sim} \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 9 \end{pmatrix} =: B$$

Also $\det A \overset{9.1.7(b)(c)}{=} \frac{1}{\frac{1}{3}} (-1) \det B = (-3)(-9) = 27$.

Bemerkung 9.1.14. Sei K ein Körper und $A \in K^{n\times n}$. Ist A in Stufenform [→5.1.10], so ist der Rang von A [→8.1.14] gleich der Anzahl der Stufen von A und die Determinante von A [→9.1.9] gleich dem Produkt der Diagonaleinträge von A [→9.1.11]. Es gilt dann

$$\operatorname{rank} A = n \iff \det A \neq 0.$$

Dies bleibt richtig für beliebiges $A \in K^{n\times n}$, denn durch Zeilenoperationen ändert sich der Rang gar nicht [→6.2.29(c)] und die Determinante nur um einen Faktor $\neq 0$ [→9.1.12]. Wegen A invertierbar $\overset{8.1.16}{\iff}$ $\operatorname{rank} A = n$ und $K^\times = K \setminus \{0\}$ gilt also auch

$$A \text{ invertierbar} \iff \det A \in K^\times,$$

was wir in §9.2 sogar beweisen werden, wenn K nur ein kommutativer Ring statt ein Körper ist.

Satz 9.1.15 (Determinantenproduktsatz)**.** *Für alle $A, B \in K^{n\times n}$ gilt*

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

Beweis. Sei $B \in K^{n\times n}$ fest und betrachte

$$\begin{aligned} f&\colon K^{n\times n} \to K,\ A \mapsto \det(AB) \qquad \text{sowie}\\ g&\colon K^{n\times n} \to K,\ A \mapsto (\det A)(\det B). \end{aligned}$$

Zu zeigen ist $f = g$. Wir zeigen $f = \delta^{(n)}_{\det B} = g$, indem wir die Eigenschaften 9.1.7(a)–(d) von $\delta^{(n)}_{\det B}$ für f und g nachweisen. Für g sind diese Eigenschaften klar. Für f rechnet man sie sofort nach, zum Beispiel (a):

$$f\begin{pmatrix} a_1\\ \vdots\\ a_{i-1}\\ b+c\\ a_{i+1}\\ \vdots\\ a_n \end{pmatrix} = \det\left(\begin{pmatrix} a_1\\ \vdots\\ a_{i-1}\\ b+c\\ a_{i+1}\\ \vdots\\ a_n \end{pmatrix} B\right) = \det\begin{pmatrix} a_1B\\ \vdots\\ a_{i-1}B\\ (b+c)B\\ a_{i+1}B\\ \vdots\\ a_nB \end{pmatrix} = \det\begin{pmatrix} a_1B\\ \vdots\\ a_{i-1}B\\ bB+cB\\ a_{i+1}B\\ \vdots\\ a_nB \end{pmatrix}$$

$$\overset{9.1.7(a)}{=} \det\begin{pmatrix} a_1B\\ \vdots\\ a_{i-1}B\\ bB\\ a_{i+1}B\\ \vdots\\ a_nB \end{pmatrix} + \det\begin{pmatrix} a_1B\\ \vdots\\ a_{i-1}B\\ cB\\ a_{i+1}B\\ \vdots\\ a_nB \end{pmatrix} = f\begin{pmatrix} a_1\\ \vdots\\ a_{i-1}\\ b\\ a_{i+1}\\ \vdots\\ a_n \end{pmatrix} + f\begin{pmatrix} a_1\\ \vdots\\ a_{i-1}\\ c\\ a_{i+1}\\ \vdots\\ a_n \end{pmatrix}$$

für alle $a_i, b, c \in K^n$. □

Definition 9.1.16. Zwei Matrizen $A, B \in K^{n\times n}$ heißen *ähnlich*, in Zeichen $A \approx B$, wenn es eine invertierbare Matrix $P \in K^{n\times n}$ gibt mit $A = P^{-1}BP$.

Proposition 9.1.17. *Ähnlichkeit ist eine Äquivalenzrelation auf $K^{n\times n}$.*

Beweis. Gemäß 1.3.1(b) ist zu zeigen:

(a) $\forall A \in K^{n\times n} : A \approx A$

(b) $\forall A, B \in K^{n\times n} : (A \approx B \implies B \approx A)$

(c) $\forall A, B, C \in K^{n\times n} : ((A \approx B \,\&\, B \approx C) \implies A \approx C)$

Zu (a). $A = I_n^{-1} A I_n$ für alle $A \in K^{n\times n}$ [→7.2.9]

Zu (b). Ist $P \in K^{n\times n}$ invertierbar mit $A = P^{-1}BP$, so ist auch P^{-1} invertierbar und $B = PAP^{-1} = (P^{-1})^{-1}AP^{-1}$.

Zu (c). Sind $P, Q \in K^{n\times n}$ invertierbar mit $A = P^{-1}BP$ und $B = Q^{-1}CQ$, so ist auch QP invertierbar mit $(QP)^{-1} = P^{-1}Q^{-1}$ (denn $P^{-1}Q^{-1}QP = I_n = QPP^{-1}Q^{-1}$) und es gilt $A = P^{-1}Q^{-1}BQP = (QP)^{-1}B(QP)$. □

Satz 9.1.18. *Sind $A, B \in K^{n\times n}$ mit $A \approx B$, so gilt* $\det A = \det B$.

Beweis. Seien $A, B \in K^{n\times n}$ mit $A \approx B$. Wähle $P \in K^{n\times n}$ invertierbar mit $A = P^{-1}BP$. Dann $\det A \overset{9.1.15}{=} (\det(P^{-1}))(\det B)(\det P)$ und

$$1 \overset{9.1.9}{=} \det I_n \overset{7.2.9}{=} \det(P^{-1}P) \overset{9.1.15}{=} (\det(P^{-1}))(\det P).$$

□

Proposition 9.1.19. *Sei f ein Endomorphismus* [→6.3.1] *eines endlichdimensionalen Vektorraums mit geordneten Basen $\underline{v}$ und $\underline{w}$. Dann $M(f, \underline{v}) \approx M(f, \underline{w})$* [→7.1.11].

Beweis. Es gilt $M(f, \underline{v}) \overset{7.2.5}{=} M(\underline{w}, \underline{v})M(f, \underline{w})M(\underline{v}, \underline{w})$ und

$$I_n \underset{7.2.10}{\overset{7.1.10}{=}} M(\underline{v}, \underline{v}) \overset{7.2.5}{=} M(\underline{w}, \underline{v})M(\underline{v}, \underline{w}),$$

das heißt $M(\underline{w}, \underline{v}) = M(\underline{v}, \underline{w})^{-1}$ [→7.2.13]. Mit $P := M(\underline{v}, \underline{w})$ gilt also

$$M(f, \underline{v}) \overset{7.2.5}{=} P^{-1}M(f, \underline{w})P.$$

□

Definition 9.1.20. Sei f ein Endomorphismus eines endlichdimensionalen Vektorraums V. Dann ist die *Determinante* von f definiert als $\det f := \det M(f, \underline{v})$, wobei $\underline{v}$ eine beliebig gewählte geordnete Basis von V ist [→6.2.22, 9.1.18]

Definition 9.1.21. Ist $A = (a_{ij})_{1\le i\le m, 1\le j\le n} \in K^{m\times n}$, so heißt

$$A^T := (a_{ji})_{1\le j\le n, 1\le i\le m} \in K^{n\times m}$$

die zu A *transponierte* Matrix.

Beispiel 9.1.22. $\begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 1 \end{pmatrix}^T = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 1 \end{pmatrix}$

Proposition 9.1.23. $\forall A \in K^{n\times n} : \det A = \det(A^T)$

Beweis. Wegen $\sigma = (\sigma^{-1})^{-1}$ für alle $\sigma \in S_n$ ist $\Phi\colon S_n \to S_n, \sigma \mapsto \sigma^{-1}$ bijektiv. Es gilt

$$\begin{aligned}
\det(A^T) &\overset{9.1.7(*)}{=} \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma) a_{\sigma(1)1}\cdots a_{\sigma(n)n} \\
&\overset{\Phi \text{ bijektiv}}{=} \sum_{\sigma\in S_n} (\operatorname{sgn}(\sigma^{-1})) a_{\sigma^{-1}(1)1}\cdots a_{\sigma^{-1}(n)n} \\
&\overset{\substack{\sigma \text{ bijektiv} \\ \text{für } \sigma\in S_n}}{=} \sum_{\sigma\in S_n} (\operatorname{sgn}(\sigma^{-1})) a_{\sigma^{-1}(\sigma(1))\sigma(1)}\cdots a_{\sigma^{-1}(\sigma(n))\sigma(n)} \\
&\overset{\substack{\operatorname{sgn}(\sigma^{-1})=\operatorname{sgn}\sigma \\ \text{da } (\operatorname{sgn}(\sigma^{-1}))(\operatorname{sgn}\sigma)=\operatorname{sgn}(\mathrm{id})=1}}{=} \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma) a_{1\sigma(1)}\cdots a_{n\sigma(n)} \overset{9.1.7(*)}{=} \det A
\end{aligned}$$

□

Bemerkung 9.1.24. Proposition 9.1.23 hat offensichtliche Konsequenzen: Die Determinantenfunktion ist auch *linear in den Spalten* [→9.1.8(a)], zur Berechnung von Determinanten kann man auch (analog zu 5.2.1 definierte) *Spaltenoperationen* heranziehen [→9.1.12], die Determinante einer Matrix in *unterer* Dreiecksgestalt [→9.1.11] ist das Produkt ihrer Diagonaleinträge und so weiter.

§9.2 Determinantenentwicklung und Komatrix

Satz 9.2.1. *Sei $A = (a_{ij})_{1\le i,j\le n} \in K^{n\times n}$ und bezeichne $A_{ij} \in K^{(n-1)\times(n-1)}$ für $i,j \in \{1,\dots,n\}$ die Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht. Dann gilt:*

(a) *Für alle $i \in \{1,\dots,n\}$ gilt $\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$ („Entwicklung nach der i-ten Zeile“)*

(b) *Für alle $j \in \{1,\dots,n\}$ gilt $\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$ („Entwicklung nach der j-ten Spalte“)*

Beweis. Wegen 9.1.23 reicht es, (a) zu zeigen. Hierzu rechnet man

$$\det A = \sum_{j=1}^{n} a_{ij} \det \begin{pmatrix} & \text{„wie in } A\text{“} & \\ 0\dots0 & 1 & 0\dots0 \\ & \text{„wie in } A\text{“} & \\ & \uparrow_{j} & \end{pmatrix} \leftarrow i \overset{9.1.12\,(a)}{=} \sum_{j=1}^{n} a_{ij} \left(\begin{array}{c|c|c} & \overset{j}{\downarrow} & \\ & 0 & \\ & \vdots & \\ & 0 & \\ \hline 0\dots0 & 1 & 0\dots0 \\ \hline & 0 & \\ & \vdots & \\ & 0 & \end{array}\right) \leftarrow i$$

$$\overset{9.1.12\,(c)}{=} \sum_{j=1}^{n} a_{ij} \underbrace{(-1)^{(n-j)+(n-i)}}_{(-1)^{i+j}} \underbrace{\det \left(\begin{array}{c|c} & 0 \\ A_{ij} & \vdots \\ & 0 \\ \hline 0\dots0 & 1 \end{array}\right)}_{\overset{9.1.11}{=}(\det A_{ij})\cdot\det(1)} = \sum_{j=1}^{n} a_{ij}(-1)^{i+j} a_{ij} \det A_{ij}$$

□

Beispiel 9.2.2.

$$\det \begin{pmatrix} 2 & 0 & 4 & 1 \\ 1 & 9 & -1 & 0 \\ 0 & 1 & 1 & -3 \\ -1 & 1 & 2 & 0 \end{pmatrix} \qquad \begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}$$

$$\overset{\text{„Entw. nach}}{\underset{\text{letzter Spalte“}}{=}} -1 \underbrace{\det \begin{pmatrix} 1 & 9 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}}_{\substack{\overset{\text{„Entw. n.}}{\underset{\text{1. Spalte“}}{=}} \det\left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right) - \det\left(\begin{smallmatrix} 9 & -1 \\ 1 & 1 \end{smallmatrix}\right) \\ =(2-1)-(9+1) \\ =-9}} - (-3) \underbrace{\det \begin{pmatrix} 2 & 0 & 4 \\ 1 & 9 & -1 \\ -1 & 1 & 2 \end{pmatrix}}_{\substack{\overset{\text{„Entw. n.}}{\underset{\text{1. Zeile“}}{=}} 2\det\left(\begin{smallmatrix} 9 & -1 \\ 1 & 2 \end{smallmatrix}\right) + 4\det\left(\begin{smallmatrix} 1 & 9 \\ -1 & 1 \end{smallmatrix}\right) \\ =2\cdot19+4\cdot10 \\ =78}}$$

$$= 9 + 3 \cdot 78 = 243$$

Definition 9.2.3. Ist $A = (a_{ji})_{1\leq i,j\leq n} \in K^{n\times n}$ und bezeichnet $A_{ij} \in K^{(n-1)\times(n-1)}$ für $i, j \in \{1, \dots, n\}$ wieder die Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht, so nennt man $\operatorname{com} A := ((-1)^{i+j} \det A_{ij})_{1\leq i,j\leq n}$ die *Komatrix* von A.

Satz 9.2.4. *Sei $A \in K^{n\times n}$. Dann $A(\operatorname{com} A)^T = (\det A)I_n = (\operatorname{com} A)^T A$.*

Beweis. Es gilt $(\operatorname{com} A)^T = ((-1)^{i+j} \det A_{ji})_{1\leq i,j\leq n}$ [→ 9.1.21, 9.2.3]. Schreibe $I_n = (\delta_{ij})_{1\leq i,j\leq n}$. Seien $i, k \in \{1 \dots, n\}$. Nach 7.2.1 ist zu zeigen:

$$\sum_{j=1}^{n} a_{ij}(-1)^{j+k} \det A_{kj} = (\det A)\delta_{ik} = \sum_{j=1}^{n} (-1)^{i+j}(\det A_{ji})a_{jk}.$$

Für $i = k$ steht $\left\{\begin{smallmatrix}\text{links}\\\text{rechts}\end{smallmatrix}\right\}$ die Entwicklung der Determinante von A nach der k-ten $\left\{\begin{smallmatrix}\text{Zeile}\\\text{Spalte}\end{smallmatrix}\right\}$.

Für $i \neq k$ steht $\left\{\begin{smallmatrix}\text{links}\\\text{rechts}\end{smallmatrix}\right\}$ die Entwicklung der Determinante eine Matrix, deren i-te und k-te $\left\{\begin{smallmatrix}\text{Zeile}\\\text{Spalte}\end{smallmatrix}\right\}$ übereinstimmen nach der k-ten $\left\{\begin{smallmatrix}\text{Zeile}\\\text{Spalte}\end{smallmatrix}\right\}$. □

Korollar 9.2.5 (in 9.1.14 schon bewiesen, falls K ein Körper)**.** *Eine Matrix $A \in K^{n\times n}$ ist invertierbar genau dann, wenn $\det A \in K^\times$. In diesem Fall gilt $A^{-1} = (\det A)^{-1}(\operatorname{com} A)^T$.*

Beweis. Ist A invertierbar [→ 7.2.9], so $1 = \det I_n = \det(AA^{-1}) \overset{9.1.15}{=} (\det A)(\det(A^{-1}))$, also $\det A \in K^\times$. Ist $\det A \in K^\times$, so $A\left(\frac{1}{\det A}(\operatorname{com} A)^T\right) \overset{9.2.4}{=} I_n \overset{9.2.4}{=} \left(\left(\frac{1}{\det A}\right)(\operatorname{com} A)^T\right) A$. □

Korollar 9.2.6 (in 7.2.13 schon bewiesen, falls K Körper)**.** *Seien $A, B \in K^{n\times n}$. Dann $AB = I_n \iff BA = I_n$.*

Beweis. Gelte $AB = I_n$. Dann $(\det A)(\det B) = 1$ nach 9.1.15. Also $\det B \in K^\times$ und nach 9.2.5 ist B invertierbar. Dann aber

$$BA = BA(BB^{-1}) = B(AB)B^{-1} = BB^{-1} = I_n$$

wie im Beweis von 7.2.13. □

Satz 9.2.7 (Cramersche Regel [Gabriel Cramer *1704, †1752])**.** *Seien $A \in K^{n\times n}$ und $x, b \in K^n$ mit $Ax = b$ [→ §7.3]. Bezeichne $A_i \in K^{n\times n}$ für $i \in \{1, \dots, n\}$ die Matrix, die aus A entsteht, indem man die i-te Spalte durch b ersetzt. Dann*

$$(\det A)x_i = \det A_i \qquad \text{für } i \in \{1, \dots, n\}.$$

Beweis. Für $X_i := \left(\begin{array}{ccc|c} 1 & & & x_1 & \\ & \ddots & & \vdots & \text{„wie in } I_n\text{“} \\ & & & x_i & \\ & & & \vdots & \ddots \\ & \text{„wie in } I_n\text{“} & & x_n & & 1 \end{array}\right)$ gilt

$$\det X_i \overset{\substack{9.1.12(b)\\9.1.24}}{=} \det \begin{pmatrix} 1 & 0 & \cdots & & 0 \\ & \ddots & & & \\ 0 & & 1 & & \vdots \\ & & & x_i & \\ \vdots & & & & 1 & 0 \\ & & & & & \ddots \\ 0 & & \cdots & & 0 & 1 \end{pmatrix} \overset{9.1.11}{=} x_i$$

sowie $AX_i \overset{\substack{Ax=b\\7.2.2(c)}}{=} A_i$ und daher

$$(\det A)x_i = (\det A)(\det X_i) = \det(AX_i) = \det(A_i) \qquad \text{für } i \in \{1, \dots, n\}$$

□

§10 Eigenvektoren

In diesem Kapitel sei stets K ein Körper.

§10.1 Charakteristisches Polynom und Eigenwerte

Notation und Wiederholung 10.1.1. [→ 7.1.6, 6.3.1] Ist V ein K-Vektorraum, so ist

$$\begin{aligned}\mathrm{End}(V) := \mathrm{Hom}(V,V) &= \{f \mid f \text{ Endomorphismus der Vektorraums } V\} \\ &= \{f \mid f : V \to V \text{ linear}\}\end{aligned}$$

ein Unterraum des K-Vektorraums V^V.

Definition 10.1.2. Sei V ein K-Vektorraum und $f \in \mathrm{End}(V)$. Es heißt $\lambda \in K$ ein *Eigenwert* (*EW*) von f, wenn es ein $v \in V$ gibt mit $v \neq 0$ und $f(v) = \lambda v$. Jedes solche v heißt *Eigenvektor* (*EV*) von f zum Eigenwert λ. Für jeden Eigenwert λ von f nennt man den Unterraum $\ker(f - \lambda \,\mathrm{id}_V) = \{v \in V \mid f(v) = \lambda v\} \subseteq V$, welcher aus dem Nullvektor und den Eigenvektoren zum Eigenwert λ besteht, den *Eigenraum* von f zum Eigenwert λ.

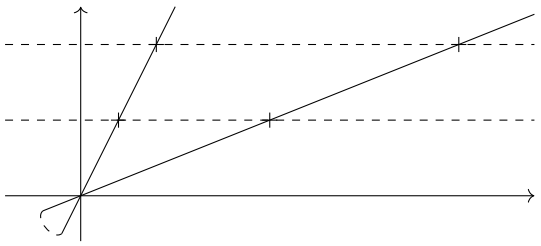
Beispiel 10.1.3. [→ 6.3.2, 7.1.4]

(a) Die Drehung R_φ um den Winkel $\varphi \in \mathbb{R}$, hat nur dann einen Eigenwert, wenn $\varphi = n\pi$ für ein $n \in \mathbb{Z}$. Ist $\varphi = n\pi$ für ein $\left\{\begin{smallmatrix}\text{gerades}\\ \text{ungerades}\end{smallmatrix}\right\}$ $n \in \mathbb{Z}$, so ist $\left\{\begin{smallmatrix}+1\\ -1\end{smallmatrix}\right\}$ der einzige Eigenwert von φ und jedes $v \in \mathbb{R}^2 \setminus \{0\}$ ein Eigenvektor zu diesem Eigenwert (und damit $\mathbb{R}^2$ der Eigenraum zu diesem Eigenwert).

(b) Die Spiegelung S hat die Eigenwerte 1 und -1. Es ist $\mathrm{span}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ der Eigenraum zum Eigenwert 1 und $\mathrm{span}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ der Eigenraum zum Eigenwert -1.

(c) Die Projektion P hat die Eigenwerte 1 und 0. Es ist $\mathrm{span}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ der Eigenraum zum Eigenwert 1 und $\mathrm{span}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ der Eigenraum zum Eigenwert 0.

(d) Der einzige Eigenwert der Scherung T_a ($a \in \mathbb{R}$) ist 1. Falls $a = 0$, so ist $\mathbb{R}^2$ der dazugehörige Eigenraum, sonst $\mathrm{span}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$.



(e) Ist $A \in K^{n\times n}$, so ist $f_A : K^n \to K^n, x \mapsto Ax$ ein Endomorphismus von K^n. Man spricht dann auch von Eigenwert, Eigenvektor und Eigenräumen *von* A statt *von* f_A. Es ist also $\lambda \in K$ ein Eigenwert von A, wenn es ein $x \in K^n$ gibt mit $x \neq 0$ und $Ax = \lambda x$.

Jedes solche x heißt ein Eigenvektor von A zum Eigenwert λ. Ist λ ein Eigenwert von A, so ist $\ker(A - \lambda I_n) \subseteq K^n$ der dazugehörige Eigenraum.

(f) Die formale Ableitung $D : K[X] \to K[X]$ hat nur den Eigenwert 0. Es ist zum Beispiel 1 ein Eigenvektor zu diesem Eigenwert. Für $K = \mathbb{R}$ ist der zugehörige Eigenraum gleich $\mathbb{R}$. Im Allgemeinen ist der Eigenraum komplizierter, denn es ist für $K = \mathbb{F}_2$ auch X^2 ein Eigenvektor zum Eigenwert 0, denn $X^2 \neq 0$ und $D(X^2) = 2X = 0 \cdot X = 0$.

(g) Die Auswertung $E_{a_1,\dots,a_n}$ ist zwar eine lineare Abbildung, aber kein Endomorphismus, weswegen die Begriffe Eigenwert und Eigenvektor dafür keinen Sinn machen.

(h) Als Endomorphismus des $\mathbb{R}$-Vektorraums $\mathbb{C}$ hat die komplexe Konjugation C die Eigenwerte 1 und -1. Es ist $\mathbb{R}$ der Eigenraum zum Eigenwert 1 und $\left\{\overset{\circ}{xi} \mid x \in \mathbb{R}\right\}$ der Eigenraum zum Eigenwert -1.

Lemma 10.1.4. *Sei V ein K-Vektorraum mit geordneten Basen $\underline{v}$ und $\underline{w}$ und $f \in \mathrm{End}(V)$. Dann sind die beiden Matrizen $M(f,\underline{v}) - XI_n, M(f,\underline{w}) - \underbrace{XI_n}_{= \begin{pmatrix} X & & 0 \\ & \ddots & \\ 0 & & X \end{pmatrix}} \in K[X]^{n\times n}$*

ähnlich [→ 9.1.16] und haben daher dieselbe Determinante.

Beweis. Nach 9.1.19 gilt $M(f,\underline{v}) \approx M(f,\underline{w})$, das heißt es gibt ein invertierbares $P \in K^{n\times n}$ mit $M(f,\underline{v}) = P^{-1}M(f,\underline{w})P$. Es folgt $P^{-1}(M(f,\underline{w}) - XI_n)P = P^{-1}M(f,\underline{w})P - P^{-1}(XI_n)P = M(f,\underline{v}) - XP^{-1}P = M(f,\underline{v}) - XI_n$ und daher $M(f,\underline{v}) - XI_n \approx M(f,\underline{w}) - XI_n$. □

Definition 10.1.5. Sei V ein K-Vektorraum, $n := \dim V < \infty$ und $f \in \mathrm{End}(V)$. Dann ist das *charakteristische Polynom* von f definiert als $\underbrace{\chi_f}_{\text{„chi“}} := \det(M(f,\underline{v}) - XI_n) \in K[X]$,

wobei $\underline{v}$ eine beliebig gewählte Basis von V ist [→ 10.1.4].

Bemerkung 10.1.6. Sei V ein K-Vektorraum, $n := \dim V < \infty$ und $f \in \mathrm{End}(V)$. Dann gibt es $a_0, \ldots, a_{n-1} \in K$ mit $\chi_f = (-1)^n X^n + a_{n-1}X^{n-1} + \cdots + a_1 X_1 + a_0$, wobei $a_0 = \det f$. Insbesondere hat χ_f den Grad n.

$$\begin{pmatrix} b_{11}X & \text{„aus } K\text{“} \\ & \ddots \\ \text{„aus } K\text{“} & b_{nn}X \end{pmatrix}$$

Proposition 10.1.7. *Sei V ein K-Vektorraum, $n := \dim V < \infty$, $f \in \mathrm{End}(V)$ und $\lambda \in K$. Dann λ Eigenwert von f [→ 10.1.2] $\iff$ $\chi_f(\lambda) = 0$ [→ 3.2.15].*
„Die Eigenwerte sind die Nullstellen [→ 4.2.9] des charakteristischen Polynoms.“

Beweis.

$$\begin{aligned}
\lambda \text{ Eigenwert von } f &\overset{10.1.2}{\iff} \exists v \in V \setminus \{0\} : f(v) = \lambda v \\
&\iff \ker(f - \lambda \mathrm{id}_V) \neq \{0\} \\
&\overset{6.2.25(b)}{\iff} \dim\ker(f - \lambda \mathrm{id}_V) \neq 0 \\
&\overset{8.1.12}{\iff} \dim\mathrm{im}(f - \lambda \mathrm{id}_V) \neq n \\
&\overset{8.1.15}{\iff} \mathrm{rank}(M(f - \lambda \mathrm{id}_V)) \neq n \\
&\overset{8.1.16}{\iff} M(f - \lambda \mathrm{id}_V, \underline{v}) \text{ nicht invertierbar} \\
&\overset{9.1.14}{\iff} \det(\underbrace{M(f - \lambda \mathrm{id}_V, \underline{v})}_{\overset{7.1.8}{=} M(f,\underline{v}) - \lambda \underbrace{M(\mathrm{id}_V,\underline{v})}_{\overset{7.2.10}{=} I_n}}) = 0 \\
&\iff \det(M(f, \underline{v}) - \lambda I_n) = 0 \\
&\underset{3.2.15}{\overset{9.1.7(*)}{\iff}} \chi_f(\lambda) = 0
\end{aligned}$$

□

Korollar 10.1.8. *Sei V ein Vektorraum, $n := \dim V < \infty$ und $f \in \mathrm{End}(V)$. Dann hat f höchstens n Eigenwerte.*

Beweis. 10.1.7, 10.1.6, 4.2.11 □

Beispiel 10.1.9. [→ 6.3.2, 7.1.4, 10.1.3]

(a) Sei $\varphi \in \mathbb{R}$.

$$\begin{aligned}
\chi_{R_\varphi} &= \det \begin{pmatrix} (\cos\varphi) - X & -\sin\varphi \\ \sin\varphi & (\cos\varphi) - X \end{pmatrix} \overset{9.1.10\,(d)}{=} ((\cos\varphi) - X)^2 + (\sin\varphi)^2 \\
&= X^2 - 2(\cos\varphi)X + \underbrace{(\cos\varphi)^2 + (\sin\varphi)^2}_{=1}
\end{aligned}$$

Für $\lambda \in \mathbb{R}$ gilt:

$$\begin{aligned} \chi_{R_\varphi}(\lambda) = 0 &\iff \lambda = \frac{2(\cos\varphi) \pm \sqrt{4(\cos\varphi)^2 - 4}}{2} \\ &\iff \lambda = \cos\varphi \in \{-1, 1\} \\ &\iff ((\exists n \in \mathbb{Z} \text{ ungerade} : \varphi = n\pi) \,\&\, \lambda = -1) \\ &\quad \text{oder } ((\exists n \in \mathbb{Z} \text{ gerade} : \varphi = n\pi) \,\&\, \lambda = 1) \end{aligned}$$

(b) $\chi_S = \det \begin{pmatrix} -1 - X & 0 \\ 0 & 1 - X \end{pmatrix} = (X - 1)(X + 1)$

(c) $\chi_P = \det \begin{pmatrix} 1 - X & 0 \\ 0 & -X \end{pmatrix} = (X - 1)X$

(d) $\chi_{T_a} = \det \begin{pmatrix} 1 - X & a \\ 0 & 1 - X \end{pmatrix} \overset{9.1.11}{=} (X - 1)^2$ für $a \in \mathbb{R}$.

(e) Ist $A \in K^{n\times n}$, so nennt man $\chi_A := \chi_{f_A} = \det(A - XI_n)$ das *charakteristische Polynom* von A.

(f) $\chi_{D^{(d)}} = (-X)^{d+1}$ für $d \in \mathbb{N}_0$

(g) $E_{a_1,\ldots,a_n}$ ist kein Endomorphismus!

(h) C hat als Endomorphismus des $\mathbb{R}$-Vektorraums $\mathbb{C}$ das charakteristische Polynom

$$\chi_C = \det \begin{pmatrix} 1 - X & 0 \\ 0 & -1 - X \end{pmatrix} = (X - 1)(X + 1)$$

alternativ:

$$\chi_C = \det \begin{pmatrix} -X & 1 \\ 1 & -X \end{pmatrix} = X^2 - 1$$

Proposition 10.1.10. *Sei V ein K-Vektorraum und $f \in \mathrm{End}(V)$. Seien $\lambda_1, \ldots, \lambda_n$ paarweise verschiedene Eigenwerte von f und v_i ein Eigenvektor zum Eigenwert λ_i für jedes $i \in \{1, \ldots, m\}$ [→ 10.1.2]. Dann sind $v_1, \ldots, v_m$ linear unabhängig in V.*

„Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig."

Beweis. Induktion nach $m \in \mathbb{N}_0$

Induktionsanfang: $\underline{m = 0}$: $\emptyset$ ist linear unabhängig.

Induktionsschritt: $\underline{m - 1 \to m (m \in \mathbb{N})}$. Seien $\mu_1, \ldots, \mu_m \in K$ mit $\sum_{i=1}^m \mu_i v_i = 0$. Zu zeigen: $\mu_1 = \ldots = \mu_m = 0$. Man hat

$$\sum_{i=1}^m \mu_i \lambda_i v_i = \sum_{i=1}^m \mu_i f(v_i) = f\left(\sum_{i=1}^m \mu_i v_i\right) = f(0) = 0$$

und

$$\sum_{i=1}^{m} \mu_i \lambda_m v_i = \lambda_m \sum_{i=1}^{m} \mu_i v_i = \lambda_m 0 = 0$$

Bildet man die Differenz, so erhält man

$$\sum_{i=1}^{m-1} \mu_i \underbrace{(\lambda_i - \lambda_m)\, v_i}_{=:w_i} = 0.$$

Für jedes $i \in \{1, \ldots, m-1\}$ gilt $w_i \neq 0$ (denn $\lambda_i - \lambda_m \neq 0$ und $v_i \neq 0$) und $f(w_i) = (\lambda_i - \lambda_m) f(v_i) = \lambda_i((\lambda_i - \lambda_m)v_i) = \lambda_i w_i$. Also ist w_i Eigenvektor zum Eigenwert λ_i für $i \in \{1, \ldots, m-1\}$. Nach Induktionsvoraussetzung gilt $\mu_1 = \ldots = \mu_{m-1} = 0$. Wegen $v_m \neq 0$ gilt dann auch $\mu_m = 0$.

□

Bemerkung 10.1.11. 10.1.8 folgt auch aus 10.1.10 und 6.2.19.

Korollar 10.1.12. *Sei f ein Endomorphismus eines endlichdimensionalen Vektorraums V. Dann ist die Summe der Eigenräume* [→ 10.1.2] *von f direkt* [→ 8.2.1]*, das heißt*

$$\sum_{\lambda \text{ Eigenwert von } f} \ker(f - \lambda \operatorname{id}_V) = \bigoplus_{\lambda \text{ Eigenwert von } f} \ker(f - \lambda \operatorname{id}_V)$$

Beweis. Bezeichne $\lambda_1, \ldots, \lambda_m$ die paarweise verschiedenen Eigenwerte von f. Zu zeigen:

$$\prod_{i=1}^{n} \ker(f - \lambda_i \operatorname{id}_V) \to \sum_{i=1}^{m} \ker(f - \lambda \operatorname{id}_V), (v_1, \ldots, v_m) \mapsto v_1 + \ldots + v_m$$

ist injektiv. Aus 10.1.10 folgt sofort, dass diese Abbildung Kern $\{0\}$ hat. □

Proposition und Definition 10.1.13. Sei $p \in K[X]$ und $\deg p = n \in \mathbb{N}_0$. Bezeichne mit $\lambda_1, \ldots, \lambda_m \in K$ die paarweise verschiedenen Nullstellen [→ 4.2.9] von p (es gilt $m \leq n$ [→ 4.2.11]. Dann gibt es eindeutig bestimmte $\alpha_1, \ldots, \alpha_m \in \mathbb{N}$ und $r \in K[X]$ derart, dass $p = (X - \lambda_1)^{\alpha_1} \ldots (X - \lambda_m)^{\alpha_m} r$ und r keine Nullstelle in K hat. Man nennt α_i die *Vielfachheit* der Nullstelle λ_i von p. Man sagt, dass p *(in Linearfaktoren) zerfällt*, wenn $r \in K$ (d.h. $\deg r = 0$). Es gilt $\alpha_1 + \ldots + \alpha_m + \deg r = n$.

Beweis. Existenz mit 4.2.10, Eindeutigkeit leicht zu sehen. □

Definition 10.1.14. Sei V ein endlichdimensionaler Vektorraum und $f \in \operatorname{End}(V)$. Die *algebraische Vielfachheit* eines Eigenwertes λ von f ist die Vielfachheit von λ als Nullstelle von χ_f.

Die *geometrische Vielfachheit* eines Eigenwertes λ von f ist die Dimension des Eigenraums von f zum Eigenwert λ.

Satz 10.1.15. *Für jeden Eigenwert eines Endomorphismus eines endlichdimensionalen Vektorraums ist seine geometrische Vielfachheit kleiner oder gleich seiner algebraischen Vielfachheit.*

Beweis. Sei V ein K-Vektorraum und $n := \dim V < \infty$, $f \in \mathrm{End}(V)$ und λ Eigenwert von f. Wähle eine Basis $(v_1, \ldots, v_m)$ von $\ker(f - \lambda \mathrm{id}_V)$ [→ 6.2.18] und ergänze sie zu einer Basis $\underline{v} = (v_1, \ldots, v_n)$ von V [→ 6.2.20]. Dann hat $M(f, \underline{v})$ die Gestalt $M(f, \underline{v}) = \left(\begin{array}{c|c} \lambda I_m & A \\ \hline 0 & B \end{array}\right)$ mit $A \in K^{m\times(n-m)}$ und $B \in K^{(n-m)\times(n-m)}$, da $f(v_i) = \lambda v_i$, also $\mathrm{coord}_{\underline{v}}(f(v_i)) = \lambda e_i$ für $i \in \{1, \ldots, m\}$ [→ 7.1.2]. Daher

$$\chi_f \overset{10.1.5}{=} \det\left(\begin{array}{c|c} (\lambda - X)I_m & A \\ \hline 0 & B - XI_{n-m} \end{array}\right)$$
$$\overset{9.1.11}{=} (\det((\lambda - X)I_m)) \det(B - XI_{n-m}) \overset{9.1.11}{=} (\lambda - X)^m r$$

für $r := \det(B - XI_{n-m}) \in K[X]$. □

§10.2 Begleitmatrix, Satz von Cayley-Hamilton und Minimalpolynom

[Arthur Cayley *1821 †1895; William Rowan Hamilton *1805, †1865]

Proposition und Sprechweise 10.2.1 (Polynomdivision mit Rest)**.** *Seien $f, g \in K[X]$ mit $g \neq 0$. Dann gibt es genau ein Paar $(q, r) \in K[X]^2$ mit $\deg r < \deg g$ und $f = gq + r$. Man nennt q den* Quotienten *und r den* Rest *bei Division von f durch g.*

Beweis. Um die Eindeutigkeit zu beweisen, seien $(q_i, r_i) \in K[X]^2$ mit $\deg r_i < \deg g$ und $f = gq_i + r_i$ für $i \in \{1, 2\}$. Dann gilt $r_1 - r_2 = g(q_2 - q_1) \in (g)$ und wegen $\deg(r_1 - r_2) < \deg g$ daher $r_1 - r_2 = 0$. Also $r_1 = r_2$. Folglich $g(q_1 - q_2) = 0$ und schließlich $q_1 = q_2$.

Die Existenz beweisen wir durch Induktion nach dem Grad von f:

Induktionsanfang: Ist $\deg f < \deg g$, so setzen wir $(q, r) := (0, f)$.

Induktionsschritt: Sei $\deg f \geq \deg g$ und die Behauptung schon bewiesen, wenn f durch ein Polynom von kleinerem Grad ersetzt wird. Wähle $a \in K^\times$ und $k \in \mathbb{N}_0$ derart, dass f und $aX^k g$ denselben Grad und denselben Leitkoeffizienten haben. Dann hat $f_0 := f - aX^k g$ einen kleineren Grad als f und es gibt nach Induktionsvoraussetzung $(q_0, r) \in K[X]^2$ mit $\deg r < \deg g$ und $f_0 = gq_0 + r$. Es folgt $f = f_0 + aX^k g = g(q_0 + aX^k) + r = gq + r$ für $q := q_0 + aX^k$. □

Satz 10.2.2. [→3.3.13] *Im Polynomring $K[X]$ ist jedes Ideal ein Hauptideal.*

Beweis. Sei I ein Ideal von $K[X]$. Ist $I = \{0\}$, so ist $I = (0)$. Also bleibt nur der Fall zu betrachten, dass es ein $g \in K[X] \setminus \{0\}$ gibt mit $g \in I$. Wir wählen ein solches g von kleinstmöglichem Grad und behaupten $I = (g)$. Die Inklusion $I \supseteq (g)$ ist klar. Um $I \subseteq (g)$ zu beweisen, sei $f \in I$. Zu zeigen ist $f \in (g)$. Wähle mit 10.2.1 $q, r \in K[X]$ mit

$\deg r < \deg g$ und $f = gq + r$. Dann gilt $r = f - gq \in I$ und nach Wahl von g muss $r = 0$ gelten. Dann aber $f = gq \in (g)$. □

Definition 10.2.3. Ein Polynom $p \in K[X]$ heißt *normiert*, wenn $p \neq 0$ und der Leitkoeffizient von p gleich 1 ist.

Korollar 10.2.4. *Sei I ein Ideal von $K[X]$ mit $I \neq \{0\}$. Dann gibt es genau ein normiertes $p \in K[X]$ mit $I = (p)$.*

Beweis. Die Existenz ist klar aus 10.2.2 durch „Normieren“. Zur Eindeutigkeit: Seien $p, q \in K[X]$ normiert mit $(p) = I = (q)$. Dann $\deg p \geq \deg(q)$ wegen $p \in (q)$ und $\deg q \geq \deg(p)$ wegen $q \in (p)$, also $\deg p = \deg q$. Weiter gilt $p - q \in (p)$ und daher $p - q = 0$ oder $\deg(p - q) \geq \deg p$. Letzteres ist unmöglich, also gilt $p = q$. □

Proposition 10.2.5. *Sei $p \in K[X]$ ein Polynom vom Grad $n \in \mathbb{N}_0$ [→3.2.6]. Dann ist das Ideal (p) [→3.3.11] des kommutativen Ringes $K[X]$ [→3.2.12] ein Unterraum des Vektorraums $K[X]$ [→6.1.4(b)] und $(\overline{1}, \overline{X}, \ldots, \overline{X^{n-1}})$ eine Basis des Quotientenvektorraums $K[X]/(p)$ [→8.1.3].*

Beweis. Nach der Definition einer Basis 6.2.1(c) ist zu zeigen:

(a) $\overline{1}, \overline{X}, \ldots, \overline{X^{n-1}}$ sind linear unabhängig in $K[X]/(p)$.

(b) $\overline{1}, \overline{X}, \ldots, \overline{X^{n-1}}$ erzeugen $K[X]/(p)$.

Zu (a). Wir benutzen 6.2.3. Seien also $a_0, \ldots, a_{n-1} \in K$ mit $\sum_{k=0}^{n-1} a_k \overline{X^k} = 0$. Zu zeigen ist $a_0 = \ldots = a_{n-1} = 0$. Schreibt man $h := \sum_{k=0}^{n-1} a_k X^k \in K[X]$, so ist $h = 0$ zu zeigen. Nun gilt $\overline{h} = \sum_{k=0}^{n-1} a_k \overline{X^k} = 0$ nach 2.3.3 und 8.1.3 und daher $h \in (p)$. Wegen $\deg h < n = \deg p$ folgt $h = 0$.

Zu (b). Sei $f \in K[X]$. Zu zeigen ist, dass es ein $r \in K[X]$ mit $\deg r < n$ und $\overline{f} = \overline{r}$ in $K[X]/(p)$ gibt. Mit 10.2.1 findet man $(q, r) \in K[X]^2$ mit $\deg r < \deg g$ und $f = pq + r$. Dann $f - r \in (p)$ und daher $\overline{f} = \overline{r}$ in $K[X]/(p)$ wie gewünscht. □

Proposition und Definition 10.2.6. Sei $p = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \in K[X]$ mit $a_0, \ldots, a_{n-1} \in K$ ein normiertes Polynom. Dann ist

$$f\colon K[X]/(p) \to K[X]/(p),\ \overline{q} \mapsto \overline{Xq} \quad (q \in K[X])$$

wohldefiniert und linear. Die Darstellungsmatrix $C_p := M(f, \underline{v})$ von f bezüglich der Basis $\underline{v} := (\overline{1}, \overline{X}, \ldots, \overline{X^{n-1}})$ von $K[X]/(p)$ nennen wir die *Begleitmatrix* von p (engl.: companion matrix). Es gilt

$$C_p = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \ddots & \vdots & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

Beweis. Da im Quotientenring $K[X]/(p)$ gilt $\overline{Xq} \overset{3.3.3}{=} \overline{X} \cdot \overline{q}$ für alle $q \in K[X]$, ist f als Abbildung wohldefiniert. Weiter ist f linear, denn für alle $q, r \in K[X]$ und $\lambda \in K$ gilt

$$f(\overline{q}+\overline{r}) = f(\overline{q+r}) = \overline{X(q+r)} = \overline{Xq+Xr} = \overline{Xq} + \overline{Xr} = f(\overline{q}) + f(\overline{r}) \qquad \text{und}$$
$$f(\lambda\overline{q}) = f(\overline{\lambda q}) = \overline{X(\lambda q)} = \overline{\lambda(Xq)} = \lambda\overline{Xq} = \lambda f(\overline{q}).$$

Nach 10.2.5 ist $\underline{v}$ eine Basis des Quotientenvektorraums $K[X]/(p)$. Die behauptete Gleichheit für C_p ergibt sich aus 7.1.2 wegen $f(\overline{X^k}) = \overline{X^{k+1}}$ für alle $k \in \{0, \ldots, n-2\}$ und $f(\overline{X^{n-1}}) = \overline{X^n} = -a_0 1 - a_1\overline{X} - \cdots - a_{n-1}\overline{X^{n-1}}$. □

Satz 10.2.7. *Sei $p \in K[X]$ ein normiertes Polynom vom Grad n. Dann ist p bis auf das Vorzeichen das charakteristische Polynom seiner eigenen Begleitmatrix, das heißt*

$$p = (-1)^n \chi_{C_p}.$$

Beweis. Ist $n = 0$, so $p = 1 \overset{9.1.10(a)}{=} (-1)^0 \det() = (-1)^0 \chi_{C_p}$. Sei also Œ $n \geq 1$. Benutze wieder die Basis $\underline{v} := (\overline{1}, \overline{X}, \ldots, \overline{X^{n-1}})$ des Quotientenvektorraums $K[X]/(p)$ [→10.2.5] und schreibe $p = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \in K[X]$ mit $a_0, \ldots, a_{n-1} \in K$. Es gilt:

$$\chi_{C_p} \overset{10.1.9(e)}{=} \det(C_p - XI_n)$$
$$= \det \begin{pmatrix} -X & 0 & \cdots & 0 & -a_0 \\ 1 & -X & & \vdots & -a_1 \\ 0 & 1 & \ddots & & -a_2 \\ 0 & 0 & & & \vdots \\ \vdots & \vdots & & -X & -a_{n-2} \\ 0 & 0 & \cdots & 1 & -a_{n-1} - X \end{pmatrix}.$$

Addiert man nun in der großen Matrix nacheinander von unten beginnend jeweils das X-fache einer Zeile zur vorherigen, dann ändert sich gemäß 9.1.12(a) die Determinante nicht und man erhält

$$\chi_{C_p} = \det \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 - a_1X - \cdots - a_{n-1}X^{n-1} - X^n \\ 1 & 0 & & \vdots & -a_1 - a_2X - \cdots - a_{n-1}X^{n-2} - X^{n-1} \\ 0 & 1 & \ddots & & -a_2 - a_3X - \cdots - a_{n-1}X^{n-3} - X^{n-2} \\ 0 & 0 & & & \vdots \\ \vdots & \vdots & & 0 & -a_{n-2} - a_{n-1}X - X^2 \\ 0 & 0 & \cdots & 1 & -a_{n-1} - X \end{pmatrix}.$$

In der ersten Zeile der großen Matrix ist also nur der letzte Eintrag verschieden von null und zwar ist dieser $-p$. Entwickelt man diese Determinante mit 9.2.1(a) nun nach der ersten Zeile, so ergibt sich $\chi_{C_p} = (-1)^{1+n}(-p)\det(I_{n-1}) = (-1)^n p$. □

Definition 10.2.8. (a) Ist f eine Selbstabbildung der Menge M, so definiert man

$$f^k := \underbrace{f \circ \cdots \circ f}_{k\text{-mal}} \in M^M$$

für jedes $k \in \mathbb{N}_0$, wobei $f^0 := \mathrm{id}_M$. Insbesondere ist $f^k \in \mathrm{End}(V)$ für jedes $k \in \mathbb{N}_0$, jeden Vektorraum V und jedes $f \in \mathrm{End}(V)$ erklärt.

(b) Ist $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$, so definiert man

$$A^k := \underbrace{A \cdots A}_{k\text{-mal}} \in K^{n\times n},$$

für jedes $k \in \mathbb{N}_0$, wobei $A^0 := I_n$.

Beispiel 10.2.9. Ist $\varphi \in \mathbb{R}$ und $k \in \mathbb{N}_0$, so gilt für die Drehung $R_\varphi \in \mathrm{End}(\mathbb{R}^2)$ [→6.3.2(a)]

$$(R_\varphi)^k = R_{k\varphi}.$$

Erinnerung 10.2.10. [→7.1.6,7.1.7] Sei V ein K-Vektorraum. Dann ist $\mathrm{End}(V) := \mathrm{Hom}(V,V)$ ein K-Vektorraum und für alle $f, g, h \in \mathrm{End}(V)$ und $\lambda \in K$ gilt

$$\begin{aligned} f \circ (g+h) &= f \circ g + f \circ h, \\ (f+g) \circ h &= f \circ h + g \circ h \quad \text{und} \\ (\lambda f) \circ g &= \lambda(f \circ g) = f \circ (\lambda g). \end{aligned}$$

Definition und Proposition 10.2.11. [→3.2.4]

(a) Sei V ein K-Vektorraum und $f \in \mathrm{End}(V)$. Dann ist

$$K[f] := \left\{ \sum_{k=0}^{n} a_k f^k \mid n \in \mathbb{N}_0, a_0, \ldots, a_n \in K \right\}$$

zusammen mit der punktweisen Addition und der Hintereinanderschaltung als Multiplikation ein kommutativer Ring.

(b) Sei $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$. Dann ist

$$K[A] := \left\{ \sum_{k=0}^{n} a_k A^k \mid n \in \mathbb{N}_0, a_0, \ldots, a_n \in K \right\}$$

zusammen mit der Addition und Multiplikation von Matrizen ein kommutativer Ring.

Beweis. **(a)** $K[f] = \mathrm{span}\{f^k \mid k \in \mathbb{N}_0\}$ ist ein Unterraum des K-Vektorraums $\mathrm{End}(V)$ und daher insbesondere bezüglich punktweiser Addition eine abelsche Gruppe. Nun sind die Abgeschlossenheit bezüglich Hintereinanderschaltung sowie $(\dot{K})$, $(\dot{A})$, $(\dot{N})$ und (D) aus 3.1.1 nachzurechnen. Mit Erinnerung 10.2.10 geht dies analog zum Beweis von 3.2.4.

(b) geht genauso unter Benutzung von 7.2.6 und 7.2.7(b). □

Satz und Definition 10.2.12. (a) *Sei V ein K-Vektorraum und $f \in \mathrm{End}(V)$. Dann gibt es genau einen Ringhomomorphismus $\psi\colon K[X] \to K[f]$ mit*

$$\psi\left(\sum_{k=0}^{n} a_k X^k\right) = \sum_{k=0}^{n} a_k f^k$$

für alle $n \in \mathbb{N}_0$ und $a_0, \ldots, a_n \in K$. Ist $p \in K[X]$, so schreibt man auch $p(f)$ statt $\psi(p)$ („p ausgewertet in f“). Dass ψ ein Ringhomomorphismus ist, heißt dann $(p+q)(f) = p(f) + q(f)$, $1(f) = \mathrm{id}_V$ und $(pq)(f) = p(f) \circ q(f)$ für alle $p, q \in K[X]$.

(b) *Sei $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$. Dann gibt es genau einen Ringhomomorphismus $\psi\colon K[X] \to K[A]$ mit*

$$\psi\left(\sum_{k=0}^{n} a_k X^k\right) = \sum_{k=0}^{n} a_k A^k$$

für alle $n \in \mathbb{N}_0$ und $a_0, \ldots, a_n \in K$. Ist $p \in K[X]$, so schreibt man auch $p(A)$ statt $\psi(p)$ („p ausgewertet in A“). Dass ψ ein Ringhomomorphismus ist, heißt dann $(p+q)(A) = p(A)+q(A)$, $1(A) = I_n$ und $(pq)(A) = (p(A))(q(A))$ für alle $p, q \in K[X]$.

Beweis. **(a)** Wende 3.2.13 an: Überprüfe zunächst, dass $\varphi\colon K \to K[f]$, $a \mapsto a\,\mathrm{id}_V$ ein Ringhomomorphismus ist. Dann erhält man $\psi\colon K[X] \to K[f]$ mit

$$\psi\left(\sum_{k=0}^{n} a_k X^k\right) = \sum_{k=0}^{n} \varphi(a_k) f^k = \sum_{k=0}^{n} \underbrace{(a_k\,\mathrm{id}_V) \circ f^k}_{\overset{10.2.10}{=} a_k(\mathrm{id}_V \circ f^k) = a_k f^k} .$$

(b) geht genauso mit $\varphi\colon K \to K[A]$ definiert durch $\varphi(a) = aI_n$ für $a \in K$. □

Beispiel 10.2.13. Sei $p \in K[X]$ ein normiertes Polynom. Dann gilt $\chi_{C_p}(C_p) = 0$. Definiert man nämlich f wie in 10.2.6, so ist dies wegen 7.1.8 und 7.2.5 äquivalent zu $\chi_{C_p}(f) = 0$ und damit wegen 10.2.7 zu $p(f) = 0$. Es gilt aber $(p(f))(\overline{q}) = \overline{pq} = 0$ für alle $q \in K[X]$, wie man sich sofort überlegt. Der folgende Satz ist eine großartige Verallgemeinerung dieses Phänomens.

Satz 10.2.14 (Cayley-Hamilton)**.** (a) *Für jeden Endomorphismus f eines endlichdimensionalen Vektorraums gilt $\chi_f(f) = 0$.*

(b) *Für jede quadratische Matrix A über einem Körper gilt $\chi_A(A) = 0$.*

Beweis. **(a)** Sei V ein endlichdimensionaler Vektorraum, $f \in \mathrm{End}(V)$ und $v \in V$. Zu zeigen ist $(\chi_f(f))(v) = 0$. Nach 6.2.19 können wir das kleinste $m \in \mathbb{N}_0$ wählen derart, dass $v, f(v), \ldots, f^m(v)$ linear abhängig sind. Dann sieht man leicht, dass es $a_0, \ldots, a_{m-1} \in K$ geben muss mit $f^m(v) + a_{m-1}f^{m-1}(v) + \cdots + a_1 f(v) + a_0 v = 0$ $(f(f(v)) = 1 \cdot f(f(v)), f(f^{m-2}(v)) = 1 \cdot f^{m-1}(v), f(f^{m-1}(v)) = f^m(v) = -a_0 v - a_1 f(v) -$

$\cdots - a_{m-1}f^{m-1}(v))$. Da $v, f(v), \ldots, f^{m-1}(v)$ linear unabhängig sind, findet man mit 6.2.20 eine Basis $\underline{v} = (v, f(v), \ldots, f^{m-1}(v), v_m, \ldots, v_n)$ von V. Setzt man nun

$$p := X^m + a_{m-1}X^{m-1} + \cdots + a_1X + a_0,$$

so ist nach 7.1.2

$$M(f, \underline{v}) = \left(\begin{array}{c|c} C_p & * \\ \hline 0 & A \end{array}\right) \text{ mit } C_p = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & & \vdots & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ 0 & 0 & \ddots & 0 & \vdots \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{m-1} \end{pmatrix}$$

für ein $A \in K^{(n-m)\times(n-m)}$. Es ergibt sich

$$\begin{aligned} \chi_f &\overset{10.1.5}{=} \det(M(f, \underline{v}) - XI_n) = \det\left(\begin{array}{c|c} C_p - XI_m & * \\ \hline 0 & A - XI_{n-m} \end{array}\right) \\ &\overset{9.1.11}{=} \det(C_p - XI_m)\det(A - XI_{n-m}) \overset{10.2.7}{=} \chi_{C_p}q = (-1)^m pq = pq = qp \end{aligned}$$

mit $q := \det(A - XI_{n-m}) \in K[X]$. Nun gilt

$$\chi_f(f) = (qp)(f) \overset{10.2.12(a)}{=} q(f) \circ p(f)$$

und daher

$$(\chi_f(f))(v) = (q(f))((p(f))(v)) = (q(f))(0) = 0,$$

da

$$\begin{aligned} (p(f))(v) &= (f^m + a_{m-1}f^{m-1} + \cdots + a_1f + a_0\,\mathrm{id}_V)(v) \\ &= f^m(v) + a_{m-1}f^{m-1}(v) + \cdots + a_1f(v) + a_0 = 0. \end{aligned}$$

(b) Für $A \in K^{n\times n}$ ist $f_A \in \mathrm{End}(K^n)$ [→6.3.2(e)] und wir haben

$$\begin{aligned} \chi_A(A) &\overset{10.1.9(e)}{=} \chi_{f_A}(A) = \chi_{f_A}(M(f_A, \underline{e})) = \\ &\overset{7.1.8}{\underset{7.2.5}{=}} M(\chi_{f_A}(f_A), \underline{e}) \overset{(a)}{=} M(0, \underline{e}) = 0. \end{aligned}$$

□

Bemerkung 10.2.15. (a) Im Beweis des Satzes von Cayley-Hamilton 10.2.14 haben wir Teil (b) sofort aus Teil (a) gewonnen. Geht man umgekehrt von Teil (b) aus, so gewinnt man daraus sofort Teil (a): Ist nämlich V ein K-Vektorraum mit Basis $\underline{v} = (v_1, \ldots, v_n)$, f ein Endomorphismus von V und $A := M(f, \underline{v})$, so gilt $\chi_f = \det(f - XI_n) = \det(A - XI_n) = \chi_A$ und daher $\chi_f(f) = \chi_A(f) = 0$, denn aus $\chi_A(A) = 0$ folgt mit 7.1.8 und 7.2.5 sofort $\chi_A(f) = 0$.

(b) Der folgende „Beweis“ des Satzes von Cayley-Hamilton ist *falsch*: „Ist $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$, so setzen wir in die Gleichung $\chi_A = \det(A - XI_n)$ für X die Matrix A ein [→10.2.12(b)] und erhalten $\chi_A(A) = \det(A - AI_n) = \det(A - A) = \det(0) = 0$.“ Es gibt viele Gründe, warum dies offensichtlich Unsinn sein muss: Zum Beispiel ist $\chi_A(A)$ eine Matrix, aber $\det(A - AI_n)$ ein Skalar. Ausserdem ist im Spezialfall $n = 0$ die Determinante von $0 \in K^{n\times n}$ nicht 0, sondern 1 [→9.1.10(a)]. Wo liegt aber genau der Fehler? Da das Einsetzen von A für X ein Ringhomomorphismus von $K[X]$ nach $K[A]$ ist [→10.2.12(b)], kann man wegen der Leibniz-Formel 9.1.7(∗) tatsächlich zuerst *in jedem Eintrag* der Matrix $A - XI_n$ die Unbestimmte X durch A ersetzen und dann erst die Determinante bilden. Aber XI_n hat nichts mit einem Matrizenprodukt zu tun, sondern es gilt [→7.2.7(b)]

$$XI_n = \begin{pmatrix} X & & 0 \\ & \ddots & \\ 0 & & X \end{pmatrix} \in K[X]^{n\times n}.$$

Einsetzen von A für X in (den Einträgen von) XI_n liefert daher nicht das Produkt der Matrizen A und I_n, sondern

$$\begin{pmatrix} A & & 0 \\ & \ddots & \\ 0 & & A \end{pmatrix} \in K[A]^{n\times n}.$$

Einsetzen von A für X in $A - XI_n$ liefert also eine Matrix in $K[A]^{n\times n}$, die im Allgemeinen nicht die Nullmatrix ist.

(c) Wir geben noch einen zweiten, sehr kurzen, aber etwas unheimlichen Beweis von Cayley-Hamilton mit Hilfe der Komatrix aus 9.2.3. Sei wieder $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$. Man sieht leicht, dass die Einträge der Komatrix von $A - XI_n$ Polynome vom Grad $\leq n - 1$ sind. Schreibe nun $\chi_A = a_nX^n + \cdots + a_1X + a_0$ mit $a_0, \ldots, a_n \in K$ und $(\mathrm{com}(A - XI_n))^T = X^{n-1}B_{n-1} + \cdots + XB_1 + B_0$ mit $B_0, \ldots, B_{n-1} \in K^{n\times n}$. Dann gilt $(A - XI_n)(X^{n-1}B_{n-1} + \cdots + XB_1 + B_0) = (A - XI_n)(\mathrm{com}(A - XI_n))^T \overset{9.2.4}{=} (\det(A - XI_n))I_n = X^n a_n I_n + \cdots + Xa_1I_n + a_0I_n$, woraus durch Vergleich der Koeffizienten der Einträge folgt:

$$\begin{aligned} -B_{n-1} &= a_nI_n \\ AB_{n-1} - B_{n-2} &= a_{n-1}I_n \\ &\vdots \\ AB_1 - B_0 &= a_1I_n \\ AB_0 &= a_0I_n. \end{aligned}$$

Multiplizieren von links mit Potenzen von A liefert

$$\begin{aligned} -A^n B_{n-1} &= a_n A^n \\ A^n B_{n-1} - A^{n-1} B_{n-2} &= a_{n-1} A^{n-1} \\ &\vdots \\ A^2 B_1 - A B_0 &= a_1 A \\ A B_0 &= a_0 I_n. \end{aligned}$$

Addiert man diese Gleichungen, so erhält man links die Nullmatrix und rechts $\chi_A(A)$.

(d) *Nous avons testé pour vous 30 démonstrations du théorème de Cayley-Hamilton* von Michel Coste ist eine Übersicht über Beweise von Cayley-Hamilton: `http://agreg-maths.univ-rennes1.fr/documentation/docs/HaCa.pdf`.

Definition 10.2.16. [→10.2.12, 3.3.15]

(a) Sei V ein K-Vektorraum und $f \in \mathrm{End}(V)$. Dann heißt der Kern $I_f := \ker \psi$ des Ringhomomorphismus $\psi\colon K[X] \to K[f],\ p \mapsto p(f)$ das Ideal der *algebraischen Identitäten* von f.

(b) Sei $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$. Dann heißt der Kern $I_A := \ker \psi$ des Ringhomomorphismus $\psi\colon K[X] \to K[A],\ p \mapsto p(A)$ das Ideal der *algebraischen Identitäten* von A.

Bemerkung 10.2.17. Der Satz von Cayley-Hamilton 10.2.14 besagt:

(a) Ist f ein Endomorphismus eines endlichdimensionalen Vektorraums, so gilt $\chi_f \in I_f$ und daher insbesondere $I_f \neq \{0\}$.

(b) Ist $A \in K^{n\times n}$, so gilt $\chi_A \in I_A$ und daher insbesondere $I_A \neq \{0\}$.

Definition 10.2.18. [→10.2.4]

(a) Sei V ein K-Vektorraum und $f \in \mathrm{End}(V)$ mit $I_f \neq \{0\}$. Dann heißt das eindeutig bestimmte normierte Polynom $\mu_f \in K[X]$ mit $I_f = (\mu_f)$ das *Minimalpolynom* von f.

(b) Sei $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$. Dann heißt das eindeutig bestimmte normierte Polynom $\mu_A \in K[X]$ mit $I_A = (\mu_A)$ das *Minimalpolynom* von A.

Bemerkung 10.2.19. [→10.2.14]

(a) Sei V ein Vektorraum mit $n := \dim V < \infty$ und $f \in \mathrm{End}(V)$. Dann gibt es $r \in K[X]$ mit $\chi_f = \mu_f r$. Insbesondere gilt $\deg(\mu_f) \leq n$.

(b) Sei $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$. Dann gibt es $r \in K[X]$ mit $\chi_A = \mu_A r$. Insbesondere gilt $\deg(\mu_A) \leq n$.

Beispiel 10.2.20. [→6.3.2, 7.1.4, 10.1.3, 10.1.9] In den folgenden Beispielen benutzen wir, dass für einen Endomorphismus f eines zweidimensionalen K-Vektorraums V offensichtlich gilt:

$$\chi_f \neq \mu_f \iff \exists \lambda \in K : f = \lambda \,\mathrm{id}_V .$$

(a) Sei $\varphi \in \mathbb{R}$. Dann gilt

$$\begin{aligned} \chi_{R_\varphi} &= X^2 - 2(\cos\varphi)X + 1 \qquad \text{und} \\ \mu_{R_\varphi} &= \begin{cases} \chi_{R_\varphi} & \text{falls } \varphi \notin \{n\pi \mid n \in \mathbb{Z}\} \\ X - 1 & \text{falls } \varphi = n\pi \text{ für ein gerades } n \in \mathbb{Z} \\ X + 1 & \text{falls } \varphi = n\pi \text{ für ein ungerades } n \in \mathbb{Z} \end{cases} . \end{aligned}$$

Nach dem Satz von Cayley-Hamilton 10.2.14 gilt

$$\begin{aligned} \underbrace{(R_\varphi)^2}_{\overset{10.2.9}{=} R_{2\varphi}} - 2(\cos\varphi)R_\varphi + \mathrm{id}_{\mathbb{R}^2} &= 0, \quad \text{also} \\ R_{2\varphi}(v) - 2(\cos\varphi)R_\varphi(v) + v &= 0 \quad \text{für alle } v \in \mathbb{R}^2. \end{aligned}$$

(b) Es gilt $\mu_S = \chi_S = (X-1)(X+1) = X^2 - 1$ und Cayley-Hamilton besagt $S^2 = \mathrm{id}_{\mathbb{R}^2}$ („zweimal spiegeln ist keinmal spiegeln“).

(c) Es gilt $\mu_P = \chi_P = X(X-1) = X^2 - X$ und Cayley-Hamilton besagt $P^2 = P$ („zweimal projizieren ist einmal projizieren“).

(d) Sei $a \in \mathbb{R}$. Dann gilt

$$\begin{aligned} \chi_{T_a} &= (X-1)^2 = X^2 - 2X + 1 \qquad \text{und} \\ \mu_{T_a} &= \begin{cases} \chi_{T_a} & \text{falls } a \neq 0 \\ X - 1 & \text{falls } a = 0 \end{cases} . \end{aligned}$$

Cayley-Hamilton sagt hier $T_a^2 = 2T_a - \mathrm{id}_{\mathbb{R}^2}$ („zweimal scheren ist scheren, verdoppeln und Ausgangsvektor abziehen“).

(e) Ist $A \in K^{n\times n}$, so ist $\chi_{f_A} = \chi_A$ und $\mu_{f_A} = \mu_A$ wegen 7.1.8 und 7.2.5.

(f) Sei $d \in \mathbb{N}_0$. Wegen $\chi_{D^{(d)}} = (-X)^{d+1}$ besagt Cayley-Hamilton hier, dass $D^{d+1}(p) = (D^{(d)})^{d+1}(p) = 0$ für alle $p \in K[X]_d$, das heißt ein Polynom vom Grad $\leq d$ wird nach $(d+1)$-maligem Ableiten das Nullpolynom.

(g) $E_{a_1,\ldots,a_n}$ ist kein Endomorphismus!

(h) Es gilt $\mu_C = \chi_C = X^2 - 1$ und Cayley-Hamilton besagt $C^2 = \mathrm{id}_{\mathbb{C}}$ („zweimal komplex konjugieren ist keinmal komplex konjugieren“).

§10.3 Diagonalisierbarkeit und Trigonalisierbarkeit

Erinnerung und Definition 10.3.1. Eine Matrix [→ 5.1.8] mit ebensoviel Zeilen wie Spalten nennt man *quadratisch* [→ 10.2.14(b)]. Eine quadratische Matrix nennt man in $\left\{\begin{array}{c}\textit{oberer Dreiecksgestalt}\\ \textit{Diagonalgestalt}\\ \textit{unterer Dreiecksgestalt}\end{array}\right\}$ oder eine $\left\{\begin{array}{c}\textit{obere Dreiecksmatrix}\\ \textit{Diagonalmatrix}\\ \textit{untere Dreiecksmatrix}\end{array}\right\}$, wenn sie von der Form $\left\{\begin{pmatrix}\lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}, \begin{pmatrix}\lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}, \begin{pmatrix}\lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n\end{pmatrix}\right\}$ ist.

Definition 10.3.2. Sei V ein Vektorraum und $f \in \mathrm{End}(V)$.

(a) f heißt $\left\{\begin{array}{c}\textit{diagonalisierbar}\\ \textit{trigonalisierbar}\end{array}\right\}$, wenn es eine geordnete Basis $\underline{v}$ von V gibt derart, dass $M(f,\underline{v})$ [→ 7.1.11] $\left\{\begin{array}{c}\text{Diagonal-}\\ \text{obere Dreiecks-}\end{array}\right\}$gestalt hat.

(b) Eine (geordnete) Basis von V heißt (geordnete) *Eigenbasis* für f, wenn sie aus Eigenvektoren [→10.1.2] von f besteht.

Satz 10.3.3. *Sei f ein Endomorphismus eines endlichdimensionalen Vektorraums V. Äquivalent sind folgende Aussagen:*

(a) *f ist diagonalisierbar.*

(b) *f besitzt eine Eigenbasis.*

(c) *χ_f zerfällt und für jeden Eigenwert von f stimme geometrische und algebraische Vielfachheit überein* [→ 10.1.14, 10.1.15].

Beweis. Sei V ein K-Vektorraum, $n := \dim V < \infty$ und $f \in \mathrm{End}(V)$.

(a) $\Longleftrightarrow$ (b). folgt aus der folgenden Tatsache:

($*$) Sei $\underline{v} = (v_1, \ldots, v_n)$ Basis von V. Dann gilt für $\lambda_1, \ldots, \lambda_n \in K$:

$$M(f,\underline{v}) = \begin{pmatrix}\lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix} \overset{7.1.1}{\Longleftrightarrow} \forall i \in \{1,\ldots,n\} : \mathrm{coord}_{\underline{v}}(v_i) = \begin{pmatrix}0\\ \vdots\\ 0\\ \lambda_i \leftarrow i\text{-te Stelle}\\ 0\\ \vdots\\ 0\end{pmatrix}$$

$$\overset{6.3.7}{\Longleftrightarrow} \forall i \in \{1,\ldots,n\} : f(v_i) = \underbrace{0\cdot v_1 + \ldots + 0\cdot v_{i-1} + \lambda_i v_i + 0\cdot v_{i+1} + \ldots + 0\cdot v_n}_{\lambda_i v_i}$$

$$\Longleftrightarrow \forall i \in \{1,\ldots,n\} : v_i \text{ ist Eigenvektor von } f \text{ zum Eigenwert } \lambda$$

(a) $\Longrightarrow$ (c). Sei f diagonalisierbar. Wähle geordnete Basis $\underline{v}$ von V mit $M(f,\underline{v}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, $\lambda_i \in K$. Dann

$$\chi_f \overset{10.1.5}{=} \det(M(f,\underline{v}) - XI_n) = \det \begin{pmatrix} \lambda_1 - X & & 0 \\ & \ddots & \\ 0 & & \lambda_n - X \end{pmatrix}$$
$$\overset{9.1.11}{=} \prod_{i=1}^{n} (\lambda_i - X) = (-1)^n \prod_{i=1}^{n} (X - \lambda_i).$$

Es zerfällt χ_f also [→ 10.1.13]. Sei nun λ ein Eigenwert von f. Nach Umnummerieren der λ_i können wir $\lambda = \lambda_1 = \ldots = \lambda_m \neq \lambda_i$ für $i \in \{m+1, \ldots, n\}$. Zu zeigen: $\dim \ker(f - \lambda \operatorname{id}_V) = m$ [→ 10.1.14]. Aus 10.1.15 wissen wird schon „$\leq$“.

„$\geq$“: Nach $(*)$ gilt $f(v_i) = \lambda v_i$ für $i \in \{1, \ldots, m\}$. Also sind $v_1, \ldots, v_m$ linear unabhängige Elemente von $\ker(f - \lambda \operatorname{id}_V)$.

(c) $\Longrightarrow$ (b). Gelte (c). Bezeichne die paarweise verschiedenen Eigenwerte von f mit $\lambda_1, \ldots, \lambda_m \in K$ und deren algebraische Vielfachheiten mit $\alpha_1, \ldots, \alpha_m \in \mathbb{N}$. Da χ_f zerfällt gilt $\alpha_1 + \cdots + \alpha_m = \deg \chi_f \overset{10.1.6}{=} n$. Wegen $\dim \ker(f - \lambda_i \operatorname{id}_V) \overset{\text{(c)}}{=} \alpha_i$ folgt [→ 10.1.12]

$$\dim \left(\bigoplus_{i=1}^{n} \ker(f - \lambda_i \operatorname{id}_V) \right) \overset{8.2.2}{=} \sum_{i=1}^{m} \alpha_i = n,$$

also $\bigoplus_{i=1}^{m} \ker(f - \lambda_i \operatorname{id}_V) = V$ nach 6.2.27, Wählt man nun für jedes $i \in \{1, \ldots, m\}$ eine Basis B_i von $\ker(f - \lambda_i)$, so ist $B_1 \cup \ldots \cup B_m$ eine Basis von V nach 8.2.2. □

Definition 10.3.4. Sei $A \in K^{n \times n}$. Dann heißt A $\left\{ \begin{matrix} \textit{diagonalisierbar} \\ \textit{trigonalisierbar} \end{matrix} \right\}$, wenn f_A [→ 6.3.2 (e)] $\left\{ \begin{matrix} \text{diagonalisierbar} \\ \text{trigonalisierbar} \end{matrix} \right\}$ [→ 10.3.2] ist. Wir nennen eine Eigenbasis für f_A auch Eigenbasis für A.

Satz 10.3.5. *Sei $A \in K^{n \times n}$. Dann sind äquivalent:*

(a) *A ist diagonalisierbar*

(b) *A ist ähnlich zu einer Diagonalmatrix*

(c) *A besitzt eine Eigenbasis*

(d) *χ_A zerfällt und für jeden Eigenwert von A stimmen geometrische und algebraische Vielfachheit überein.*

Beweis. (a) $\Longleftrightarrow$ (c) $\Longleftrightarrow$ (d) ist klar, da dasselbe für f_A gilt [→ 10.3.3].

(a) $\Longrightarrow$ (b). Sei f_A diagonalisierbar, etwa $\underline{v}$ eine geordnete Basis von K^n mit $D := M(f_A, \underline{v})$ in Diagonalgestalt. Dann gilt

$$A \overset{7.1.4(e)}{=} M(f_A, \underline{e}) \overset{9.1.19}{\approx} M(f_A, \underline{v}) = D.$$

(b) $\implies$ (a). Sei $P \in K^{n\times n}$ invertierbar und $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in K^{n\times n}$ mit $A = P^{-1}DP$. Dann bilden $v_1, \ldots, v_n \in K^n$ mit $v_i := P^{-1}e_i$ eine Basis von V mit

$$f_A(v_i) = Av_i = P^{-1}DPP^{-1}e_i = P^{-1}De_i = P^{-1}\lambda_i e_i = \lambda_i P^{-1}e_i = \lambda_i v_i.$$

Also $M(f, \underline{v}) = D$ mit $\underline{v} := (v_1, \ldots, v_n)$. □

Proposition 10.3.6. *Sei f ein Endomorphismus des Vektorraum V. Ein Unterraum U von V heißt f-*invariant*, wenn $f(U) \subseteq U$. Ist U f-invariant, so sind $f|_U : U \to U, v \mapsto f(v)$ und $\overline{f}^U : V/U \to V/U, \overline{v}^U \mapsto \overline{f(v)}^U$ wohldefinierte Endomorphismen.*

Beweis. klar für $f|_U$.

für $\overline{f}^U : V \xrightarrow{f} V \xrightarrow{v\mapsto \overline{v}^U} V/U$ sind linear [→ 8.1.8], also nach 6.3.3 $g : V \to V/U, v \mapsto \overline{f(v)}^U$. Es gilt $U \subseteq \ker g$, denn für $v \in U$ gilt $f(v) \in U$ und daher $g(v) = \overline{f(v)}^U = 0$. Nach Homomorphiesatz 8.1.9 ist $\overline{f}^U = \overline{g}$ wohldefiniert und linear. □

Lemma 10.3.7. [→ 8.1.11] *Seien V ein Vektorraum, $f \in \mathrm{End}(V)$ und U ein f-invarianter Unterraum von V. Weiter seien $m, n \in \mathbb{N}_0$ mit $m \leq n$ und $v_1, \ldots, v_n$ derart, dass $\underline{u} := (v_1, \ldots, v_m)$ eine Basis von U und $\underline{w} := (\overline{v_{m+1}}, \ldots, \overline{v_n})$ eine Basis von V/U. Dann $\underline{v} := (v_1, \ldots, v_n)$ eine Basis von V und $M(f, \underline{v})$ von der Gestalt* $M(f, \underline{v}) = \left(\begin{array}{c|c} M(f|_U) & * \\ \hline 0 & M\left(\overline{f}^U, \underline{w}\right) \end{array}\right)$.

Beweis. $\underline{v}$ ist eine Basis von V nach 8.1.11. Sei nun $j \in \{1, \ldots, n\}$ und $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ die j-te Spalte von $M(f, \underline{v})$. Dann $f(v_j) = \sum_{i=1}^n \lambda_i v_i$ [→ 7.1.1]. Ist $j \in \{1, \ldots, m\}$, so ist $v_j \in U$, also $f(v_j) \in I$ und daher $\lambda_{m+1} = \ldots = \lambda_n = 0$. Ist $j \in \{m+1, \ldots, n\}$, so ist

$$\overline{f}^U(v_j) \overset{10.3.6}{=} \overline{f(v_j)}^U = \overline{\sum_{i=1}^n \lambda_i v_i}^U = \sum_{i=1}^n \lambda_i \overline{v_i}^U = \sum_{i=m+1}^n \lambda_i \overline{v_i}^U.$$

□

Korollar 10.3.8. [→ 9.1.11] *Ist f ein Endomorphismus eines endlichdimensionalen Vektorraums V und U ein f-invarianter Unterraum von V, so $\chi_f = \chi_{f|_U}\chi_{\overline{f}^U}$.*

Satz 10.3.9. [→ 10.3.3] *Sei f ein Endomorphismus eines endlichdimensionalen Vektorraums V. Dann sind äquivalent:*

(a) *f trigonalisierbar*

(b) *χ_f zerfällt*

(c) *Es gibt eine geordnete Basis $\underline{v}$ von V mit $M(f,\underline{v})$ in unterer Dreiecksgestalt.*

Beweis. (a) $\Longrightarrow$ (b). Ist $\underline{v}$ geordnete Basis von V mit $M(f,\underline{v}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, so

$$\chi_f \overset{9.1.11}{=} \prod_{i=1}^{n}(\lambda_i - X) = (-1)^n \prod_{i=1}^{n}(X - \lambda_i).$$

(a) $\Longleftrightarrow$ (c). Ist $\underline{v} = (v_1, \ldots, v_n)$ Basis von V, so auch $\underline{w} := (v_n, \ldots, v_1)$ und es gilt $M(f,\underline{v})$ hat obere Dreiecksgestalt $\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ $\Longleftrightarrow$ $M(f,\underline{w})$ hat untere Dreiecksgestalt $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{pmatrix}$.

(b) $\Longrightarrow$ (a). Induktion nach $n := \dim V$.

Induktionsanfang: $\underline{n = 0}$: Die 0×0-Matrix ist eine obere Dreiecksmatrix.

Induktionsschritt: $n - 1 \to n$ $(n \in \mathbb{N})$. Sei V ein K-Vektorraum mit $\dim V = n$ und $f \in End(V)$. Es zerfalle χ_f. Wegen $n \geq 1$ hat dann χ_f eine Nullstelle $\lambda \in K$. Wähle Eigenvektor v_1 von f zum Eigenwert λ [$\to$ 10.1.7]. Dann ist $U := \operatorname{span}(v_1)$ f-invariant, $\underline{u} := (v_1)$ eine Basis von U, $M(f|_U, \underline{u}) = (\lambda) \in K^{1\times 1}$ und $\chi_{f|_U} = \lambda - X$. Nach 10.3.8 $\chi_f = \chi_{f|_U}\chi_{\overline{f}^U}$, weshalb auch $\chi_{\overline{f}^U}$ zerfällt. Wegen $\dim(V/U) \overset{8.1.11}{=} n-1$ gibt es dann nach IV eine Basis $\underline{w} = (\overline{v_2}, \ldots, \overline{v_n})$ von V/U (mit $v_2, \ldots, v_n \in V$) derart, dass $M\left(\overline{f}^U, \underline{w}\right)$ obere Dreiecksgestalt hat. Nach 10.3.7 ist dann $\underline{v} := (v_1, \ldots, v_n)$ eine Basis von V. und $M(f,\underline{v}) = \left(\begin{array}{c|c} \lambda & * \\ \hline 0 & M\left(\overline{f}^U, \underline{w}\right) \end{array}\right)$ hat obere Dreiecksgestalt.

□

Satz 10.3.10. [$\to$ 10.3.5] *Sei $A \in K^{n\times n}$. Dann sind äquivalent:*

(a) *A ist trigonalisierbar.*

(b) *A ist ähnlich zu einer oberen Dreiecksmatrix.*

(c) *A ist ähnlich zu einer unteren Dreiecksmatrix.*

(d) *χ_A zerfällt.*

Beweis. ähnlich wie 10.3.5 (mit 10.3.9 statt 10.3.3). □

§11 Vektorräume mit Skalarprodukt

In diesem Kapitel sei stets $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Ist $a \in \mathbb{K}$, so schreiben wir wieder a^* für die komplex-konjugierte Zahl [→4.2.7] (falls $\mathbb{K} = \mathbb{R}$, so gilt natürlich $a^* = a$). Allgemeiner schreiben wir A^* für die komplex-konjugierte transponierte Matrix $(a^*_{ji})_{1\le j\le n, 1\le i\le m} \in \mathbb{K}^{n\times m}$ einer Matrix $A = (a_{ij})_{1\le i\le m, 1\le j\le n} \in \mathbb{K}^{m\times n}$ (falls $\mathbb{K} = \mathbb{R}$, so gilt natürlich $A^* = A^T$ [→9.1.21]). Zum Beispiel gilt

$$\begin{pmatrix} 1+2\mathring{i} & 0 & 1 \\ 0 & -\mathring{i} & 2 \end{pmatrix}^* = \begin{pmatrix} 1-2\mathring{i} & 0 \\ 0 & \mathring{i} \\ 1 & 2 \end{pmatrix}.$$

§11.1 Skalarprodukte

Definition 11.1.1. Sei V ein $\mathbb{K}$-Vektorraum. Ein *Skalarprodukt* auf V ist eine Abbildung $V \times V \to \mathbb{K},\ (v, w) \mapsto \langle v, w\rangle$ derart, dass für alle $u, v, w \in V$ und $\lambda \in \mathbb{K}$ gilt:

(1)	$\langle u+v, w\rangle$	$=$	$\langle u, w\rangle + \langle v, w\rangle$
(2)	$\langle \lambda v, w\rangle$	$=$	$\lambda^* \langle v, w\rangle$
(3)	$\langle u, v+w\rangle$	$=$	$\langle u, v\rangle + \langle u, w\rangle$
(4)	$\langle v, \lambda w\rangle$	$=$	$\lambda \langle v, w\rangle$
(5)	$\langle v, w\rangle$	$=$	$\langle w, v\rangle^*$
(6)	$v \neq 0$	$\Longrightarrow$	$\langle v, v\rangle > 0$

Für $\mathbb{K} = \mathbb{R}$ sagt man, dass die Abbildung *bilinear* [(1)–(4)] (nämlich *linear im ersten* [(1),(2)] und *zweiten Argument* [(3),(4)]), *symmetrisch* [(5)] und *positiv definit* [(6)] ist. Für $\mathbb{K} = \mathbb{C}$ sagt man, dass die Abbildung *sesquilinear*[1] [(1)–(4)] (nämlich *semilinear oder auch antilinear im ersten* [(1),(2)] und *linear im zweiten Argument* [(3),(4)]), *hermitesch* [(5)] und *positiv definit* [(6)] ist.

Bemerkung 11.1.2. (a) Sei V ein K-Vektorraum und $V \times V \to \mathbb{K},\ (v, w) \mapsto \langle v, w\rangle$ eine Abbildung, die 11.1.1(4),(5) erfüllt. Dann gilt $\langle v, 0\rangle = \langle v, 0_{\mathbb{K}} \cdot 0\rangle = 0_{\mathbb{K}} \langle v, 0\rangle = 0_{\mathbb{K}}$

[1] „sesqui“ kommt aus dem Lateinischen und bedeutet „anderthalb“.

und daher $\langle 0, v\rangle = \langle v, 0\rangle^* = 0_{\mathbb{K}}^* = 0_{\mathbb{K}}$. Beachte auch, dass nach (5) gilt $\langle v, v\rangle \in \mathbb{R}$ für alle $v \in V$. Es ist dann 11.1.1(6) äquivalent dazu, dass für alle $v \in V$ gilt

(6') $\qquad \langle v, v\rangle \geq 0 \qquad$ und $\qquad (\langle v, v\rangle = 0 \implies v = 0)$.

(b) Manche Autoren fordern in der Definition eines Skalarprodukts auf einem $\mathbb{C}$-Vektorraum die Linearität im ersten und die Semilinearität im zweiten Argument. Dies ist kein Problem: Um zwischen den beiden Literaturen hin und her zu springen, muss man lediglich die beiden Argumente vertauschen.

Beispiel 11.1.3. (a) $\mathbb{K} \times \mathbb{K} \to \mathbb{K},\ (x, y) \mapsto \langle x, y\rangle$ definiert durch

$$\langle x, y\rangle := \sum_{i=1}^{n} x_i^* y_i = x^* y \qquad \text{für } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{K}^n$$

ist ein Skalarprodukt auf $\mathbb{K}^n$, das *Standardskalarprodukt* auf $\mathbb{K}^n$, wobei 11.1.1(6) folgt aus $\langle x, x\rangle = \sum_{i=1}^n |x_i|^2$ für alle $x \in \mathbb{K}^n$ [→4.2.8].

(b) $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R},\ (x, y) \mapsto \langle x, y\rangle$ definiert durch

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle := x_1 y_1 + 5x_1 y_2 + 5x_2 y_1 + 26 x_2 y_2 \qquad \text{für } x_1, x_2, y_1, y_2 \in \mathbb{R}$$

ist ein Skalarprodukt auf $\mathbb{R}^2$, denn $x_1^2 + 10x_1x_2 + 26x_2^2 = (x_1 + 5x_2)^2 + x_2^2 > 0$ für $\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) \in \mathbb{R}^2 \setminus \{0\}$.

(c) Es ist $C([0,1], \mathbb{K}) := \{f \mid f\colon [0,1] \to \mathbb{K} \text{ stetig}\}$ ein Unterraum des $\mathbb{K}$-Vektorraums $\mathbb{K}^{[0,1]}$ [→7.1.5] und

$$C([0,1], \mathbb{K}) \times C([0,1], \mathbb{K}) \to \mathbb{K},\ (f, g) \mapsto \langle f, g\rangle := \int_0^1 f(x)^* g(x) dx$$

ein Skalarprodukt auf $C([0,1], \mathbb{K})$, denn $\langle f, f\rangle = \int_0^1 f(x)^* f(x) dx = \int_0^1 |f(x)|^2 dx > 0$ falls $f \in C([0,1], \mathbb{K}) \setminus \{0\}$.

Definition 11.1.4. Sei V ein $\mathbb{K}$-Vektorraum. Eine *Norm* auf V ist eine Abbildung $V \to \mathbb{R},\ v \mapsto \|v\|$ derart, dass für alle $v, w \in V$ und $\lambda \in \mathbb{K}$ gilt

$$\begin{array}{rcll} \|v + w\| & \leq & \|v\| + \|w\| & \text{„Dreiecksungleichung“} \\ \|\lambda v\| & = & |\lambda| \|v\| & \text{„absolute Homogenität“} \\ v \neq 0 & \implies & \|v\| > 0 & [\text{beachte auch } \|0\| = \|0_{\mathbb{K}} 0\| = |0_{\mathbb{K}}| \|0\| = 0_{\mathbb{K}}] \end{array}$$

Definition 11.1.5. Einen ($\mathbb{K}$-)Vektorraum zusammen mit einem Skalarprodukt oder einer Norm auf V nennt man einen ($\mathbb{K}$-) *Vektorraum mit Skalarprodukt* beziehungsweise einen *normierten* ($\mathbb{K}$-) *Vektorraum*.

Bemerkung 11.1.6. (a) Einen $\mathbb{K}$-Vektorraum mit Skalarprodukt nennt man einen $\mathbb{K}$-*Prähilbertraum*, im Fall $\mathbb{K} = \mathbb{R}$ auch einen *euklidischen Raum* und im Fall $\mathbb{K} = \mathbb{C}$ auch einen *unitären Raum*. Altmodische Autoren sprechen jedoch sowohl im reellen als auch im komplexen Fall von einem unitären Raum, während neumodische Autoren in beiden Fällen von einem euklidischen Raum sprechen.

(b) Formal sind Vektorräume mit Skalarprodukt und normierte Räume bei uns 7-Tupel, da Vektorräume 6-Tupel sind [→6.1.1]. So wie wir in einer abelschen Gruppe die Addition fast immer mit + notieren [→2.1.2(d)], schreiben wir das Skalarprodukt in einem Vektorraum mit Skalarprodukt fast immer mit $\langle .,. \rangle$ und die Norm in einem normierten Vektorraum fast immer mit $\|.\|$.

Lemma 11.1.7. *Seien V ein Vektorraum mit Skalarprodukt und $v, w \in V$. Dann*

$$\langle v+w, v+w\rangle = \langle v, v\rangle + 2\,\mathrm{Re}(\langle v, w\rangle) + \langle w, w\rangle.$$

Beweis. $\langle v+w, v+w\rangle = \langle v, v+w\rangle + \langle w, v+w\rangle = \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle$
$= \langle v, v\rangle + \langle v, w\rangle + \langle v, w\rangle^* + \langle w, w\rangle = \langle v, v\rangle + 2\,\mathrm{Re}(\langle v, w\rangle) + \langle w, w\rangle$ [→4.2.8] □

Satz 11.1.8 (Cauchy-Schwarz-Ungleichung)**.** [Augustin Louis, baron Cauchy *1789 †1857, Hermann Amandus Schwarz *1843 †1921] *Seien V ein Vektorraum mit Skalarprodukt und $v, w \in V$. Dann gilt*

$$|\langle v, w\rangle|^2 \leq \langle v, v\rangle\langle w, w\rangle$$

mit Gleichheit genau dann, wenn v und w linear abhängig sind.

Beweis. Bezeichne $\mathbb{K}$ den Grundkörper von V und setze $S := \{\zeta \in \mathbb{K} \mid |\zeta| = 1\}$. Dann gilt für $x \in \mathbb{R}$ und $\zeta \in S$ nach Lemma 11.1.7

$$0 \leq \langle v + x\zeta w, v + x\zeta w\rangle = \langle v, v\rangle + 2x\,\mathrm{Re}(\zeta\langle v, w\rangle) + x^2\langle w, w\rangle.$$

Wähle für jedes $x \in \mathbb{R}$ ein $\zeta_x \in S$ mit $\mathrm{Re}(\zeta_x\langle v, w\rangle) = |\langle v, w\rangle|$ (nehme $\zeta_x := \frac{\langle v,w\rangle^*}{|\langle v,w\rangle|} \in S$ falls $\langle v, w\rangle \neq 0$). Dann gilt für alle $x \in \mathbb{R}$

$$0 \leq \langle v + x\zeta_x w, v + x\zeta_x w\rangle = \langle v, v\rangle + 2x|\langle v, w\rangle| + x^2\langle w, w\rangle.$$

Ist $\langle w, w\rangle = 0$, so gilt also $\langle v, w\rangle = 0$ und $w = 0$ ist linear abhängig. Sei also $\langle w, w\rangle > 0$. Die Diskriminante $(2|\langle v, w\rangle|)^2 - 4\langle w, w\rangle\langle v, v\rangle$ ist dann nicht positiv und genau dann null, wenn v ein skalares Vielfaches von w ist. □

Satz 11.1.9. *Sei V ein Vektorraum mit Skalarprodukt. Dann ist*

$$V \to \mathbb{R},\ v \mapsto \|v\| := \sqrt{\langle v, v\rangle}$$

eine Norm auf V. Jeder Vektorraum mit Skalarprodukt ist also auf diese Weise ein normierter Raum.

Beweis. Es ist alles klar bis auf die Dreiecksungleichung: Für alle $v, w \in V$ gilt

$$\begin{aligned} \|v+w\|^2 &= \langle v+w, v+w\rangle \\ &\overset{11.1.7}{=} \langle v,v\rangle + 2\,\mathrm{Re}(\langle v,w\rangle) + \langle w,w\rangle \\ &\overset{4.2.8}{\leq} \langle v,v\rangle + 2|\langle v,w\rangle| + \langle w,w\rangle \\ &\overset{11.1.8}{\leq} \langle v,v\rangle + 2\|v\|\|w\| + \langle w,w\rangle \\ &= (\|v\| + \|w\|)^2. \end{aligned}$$

□

Satz 11.1.10 (Polarisationsformel)**.** *Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Dann gilt für alle $v, w \in V$:*

$$4\langle v,w\rangle = \|v+w\|^2 - \|v-w\|^2 \text{ falls } \mathbb{K} = \mathbb{R} \text{ und}$$
$$4\langle v,w\rangle = \|v+w\|^2 - \|v-w\|^2 - \mathring{i}\|v+\mathring{i}w\|^2 + \mathring{i}\|v-\mathring{i}w\|^2 \text{ falls } \mathbb{K} = \mathbb{C}.$$

Beweis. Es gilt $\langle v+w, v+w\rangle - \langle v-w, v-w\rangle = 2(\langle v,w\rangle + \langle w,v\rangle)$ für alle $v, w \in V$. Im Fall $\mathbb{K} = \mathbb{R}$, folgt hieraus schon die Behauptung. Im Fall $\mathbb{K} = \mathbb{C}$ folgt hieraus $-\mathring{i}(\langle v+\mathring{i}w, v+\mathring{i}w\rangle - \langle v-\mathring{i}w, v-\mathring{i}w\rangle) = -2\mathring{i}(\langle v,\mathring{i}w\rangle + \langle \mathring{i}w, v\rangle) = 2(\langle v,w\rangle - \langle w,v\rangle)$ für alle $v, w \in V$ und man braucht dies nur zur obigen Gleichung addieren. □

Definition und Proposition 11.1.11. Seien V ein $\mathbb{R}$-Vektorraum mit Skalarprodukt und $v, w \in V \setminus \{0\}$. Dann existiert eine eindeutig bestimmte Zahl $\alpha \in \mathbb{R}$ mit $0 \leq \alpha \leq \pi$ und

$$\cos\alpha = \frac{\langle v,w\rangle}{\|v\|\|w\|}.$$

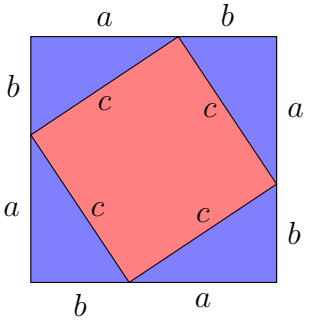
Diese Zahl nennt man den *Winkel* $\measuredangle(v,w)$ zwischen v und w.

Beweis. Aus der Analysis weiß man, dass $[0,\pi] \to [-1,1],\ \alpha \mapsto \cos(\alpha)$ eine Bijektion ist (die Injektivität folgt daraus, dass diese Funktion streng monoton fällt, und die Surjektivität aus dem Zwischenwertsatz). Aus der Injektivität dieser Funktion folgt die Eindeutigkeit von α. Aus der Surjektivität folgt die Existenz von α, denn nach Cauchy-Schwarz 11.1.8 gilt

$$-1 \leq \frac{\langle v,w\rangle}{\|v\|\|w\|} \leq 1.$$

□

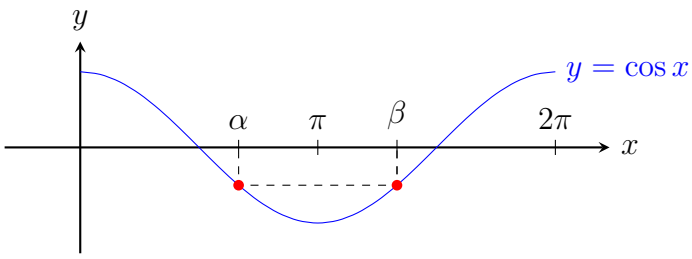
Bemerkung 11.1.12. (a) Die durch das Standardskalarprodukt auf dem $\mathbb{R}^2$ [→11.1.3(a)] induzierte Norm $\mathbb{R}^2 \to \mathbb{R},\ \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \mapsto \sqrt{a^2+b^2}$ [→11.1.9] gibt gerade die anschauliche Länge eines Vektors im $\mathbb{R}^2$ wieder, wie man leicht sieht, indem man in folgender Zeichnung den Flächeninhalt des großen Quadrats auf zwei verschiedene Weisen berechnet:



$$(a+b)^2 = c^2 + 4\tfrac{ab}{2}$$
$$a^2 + b^2 = c^2$$

Diesen Sachverhalt bezeichnet man als den Satz von Pythagoras [Pythagoras von Samos $*\approx$-570 $\dagger\approx$-510]. Obiges Bild stellt einen „geometrischen Beweis“ dar. Es handelt sich um keinen Beweis in unserem Sinne, denn es wird dort anschaulich argumentiert. Es wäre befriedigender, einige Axiome aufzustellen, die noch unstreitbarer unsere geometrische Anschauung widerspiegeln und den Satz von Pythagoras dann formal aus diesen Axiomen abzuleiten. Letztlich kann man aber ohnehin nicht vermeiden, gewisse geometrische Tatsachen einfach als gegeben vorauszusetzen (genauer gesagt axiomatisch einzuführen). Man könnte da weiter unten ansetzen, aber das würde viel Zeit kosten und es stellt sich die Frage, warum man es machen sollte. In unserem Rahmen ist daher das obige Bild kein formaler Beweis, aber ein „Argument“, das den Leser überzeugen soll, dass er sich unter der Norm eines Punktes im $\mathbb{R}^2$ (mit Standardskalarprodukt) den anschaulichen Abstand zum Nullpunkt vorstellen soll.

(b) Definition 11.1.11 von Winkeln stimmt überein mit dem anschaulichen Winkelbegriff im $\mathbb{R}^2$ (ausgestattet mit dem Standardskalarprodukt). Wie in (a) argumentieren wir wieder anschaulich: Wegen (a) verändert die Drehung R_φ ($\varphi \in \mathbb{R}$) [→6.3.2(a)] die Norm eines Vektors nicht, woraus mit der Polarisationsformel 11.1.10 folgt, dass R_φ auch das Skalarprodukt und damit Winkel im Sinne von Definition 11.1.11 nicht verändert. Es reicht also $\measuredangle(v,w)$ für den Fall zu betrachten, dass einer der beiden Vektoren $v, w \in \mathbb{R}^2\backslash\{0\}$ auf der ersten Koordinatenachse liegt (sonst drehe geeignet). Weiter kann man $\|v\| = \|w\| = 1$ annehmen, also $v = \binom{1}{0}$ und $w = \binom{\cos\beta}{\sin\beta}$ mit $\beta \in [0, 2\pi)$. Dann gilt für $\alpha := \measuredangle(v,w)$, dass $\cos\alpha = \langle \binom{1}{0}, \binom{\cos\beta}{\sin\beta} \rangle = \cos\beta$. Also gilt $\alpha = \beta$ falls $0 \le \beta \le \pi$ und $\alpha = \pi - (\beta - \pi) = 2\pi - \beta$ falls $\pi \le \beta \le 2\pi$, das heißt α ist der anschauliche Winkel zwischen v und w.



(c) Die durch das Standardskalarprodukt auf dem $\mathbb{R}^3$ [→11.1.3(a)] induzierte Norm

$\mathbb{R}^3 \to \mathbb{R}$, $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \sqrt{a^2+b^2+c^2}$ [→11.1.9] gibt gerade die anschauliche Länge eines Vektors im $\mathbb{R}^2$ wieder, denn nach Pythagoras aus Teil (a) ist die Länge von $\begin{pmatrix} a \\ b \\ 0 \end{pmatrix}$ gleich $\sqrt{a^2+b^2}$ und wieder nach Pythagoras ist die Länge von $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ daher $\sqrt{(\sqrt{a^2+b^2})^2+c^2} = \sqrt{a^2+b^2+c^2}$ (male ein Bild!).

(d) Durch Drehung im $\mathbb{R}^3$ sieht man nun wie in (b), dass auch im $\mathbb{R}^3$ der von uns in 11.1.11 eingeführte Winkelbegriff mit dem üblichen übereinstimmt.

(e) Sind $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$, so gilt im $\mathbb{C}^n$

$$\left\| \begin{pmatrix} a_1 + \mathring{i}b_1 \\ \vdots \\ a_n + \mathring{i}b_n \end{pmatrix} \right\| = \sqrt{(a_1 - \mathring{i}b_1)(a_1 + \mathring{i}b_1) + \cdots + (a_n - \mathring{i}b_n)(a_n + \mathring{i}b_n)}$$
$$= \sqrt{a_1^2 + b_1^2 + \cdots + a_n^2 + b_n^2}.$$

§11.2 Orthogonalität

Definition 11.2.1. Seien V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt und $v, w \in V$. Es heißen v und w *orthogonal* oder *senkrecht zueinander* (in Zeichen: $v \perp w$), wenn $\langle v, w\rangle = 0$.

Bemerkung 11.2.2. Seien V ein $\mathbb{R}$-Vektorraum mit Skalarprodukt und $v, w \in V$. Dann $v \perp w \iff \measuredangle(v,w) = \frac{\pi}{2}$ nach der Definition von Winkeln 11.1.11, denn $\cos\frac{\pi}{2} = 0$. Insbesondere stimmt unsere Definition von Senkrechtstehen in $\mathbb{R}^2$ und $\mathbb{R}^3$ nach 11.1.12 mit unserer geometrischen Anschauung überein.

Satz 11.2.3 (Satz von Pythagoras)**.** *Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt und seien $v, w \in V$ mit $v \perp w$. Dann $\|v\|^2 + \|w\|^2 = \|v+w\|^2$* [→11.1.9].

Beweis. $\|v+w\|^2 = \langle v+w, v+w\rangle = \langle v,v\rangle + \underbrace{\langle v,w\rangle}_{=0} + \underbrace{\langle w,v\rangle}_{=\langle v,w\rangle^*=0} + \langle w,w\rangle = \|v\|^2 + \|w\|^2$. □

Bemerkung 11.2.4. Die Kürze des Beweises des Satzes von Pythagoras in unserem Rahmen, zeigt in eindrucksvoller Weise, wie einfach man geometrische Sachverhalte mit Hilfe von Skalarprodukten erklären kann. Tatsächlich haben wir aber die Gültigkeit des Satzes von Pythagoras schon mehr oder weniger in den Begriff des Skalarprodukts hineinkodiert (und dies in 11.1.12 gerechtfertigt). Man könnte daher sagen, dass der *eigentliche* Beweis des Satzes von Pythagoras in 11.1.12(a) steht. Es stellt aber 11.1.12(a) nur die Verbindung zur Anschauung her. Da wir innerhalb unseres Formalismus niemals anschaulich argumentieren, sondern die Anschauung „nur“ als Inspiration zum Auffinden formaler Beweise benutzen, wird hier an keiner Stelle gemogelt.

Definition 11.2.5. [→6.2.1] Sei V ein Vektorraum mit Skalarprodukt. Seien $m \in \mathbb{N}_0$ und $v_1, \ldots, v_m \in V$. Dann heißt $(v_1, \ldots, v_m)$ ein *Orthonormalsystem (ONS)* (in V), wenn $v_i \perp v_j$ für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j$ und $\|v_i\| = 1$ für alle $i \in \{1, \ldots, m\}$. Ein ONS, welches V aufspannt, heißt *Orthonormalbasis (ONB)* von V.

Beispiel 11.2.6. Die Standardbasis des $\mathbb{K}^n$ [→6.2.2] ist eine ONB des $\mathbb{K}^n$ (versehen mit dem Standardskalarprodukt [→11.1.3(a)]).

Proposition 11.2.7. *Sei V ein Vektorraum mit Skalarprodukt und $(v_1, \ldots, v_m)$ ein ONS in V. Seien $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$ und $v := \sum_{i=1}^m \lambda_i v_i$. Dann $\lambda_i = \langle v_i, v \rangle$ für alle $i \in \{1, \ldots, m\}$.*

Beweis. $\langle v_j, \sum_{i=1}^m \lambda_i v_i \rangle = \sum_{i=1}^m \lambda_i \langle v_j, v_i \rangle = \lambda_j \underbrace{\langle v_j, v_j \rangle}_{=\|v_j\|^2=1} = \lambda_j$ für alle $j \in \{1, \ldots, m\}$ □

Korollar 11.2.8. *Sei V ein Vektorraum mit Skalarprodukt. In V ist jedes ONS linear unabhängig. Insbesondere ist jede ONB von V eine Basis von V.*

Definition und Proposition 11.2.9. Sei V ein Vektorraum mit Skalarprodukt und U ein Unterraum von V. Das *orthogonale Komplement* von U in V ist definiert durch

$$U^\perp := \{v \in V \mid \forall u \in U : v \perp u\}$$

und ist selber wieder ein Unterraum von V mit $U \cap U^\perp = \{0\}$ und $U \subseteq (U^\perp)^\perp$. Ist $U = \operatorname{span}(E)$ für ein $E \subseteq V$, so gilt $U^\perp = \{v \in V \mid \forall u \in E : v \perp u\}$.

Beweis. Sehr einfache Übung. □

Definition und Proposition 11.2.10. Seien V ein Vektorraum mit Skalarprodukt, U ein Unterraum von V und $v, w \in V$. Dann gibt es höchstens ein $w \in U$ mit $v - w \in U^\perp$. Falls existent, nennt man dieses w die *orthogonale Projektion* von v auf U.

Beweis. Seien $w, w' \in U$ mit $v - w, v - w' \in U^\perp$. Dann

$$\langle w - w', w - w' \rangle = \langle w - w', (v - w') - (v - w) \rangle = \langle \underbrace{w - w'}_{\in U}, \underbrace{v - w'}_{\in U^\perp} \rangle - \langle \underbrace{w - w'}_{\in U}, \underbrace{v - w}_{\in U^\perp} \rangle = 0.$$

□

Beispiel 11.2.11. (a) Sei V ein Vektorraum mit Skalarprodukt und U ein Unterraum von V. Ist $v \in U$, so ist v selber die orthogonale Projektion von v auf U. Ist $v \in U^\perp$, so ist sie der Nullvektor.

(b) Betrachte den $\mathbb{R}$-Vektorraum $\mathbb{R}^{\mathbb{N}}$ der reellen Folgen [→6.1.5, 7.1.5] und darin den Unterraum $V := \{f \mid f\colon \mathbb{N} \to \mathbb{R}, \exists c \in \mathbb{R} : \forall n \in \mathbb{N} : |f(n)| \leq c\}$ der *beschränkten* Folgen sowie den Unterraum U der Folgen mit endlichem Träger [→6.2.8(f)]. Dann ist V vermöge $\langle f, g \rangle := \sum_{i=1}^\infty \frac{1}{2^i} f(i) g(i)$ $(f, g \in V)$ ein Vektorraum mit Skalarprodukt und U ein Unterraum von V. Das orthogonale Komplement $U^\perp$ von U in V besteht offenbar nur aus der Nullfolge. Ist $f \in V$, so existiert die orthogonale Projektion von f auf U also genau dann, wenn $f \in U$ (und in diesem Fall ist sie f).

Proposition 11.2.12. *Sei V ein Vektorraum mit Skalarprodukt, $(v_1, \dots, v_m)$ ein ONS in V, $U := \operatorname{span}(v_1, \dots, v_m)$ und $v \in V$. Dann ist $\sum_{i=1}^m \langle v_i, v\rangle v_i$ die orthogonale Projektion von v auf U. Insbesondere existiert diese.*

Beweis. $w := \sum_{i=1}^m \langle v_i, v\rangle v_i \in U$. Um $v - w \in U^\perp$ zu zeigen, reicht es $\langle v_j, v - w\rangle = 0$ für $j \in \{1, \dots, m\}$ zu zeigen. Sei also $j \in \{1, \dots, m\}$. Dann

$$\langle v_j, v - w\rangle = \langle v_j, v - \sum_{i=1}^m \langle v_i, v\rangle v_i\rangle = \langle v_j, v\rangle - \sum_{i=1}^m \langle v_i, v\rangle \underbrace{\langle v_j, v_i\rangle}_{\in\{0,1\}} = \langle v_j, v\rangle - \langle v_j, v\rangle = 0.$$

□

Korollar 11.2.13. *Sei V ein Vektorraum mit Skalarprodukt und ONB $\underline{v} = (v_1, \dots, v_n)$. Dann gilt*

$$\operatorname{coord}_{\underline{v}}(v) = \begin{pmatrix} \langle v_1, v\rangle \\ \vdots \\ \langle v_n, w\rangle \end{pmatrix} \qquad \textit{für jedes } v \in V.$$

Beweis. Die orthogonale Projektion eines $v \in V$ auf V ist v selber [→11.2.11(a)]. Also gilt nach 11.2.12 $v = \sum_{i=1}^n \langle v_i, v\rangle v_i$ für alle $v \in V$. □

Proposition 11.2.14 (Gram-Schmidtsches Orthogonalisierungsverfahren)**.** *Sei V ein Vektorraum mit Skalarprodukt, $(v_1, \dots, v_m)$ ein ONS in V, $U := \operatorname{span}(v_1, \dots, v_m)$ und $v \notin U$. Sei dann $w := \sum_{i=1}^m \langle v_i, v\rangle v_i$ die orthogonale Projektion von v auf U* [→11.2.12]*. Dann ist $(v_1, \dots, v_m, \frac{v-w}{\|v-w\|})$ ein ONS und es gilt*

$$\operatorname{span}(v_1, \dots, v_m, v) = \operatorname{span}\left(v_1, \dots, v_m, \frac{v - w}{\|v - w\|}\right).$$

Beweis. Sehr einfache Übung. □

Satz 11.2.15. [→6.2.18] *Jeder endlichdimensionale Vektorraum mit Skalarprodukt besitzt eine ONB.*

Beweis. Induktion nach der Dimension, wobei der Induktionsschritt mit Gram-Schmidt 11.2.14 bewerkstelligt wird. □

Korollar 11.2.16. *Sei V ein Vektorraum mit Skalarprodukt und U ein endlichdimensionaler Unterraum von V. Dann gibt es für jedes $v \in V$ die orthogonale Projektion $P_U(v)$ von v auf U und dadurch wird eine lineare Abbildung $P_U \colon V \to V$ definiert, deren Kern $U^\perp$ und deren Bild U ist.*

Beweis. Wähle mit 11.2.15 eine ONB $(v_1, \dots, v_m)$ von U. Nach 11.2.12 haben wir dann $P_U(v) = \sum_{i=1}^m \langle v_i, v\rangle v_i$ für alle $v \in U$. Mit der Linearität des Skalarprodukts im zweiten Argument 11.1.1(3),(4) rechnet man nun sofort die Linearität von P_U nach. Weiter gilt

$$\ker P_U = \left\{ v \in V \mid \sum_{i=1}^m \langle v_i, v\rangle v_i = 0 \right\} \overset{11.2.8}{=} \{v \in V \mid \langle v_1, v\rangle = \ldots = \langle v_m, v\rangle = 0\}$$
$$\overset{11.2.9}{=} (\operatorname{span}(v_1, \dots, v_m))^\perp = U^\perp,$$

$U \subseteq \operatorname{im} P_U$ nach 11.2.11(a) und selbstverständlich $\operatorname{im} P_U \subseteq U$ nach 11.2.10. □

Proposition 11.2.17. *Sei U ein endlichdimensionaler Unterraum des Vektorraums mit Skalarprodukt V. Dann gilt $V = U \oplus U^\perp$ [→8.2.1]. Für endlichdimensionales V gilt insbesondere* $\dim(U) + \dim(U^\perp) = \dim(V)$ *[→8.2.3].*

Beweis. Für jedes $v \in V$ gilt $v = \underbrace{P_U(v)}_{\in U} + \underbrace{(v - P_U(v))}_{\in U^\perp}$, also $V = U + U^\perp$. Weiter ist die lineare Abbildung $U \times U^\perp \to U + U^\perp$, $(u, v) \mapsto u + v$ injektiv, denn ist $(u, v) \in U \times U^\perp$ mit $u + v = 0$, so folgt $u = -v \in U \cap U^\perp \overset{11.2.9}{=} \{0\}$. □

Korollar 11.2.18. *Sei U ein Unterraum des endlichdimensionalen $\mathbb{K}$-Vektorraums mit Skalarprodukt V. Dann $U = (U^\perp)^\perp$.*

Beweis. Nach 11.2.9 gilt $U \subseteq (U^\perp)^\perp$ und mit 11.2.17 angewandt auf U und auf $U^\perp$ haben wir $\dim(U) = \dim(V) - \dim(U^\perp) = \dim((U^\perp)^\perp)$. Benutze nun 6.2.27. □

Definition 11.2.19. Seien V und W Vektorräume mit Skalarprodukt. Dann heißt eine Abbildung $f\colon V \to W$ ein *Homomorphismus von Vektorräumen mit Skalarprodukt* (auch *orthogonal* oder *unitär*, ersteres vorwiegend im Fall $\mathbb{K} = \mathbb{R}$ und letzteres vorwiegend im $\mathbb{K} = \mathbb{C}$), wenn f linear ist und für alle $v, w \in V$ gilt $\langle f(v), f(w)\rangle = \langle v, w\rangle$. Ist sie zusätzlich bijektiv, so heißt sie *Isomorphismus von Vektorräumen mit Skalarprodukt* [beachte, dass die Injektivität automatisch erfüllt ist, da aus $f(v) = 0$ folgt $\langle v, v\rangle = \langle f(v), f(v)\rangle = 0$ und daher $v = 0$].

Bemerkung 11.2.20. Aus der Polarisationsformel 11.1.10 folgt, dass man in dieser Definition die Bedingung $\forall v, w \in V : \langle f(v), f(w)\rangle = \langle v, w\rangle$ ersetzen kann durch

$$\forall v \in V : \|f(v)\| = \|v\|.$$

Beispiel 11.2.21. Wie in 11.1.12 bereits bemerkt ist $R_\varphi \in \mathrm{End}(\mathbb{R}^2)$ für jedes $\varphi \in \mathbb{R}$ ein *Automorphismus des $\mathbb{R}^2$ mit Standardskalarprodukt.*

Satz 11.2.22. *Seien V und W $\mathbb{K}$-Vektorräume mit Skalarprodukt und sei $\underline{v} = (v_1, \ldots, v_n)$ eine ONB von V. Sei $f\colon V \to W$ linear. Dann ist f genau dann ein Isomorphismus von Vektorräumen mit Skalarprodukt, wenn $(f(v_1), \ldots, f(v_n))$ eine ONB von W ist.*

Beweis. Die eine Richtung ist klar. Für die andere sei $(f(v_1), \ldots, f(v_n))$ eine ONB von W. Nach 6.3.8 ist f ein Isomorphismus von Vektorräumen. Seien nun $v, w \in V$, etwa $v = \sum_{i=1}^n \lambda_i v_i$ und $w = \sum_{i=1}^n \mu_i v_i$ mit $\lambda_i, \mu_i \in \mathbb{K}$. Zu zeigen ist $\langle f(v), f(w)\rangle = \langle v, w\rangle$. Es gilt

$$\begin{aligned}\langle f(v), f(w)\rangle &= \left\langle \sum_{i=1}^n \lambda_i f(v_i), \sum_{j=1}^n \mu_j f(v_j) \right\rangle = \sum_{i,j=1}^n \lambda_i^* \mu_j \langle f(v_i), f(v_j)\rangle \\ &= \sum_{i=1}^n \lambda_i^* \mu_i = \sum_{i,j=1}^n \lambda_i^* \mu_j \langle v_i, v_j\rangle = \left\langle \sum_{i=1}^n \lambda_i v_i, \sum_{j=1}^n \mu_j v_j \right\rangle = \langle v, w\rangle.\end{aligned}$$

□

Korollar 11.2.23. *Sei $n \in \mathbb{N}_0$. Je zwei n-dimensionale $\mathbb{K}$-Vektorräume mit Skalarprodukt sind als solche isomorph.*

Beweis. Seien V und W n-dimensionale $\mathbb{K}$-Vektorräume mit Skalarprodukt. Wähle Orthonormalbasis $(v_1, \ldots, v_n)$ von V und $(w_1, \ldots, w_n)$ von W. Dann ist die lineare Abbildung $f\colon V \to W$ mit $f(v_i) = w_i$ für $i \in \{1, \ldots, n\}$ [→6.3.4] ein Isomorphismus von Vektorräumen mit Skalarprodukt. □

Korollar 11.2.24. *Sei V ein Vektorraum mit Skalarprodukt und $\underline{v} = (v_1, \ldots, v_n)$ eine Basis von V. Dann sind äquivalent:*

(a) *$\underline{v}$ ist ONB von V,*

(b) *$\mathrm{vec}_{\underline{v}}\colon \mathbb{K}^n \to V$ ist ein Isomorphismus von Vektorräumen mit Skalarprodukt.*

(c) *$\mathrm{coord}_{\underline{v}}\colon V \to \mathbb{K}^n$ ist ein Isomorphismus von Vektorräumen mit Skalarprodukt.*

Definition 11.2.25. Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *orthogonal* (vor allem wenn $\mathbb{K} = \mathbb{C}$ manchmal auch *unitär*), wenn f_A ein Isomorphismus des Vektorraums $\mathbb{K}^n$ mit dem Standardskalarprodukt ist.

Satz 11.2.26. *Seien V und W Vektorräume mit Skalarprodukt und Orthonormalbasis $\underline{v} = (v_1, \ldots, v_n)$ und $\underline{w} = (w_1, \ldots, w_n)$. Sei $f\colon V \to W$ linear. Dann gilt:*

f ist Isomorphismus von Vektorräumen mit Skalarprodukt $\iff$ $M(f, \underline{v}, \underline{w})$ orthogonal.

Beweis. Nach 7.1.1 gilt $f = \mathrm{vec}_{\underline{w}} \circ f_{M(f,\underline{v},\underline{w})} \circ \mathrm{coord}_{\underline{v}}$ und daher $f_{M(f,\underline{v},\underline{w})} = \mathrm{coord}_{\underline{w}} \circ f \circ \mathrm{vec}_{\underline{v}}$. Da $\mathrm{vec}_{\underline{w}}$, $\mathrm{coord}_{\underline{v}}$, $\mathrm{coord}_{\underline{w}}$ und $\mathrm{vec}_{\underline{v}}$ nach 11.2.24 Isomorphismen von Vektorräumen mit Skalarprodukt sind, ist f ein solcher genau dann, wenn $f_{M(f,\underline{v},\underline{w})}$ einer ist. □

Satz 11.2.27. *Sei $A \in \mathbb{K}^{n\times n}$. Dann sind äquivalent:*

(a) *A ist orthogonal.*

(b) *Die Spalten von A bilden eine ONB des $\mathbb{K}^n$.*

(c) *Die Zeilen von A bilden eine ONB des $\mathbb{K}^n$.*

(d) $A^*A = I_n$

(e) $AA^* = I_n$

(f) *A ist invertierbar mit $A^{-1} = A^*$.*

Beweis. Aus 11.2.25, 11.2.6 und 11.2.22 folgt (a) $\iff$ (b), da $f_A(e_1), \ldots, f_A(e_n)$ die Spalten von A sind. Direkt aus der Definition der Matrizenmultiplikation 7.2.1 folgen (b) $\iff$ (d) und (c) $\iff$ (e). Schließlich gilt (d) $\iff$ (e) $\iff$ (f) wegen 7.2.13. □

Beispiel 11.2.28. [→7.1.4(a)] $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ ist für jedes $\varphi \in \mathbb{R}$ orthogonal.

§11.3 Diagonalisierung symmetrischer und hermitescher Matrizen

Definition 11.3.1. Sei V ein Vektorraum mit Skalarprodukt und $f \in \mathrm{End}(V)$. Dann heißt f *selbstadjungiert* (auch *symmetrisch* oder *hermitesch*, ersteres vorwiegend im Fall $\mathbb{K} = \mathbb{R}$ und letzteres vorwiegend im $\mathbb{K} = \mathbb{C}$), wenn

$$\langle f(v), w\rangle = \langle v, f(w)\rangle \qquad \text{für alle } v, w \in V.$$

Beispiel 11.3.2. Die orthogonale Projektion $P_U \in \mathrm{End}(V)$ auf einen endlichdimensionalen Unterraum U eines Vektorraums mit Skalarprodukt V [→11.2.16] ist selbstadjungiert. In der Tat: Wegen $V = U + U^\perp$ [→8.2.1] reicht es zu beobachten, dass für alle $u_1, u_2 \in U$ und $v_1, v_2 \in U^\perp$ gilt $\langle P_U(u_1 + v_1), u_2 + v_2\rangle = \langle u_1, u_2 + v_2\rangle = \langle u_1, u_2\rangle$ und $\langle u_1 + v_1, P_U(u_2 + v_2)\rangle = \langle u_1 + v_1, u_2\rangle = \langle u_1, u_2\rangle$.

Definition 11.3.3. Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *selbstadjungiert* (im Fall $\mathbb{K} = \mathbb{R}$ auch *symmetrisch*, im Fall $\mathbb{K} = \mathbb{C}$ auch *hermitesch*), wenn $f_A \in \mathrm{End}(\mathbb{K}^n)$ selbstadjungiert ist.

Satz 11.3.4. *Sei V ein Vektorraum mit Skalarprodukt und ONB $\underline{v} = (v_1, \ldots, v_n)$. Sei $f \in \mathrm{End}(V)$. Dann gilt: f selbstadjungiert $\iff$ $M(f, \underline{v})$ selbstadjungiert.*

Beweis. Es gilt $f = \mathrm{vec}_{\underline{v}} \circ f_{M(f,\underline{v})} \circ \mathrm{coord}_{\underline{v}}$ [→7.1.1] und daher $f \circ \mathrm{vec}_{\underline{v}} = \mathrm{vec}_{\underline{v}} \circ f_{M(f,\underline{v})}$. Da $\underline{v}$ eine ONB ist, ist $\mathrm{vec}_{\underline{v}}$ nach 11.2.24 ein Isomorphismus von Vektorräumen mit Skalarprodukt. Es gilt daher

$$\begin{aligned}
&f \text{ selbstadjungiert}\\
\iff &\forall v, w \in V\colon \langle f(v), w\rangle = \langle v, f(w)\rangle\\
\iff &\forall x, y \in \mathbb{K}^n\colon \langle f(\mathrm{vec}_{\underline{v}}(x)), \mathrm{vec}_{\underline{v}}(y)\rangle = \langle \mathrm{vec}_{\underline{v}}(x), f(\mathrm{vec}_{\underline{v}}(y))\rangle\\
\iff &\forall x, y \in \mathbb{K}^n\colon \langle \mathrm{vec}_{\underline{v}}(M(f,\underline{v})x), \mathrm{vec}_{\underline{v}}(y)\rangle = \langle \mathrm{vec}_{\underline{v}}(x), \mathrm{vec}_{\underline{v}}(M(f,\underline{v})y)\rangle\\
\iff &\forall x, y \in \mathbb{K}^n\colon \langle M(f,\underline{v})x, y\rangle = \langle x, M(f,\underline{v})y\rangle\\
\iff &M(f,\underline{v}) \text{ selbstadjungiert}
\end{aligned}$$

□

Proposition 11.3.5. *Seien K ein kommutativer Ring, $m, n, r \in \mathbb{N}_0$, $A \in K^{m\times n}$ und $B \in K^{n\times r}$. Dann gilt:*

(a) $(AB)^T = B^T A^T$

(b) $(AB)^* = B^* A^*$ *falls* $K = \mathbb{K}$

Beweis. (a) Schreibt man $A = (a_{ij})_{1\le i\le m, 1\le j\le n}$, $B = (b_{jk})_{1\le j\le n, 1\le k\le r}$, so gilt

$$(AB)^T \overset{\substack{7.2.1\\ 9.1.21}}{=} \left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)_{1\le k\le r, 1\le i\le m} = \left(\sum_{j=1}^{n} b_{jk}a_{ij}\right)_{1\le k\le r, 1\le i\le m} \overset{\substack{7.2.1\\ 9.1.21}}{=} B^T A^T.$$

(b) Ist $K = \mathbb{K}$, so gilt

$$(AB)^* \overset{7.2.1}{\underset{9.1.21}{=}} \left(\sum_{j=1}^{n}(a_{ij}b_{jk})^*\right)_{1\le k\le r, 1\le i\le m} \overset{4.2.7}{=} \left(\sum_{j=1}^{n} b_{jk}^* a_{ij}^*\right)_{1\le k\le r, 1\le i\le m} \overset{7.2.1}{\underset{9.1.21}{=}} B^*A^*.$$

□

Lemma 11.3.6. *Sei $A \in \mathbb{K}^{n\times n}$ und $x, y \in \mathbb{K}^n$. Dann gilt $\langle A^*x, y\rangle = \langle x, Ay\rangle$.*

Beweis. $\langle A^*x, y\rangle \overset{11.1.3(a)}{=} (A^*x)^*y \overset{11.3.5}{=} x^*(A^*)^*y = x^*Ay \overset{11.1.3(a)}{=} \langle x, Ay\rangle$ □

Proposition 11.3.7. *Für $A \in \mathbb{K}^{n\times n}$ gilt*

$$A \text{ selbstadjungiert} \iff A^* = A.$$

Beweis.

$$\begin{aligned} A \text{ selbstadjungiert} &\overset{\substack{11.3.3\\11.3.1}}{\iff} \forall x, y \in \mathbb{K}^n\colon \langle Ax, y\rangle = \langle x, Ay\rangle \\ &\overset{11.3.6}{\iff} \forall x, y \in \mathbb{K}^n\colon \langle Ax, y\rangle = \langle A^*x, y\rangle \\ &\iff A = A^* \end{aligned}$$

□

Lemma 11.3.8. *Sei $A \in \mathbb{K}^{n\times n}$ selbstadjungiert und $\lambda \in \mathbb{C}$ mit $\chi_A(\lambda) = 0$. Dann gilt $\lambda \in \mathbb{R}$.*

Beweis. Wir können $\mathbb{K} = \mathbb{C}$ annehmen. Dann ist λ ein Eigenwert von A [→10.1.2, 10.1.3(e)], das heißt es gibt $x \in \mathbb{C}^n \setminus \{0\}$ mit $Ax = \lambda x$. Es folgt

$$\lambda\langle x, x\rangle \overset{11.1.1(4)}{=} \langle x, \lambda x\rangle = \langle x, Ax\rangle \overset{\substack{11.3.3\\11.3.1}}{=} \langle Ax, x\rangle \overset{11.1.1(2)}{=} \langle \lambda x, x\rangle = \lambda^*\langle x, x\rangle$$

und daher $\lambda = \lambda^*$ nach 11.1.1(6). □

Satz 11.3.9. [2] *Sei V ein endlichdimensionaler Vektorraum mit Skalarprodukt und f ein selbstadjungierter Endomorphismus von V. Dann gibt es eine ONB von V, die aus Eigenvektoren von f zu reellen Eigenwerten besteht. Insbesondere ist f diagonalisierbar* [→10.3.3(b)].

Beweis. Induktion nach $n := \dim V$.
$\underline{n = 0}$ nichts zu zeigen
$\underline{n - 1 \to n\ (n \in \mathbb{N})}$ Wir zeigen zunächst mit Hilfe von 10.1.7, dass f einen *reellen* Eigenwert λ besitzt: Wählt man nämlich eine ONB $\underline{w}$ von V [→11.2.15] und setzt $A := M(f, \underline{w})$, so ist A nach 11.3.4 selbstadjungiert und daher jedes $\lambda \in \mathbb{C}$ mit $\chi_A(\lambda) = 0$

[2] Im Beweis dieses Satzes benutzen wir den Fundamentalsatz der Algebra 4.2.12 [→4.2.13(g)].

nach Lemma 11.3.8 sogar aus $\mathbb{R}$. Nun gilt aber $\chi_f = \chi_A$ [→10.1.5, 10.1.9(e)] und es gibt ein solches λ wegen $\deg \chi_f \overset{10.1.6}{=} n \geq 1$ nach dem Fundamentalsatz der Algebra 4.2.12.

Wir wählen nun einen Eigenvektor $u \in V$ zu diesem Eigenwert $\lambda \in \mathbb{R}$. Setze $U := \operatorname{span}(u)$. Da f selbstadjungiert ist, gilt $f(U^\perp) \subseteq U^\perp$, denn ist $v \in U^\perp$, so gilt $\langle f(v), u\rangle = \langle v, f(u)\rangle = \lambda\langle v, u\rangle = 0$. Nun ist $f|_{U^\perp} \colon U^\perp \to U^\perp, v \mapsto f(v)$ ein selbstadjungierter Endomorphismus des Vektorraums mit Skalarprodukt $U^\perp$. Es gilt $1 + \dim(U^\perp) = \dim(U) + \dim(U^\perp) \overset{11.2.17}{=} \dim(V) = n$, also $\dim(U^\perp) = n - 1$. Nach Induktionsvoraussetzung gibt es eine ONB $(v_2, \ldots, v_n)$ von $U^\perp$, die aus Eigenvektoren von f zu reellen Eigenwerten besteht. Setzt man $v_1 := \frac{u}{\|u\|}$, so erhält man eine ONB $(v_1, \ldots, v_n)$ von V, die aus Eigenvektoren von f zu reellen Eigenwerten besteht. □

Korollar 11.3.10. [3] *Sei $A \in \mathbb{K}^{n\times n}$ selbstadjungiert. Dann gibt es eine* reelle *Diagonalmatrix $D \in \mathbb{R}^{n\times n}$ und eine orthogonale Matrix $P \in \mathbb{K}^{n\times n}$ mit $A = P^*DP$. Insbesondere ist A diagonalisierbar.*

Beweis. Wähle mit Satz 11.3.9 eine ONB $\underline{v}$ von $\mathbb{K}^n$, die aus Eigenvektoren von A zu reellen Eigenwerten besteht. Nach Satz 11.2.26 ist $P := M(\underline{e}, \underline{v})$ [→7.1.10] dann orthogonal und daher

$$P^* \overset{11.2.27(f)}{=} P^{-1} \overset{7.2.11}{=} M(\underline{v}, \underline{e}).$$

Also

$$A \overset{7.1.4(e)}{=} M(f_A, \underline{e}) \overset{7.2.5}{=} M(\underline{v}, \underline{e})M(f_A, \underline{v})M(\underline{e}, \underline{v}) = P^*DP,$$

wobei $D := M(f_A, \underline{v})$ eine reelle Diagonalmatrix ist. □

[3] Im Beweis dieses Korollars benutzen wir den Fundamentalsatz der Algebra 4.2.12 [→4.2.13(g)].

Index